

BIOMOLECULAR SIGNATURES OF DISEASE VIA ION MOBILITY AND
MASS SPECTROMETRY TECHNIQUES

By

Kelly M. Hines

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, 2014

Nashville, Tennessee

Approved:

John A. McLean, Ph.D.

David E. Cliffel, Ph.D.

Donna J. Webb, Ph.D.

Eva M. Harth, Ph.D.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my dissertation advisor Dr. John A. McLean for allowing me to explore my own interests, and for providing encouragement and direction in times when I was uncertain in my abilities or the merits of my research. The advice to not “let perfection get in the way of good enough,” has been particularly helpful. Thank you for being an excellent mentor.

I would like to thank my dissertation committee members Dr. David E. Cliffler, Dr. Donna J. Webb, and Dr. Eva M. Harth for their guidance and advice, and for challenging me to see the broader impacts of my work.

I would like to acknowledge the many collaborators who have contributed to my dissertation research: Dr. John P. Wikswo, Dr. Dana R. Marshall, Dr. Donna J. Webb, Dr. Kristie L. Rose, Dr. Jeffrey M. Davidson, and Dr. Susan R. Opalenik. I learned many new things through our work together.

I would like to thank my colleagues, past and present, in the McLean lab for being good co-workers and good friends. In particular, I would like to thank Dr. Jeffrey R. Enders, Dr. Michal Kliman and Dr. Cody Goodwin for their help, advice and patience. Special thanks to Dr. Jay G. Forsythe for being a great friend in and out of lab.

Many thanks to all my friends and family for their love and support. I would have never made it to this point without you. Thank you to my parents and siblings for always being there to listen, no matter the time of day or night, for their continual encouragement and advice, and for reminding me everything will work out when I couldn't see it for myself. To my best friends Sarah Moravia, Amelia Null, and Carlee Murphy: thank you

for befriending me in French class and still being here all these years later (Go Gators!). Finally, thanks to Bert Stone for being my rock, and for all the happiness and love you have brought me.

Lastly, I would like to acknowledge the entities which provided financial support for this work: the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institute of Chemical Biology, the Vanderbilt Institute of Integrative Biosystems Research and Education, the U.S. Defense Threat Reduction Agency (HDTRA-09-1-0013 and DTRA100271 A-5196), and the National Institutes of Health (R01GM092218, RC2DA028981 and UH2TR000491).

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS.....	x
Chapter	
I. MULTIDIMENSIONAL SEPARATIONS BY ION-MOBILITY- MASS SPECTROMETRY.....	1
1.1. IM-MS: Multidimensional Gas Phase Separations.....	1
1.1.1. The Evolution of IM Technology	2
1.1.2. The Correlation of IM and MS Dimensions	7
1.1.2.1. Tailoring the Configuration of Multidimensional Separations	9
1.1.2.2. Advantages for Complex Biological Sample Analysis.....	12
1.1.3. Structural Information from 2D Conformation Space	14
1.1.3.1. Relating Mobility Drift Time to CCS	15
1.1.3.2. Correlating Empirical and Computational CCSs.....	17
1.2. Instrumental Configurations of Modern IM-Mass Spectrometers.....	18
1.2.1. Impact of Ion Source Selection on IM Arrangement.....	19
1.2.2. IM-Time-of-Flight Mass Spectrometers	21
1.2.2.1. MALDI-IM-TOFMS	21
1.2.2.2. ESI-IM-TOFMS	24
1.2.2.3. Traveling Wave IM-TOFMS.....	24
1.3. Applications of IM-MS.....	27
1.3.1. Conformation Space Analysis.....	27
1.3.1.1. Mapping Conformation Space of Biomolecules	28
1.3.1.2. Utilizing Conformation Space with IM shift reagents	32
1.3.2. Integrating Omics Analysis by IM-MS.....	37
1.3.2.1. Proteomics.....	38
1.3.2.2. Glycomics.....	42

1.3.2.3. Lipidomics.....	46
1.3.3. Structural Biology by IM-MS.....	49
1.4. Conclusions.....	53
1.5. Summary and Objectives of Dissertation Research.....	54
1.6. Acknowledgements.....	56
1.7. References.....	57
II. BIOMOLECULAR SIGNATURES OF DIABETIC WOUND HEALING BY STRUCTURAL MASS SPECTROMETRY	76
2.1. Introduction.....	76
2.2. Experimental.....	79
2.3. Results and Discussion	85
2.4. Conclusions.....	100
2.5. Acknowledgements.....	102
2.6. References.....	103
III. STRUCTURAL MASS SPECTROMETRY OF TISSUE EXTRACTS TO DISTINGUISH CANCEROUS AND NON-CANCEROUS BREAST DISEASES	108
3.1. Introduction.....	108
3.2. Experimental.....	112
3.3. Results and Discussion	118
3.4. Conclusions.....	132
3.5. Acknowledgements.....	133
3.6. References.....	134
IV. PHOSPHORYLATION OF SERINE 106 IN ASEF2 REGULATES CELL MIGRATION AND ADHESION TURNOVER.....	140
4.1. Introduction.....	140
4.2. Experimental.....	142
4.3. Results and Discussion	150
4.4. Conclusions.....	163
4.5. Acknowledgements.....	164
4.6. References.....	165

V.	PERSPECTIVES ON EMERGING AND FUTURE DIRECTIONS	169
5.1.	Summary	169
5.2.	Future Directions	171
	5.2.1. Biomolecular Signatures of Diabetic Wound Healing.....	171
	5.2.2. Distinguishing Cancerous and Noncancerous Breast Diseases	171
	5.2.3. Characterizing the Phosphorylation of Asef2	172
5.3.	Conclusions.....	173

Appendix

A.	References of Adaptation for Chapters.....	175
B.	Supporting Information for Chapter II.....	176
C.	Supporting Information for Chapter III	183
D.	Supporting Information for Chapter IV	232
E.	Curriculum Vitae	256

LIST OF TABLES

Table 3.1	Select Features from Figure 3.2 Loadings Plot.....	124
Table 4.1	Phosphorylation Sites of Asef2 Identified by MS Analysis	155

LIST OF FIGURES

Figure 1.1	Timeline of IM and MS advances.....	4
Figure 1.2	Trends in IM-MS Publications and Patents	6
Figure 1.3	Dimensions of IM-MS Conformation Space: Cytochrome C Digest	8
Figure 1.4	Arrangements of IM-MS Instruments.....	10
Figure 1.5	Depiction of Biomolecular Class Separation in Conformation Space.....	13
Figure 1.6	Schematics of MALDI-DTIM-TOFMS and ESI-TWIM-TOFMS.....	22
Figure 1.7	Conformation Space Occupied by Biomolecular Classes	29
Figure 1.8	Deviations of Biomolecule Sub-classes from Average Trendlines	31
Figure 1.9	Graphical Depiction of High vs. Low-Density IM Shift Reagents.....	34
Figure 1.10	Lanthanide Shift Reagents for Phosphopeptide Separations	36
Figure 1.11	Separation of Isobaric Glycan Isomers by IM-MS	44
Figure 1.12	Simultaneous Glycoproteomic IM-MS Analysis of RNaseB	47
Figure 1.13	Approaches for Generating Model Structures for CCS Comparison.....	51
Scheme 2.1	Work-Flow for IM-MS Analysis of Wound Fluids	81
Figure 2.1	IM-MS Spectra for Control and Diabetic Wound Fluids.....	87
Figure 2.2	IM-extraction MS/MS Spectrum of S100-A8 Fragmentation	90
Figure 2.3	PLS-DA Score Plots and OPLS-DA S-Plots	92
Figure 2.4	Post-Mobility MS/MS of m/z 544.4.....	94
Figure 2.5	IM-MS Spectra of m/z 355.3 and 373.3.....	96
Figure 2.6	UPLC-ESI-IM-MS Analysis of Diabetic Day 2 Wound Fluid.....	98
Figure 2.7	Post-Mobility MS/MS of Cholic Acid Standard.....	99

Scheme 3.1	Workflow for Preparation and Analysis of Breast Tissues.....	113
Figure 3.1	Extraction of Thymosins β 4 and β 10 in UPLC and IM Dimensions.....	119
Figure 3.2	Multivariate Statistical Analysis of UPLC-IM-MS/MS Data.....	122
Figure 3.3	Demonstration of IM Separation of Isobaric Species.....	126
Figure 3.4	Interpretation of Post-Mobility MS/MS Spectrum: m/z 613.16.....	127
Figure 4.1	Isolation and Purification of Asef2 Protein	151
Figure 4.2	Localization of Phosphorylated Residues in Asef2	153
Figure 4.3	MS/MS Spectrum Showing Confirmation of pS106 Site.....	156
Figure 4.4	Phosphorylation of S106 Stimulates Asef2 GEF Activity.....	158
Figure 4.5	Phosphorylation of S106 Critical for Asef2-promoted Cell Migration ...	160
Figure 4.6	S106 Phosphorylation Regulates Adhesion Turnover	162

LIST OF ABBREVIATIONS

2D	Two-dimensional
2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
ABR	APC-binding region
ACN	Acetonitrile
APC	Adenomatous polyposis coli
CCS	Collision cross section
CE	Capillary electrophoresis
CHO	Choline
CID	Collision induced dissociation
DH	Dbl homology
DTIM	Drift tube ion mobility
DTT	Dithiothreitol
ESI	Electrospray ionization
FAIMS	Field asymmetric ion mobility spectrometry
FWHM	Full width at half maximum
GAP	GTP-ase activating protein
GC	Gas chromatography
GEF	Guanine nucleotide exchange factor
GPC	Glycerophosphocholine
GSH	Glutathione, reduced
GSSG	Oxidized glutathione

GST	Glutathione-S-transferase
HPLC	High-performance liquid chromatography
IAM	Iodoacetamide
ICAT	Isotope-coded affinity tagging
IM	Ion mobility
IM-MS	Ion mobility-mass spectrometry
IMS	Ion mobility spectrometry
LC	Liquid chromatography
LPC	Lysophosphatidylcholine
MALDI	Matrix-assisted laser desorption/ionization
MEOH	Methanol
MM/MD	Molecular mechanics/molecular dynamics
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MudPIT	Multidimensional protein identification technology
m/z	Mass-to-charge
OPLS-DA	Orthogonal partial least-squares discriminant analysis
PAK	p21-activated kinase
PBD	PAK binding domain
PC	Phosphatidylcholine (class of lipids)
PC	Phosphocholine (lipid head group)
PC	Principal component (statistical descriptor)
PCA	Principal components analysis

PH	Pleckstrin homology
PLS-DA	Partial least-squares discriminant analysis
ppm	Parts-per-million
PS	Phosphatidylserine
PTM	Post-translational modification
SH3	Src homology 3
SID	Surface induced dissociation
SM	Sphingomyelin
S/N	Signal-to-noise
SSISP	Sequence-specific intrinsic size parameters
tCHO	Total choline
TOF	Time-of-flight
TWIM	Traveling wave ion mobility
UPLC	Ultra-performance liquid chromatography

CHAPTER I

MULTIDIMENSIONAL SEPARATIONS BY ION-MOBILITY- MASS SPECTROMETRY

1.1. Ion Mobility-Mass Spectrometry: Multidimensional Gas Phase Separations

Mass spectrometry (MS) is an increasingly powerful analytical technique with applications in fields ranging from the life sciences to nanotechnology. MS analyses are extraordinarily rapid and well-suited for large-scale studies requiring high-throughput. As a highly sensitive technique MS requires only pico to femtomole quantities of sample, making it a preferred method in situations where a limited amount of sample is available. Combined with the availability of MS databases and informatics, unknown species are often readily identified even from complex matrices. For the life sciences in particular, MS holds significant advantages in the analysis of complex samples, but is limited due to complications arising from isobaric (*i.e.* same mass) species resulting in endogenous or exogenous sources of noise. Such complications can be surmounted by combining MS with additional dimensions of analyte selectivity.

The pairing of pre-ionization separations with MS is widely used to increase the dimensionality of data, thereby increasing the information obtained. Most commonly these include condensed-phase separations (*e.g.* high performance liquid chromatography (HPLC)) or gas-phase separations (*e.g.* gas chromatography (GC)), respectively. The fields of metabolomics and proteomics have advanced greatly in the recent decades through the interfacing of mass spectrometry with GC and LC separations, as well as by

the improved capabilities in data analysis that such multidimensional separations demanded. GC-MS is a widely-used tool in the quantitation of metabolite levels for metabolic flux analyses of microorganisms and cells.^{1,2} Identification of metabolites or proteins from complex biological samples, such as serum or plasma, is now routinely performed by LC-MS or LC-MS/MS.³⁻⁶ In contrast to separations based on bi-phasic partitioning, *i.e.* chromatography, gas phase electrophoretic separations by ion mobility (IM) are significantly more rapid, occurring within μ s-ms versus the min-hrs required for GC and LC separations. The nature of IM separations offers additional information as direct structural information on analytes can be obtained nearly regardless of the complexity of the sample. As a post-ionization technique, IM-MS is also amenable to the addition of pre-ionization separations for even greater data dimensionality.

The aim of this chapter is to provide a survey of the key aspects of IM-MS for those seeking an introductory discussion and is not intended to be a comprehensive review. The merits and motivations of coupling IM and MS and the transformation of IM-MS measurements to structural information are discussed in Section 1.1. A survey of the arrangements of modern IM-MS instruments and the practical aspects of coupling IM and MS are presented in Section 1.2. Section 1.3 highlights the diverse range of applications of IM-MS, with particular emphasis on applications to the life sciences and nanotechnology. Lastly, Section 1.4 provides a summary and the objectives of the dissertation research.

1.1.1. The Evolution of Ion Mobility Technology

Although the fundamentals of both MS and ion mobility spectrometry (IMS) are

rooted in the late 1890s, with the discovery of X-rays by Röntgen and the first gas-phase IMS experiments (Figure 1.1), progress in the development of these technologies did not occur in parallel.^{7,8,9} Since the construction of the first mass spectrometer in the early 1900s, MS has experienced consistent growth in the development of new techniques for mass analysis. For example, the precursor of orbitrap technology, the Kingdon trap, was described in a 1923 publication and has since become a powerful tool for high-resolution mass analysis.^{10,11} Conceptualization of the time-of-flight (TOF) mass analyzer, a staple instrument in modern mass spectrometry labs, is attributed to Stephens in 1946 and further developments by Mamyrin led to the reflectron TOF-MS for high resolution measurements twenty years later.^{12,13-15} The late 1980s marked a time of significant advances in MS for applications to the life sciences. With the development of soft ionization techniques such as matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI), MS analyses of proteins up to 100 kDa by MALDI-MS and 130 kDa proteins by ESI-MS were reported for the first time.^{16,17,18} By the mid-1990s, the advantages of ESI- and MALDI-MS for fields such as proteomics were evident and it has since become the primary tool for such work.

In the development of IMS, the late 1890s to mid-1920s marked a period of interest in fundamental physics studies of the movement of ions in the gas phase and these works would later be regarded as the foundations of modern day IMS. Construction of the first IM spectrometer is attributed to Zeleny, who in 1898 developed an apparatus for the purpose of measuring the ratio of the velocities of positive and negative ions in an electrical field through several gases.⁹ Nearly 30 years later, precise measurements of the mobilities of ions in air were performed by Tyndall.¹⁹⁻²³ Despite these early

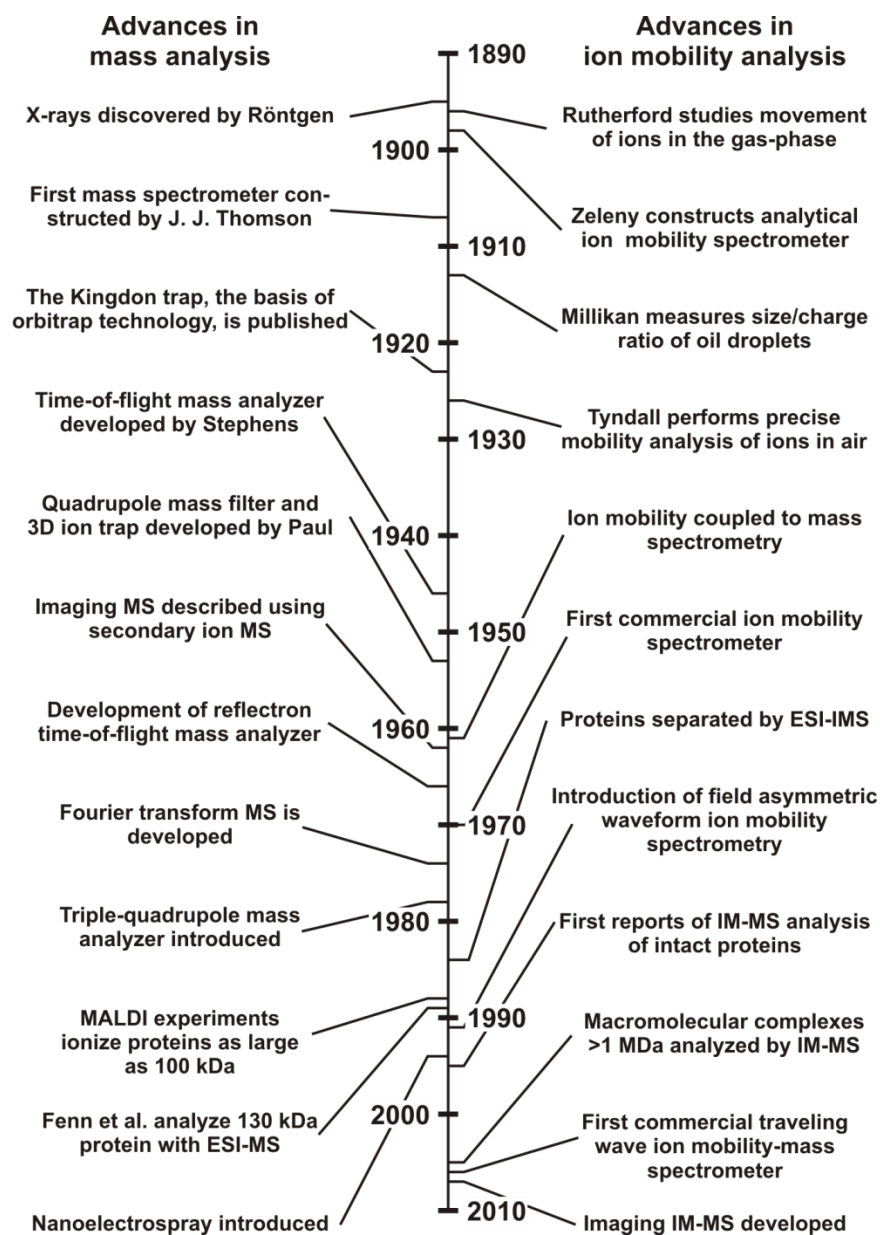


Figure 1.1. Timeline highlighting some of the significant advances in mass spectrometry (left) and ion mobility (right) from the 1890s to 2010.

developments, little progress was made in the field of IMS over the next three decades.

Renewed interest in IM separations did not occur until the early 1960s, when the first reports of IMS and MS analyses performed in tandem were published by McDaniel, Edelson, and colleagues.^{24,25} It was nearly a decade later that the first commercially produced IM spectrometer was released in 1970.²⁶ Referred to as “plasma chromatography” at the time, IMS developed as an analytical tool for gas phase separations of organic molecules in the 1960s-70s with applications in the detection of drugs and explosives.²⁷⁻²⁹

Similar to mass spectrometry, the applications of IMS to the biological sciences began to appear in the mid-1980s with the utility of ESI-IMS for the separation of multiply charged proteins demonstrated by Dole.³⁰ IM-MS separations of peptides were first reported in the mid-1990s by Bowers and colleagues.^{31,32} At the same time, studies of the gas-phase conformations of intact proteins were performed by Jarrold, Clemmer and colleagues.^{33,34} These works and others mark a significant milestone in the development of IM-MS as a multidimensional separations tool for biologically-relevant applications such as complex biological sample analysis and structural studies of biomolecules in the gas phase.

Over the past 50 years, the fields of IMS and IM-MS research have experienced an extraordinarily rapid growth. With applications spanning biological and physical sciences, the number of IM-related publications per year (Figure 1.2, top) has increased at a nearly exponential rate since the 1960s and technological innovations for IM-MS have experienced similar growth (Figure 1.2, bottom). With the recent availability of

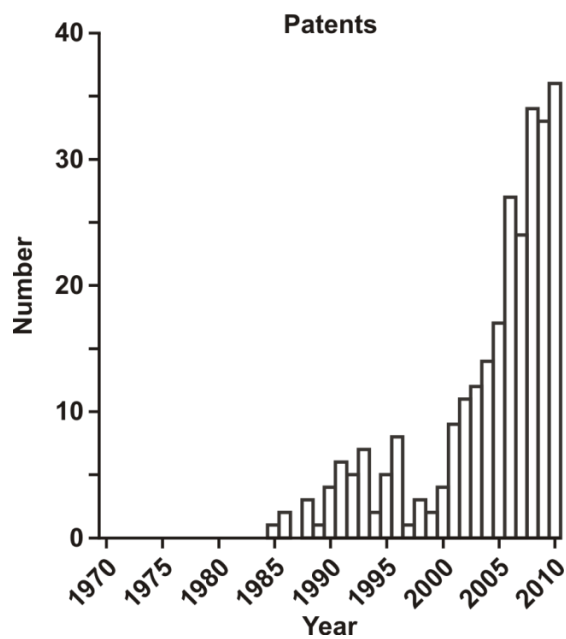
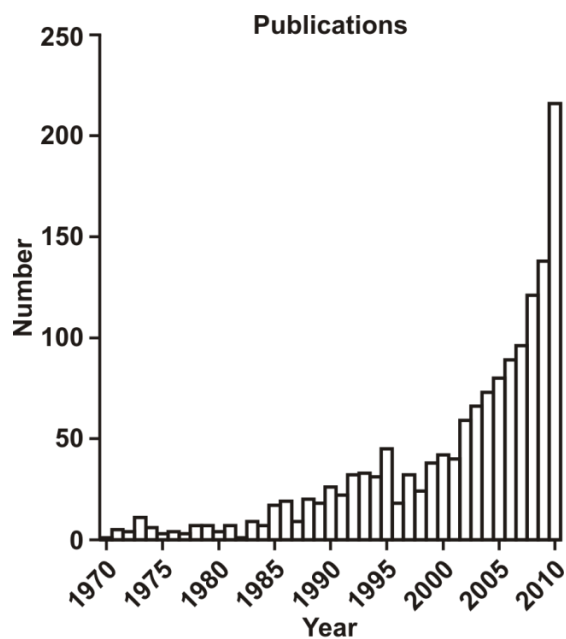


Figure 1.2. Trends in the numbers of IM-MS publications (top) and patents (bottom) from 1970-2010. The data displayed above was obtained by searching the phrase “ion mobility with mass spectrometry” in SciFinder.

commercial instruments and demonstration of its utility for imaging³⁵ and macromolecular complex analysis,³⁶ the field of IM-MS is expected to continue its exponential expansion in applications and innovations for the foreseeable future.

1.1.2. The Correlation of Ion Mobility and Mass Spectrometry Dimensions

The pairing of IM and MS separations provides structural information derived from the ion-neutral collision cross section in the IM dimension as well as accurate mass information in the form of the ion's mass-to-charge (m/z) in the MS dimension, thereby providing an increased dimensionality in the data. Figure 1.3 shows a typical representation of 2D IM-MS data, which is demonstrated with the ESI-IM-MS analysis of a cytochrome c tryptic digest. Panel (a) contains the plot of drift time (ms) versus m/z , referred to as conformation space. The intensity of signals is represented by a color scale, where signal intensity increases from white (no signal) to grey (low intensity) to black (high intensity). The full m/z profile (b) for the data can be produced by integrating the IM-MS plot (a) over all drift time space. Alternatively, integrating the IM-MS plot (a) over all m/z space yields the full IM drift time profile, as shown in panel (c). For a given signal in the IM-MS plot (a), such as the singly charged signal corresponding to the peptide "GITWK", integrating over a defined region of drift time- m/z space about the signal (indicated by the black rectangle in (a)) yields both the m/z and drift time profiles for the individual signal, as shown in panels (d) a (e), respectively.

For the analysis of biomolecules, which are mainly comprised of the atoms C, H, O, N, P and S, the mass of such species scales as length cubed, or volume, while in the

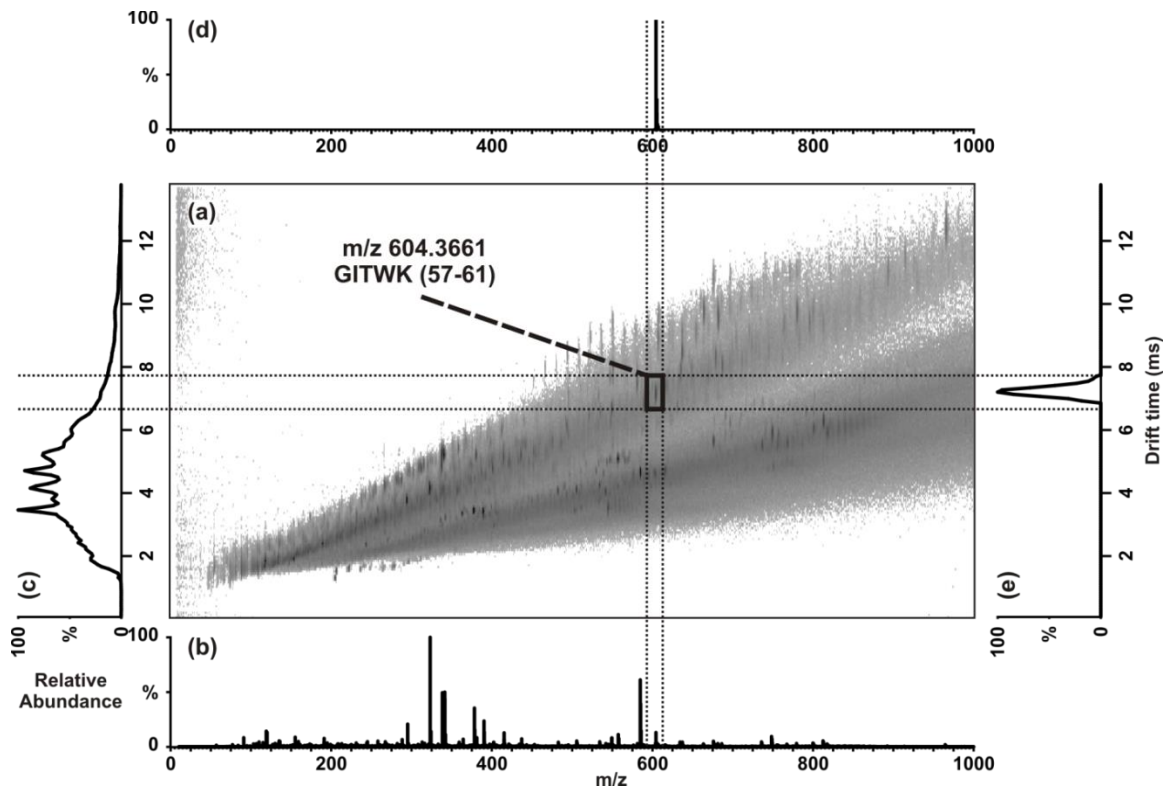


Figure 1.3. (a) A 2D ESI-IM-MS plot of conformation space obtained for the tryptic digest of cytochrome c. (b) The full m/z profile produced by integrating (a) over all drift time space. (c) Integrating (a) over all m/z space yields the full IM drift time profile. For the singly charged signal corresponding to the peptide GITWK, integrating about a define region of drift time- m/z conformation space (black rectangle in (a)) produces both the m/z (d) and drift time (e) profiles for the individual signal.

IM dimension the collision cross section, or surface area, scales as length squared. As both scale by length, the IM-MS separation dimensions are highly correlated due to the limited range of densities in which biomolecules exist. This high degree of correlation can present a significant advantage and some challenges when compared to multidimensional separations with more orthogonality.³⁷ As discussed in the section below (1.1.2.2), the classes of biomolecules, such as peptides and carbohydrates, each display different average densities or gas-phase packing efficiencies which can be a great advantage in complex sample analysis. In contrast, the high correlation of IM and MS dimensions observed for a particular class of biomolecule results in a decreased peak capacity (approx. $10^3 - 10^4$) in comparison to highly orthogonal multidimensional separations such as LC-FT-MS (approx. $10^7 - 10^8$), however IM-MS benefits from an extraordinarily high peak capacity production rate (10^6 s^{-1}) due to the rapid nature of the gas-phase separation.³⁸⁻⁴¹

1.1.2.1. Tailoring the Configuration of Multidimensional Separations

Unlike multidimensional separations with a high degree of orthogonality, the correlation between IM and MS allows versatility in the configurations of multidimensional separations, which may be tailored to fit the experimental goals. The most common configuration, as depicted in Figure 1.4(a), consists of an ion source, usually ESI or MALDI, from which ions are injected directly into the IM drift region. Mobility separated ions exiting the IM region are transferred to the mass analyzer. Depending on the goals of the experiment, full mass analysis of the mobility separated analyte ions can be performed with an orthogonal TOFMS (oTOFMS) or a quadrupole

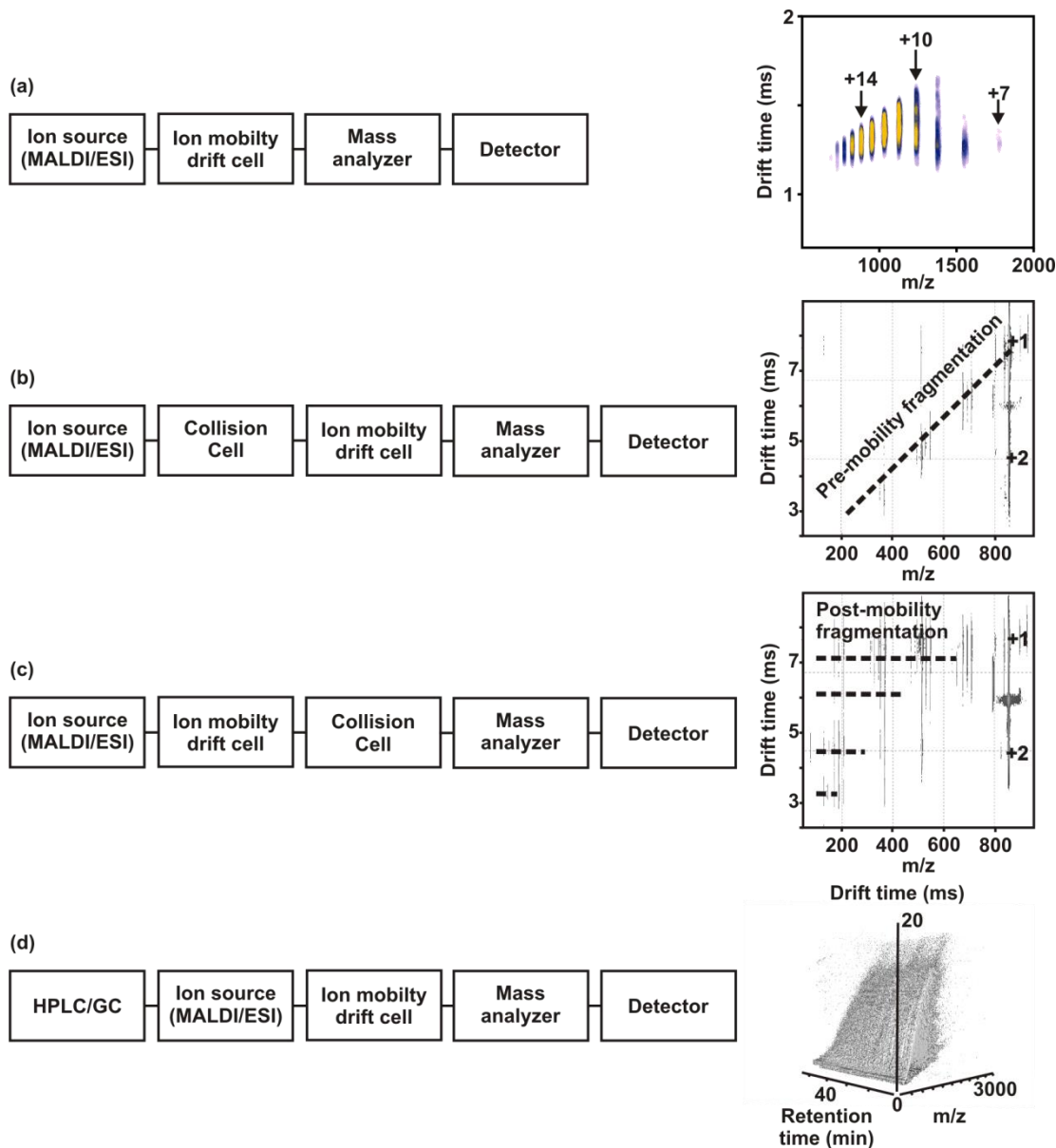


Figure 1.4. Arrangements of IM-MS instruments (left) and examples of the data obtained from each (right). The simplest arrangement (a) consists of an ionization source, typically MALDI or ESI, coupled to an IM drift cell, followed by a mass analyzer and detector. Placement of the collision cell prior the IM drift cell (b) provides mobility separated fragment and precursor ions. Fragmentation after IM separation (c) yields fragment ions which with the drift time of the precursor ion. IM-MS is compatible with pre-ionization separations (HPLC or GC) for additional data dimensionality (d).

MS may be used for the selection of a single m/z . Due to the complementary time scales of IM and TOFMS separations (ms vs. μ s, respectively), an oTOFMS is typically preferred for IM-MS of biomolecules to achieve multiple sampling across the analyte IM profile.

Additional structural information can be obtained from the MS dimension by performing IM-MS/MS, in which a collision cell is placed prior to the IM drift cell (Figure 1.4b). The IM region then behaves similarly to the MS¹ dimension in a tandem MS experiment, but disperses ions by ion-neutral collision cross section in time as opposed to dispersal by m/z . Pre-fragmentation mass selection, typically required for scanning MS/MS methods, is not required in this arrangement. Placement of the collision cell after the IM region (Figure 1.4c) produces mobility organized precursor and fragment ions, such that the fragment ions occur with the drift time as the precursor ion from which they were formed and both are determined in the MS region. This correlation of precursor and fragment ions allows fragmentation to be performed on nearly all ions, providing a multiplexed MS/MS experiment.⁴² Further tailoring of the instrumental arrangement may include additional regions of MS, MS/MS, IM or IM/IM to suit the goals of the experiment. For example, the addition of a mass filter (*e.g.* a quadrupole MS) for precursor ion selection and placement of the ion activation region prior to the IM drift cell enables IM structural information to be obtained for the fragment ions. IM-MS is also amendable to the addition of pre-ionization separation techniques (Figure 1.4d) such as GC, HPLC, capillary electrophoresis (CE), etc. for enhanced peak capacity and data dimensionality in the analysis of complex samples.^{38-40,43}

1.1.2.2. Advantages for Complex Biological Sample Analysis

Biologically derived complex samples natively contain multiple classes of biomolecules, such as proteins, lipids, carbohydrates, etc. The analysis of such samples by MS requires steps to enrich for the species of interest (*i.e.* peptides and proteins for proteomic experiments) by depleting the other components from the sample. This is necessary to prevent complications arising from ion suppression effects and interference from chemical noise, and to obtain mass spectra from which identifications or quantifications can be determined with high confidence. While condensed-phase pre-ionization separations are commonly used for these purposes, gas-phase separations on the basis of analyte structure by post-ionization IM-MS achieve the same goals of improved sensitivity and limit of detection but require separation times nearly 4 to 5 orders of magnitude faster.

The separation of complex samples in the IM dimension is based in the differences in gas-phase packing efficiencies amongst the classes of biomolecules, which decrease in the order: oligonucleotides > carbohydrates > peptides/proteins > lipids.⁴⁴⁻⁴⁸ Ultimately, the gas phase packing efficiency of a biomolecule dictates its conformation and is therefore inherent to its collision cross section as well. The result is the observation of unique regions of collision cross section- m/z correlation, as depicted in Figure 1.5, for each of the classes of biomolecules in the 2D IM-MS conformation space, the general order of which is conserved despite the separation parameters. This consistency in the order of separation simplifies the general assignment of signals in a complex sample to their respective biomolecular classes and provides increased confidence in the validity of

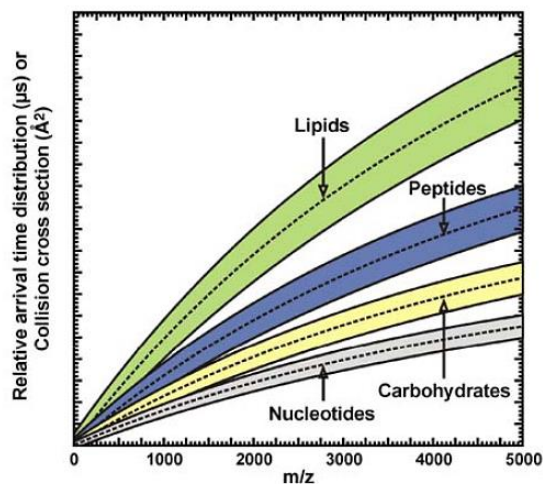


Figure 1.5. A hypothetical depiction of the correlation lines observed in IM-MS conformation space for each of the classes of biomolecular species. Adapted from L.S. Fenn and J.A. McLean, *Analytical and Bioanalytical Chemistry* **2008**, *391*, 905-909, Fig. 2(a), with permission from Springer Science+Business Media.

signal identifications. Fine-grained structural details within a biomolecular class can be visualized upon closer inspection of its correlation region (discussed in Section 1.3.1).

The structural separation of molecular classes by IM has practical advantages for many analyses typically performed by MS methods. In proteomics experiments performed by IM-MS, non-peptide interferences can be removed by extracting the 2D region of conformation space in which the peptide signals reside, providing higher confidence in the identifications provided by proteomics databases and an improved dynamic range.³⁸ Isobaric peptides with conformational isomers or amino acid sequence permutations can be resolved with IM on the basis of their unique preferred structures but would be challenging to observe by MS alone.^{47,49-52}

1.1.3. Structural Information from Two-Dimensional Conformation Space

In the analysis of biological samples, ion mobility separations are generally utilized in two distinct, but complementary ways: (i) as a structural separations technique on the basis of collision cross section, and (ii) to derive low resolution structural data that is interpreted by complementary molecular simulation strategies. There exist a number of means to perform ion mobility using electrostatic and electrodynamic drift fields. However, these two approaches differ in that absolute structural information can be obtained from first principles of gas kinetic theory using electrostatic fields, and relative structural can be obtained by comparison with structural standards using electrodynamic fields, respectively. The present discussion focuses deriving structural information exclusively using electrostatic, or uniform-field, separation devices.

1.1.3.1. Relating Mobility Drift Time to Collision Cross Section

The ion-neutral collision cross section (CCS) is effectively the rotationally-averaged surface area (in \AA^2) presented by the ion during the collision with a neutral drift gas molecule, and is directly related to analyte structure. If these collisions are considered to be brief and elastic, an equation can be derived from the kinetic model of gases for the calculation of collision cross sections from ion mobility separation parameters.

Under the influence of a weak electrostatic field (E), the partitioning of ions in a drift cell of fixed length (L) is given by the ion drift velocity (v_d) and is related to E by the ion's mobility constant, K , in the neutral gas:

$$v_d = \frac{L}{t_d} = KE \quad [1]$$

The velocity of a packet of ions is derived from the time required for the ions to traverse the IM drift cell, or drift time (t_d), which is empirically measured and used to evaluate the CCS. Under “low-field”, or weak electrostatic field, conditions, the velocity of ions in the neutral drift gas will be attributed mainly to the random motion of ions at the temperature of the drift gas. This requires the assumption that molecular or ion velocities in thermodynamic equilibrium can be described by a Maxwellian distribution function, where the mean thermal velocity of the gas is:

$$v_{mean} = \left(\frac{8k_B T}{\pi M} \right)^{\frac{1}{2}} \quad [2]$$

where T (Kelvin) is the temperature of the gas, k_B is the Boltzmann constant, and M is the molar mass of the gas. In addition to the mean thermal velocity, a small component of velocity is added in the direction of the electrostatic field. As the ion packet's drift velocity depends on the temperature of the separation, as well as the pressure (p , Torr), K

is conventionally normalized to standard conditions of 0 °C and 760 Torr (STP) and reported as the reduced mobility, K_0 :

$$K_0 = K \frac{p}{760} \frac{273}{T} \quad [3]$$

Alternatively, at high field conditions the distribution of ion velocities is less strongly dependent on the separation temperature and K is no longer constant, but depends on the ratio of the electrostatic field to the gas number density (E/N).

Under low field conditions, in which K is constant, the mobility of an ion is inversely related to the ion-neutral collision cross section, effectively, the surface area, Ω :

$$K_0 = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_B T)^{1/2}} \left[\frac{1}{m_i} + \frac{1}{m_n} \right]^{1/2} \frac{760}{p} \frac{T}{273} \frac{1}{N_0} \frac{1}{\Omega} \quad [4]$$

where K_0 is also dependent upon the charge of the ion (ze), the drift gas number density at STP (N_0), and the masses of the ion (m_i) and neutral gas (m_n) in the form of the reduced mass of the ion-neutral collision pair. To solve for the collision cross section, Eqn. [1] can be substituted for K_0 and Eqn. [4] can be rearranged to yield the form commonly used to calculate collision cross sections from IM separations:

$$\Omega = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_B T)^{1/2}} \left[\frac{1}{m_i} + \frac{1}{m_n} \right]^{1/2} \frac{t_d E}{L} \frac{760}{p} \frac{T}{273} \frac{1}{N_0} \quad [5]$$

The dimensionality of units used in Eqns. [4] and [5] should be closely monitored and the units of E in particular must be expressed the cgs Gaussian units statvolts cm^{-1} .

As previously discussed, collisions which occur during the IM separation are approximated to be brief and elastic, *i.e.* the total translational kinetic energy of the ion and the molecule are conserved during the collision. However, it has been shown by the comparison of experimental and theoretical collision cross sections that the elastic, or “hard sphere,” approximation hold best for analytes 1000 Da or larger.⁵³ These

approximations are also challenged by analytes which are similar in size to the neutral separation gas, resulting in long-range interactions between the collision pair.⁵⁴⁻⁵⁶ To obtain accurate collision cross sections for such analytes, the long-range interaction potential should be accounted for. For biological IM studies, the analytes of interest typically do not fall under these limitations and accurate collision cross sections can be obtained using Eqn. [5].

1.1.3.2. Correlating Empirical and Computational Collision Cross Sections

Collision cross sections derived from experimental IM-MS data provide a very general view of analyte ion structure in the form of rotationally averaged surface area. To gain insight into the analyte structure, empirically derived collision cross sections are often paired with computationally determined results. Structures arrived at by this process are useful in identifying the dominant structural motifs as well as eliminating those which are in disagreement with empirical data, but results should not be over interpreted as providing fine-course details such as atomic coordinates.

In general, the work flow for comparing empirically and computationally-derived structures consists of five procedures. As the first step, *in silico* model structures of the analyte ion are generated by molecular modeling software. The conformational landscape of the model structure is then sampled with methods such as simulated annealing, molecular mechanics/molecular dynamics (MM/MD), or distance geometry. Of the conformational sampling approaches, MM/MD methods are most commonly used for the structural study of large biomolecules. However, MM/MD approaches experience difficulties in the sampling of compact conformations and MM/MD force fields must

exist that are properly parameterized for the analyte.⁵⁷⁻⁵⁹ Collision cross sections are then calculated for each of the conformations obtained by the sampling technique. This procedure is widely performed with the MOBCAL software developed by Jarrold and coworkers.^{53,54,60} The relative energies of the conformations are plotted against their computationally derived collision cross sections. The structures of most interest are the lowest energy conformations with collision cross sections consistent with the experimental results. A web-based tutorial on the procedures and considerations for comparing empirical and computation collision cross sections, with an emphasis on simulated annealing methods, is also available from Bowers and colleagues.⁶¹ The specific IM-MS arrangement used does have implications for the calculation of collision cross sections, and several IM-MS arrangements used for determining CCSs are detailed in the following section.

1.2. Instrumental Configurations of Modern Ion Mobility-Mass Spectrometers

The flexibility available in pairing IM and MS has led to a number of unique instrumental arrangements designed around the needs of specific experiments or a preferred type of drift cell and/or mass analyzer. Early IM-MS instruments were comprised of drift tube ion mobility (DTIM) spectrometers combined with quadrupoles²⁴ or TOFMS²⁵ and such arrangements have remained popular throughout the years.⁶²⁻⁶⁷ Additionally, IM spectrometers have been interfaced with mass analyzers such as double-focusing sector fields,⁶⁸ quadrupole ion traps,^{69,70} and Fourier transform ion cyclotron resonance.⁷¹ IM spectrometers are also available in a variety of designs and IM-MS instruments have incorporated segmented quadrupoles,⁷²⁻⁷⁴ field asymmetric ion mobility

spectrometers (FAIMS),^{75,76} and traveling-wave ion mobility (TWIM).^{36,77,78} This section highlights a few fundamental considerations necessary in choosing an appropriate IM-MS arrangement and the configurations of several modern IM-MS instruments.

1.2.1. Impact of Ion Source Selection on Ion Mobility Arrangement

The development of MALDI^{16,17} and ESI¹⁸ greatly expanded the applicability of both MS and IM-MS to biologically-relevant analyses. IM-MS instruments have been successfully interfaced with both ESI^{65,66,69,78-83} and MALDI^{67,68,84} ion sources, although ESI remains the more widely used ionization method. The fundamental differences in the ionization processes of MALDI and ESI dictate the selection of ion source for certain experiments, and therefore the particular design of the IM-MS instrument. MALDI typically produces singly charged ions, while ESI produces a series of ions at multiple charge states depending on the analyte's size and the number of basic sites available. For IM-MS, the ion's charge state and the number of charges per ion are key considerations as the mobility of the ion is directly proportional to its charge state. In the analysis of complex samples, the multiplicity of signals produced for an ion existing in multiple charge states may lead to overly congested spectra while also reducing the sensitivity as the analyte signals are partitioned into several channels. For structural analyses in which collision cross sections CCS are experimentally calculated, charge repulsion due to the low dielectric of the IM cell results in CCS larger than predicted solution phase structure. Despite these factors, the multiple charges incurred during ESI may sometimes be necessary to produce ions within the appropriate m/z range for the mass analyzer (*e.g.* quadrupoles, which are typically limited to a lower m/z range). The detection scheme

may also dictate the use of ESI, as the higher charge state ions provide higher sensitivity for micro-channel plate or image current detectors. Fragmentation studies performed by IM-MS/MS also benefit from multiply charged ions as they provide higher fragmentation efficiencies.

In regards to structural IM-MS studies, it should be noted that the equivalency of structures generated by MALDI and ESI, as well as the equivalency of gas-phase versus solution-phase structures, remains to be determined on a case-by-case basis. The presence of acids or organic modifiers for ESI and the co-crystallization with small organic matrix molecules for MALDI both to some extent produce denaturing conditions during sample preparation. For both MALDI and ESI, evaporative cooling from matrix cluster evaporation and solution evaporation, respectively, can alleviate the additional internal energy transferred to the analyte during the ionization process. In the absence of intermolecular solvation, the anhydrous ions generated by both ionization methods should adopt the minimum energy conformation available via intramolecular folding forces. An instrumental approach to this issue is the development of a dual source instrument in which the IM-MS operates independent of the source in use, allowing for direct comparison of collision cross sections of ion produced by MALDI and ESI. Using this arrangement, it has been shown that the collision cross sections of +1 MALDI and ESI ions are in good agreement, while the +2 ions generated by ESI tend to be larger than +1 MALDI ions.⁸⁵ The combination of IM-MS with structural analyses such as X-ray crystallography and NMR has provided some insight into the correlation of gas and solution phase structures.^{34,86-88} Such work has typically been limited to relatively small model systems but recently advances have been made towards much larger systems.⁸⁹

Further research into the structural differences between MALDI and ESI-generated ions, as well as between gas and solution phase structures, would be of great benefit to MS-based studies of protein structure and protein complexes.^{90,91}

1.2.2. Ion Mobility-Time-of-Flight Mass Spectrometers

Similar to IM separations, time-of-flight mass analyzers disperse ions in time which allows the interfacing of IM and TOFMS to be relatively straightforward. This arrangement further benefited from the complementary timescales of analyte elution from the IM drift cell (100s μ s-ms) and mass analysis in the TOFMS (10s μ s), ensuring the nearly simultaneous mass analysis of the analytes as they elute from the IM drift cell. For these reasons, TOFMS and orthogonal TOFMS (oTOFMS) are widely used in IM-MS arrangements. Several arrangements of TOFMS-based IM-MS instruments based on ESI and MALDI ionization sources and drift tube and traveling wave IM are discussed, as well as an arrangement for imaging MALDI-IM-MS.

1.2.2.1. MALDI-IM-TOFMS

Ions produced by MALDI are both spatially and temporally focused, which allows for the rapid separations and high sensitivity necessary for high throughput studies such as proteomic analyses. As the MALDI event itself is a pulsed process, it is unnecessary to temporally gate the ions prior to injection into the drift cell, a limiting factor for resolution.⁹²⁻⁹⁵ Additionally, the pulsed nature of MALDI provides an inherent t_0 for the IM drift time measurement. A schematic diagram of a moderate pressure

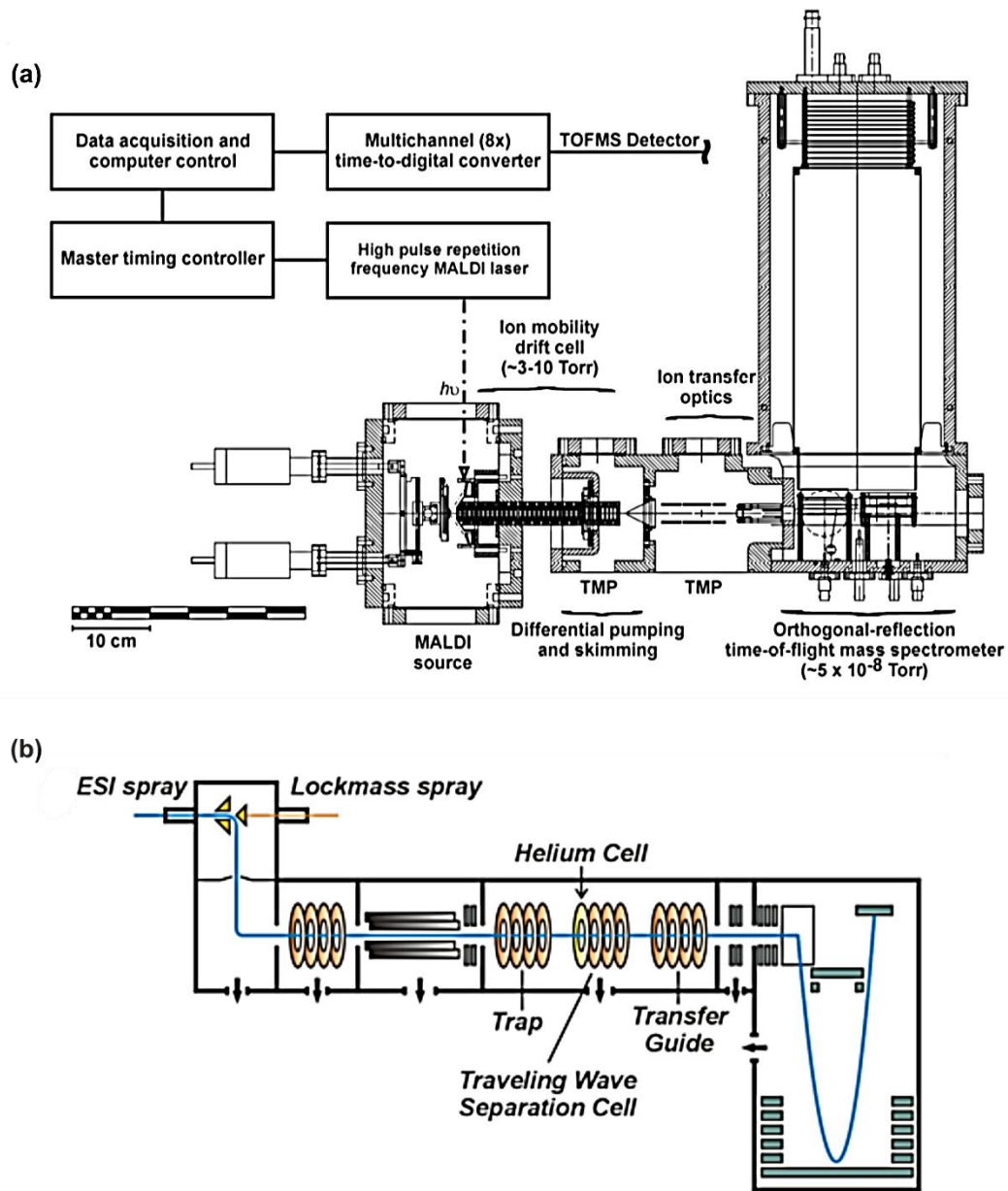


Figure 1.6. Schematic diagrams of an IM-MS instrument with a drift tube ion mobility cell, MALDI ionization source and TOF mass analyzer (a), and a traveling wave-based IM-MS instrument with an ESI source and TOF mass analyzer (b). Panel (a) is adapted with permission from John Wiley & Sons, Inc., J.A. McLean, J.A. Schultz, A.S. Woods, *Ion Mobility-Mass Spectrometry for Biological and Structural Mass Spectrometry*, in *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practices and Biological Applications*, Richard B. Cole (Ed.), John Wiley & Sons, Inc., New York, 2010, pp. 411-439.

MALDI-DTIM-TOFMS is presented in Figure 1.6(a). Ions are produced at the same pressure (3-10 Torr) as the IM drift cell with high repetition rate MALDI (approx. 5000 Hz). The moderate pressure MALDI source allows for a simplification in the design of the instrument, which would otherwise require a pressure transition from high vacuum to moderate pressure to accommodate a high vacuum MALDI source.⁶⁷ Under moderate pressure, successive pulses of the laser will experience higher ionization efficiencies due to neutral plume stagnation. High repetition rate MALDI is advantageous under such conditions as it provides higher throughput and improved limits of detection (fmol-amol).⁹⁶ The ions then enter a dc-only ion guide drift cell. This design radially refocuses the ions to the center of the drift cell to prevent losses due to radial diffusion of the ions as they migrate through the cell, thereby improving the sensitivity.⁹⁷ As ions elute from the drift cell, they are focused through a region of differential pumping and skimming and transferred to the entrance of an oTOFMS (approx. 10^{-8} Torr). Overall, this arrangement yields ATDs of 100s μ s for peptides and approximately 1-2 ms for proteins.

While interfacing two time-dispersive measurements may seem straightforward at first glance, matching the time resolution of each dimension may be complicated. For an oTOFMS operating at a 10 kHz pulse frequency, an MS spectrum would be collected ever 100 μ s. The resulting IM profile would be 1 ms long and composed of 10 points at 100 μ s time resolution. The peak profile widths (FWHM) would be approximately 10-20 μ s for an IM operating at a resolution of 30-50. Consequently, the ion packets exiting the drift cell during the 100 μ s MS cycle will not be recorded by the oTOFMS. The solution to this challenge is a specialized data acquisition scheme in which the time resolutions of both dimensions are independent and interleaved post-acquisition.^{38,98} Sequential MS

spectra are acquired at the rate of the TOFMS, but after each subsequent MALDI event the TOFMS acquisition is offset by a desired amount. These offset MS spectra are interleaved, or stitched together, after a sufficient number of ion injections into the drift cell. This provides a two-dimensional IM-MS plot with independent resolution (*i.e.* number of time points) in the IM and MS dimensions.

1.2.2.2. ESI-IM-TOFMS

ESI ion sources provide a continuous source of ions in contrast to the pulsed generation of ions provided by a MALDI ion source. For injection into the IM region, a continuous ion source such as ESI requires modulation of the ion beam by storing ions and gating the injection. Ion funnels^{80,99,100} and pulse ion traps^{69,77,82} are typically used for these functions. The performance characteristics of modern devices provide very little ion loss between the source and injection into the drift region, and ESI-IM-MS instruments can achieve sensitivities comparable to commercial ESI-MS instruments.⁹⁹ The inclusion of an ESI source is advantageous as it enables the interfacing of pre-ionization liquid-phase separations with IM-MS with relative ease. The development and utility of LC-ESI-IM-MS has been widely explored for proteomics applications by Clemmer and colleagues.^{43,101-105} A detailed discussion of an ESI-IM-TOFMS instrument utilizing traveling wave IM is provided in the following section.

1.2.2.3. Traveling Wave IM-TOFMS

The first commercially available IM-MS instrument based on traveling wave ion mobility was released in the late 2000s (Waters Corporation) and has since become

widely used. The basic principles of TWIM are similar to those of DTIM in that ions travel through the drift region under the influence of an electric field and their motion is impeded to some extent by collisions with a neutral drift gas. A key difference is the nature of the electric field used to propel ions through the drift cell. In contrast to uniform field DTIM instruments, TWIM drift cells utilize a dynamic electric field. The TWIM drift cell is comprised of a series of planar ring electrodes and the opposite phases of an RF voltage are applied to adjacent ring electrodes to radially confine ion motion to the center of the ring electrodes.⁷⁷ In addition to the RF voltages, a dc voltage is briefly applied to a pair of neighboring electrodes in a repeating pattern across the length of the drift cell (typically comprised of 61 electrode pairs).⁷⁸ The dc pulse generates a potential hill and, when applied in the repeating sequence, results in a series of traveling waves which propel ions away.^{77, 78} Ions of different mobilities are then separated based on their abilities to travel with, or “surf,” a traveling wave pulse, which is dependent upon the traveling wave height and velocity and the neutral drift gas pressure.⁷⁷ High mobility ions tend to surf the front of the traveling wave and exit the drift cell quickly. In contrast, low mobility ions cannot keep up with the traveling wave and will roll over the top of the wave. The low mobility ions, on average, will be propelled through the drift cell and exit at a time dependent upon the number of roll over events.⁷⁷

A schematic diagram of an ESI-TWIM-TOFMS instrument is shown in Figure 1.6(b). Following ionization, ions are directed by a traveling wave ion guide into the quadrupole, which typically operates in RF-only mode for full mass analysis. Ions then enter the IM region of the instrument, which is comprised of three stacked ring electrode ion guides capable of traveling wave separations. The trap region stores ions during ion

mobility separation, as necessary to modulate the continuous ion beam from the ESI source. This region typically operates without the traveling wave and the gating and release of stored ions into the IMS cell is performed by the last ring electrode, which is operated as dc-only. As discussed above, the IMS cell utilizes the traveling wave and separation is typically performed in nitrogen gas (N_2). The transfer ion guide operates with a small traveling wave to maintain the mobility separation of the ions. Ions eluting from the IM region travel through a series of focusing lenses prior to entering the source of the oTOFMS. The TOFMS pusher is synchronized to the release of ions from the trap and up to 200 mass spectra are recorded per cycle of the TOF pusher. Note that in present instrumentation there is no interleaving of data, so time resolution in both the IM and MS dimensions are not independent.

The ESI-TWIM-MS arrangement shown in Figure 1.6(b) offers an important advantage in that several types of fragmentation studies can be performed. The quadrupole, when operated as a mass filter, enables the mass selection of precursor ions. Both the trap and transfer regions are capable of fragmentation via collision induced dissociation (CID). Fragmentation within the trap occurs prior to IM analysis (termed pre-mobility fragmentation) and product, as well as precursor, ions undergo mobility separation. Thus, structural information for both the precursor and fragment ions is obtained simultaneously. Alternatively, fragmentation in the transfer region occurs post-mobility and product ions are aligned in drift time with their precursor ion. This scheme is particularly useful where isobaric, but structurally distinct species are present, as it would be straightforward to distinguish which product ions correspond to a given

precursor ion. It is also possible to perform multiplexed fragmentation by operating the quadrupole in full-MS mode and utilizing post-mobility CID in the transfer region.

As discussed above (Section 1.1.1.3) the electrodynamic nature of separations in TWIM do not provide a convenient means for calculation of absolute collision cross sections by Eqn. [5]. Nevertheless, relative collision cross sections can be determined but the TWIM drift times must be calibrated against absolute collision cross sections determined by DTIM. Several detailed protocols are available for performing the calibration of TWIM drift times to collision cross sections.^{106, 107}

1.3. Applications of IM-MS

As the availability and popularity of IM-MS technologies has grown, so too has the number and diversity of its applications. IM-MS has become a key research tool in fields ranging from the life sciences to nanotechnology. In addition to emerging IM applications, IM-MS has also been demonstrated as an advantageous technique in established fields, such as proteomics and systems biology. This section serves as an introduction to a few of the many applications of IM-MS technologies, with focuses on conformation space analysis, “omics” analyses, and macromolecular protein complexes.

1.3.1. Conformation Space Analysis

The area formed by the 2D plot of IM drift time versus m/z , a typical data output from IM-MS experiments, is commonly referred to as conformation space. As previously discussed (Section 1.1.2.), the correlation of m/z and drift time (or mass and size, respectively) results in a low peak capacity for IM-MS and signals do not occupy every

region of the conformation space. Instead, signals form into regions of correlation, or trend lines, based on their chemical composition. The comprehensive mapping of the regions of conformation space typically occupied by biomolecules also revealed generalized strategies for ion mobility labeling approaches to utilize those regions typically void of signals and increase information content.

1.3.1.1. Mapping the Conformation Space of Biomolecules

For biological analytes, it is well known that the gas-phase packing efficiencies differ significantly amongst the different classes of biomolecules due to their differing degrees of freedom. The relative packing efficiencies of the biomolecular classes follow the trend in order of decreasing packing efficiency or density: oligonucleotides > carbohydrates > peptides/proteins > lipids.⁴⁴⁻⁴⁸ This ordering is observed in the IM-MS conformation space in the form of trend lines unique to each biomolecular class, as illustrated in Figure 1.5 in Section 1.1.2.2. The ordered separation of biomolecular species in IM-MS conformation space can be utilized as a predictive tool for identifying the biomolecular class to which unknown species belongs. A map of IM-MS conformation space developed from a large dataset of singly-charged lipids ($n = 53$), peptides ($n = 610$), carbohydrates ($n = 192$), and oligonucleotides ($n = 96$) is presented in Figure 1.7a. In contrast to typical plots of IM-MS conformation space, the drift times have been transformed to collision cross sections and then plotted against m/z . The primary purpose for the transformation from drift time to collision cross section is to normalize for experimental conditions, but the nonlinear nature of the transformation has the added benefit of enhancing the separation of the biomolecular trend lines.⁴⁸ The data

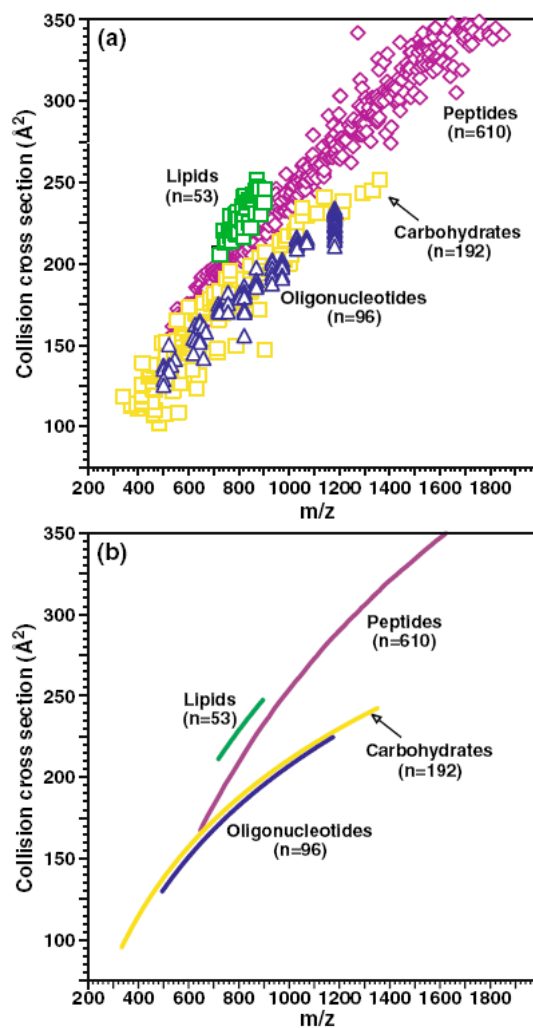


Figure 1.7. (a) A plot of collision cross section versus m/z for the different classes of biomolecular species: lipids ($n = 52$), peptides ($n = 610$), carbohydrates ($n = 192$), and oligonucleotides ($n = 96$). All species are singly charged ions generated by MALDI, and the data point generally contains error $\pm 1\sigma$. (b) For each biomolecular class, the average collision cross section and m/z data in (a) has been fitted to a logarithmic regression. Adapted from L.S. Fenn, M. Kliman, A. Mahshutt, S.R. Zhao, and J.A. McLean, *Analytical and Bioanalytical Chemistry* **2009**, 394, 235-244, Fig. 1(a,b), with permission from Springer Science+Business Media.

in Figure 1.7a has been fitted to a logarithmic function and the average logarithmic correlation for each biomolecular class is presented in Figure 1.7b. The approximation of collision cross section as a logarithmic function of m/z is appropriate as the increase in collision cross section with increased mass would be expected to approach a limit at high mass depending on the average packing density of a molecular class.⁴⁷ Inspection of the distribution of collision cross section values about the average correlations for each class reveals that the *breadth* of conformation space occupied increases in the order: lipids < oligonucleotides < peptides < carbohydrates.⁴⁷ This ordering can be explained by examination of the structural components from which each biomolecular class is composed as well as their prevailing intramolecular forces (*e.g.* hydrogen bonding, π - π interactions, van der Waals interactions, etc.). Oligonucleotides are linear biopolymers comprised of only four monomeric units, and therefore the narrowness of their distribution can be explained by the limited arrangements available and the strong π - π base pair stacking intramolecular forces. Alternatively, peptides and carbohydrates have greater numbers of monomeric units and greater potential for structural diversity due to branching. The small distribution of the lipids can be rationalized by the highly conserved structure of the class, in which most species are comprised of a small head group and two large fatty acid tails; however, it should also be noted that the small deviation may be due to under sampling.⁴⁷

In addition to course-grained structural information, closer inspection of the correlations observed in conformational space can reveal fine structural details. The different sub-classes or structural motifs within a biomolecular class can be visualized in the form of deviations from the average correlation for the class, as depicted in Figure

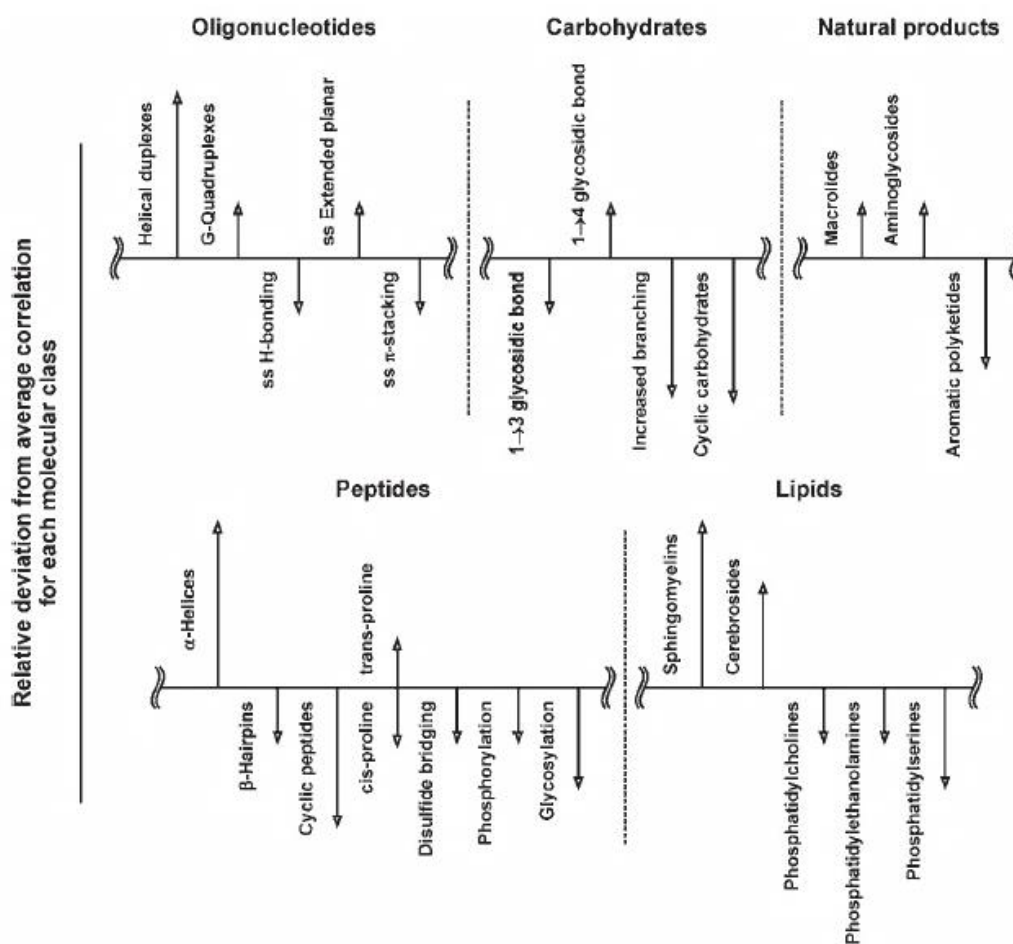


Figure 1.8. A plot demonstrating the deviations about the average correlations of oligonucleotides, carbohydrates, natural products, peptides and lipids resulting from the structural motifs or sub-classes present within the biomolecular classes. The length of the arrows indicated the relative magnitude of the deviation and the direction up or down indicates their location above or below the average correlation line of the class. Reprinted from *J. Am. Soc. Mass Spectrom.*, 20, J.A. McLean, The mass-mobility-correlation redux: the conformational landscape of anhydrous biomolecules, 1775-1781, 2009, with permission from Elsevier.

1.8. For example, the dominant tertiary structural motif of a peptide results in a deviation above or below the average peptide correlation, where elongated α -helical peptides have a larger deviation than more compact β -hairpin peptides.^{38,47,48} Similarly, post-translational modifications (PTMs) of peptide sequence such as phosphorylation and glycosylation result in a deviation below the average peptide correlation.^{38,48,108} The sub-classes of lipid species display a clear bimodal distribution about the average lipid correlation, as shown in Figure 1.8. These sub-classes can be further grouped into the sphingolipid and glycerophospholipid families, which differ in the linkage (sphingosine and glycerol, respectively) of the phosphate and head group to the fatty acyl tails. The sphingomyelin and cerebroside lipids, belonging to the sphingolipid family, have much more extended structures than the more compact glycerophospholipid family (*e.g.* phosphatidylcholines, phosphatidylserines, etc.). This is clearly indicated in Figure 1.8, as the sphingolipids tend to have large relative deviations above the average lipid correlation while glycerophospholipids have smaller relative deviations below the average correlation.^{47, 48, 109}

1.3.1.2. Utilizing Conformation Space with Ion Mobility Shift Reagents

The correlation of the mobility and mass dimensions leads to only a portion of the total conformation space being occupied by signals. For highly complex samples such as plasma or cell lysates, this area can become congested and detection of low abundance species can be challenging due to overlap. With knowledge of the regions of IM-MS conformation space which are expected to be occupied by biomolecules, it becomes possible to utilize those regions predicted to be unoccupied. With the assistance of ion

mobility shift reagents, analytes containing specific chemical functionalities (*i.e.* cysteine residues, phosphorylations, etc.) can be shifted from the correlation region for their biomolecular class to an unoccupied area of conformation space. These empty regions of conformation space occur both above and below the correlation region, giving rise to two possibilities for shifting the analyte. To shift analytes to an area above the correlation region for their biomolecular class, the IM shift reagent must have a lower density than that biomolecular class, as depicted by the parachutes in Figure 1.9. In contrast, higher density IM shift reagents, which behave like anchors (Figure 1.9), must be used to shift analytes below their correlation region.

One approach to the development of IM shift reagent labeling has been focused on the formation of noncovalent ion-molecule complexes in the gas phase, in which chemical complexation occurs at a specific functional group in the ion. Crown ethers, in particular, have been widely studied due to their ability to form ion-crown complexes in the gas-phase.¹¹⁰⁻¹¹⁷ The utility of crown ethers to form complexes with primary amines and amino acids was described by Creaser and colleagues.^{116, 117} In a series of IM-MS experiments, they demonstrated the ability to separate isomeric amines and peptides based on the increase in collision cross section which occurred upon complexation with the crown ethers.^{116, 117} The utility of crown ethers as low density IM shift reagents for peptides was further developed by Clemmer and colleagues, who demonstrated the sequence selectivity of crown ether-peptide complexation, as the crown ethers preferentially formed complexes with the basic sites in the peptides (*i.e.* arginine and lysine).¹¹⁸ It was also demonstrated that the peptide-crown ether complexes could be dissociated by collisional activation to reveal the free peptide, enabling mass analysis of

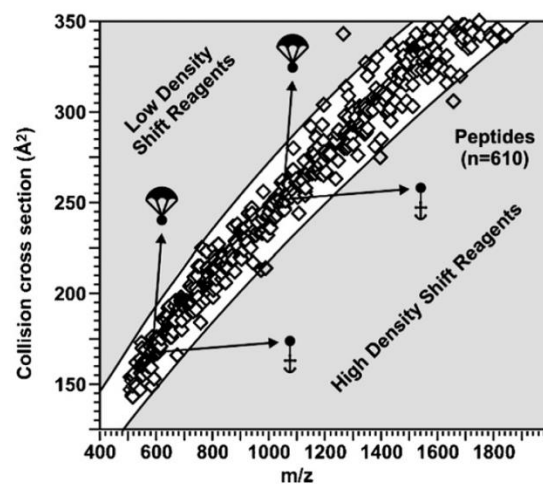


Figure 1.9. A graphical depiction of the peptide mobility-mass correlation in IM-MS conformation space based on the collision cross sections of *ca.* 600 singly-charged peptides. The effects of low density and high density shift reagents are illustrated as parachutes and anchors, respectively. Reprinted from *Int. J. Mass Spectrom.*, 307, T.J. Kerr, R.L. Gant-Branum, J.A. McLean, Multiplexed analysis of peptide functionality using lanthanide-based structural shift reagents, 28-32, **2011**, with permission from Elsevier.

the free peptide ions rather than the peptide-crown ether complexes. Although there are benefits to the use of noncovalent complexes as IM shift reagents, the complexity arising from the presence of both complexed and free target species is not ideal as the approach cannot be used quantitatively.

Labeling via covalent chemical modification is an alternative approach to ion mobility shift reagents which allows relative quantitation information to be obtained. Originally described for MS-based relative quantitation of peptides and proteins^{119,120,121}, lanthanide-based labeling strategies have recently been demonstrated as high-density IM shift reagents.^{122,123} The tags function as high density IM shift labels as the chelated lanthanide metal contributes a large mass increase without significantly increasing the collision cross section of the labeled peptide.^{122, 123} Subsequently, the labeled peptide is shifted below the predicted peptide correlation line to a region of IM-MS conformation space where such signals are not predicted to occur. The label generally consists of a lanthanide metal chelating moiety, a linker region and a chemically selective moiety which covalently modifies a specific functional group in the peptide or protein. A variety of chemical selective moieties have been developed for the selective modification of peptide functional groups such as primary amines, cysteine residues, and phosphorylated residues.¹²³ Incorporation of two different lanthanide metals in a known ratio allows for relative quantitation in a manner analogous to stable isotope labeling strategies such as isotope-coded affinity tagging (ICAT).

The use of lanthanide-based labeling for relative protein quantitation is illustrated in Figure 1.10(a). The cysteine residue of the peptide CLRRASLG was selectively labeled with a 1:2 mixture of the lanthanide metals Tb and Ho, respectively. Relative

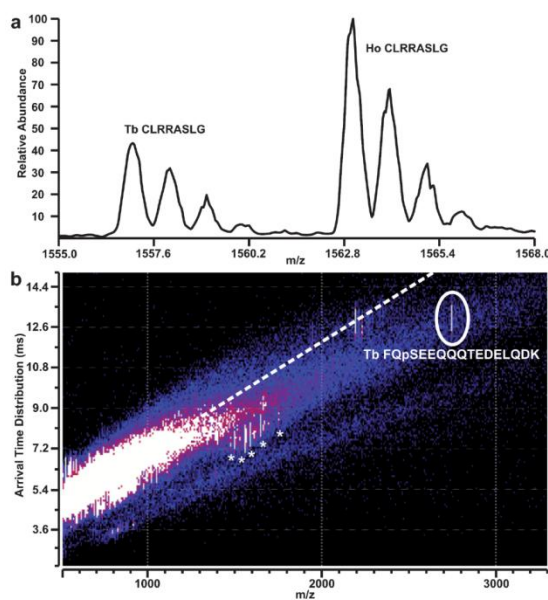


Figure 1.10. (a) The mass spectrum from a MALDI-TOFMS relative quantitation experiment in which the peptide CLRRASLG was selectively labeled at the cysteine residue with a 1:2 Tb:Ho mixture of lanthanide-based tags. Relative quantitation information is obtained by calculating the ratio of the isotopic envelopes for the Ho and Tb species. (b) The utility of lanthanide-based tags as ion mobility shift reagents is demonstrated. The model phosphoprotein β -casein (bovine) was digested with trypsin and labeled with a phosphopeptide selective lanthanide-based labeling reagent incorporating Tb. The peptide correlation line (white dashed line) is formed by the unlabeled tryptic peptides. The labeled phosphorylated peptide is shifted to a region of IM-MS conformation space which is not predicted to contain peptide signals. The peaks marked with an asterisk “*” correspond to excess reagents remaining from the labeling process. R.L. Gant-Branum, T.J. Kerr, J.A. McLean, Labeling Strategies in mass spectrometry-based protein quantitation, *Analyst*, **2009**, *134*, 1525-1530. – Reproduced by permission of The Royal Society of Chemistry.

quantitation was performed by calculating the ratio of the isotopic envelopes for the Tb and Ho-labeled species, whereby the measured Tb:Ho ratio for the data presented in Figure 1.10(a) was determined to be 1:2.04.¹²² Figure 1.10(b) demonstrates the utility of phosphopeptide selective lanthanide-based tags as IM shift reagents. The model phosphoprotein β -casein was proteolytically digested with trypsin and the resulting phosphopeptide, FQpSEEQQQTEDELQDK, was selectively labeled with a tag incorporating Tb. The unlabeled tryptic peptides were then used to establish the peptide correlation line, as indicated by the white dashed line in Figure 1.10(b). The lanthanide-labeled phosphopeptide is shifted to a region significantly below the peptide correlation line, providing a rapid means to locate phosphorylated peptides and improve accuracy in peak area analysis for relative quantitation experiments.^{122, 123}

1.3.2. Integrating Omic Analyses by IM-MS

An important feature of the different conformation space regions in which classes of molecules reside is the ability to integrate various omics measurements into the same measurement. Contemporary MS-based strategies in proteomics, glycomics, lipidomics, etc. involve extensive sample purification and preparation prior to the analysis to limit the identifications that are made to the molecular class of interest. In this manner, all other types of molecules are considered undesirable chemical noise. Conformation space analysis allows the simultaneous separation and identification of these various classes without these time consuming and potentially sample altering protocols. In the sections that follow, proteomics, glycomics, and lipidomics are treated separately, but it should be noted that conformational analysis provides a well suited tool for the comprehensive

characterization of these different molecular classes in the same analysis as necessary in nascent fields such as systems biology.

1.3.2.1. Proteomics

Mass spectrometry-based approaches such as of top-down¹²⁴ and bottom-up¹²⁵ proteomics have become standard practices in the pursuit of protein identifications from biological systems.^{126,127} Due to the broad range of protein concentrations in the biological samples (*e.g.* 10^{-3} to 10^{-24} M for human plasma¹²⁸) typically submitted for proteomic analysis, an additional dimension of separation is often implemented prior to detection by mass spectrometry to decrease the complexity of the sample and improve the dynamic range.¹²⁹⁻¹³¹ Among the commonly used separation techniques are two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and LC. Utilizing multiple dimensions of LC separation, techniques such as multidimensional protein identification technology (MudPIT) have been developed to further resolve sample complexity prior to MS in an online manner.^{132, 133} However, liquid-phase chromatography typically requires minutes to hours to achieve good separation, limiting the throughput for proteomic analyses.

Two dimensional gas-phase IM-MS separations are an alternative technique for reducing the complexity of samples for protein analysis and offer unique advantages for proteomic applications. In contrast to liquid-phase separations, gas-phase IM separations are significantly more rapid (μ s-ms) and more closely match the timescale of the MS separation.³⁸ Although the peak capacity (Φ) of IM-MS (5.5×10^3 for MALDI-IM-MS⁴⁰) is limited relative to more orthogonal two-dimensional separation techniques (6×10^7 for

LC-FTICR-MS⁴¹), IM-MS benefits from an exceptionally high peak capacity production rate ($2.8 \times 10^6 \Phi \text{ s}^{-1}$ for MALDI-IM-MS⁴⁰ vs. $1.25 \times 10^4 \Phi \text{ s}^{-1}$ for LC-FTICR-MS⁴¹). As discussed in Section 1.3.1.1., separation on the basis of structure in IM results in the formation of trendlines for each class of biomolecular species (see Figure 1.7). The presence of a peptide trendline in proteomic IM-MS analyses improves the confidence in identification while also increasing signal-to-noise of peptide species. Signals due to chemical noise (*i.e.* those corresponding to surfactants, MALDI matrices, or internal calibrants) occur in a different region of conformation space than the peptide signals and can be removed by extracting only the defined 2D region of correlation space containing the peptide trendline.³⁸ The location of a peptide signal about the average peptide correlation line in conformation space can provide insight into its secondary structure (*i.e.* α -helix, β -sheet, etc.) or the presence of PTMs, as discussed in Section 1.3.1.1. (see Figure 1.8).^{38,47,48,108,134,135} Additionally, it has been demonstrated that peptides with structural or sequence isomers can be resolved in the IM dimension due to differences in their adopted gas-phase structures.^{33, 136}

Much work has been done to further enhance the utility of IM-MS techniques for proteomic applications. One such goal has been the development of peptide drift time and collision cross section prediction algorithms based on intrinsic factors such as amino acid sequence and amino acid size.¹³⁷⁻¹⁴⁴ From the analysis of 113 lysine-terminated tryptic peptides, Clemmer and colleagues determined that collision cross section correlated with the amount of nonpolar or polar residues in the peptides, where nonpolar residues contributed more to the collision cross section than polar residues.¹³⁸ They demonstrated highly accurate calculations of collision cross sections from intrinsic amino acid size

parameters, with greater than 90% of calculated CCS values falling within 2% of the experiment CCS measurements.¹³⁸ An improvement on this approach was later developed by Clemmer and colleagues in which the position of the amino acid within the peptide sequence is taken into account by the size parameters, referred to as sequence-specific intrinsic size parameters (SSISPs).¹³⁹ Most recently, a method based on intrinsic amino acid size parameters has been proposed in which the ion mobility predictions can be utilized as a ranking tool to improve candidate peptide ion assignment.¹⁴³

From an instrumental perspective, the incorporation of a means by which peptide sequencing can be performed has been a key advance for IM-MS as a tool for proteomic analyses as primary sequence information is highly desirable for accurate protein identifications. IM-MS platforms have been developed around dissociation methods such as surface-induced dissociation (SID)¹⁴⁵ and collision-induced, or activated, dissociation (CID or CAD, respectively).^{42,84,103,146-149} As discussed in Section 1.1.2.1, placement of the collision cell between the IM and MS enables multiplexed fragmentation to be performed without precursor mass selection, as the fragment ions are essentially labeled with the drift time of the precursor ion.⁴² For greater separation of complex protein mixtures, a number of three dimensional LC-IM-MS and LC-IM-MS/MS platforms have been described.^{43,101,103,104,136,150-152} Relative to LC-MS techniques, incorporation of a dimension of IM separation enhances component resolution by more than one order of magnitude without increasing analysis times.¹⁵²

The viability and merits of IM-MS for proteomic applications are best demonstrated by its application proteome profiling experiments comparable to those typically performed by LC-MS- based platforms. One of the earliest examples was the

analysis of the human urinary proteome by LC-IM-MS and LC-IM-MS/MS techniques. Precursor and fragmentation data were collected separately and combined post-acquisition by alignment of retention and drift times (t_R and t_d , respectively). This data set of precursor and fragment ions was searched against a protein database and 13 proteins were identified from 27 peptides.¹⁰³ Importantly, this study demonstrated the benefit of IM separations in that peptide signals unresolved in the LC dimension were successfully resolved in the mobility dimension.

Several proteomic LC-IM-MS studies have focused on the model organism *Drosophila melanogaster*.^{104,105,150,153} In one such study, the use of SCX-LC-IM-MS analyses was described as a means to build proteome maps, in which peptides corresponding to a specific protein are reproducibly located within a multidimensional space.¹⁵⁰ This approach was applied in a study of the developmental phases of *D. melanogaster*, the whole embryo and the adult head, to identify the proteins present in both stages. Comparison of the proteome maps, containing the values of t_d , t_R , t_{SCX} , m/z and intensity for each feature, for the two development stages revealed that 307 of the 1133 proteins identified were observed in both samples.¹⁵⁰

Most impressively, mapping of the human plasma proteome has also been attempted with LC-IM-MS and SCX-LC-IM-MS techniques.^{102, 152, 154} For studies of such highly complex samples, the rapid nature of IM-MS techniques is greatly beneficial for sample throughput as multiple experiments are likely required to obtain maximal proteome coverage. For example, a 50 hr (10 runs, cumulative) LC-MS analysis of human plasma resulted in the identification of 300 proteins.¹⁵⁵ In contrast, 438 protein identifications were reported by Clemmer and colleagues from a single 3.3 hr LC-IM-

CID-MS analysis of human plasma.¹⁵² Proteomic profiling of plasma is particularly challenging as protein concentration range is approximately $10^{21,128}$. Frequently the most abundant proteins must be depleted to improve the experimental dynamic range for the detection of low abundance species.¹⁵⁶ With the use of 2D-LC-IM-MS techniques such as SCX-LC-IM-MS depletion of the most abundant plasma proteins is not typically necessary, as the mobility dimension can resolve the lower abundance species from coeluting high abundance species.^{151, 152}

1.3.2.2. Glycomics

The carbohydrates which form glycoconjugates such as glycoproteins and glycolipids derive their functions from the monosaccharide units from which they are composed and the structures those monomers collectively form.¹⁵⁷ The role of glycosylation in biological processes such as cell signaling and communication and its implication in certain types of cancer and in Alzheimer's disease have underscored the need of analytical techniques for rapid determination of the compositions and structures of the carbohydrates forming glycoprotein and glycolipid glycoconjugates.¹⁵⁷⁻¹⁶⁰ A number of MS-based approaches to glycoconjugate and carbohydrate characterization have been described.¹⁶¹⁻¹⁶⁶ This approach, however, can be challenging due to the complex branching patterns of glycans and the potential for numerous isobaric positional and/or structural isomers.^{161, 167, 168}

IM-MS techniques have been evaluated as an approach to the characterization of carbohydrates, as isobaric and isomeric carbohydrates can be resolved in the IM dimension due to differences in gas-phase conformation.^{158, 169-178} Early works focused on

small linear and cyclic (*e.g.* cyclodextrins) oligosaccharides, for which experimental collision cross sections of free or metal-coordinated oligosaccharides were determined.^{169-172,174} Comparison of these experimental collision cross sections with computationally derived values provided insight into the preferred conformations of metal-coordinated oligosaccharides, notably with alkali metals.^{170,172} Gas-phase structural separations (FAIMS and IM-MS) have also been demonstrated with mono and disaccharides as means to resolve carbohydrate anomeric configurations and stereoisomers.^{173, 179}

IM-MS has also been demonstrated for the separation of larger branched carbohydrates and their isobaric positional or structural isomers.^{158,175-177} Positional isomers, which differ only in the specific glycosidic linkage arrangement (*e.g.* 1→3 vs. 1→4), and structural isomers, which differ in branching patterns, have both been resolved by IM-MS.^{158, 177} The differences in the structures of the isomers results in differences in their gas-phase mobilities, and therefore drift times. This is depicted in Figure 1.11, which shows the arrival time distributions for three sets of isomeric carbohydrates. Overall, it is clear from panels (a) – (c) that smaller and more compact carbohydrates eluted from the drift cell more rapidly than the larger structures. Structures (1) – (4), shown on the left, are positional isomers, where carbohydrates (1) and (2) differ by the shift of both 1→3 glycosidic linkages to 1→4 linkages and carbohydrates (3) and (4) differ by the shift of one 1→3 linkage to a 1→4 linkage. As indicated in Figure 1.11a, the positional isomer (1) displayed a shorter drift time than that of positional isomer (2). For positional isomers (3) and (4), carbohydrate (3) showed a shorter drift time (Figure 1.11b). The shorter drift times of carbohydrates (1) and (3) relative to their respective

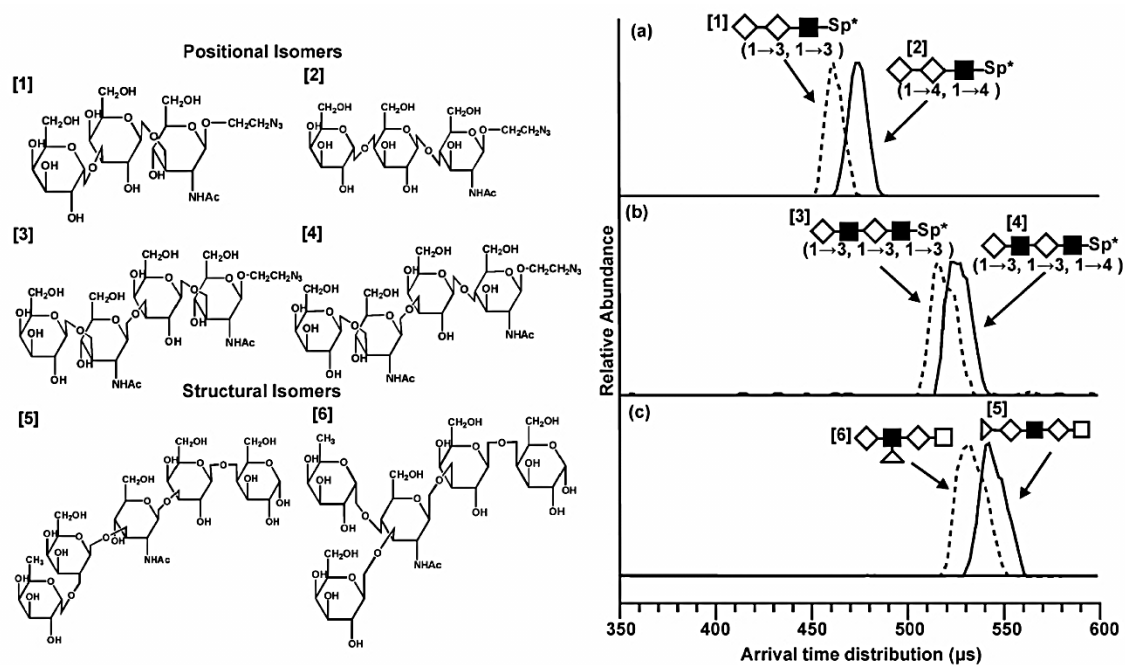


Figure 1.11. (Left) Three sets of isobaric positional or structural isomers. Positional isomers (1) and (2) differ by the change of both 1→3 glycosidic linkages in glycan (1) to 1→4 linkages in (2). Only one of the 1→3 linkage in (3) has been varied to a 1→4 linkage to form structure (4). Structures (5) and (6) represent structural isomers, in which a fucose monomer has been transferred from the galactose in (5) to the N-acetylglucosamine to form (6). (Right) The drift time profiles for the pairs of isomers. (a) Structure (1) (dotted line) has a shorter drift time than structure (2) (solid line) as the 1→4 linkages result in a more elongated structure. (b) The presence of only one 1→4 linkage increases the drift time of glycan (4) (solid line) relative to glycan (3) (dashed line), which has only 1→3 linkages. (c) The increased branching of glycan (6) (dashed line) results in a shorter drift time than the linear glycan (5) (solid line). L.S. Fenn and J.A. McLean, Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2196-2205. – Reproduced by permission of the PCCP Owner Societies.

positional isomers is indicative of a general trend in which carbohydrates with predominantly 1→3 glycosidic linkages have more compact structures than those with 1→4 linkages.¹⁵⁸

Carbohydrates (5) and (6), as shown in the bottom left of Figure 1.11, are isobaric structural isomers which are comprised of the same five monomers and have identical glycosidic linkages but differ in the placement of a fucose monomer. The shift of the fucose monomer from the non-reducing end of the linear carbohydrate (5) to the secondary N-acetylglucosamine creates a branched carbohydrate (6) with two non-reducing ends. The drift time profiles of carbohydrates (5) and (6) (Figure 1.11(c)) reveal that the branched carbohydrate (6) shows a shorter drift time than its linear structural isomer (5), indicating that (6) can adopt a more compact structure than (5) due to the site of branching.¹⁵⁸

For complex mixtures of glycans, such as those derived from glycoconjugates in biological samples, the ability to resolve isobaric isomeric species in the IM dimension significantly reduces spectral complexity and improves *de novo* sequencing capabilities.^{158,175,176} An IM-MS analysis of the N-linked glycans enzymatically released from model glycoprotein ovalbumin identified a total of 42 distinct isomers and/or conformers from 19 *m/z* values corresponding different glycan structures.¹⁸⁰ Large scale studies of complex glycan mixtures from biological samples such as serum and urine have been performed by LC-IM-MS to demonstrate its potential for revealing disease-related differences in glycan expression.^{175, 176}

An IM-MS approach for the simultaneous analysis of glycoproteins has also been described.¹⁷⁸ Typical MS-based techniques for glycoprotein analysis require that the

carbohydrates and peptides be separated and analyzed individually, requiring extensive and time consuming sample preparation.^{161,166} Based on the separation of different classes of biomolecules in IM-MS conformation space,^{46,47} it is possible to simultaneously characterize both the deglycosylated peptides and their associated glycans.¹⁷⁸ The glycoprotein is proteolytically digested and the N-linked glycans are released from the resulting glycopeptides by addition of PNGase F.¹⁷⁸ The mixture of glycans and peptides is then analyzed by IM-MS. A plot of conformation space from the MALDI-IM-MS simultaneous glycomic analysis of bovine ribonuclease B (RNase B) is presented in Figure 1.12(a). The tryptic peptides adopt more elongated structures than the glycans and are structurally separated into their respective mobility-mass correlations,^{46,47} as indicated by the white and black dashed lines, respectively. The mass spectra corresponding to the two mobility-mass correlations (Figure 1.12b,c) can be isolated from each other, allowing for simultaneous analysis of both the peptides and the glycans without interference from overlapping isobaric species.¹⁷⁸

1.3.2.3. Lipidomics

Lipids are integral components of cells which function in roles such as cell structure, energy storage, and cell signaling.^{181,182} The recent realizations of their involvement in diabetes,¹⁸³ atherosclerosis,¹⁸⁴ and neurological disorders such as Alzheimer's disease¹⁸⁵ have propelled the development of MS-based lipidomics for lipid profiling in biological systems.^{186,187} However, the analysis of lipids by MS-based approaches poses several analytical challenges. Naturally occurring lipids occupy a narrow range of m/z space and interference from isobaric lipid species is common.¹⁰⁹ For

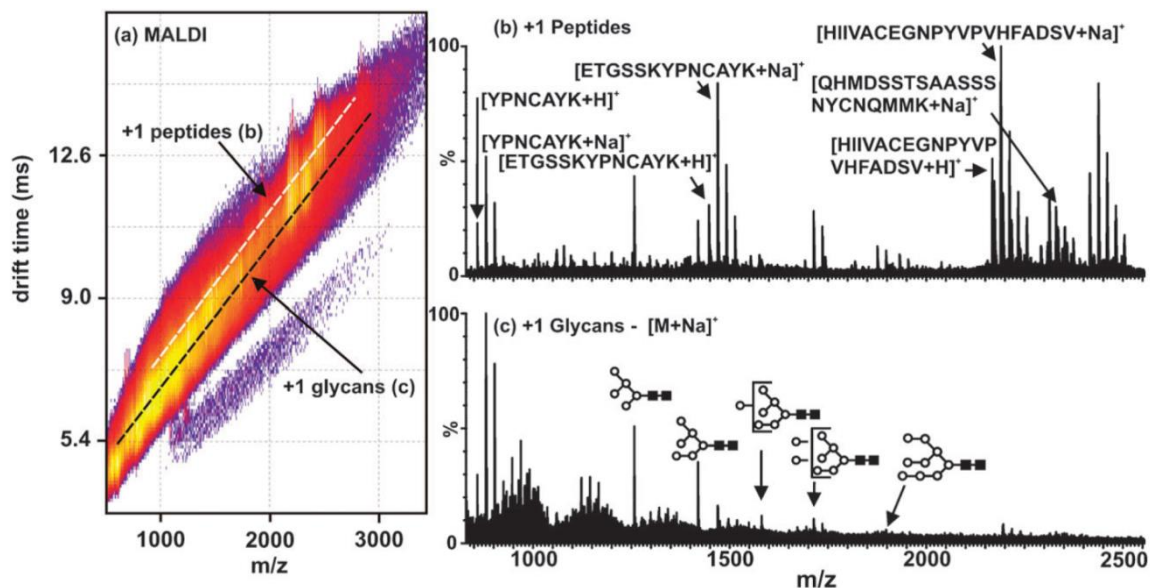


Figure 1.12. (a) A 2D plot of IM-MS conformation space from the MALDI-IM-MS analysis of digested and deglycosylated RNase B. The mobility-mass correlations for the singly charged peptides and glycans are indicated by the white and black dashed lines, respectively. (b) The mass spectrum corresponding to the tryptic peptides is obtained by extracting along the peptide correlation line (white dashed line in (a)). (c) The extracted mass spectrum corresponding to the glycans, obtained by extracting along the glycan correlation line (black dashed line in (b)). The open circles correspond to mannose and filled boxes correspond to *N*-acetylglucosamine. L.S. Fenn and J.A. McLean, Simultaneous glycoproteomics on the basis of structure using ion mobility-mass spectrometry, *Mol. BioSyst.* **2009**, 5, 1298-1302. – Reproduced by permission of The Royal Society of Chemistry.

the direct analysis of lipids from complex samples, the presence of exogenous and endogenous chemical noise, including some highly-abundant lipid species,¹⁸⁸ limits the detection and identification of low abundance lipid signals.^{109,189} Furthermore, the ionization preference for different classes of lipids are highly dependent on the chemical composition of the polar headgroup, making a single optimum of broad lipid ionization elusive

IM-MS-based approaches for lipidomics are highly amenable due to the diverse structures of the different sub-classes of lipids. A detailed study of the separation of phospholipids in IM-MS conformation space has been performed by Woods and colleagues.¹⁹⁰ Standard extracts of phosphatidic acid (PA), phosphatidylcholine (PC), -ethanolamine (PE), -glycerol (PG), -serine (PS), -inositol (PI) and sphingomyelin (SM) were analyzed by MALDI-IM-MS to obtain their mobility-mass correlations. Although both phosphatidylcholines and sphingomyelins contain choline head groups, these two species were consistently separated, with SM species generally having longer drift times, due to the differences in backbone structure. The effect of cationization was also investigated for its utility as lipid IM-MS shift reagents. The addition of an alkali ion was demonstrated to significantly reduce IM drift times, with the exception of cesium adduct formation which greatly shifted both the m/z and drift time. In general, the different phospholipid species occupy different regions of IM-MS conformation space due to the differences in head group, acyl chain degree of unsaturation and length, and adduct formation.^{47,48,109,190}

The effects of phosphatidylcholine acyl chain degree of unsaturation was further investigated by Kim and coworkers with ESI-IM-MS.¹⁹¹ The addition of a single double

bond in the acyl chain resulted in a 5% decrease in drift time, with a further reduction by 1% per each additional double bond. To understand the structural source of the observed reductions in drift time, molecular dynamics simulations were performed for several saturated and unsaturated PCs. Evaluation of the computationally generated structures indicated an increase in the structural rigidity of the unsaturated acyl chain due to the presence of the cis double bond, resulting in a decrease of the collision cross section of the unsaturated PCs relative to the saturated PCs.¹⁹¹

A number of studies have utilized the IM-MS approach for the direct analysis of lipids from tissues and, in particular, brain tissues.^{35,189,192,193} Lipid changes in a *Drosophila* model of epilepsy were measured directly from microdissected *Drosophila* brains with MALDI-IM-MS by Kliman et al.¹⁹² Signals corresponding to lipids were extracted from IM-MS conformation space, reducing the chemical noise due to MALDI matrix signals. Spectral alignment and statistical analysis of the lipid MS data revealed significant changes in multiple biosynthetically related lipid species between the mutant and wild type control.¹⁹² MALDI-IM-MS imaging of lipids from directly tissues has also been demonstrated.^{35, 193}

1.3.3. Structural Biology by IM-MS

Noncovalent assemblies of multiple proteins, or protein complexes, act as machines to collectively perform of an array of cellular functions and the particular function a protein complex performs is highly correlated to its overall structure.^{194, 195} Electrospray ionization-mass spectrometry (ESI-MS) has proven to be an informative tool for the characterization of the structures of protein complexes, in that soft ionization

by ESI can leave the noncovalent assemblies intact and generally can conserve their condensed-phase quaternary structure.^{91,196-198} With molecular mass and charge state distribution information available from ESI-MS, insight into the stoichiometries and conformations of protein complexes can be derived.^{91, 196} Over the past decade, IM-MS has emerged as a complementary technique for studying the shape and conformations of protein complexes due to the direct structural information available from the IM dimension. Critical works in the development of IM-MS for protein complex characterization set out to determine the stability of noncovalent complexes in the gas-phase and were often paired with complementary data from classical methods of structural biology (*e.g.* X-ray crystallography, nuclear magnetic resonance, electron microscopy) to evaluate if condensed-phase structure was conserved in the absence of bulk solvent (Figure 1.13).^{36,89,106,199-202} In general, it has been shown that protein complexes do maintain their solution-phase structure under the condition of minimal collisional heating, or low activation voltages,¹⁹⁸ however, the application of higher activation voltages is often necessary for desolvation of the protein complex and optimal mass accuracy.^{203, 204} These competing parameters for optimal accuracy in both IM and MS data requires special considerations in the experimental approach to studying protein complexes. A detailed protocol addressing such issues is available from Robinson and colleagues.¹⁰⁶

At high activation voltages, IM-MS can be utilized to study the conformational transitions protein complexes undergo as they dissociate. To determine if dissociation of the tetrameric transthyretin complex occurs via unfolded intermediates, Ruotolo et al. measured the arrival time distribution of the tetramer at increasing activation energies.²⁰⁵

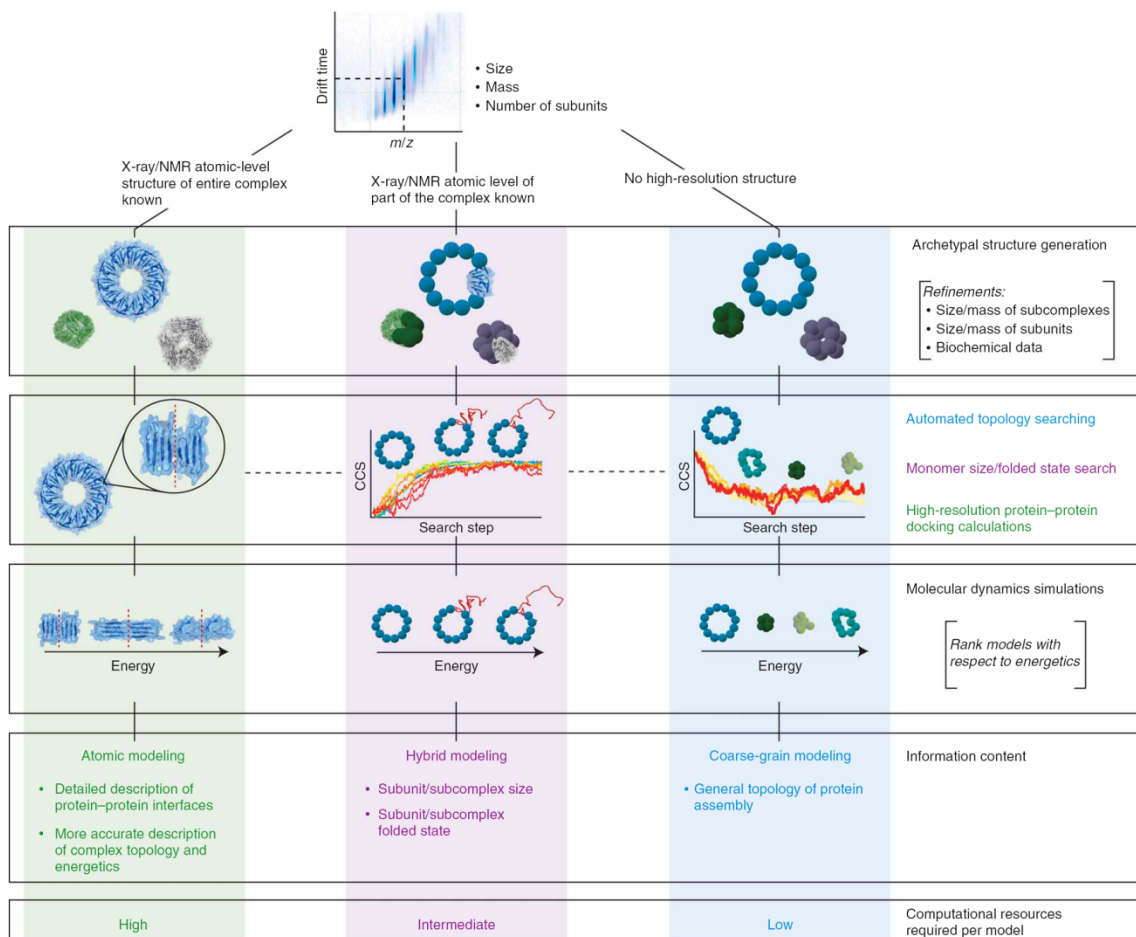


Figure 1.13. A flow diagram depicting three approaches for generating model structures of protein complexes, which can then be used to compare experimental and computational collision cross section measurements. The information available on the protein complex in the form of X-ray crystallography or NMR structures, as well as the computation resources available, dictates the detail of the model. Atomic and hybrid modeling require X-ray or NMR data for the entire complex or part of the complex, respectively, but are also highly demanding of computation resources. In contrast, coarse-grained modeling is appropriate when no such data is available. B.T. Ruotolo, J.L.P. Benesch, A.M. Sandercock, S.J. Hyung, and C.V. Robinson, Ion mobility-mass spectrometry analysis of large protein complexes, *Nature Protocols*, **2008**, *3*, 1139-1152. – Reproduced by permission of Nature Publishing Group.

It was observed that the activated protein complex had a significantly broadened arrival time distribution relative to the ATD observed without activation, suggesting the presence of extended intermediate structures resulting from activation. For the 15+ charge state, three populations of structures with collision cross sections (CCSs) greater than that of the unactivated species were observed at medium activation energy. Higher activation energy resulted in a single population of even larger CCS. These results were compared with CCSs calculated for a series of increasingly unfolded structures generated by molecular modeling simulations. Experimental CCSs for the 15+ charge state of the activated tetramer were found to closely match model structures in which one monomer was up to 80% unfolded and model structures in which multiple monomers had partially unfolded. These results indicated that dissociation of activated protein complexes occurs through a mechanism which generates partially unfolded intermediate structures.²⁰⁵

Additionally works have focused on the dependence of complex dissociation upon factors such as charge state and subunit packing.²⁰⁶⁻²¹⁰ Collision induced dissociation (CID) of the dodecameric and toroidially-shaped heat shock protein 16.9 (HSP16.9) and stable protein 1 (SP-1) protein complexes demonstrated that they dissociated through different pathways.²⁰⁷ While HSP16.9 dissociates by ejected unfolded monomers, SP-1 dissociates ejects monomers as well as dimers and multimers of the dimers. Investigation of these alternate dissociation pathways of similarly shaped protein complexes by IM-MS/MS revealed a charge state dependence in the dissociation of multiple dimers from SP-1, where the occurrence of multimer loss increased with higher charge states. The unfolding and dissociation of HSP16.9 and SP-1 were monitored as a function of collision energy. For HSP16.9, dissociation followed the pattern most typically observed

for multiprotein, in which unfolding begins to occur at low energies and is followed by dissociation at higher energies. In contrast, the behavior of SP-1 was more atypical in that it simultaneously undergoes unfolding and dissociation at the same energy. From inspection of the SP-1 crystal structure, it was concluded that the resistance to unfolding prior to dissociation observed in SP-1 is due to the inability of terminal residues to initiate unfolding because they are not exposed on the surface of the complex.²⁰⁷ Such studies have demonstrated the critical influence of quaternary structure and charge density on the dissociation pathways of macromolecular protein assemblies, and initiated further studies into the effects of charge state manipulation (*e.g.* charge reduction and supercharging) on dissociation pathways.^{206-208, 211}

1.4. Conclusions

The historical evolution of IM and IM-MS is characterized by periods of rapid and explosive growth in specific application areas interspersed with periods of slow maturation in these areas. The record is clear that IM-MS is experiencing a renaissance of this type as a multidimensional separations strategy in life sciences research, both for the analysis of complex biological samples and as a new tool for structural biology studies. The rapid separations afforded by IM-MS and the reduction of chemical noise in omic applications provide compelling reasons for adopting its utilization in systems biology research. Higher confidence assignments in allied omic areas like proteomics and lipidomics can be obtained by incorporating retention time, drift time, and accurate mass measurements. Similar to the advances in MS-based approaches to these endeavors, considerable research in smoothly integrating bioinformatics and biostatistics with these

multidimensional workflows are ongoing research goals. Our proclivity for generating vast amounts of data for biological systems is nearly unsurpassed, but translation of that data to desired biological information remains the rate limiting step.

Technological advances will certainly continue in this vibrant research area and wider adoption of these techniques in shotgun omics, spatial imaging IM-MS, and temporal studies of biological phenomena will presumably be forthcoming. In part this is facilitated by broader commercial offerings of these techniques to a wider biological community. Nevertheless, it is important to be vigilant of the past and the unique physical aspects of ion mobility in contrast with its bigger, but younger, sibling mass spectrometry.

1.5. Summary and Objectives of Dissertation Research

The identification of biomolecular signatures of disease from complex biological samples holds great promise for the discovery of potential diagnostic or prognostic markers as well as new avenues of therapeutics and preventative treatments.²¹² While the variety and complexity of diseases has led to a multitude of analytical methodologies for the study of complex biological samples, the scope of research has often been reduced to a singular class of biomolecules, resulting in the specialized fields of glycomics, lipidomics, metabolomics, and proteomics. Such reductionist approaches to complex biological sample analysis require rigorous and time-consuming sample preparation procedures to isolate the desired molecular species and significantly alter the original state of the biological sample in the process. While these approaches have greatly increased the knowledge of disease processes, the reduction of biological problems such

as disease to a single class of biomolecular species provides only a narrowed view of a naturally broad and intricate system.²¹³⁻²¹⁵

As described in sections *1.1.2.2 Advantages for Complex Biological Sample Analysis* and *1.3.1 Mapping the Conformation Space of Biomolecules*, ion mobility-mass spectrometry is well-suited for the analysis of complex biological samples given its ability to structurally resolve molecules arising from different biomolecular classes. My dissertation research in the McLean laboratory has focused on developing workflows for applying ion mobility and mass spectrometry-based techniques to the untargeted identification of biomolecular signatures of disease. Ion mobility-mass spectrometry approaches for revealing such biomolecular signatures from complex samples offer advantages including: (i) reduced sample preparation and purification, and preservation of the native sample complexity; (ii) improved throughput by combining multiple “omics” experiments into one analysis; and (iii) direct observation of the interdependencies between different types of biomolecules, *i.e.* proteins and their metabolites. Methodologies which combine IM-MS, chromatographic separations, specialized sample preparation, bioinformatics strategies, and statistically-directed transitions from untargeted to targeted analyses have been developed to reveal the biomolecular signatures most significantly distinguishing disease from non-disease states. Chapter II describes the initial effort toward these goals using an animal model of wound healing in diabetes, in which biomolecular species ranging from proteins to lipids were distinguished diabetic from non-diabetic wound fluids over a two-point time course experiment. Chapter III expands upon the strategies developed in Chapter II to distinguishing cancerous, benign and healthy breast tissues based on the differential

expression of biomolecular species, including a number of which are well known to be aberrantly overexpressed in cancer. Chapter IV demonstrates a more established approach to biomolecular signatures of disease by extensively mapping the phosphorylation of guanine nucleotide exchange factor Asef2 using multiple proteolysis strategies, and characterizing the effects of phosphorylation at S106 on adhesion turnover and cell migration. Finally, Chapter V closes with perspectives on future directions for both the works described here and, more broadly, for the development of IM-MS as a platform for biomolecular signature identification.

1.6. Acknowledgements

This chapter contains the invited book chapter for the Encyclopedia of Analytical Chemistry (online): “Multidimensional Separations by Ion Mobility-Mass Spectrometry,” by Kelly M. Hines, Jeffrey R. Enders, and John A. McLean. Robert Myers and David Muddiman, Eds. John Wiley and Sons, 2012.

We specifically thank Prof. Jody C. May (Vanderbilt University, Department of Chemistry) for insightful discussions in the preparation of this manuscript and to the many colleagues at Vanderbilt University and abroad for their part in works discussed herein. Financial support for this work was provided by the U.S. Defense Threat Reduction Agency (HDTRA-09-1-0013), the National Institutes of Health (R01GM092218 and RC2DA028981), the Vanderbilt University College of Arts and Sciences, the Vanderbilt Institute of Chemical Biology, and the Vanderbilt Institute of Integrative Biosystems Research and Education.

1.7. References

1. Sauer, U., High-throughput phenomics: experimental methods for mapping fluxomes. *Current Opinion in Biotechnology* **2004**, *15* (1), 58-63.
2. Tang, Y. J.; Martin, H. G.; Myers, S.; Rodriguez, S.; Baidoo, E. E. K.; Keasling, J. D., Advances in analysis of microbial metabolic fluxes via (13)C isotopic labeling. *Mass Spectrometry Reviews* **2009**, *28* (2), 362-375.
3. Metz, T. O.; Zhang, Q.; Page, J. S.; Shen, Y.; Callister, S. J.; Jacobs, J. M.; Smith, R. D., Future of liquid chromatography-mass spectrometry in metabolic profiling and metabolomic studies for biomarker discovery. *Biomarkers in Medicine* **2007**, *1* (1), 159-185.
4. Lu, X.; Zhao, X.; Bai, C.; Zhao, C.; Lu, G.; Xu, G., LC-MS-based metabonomics analysis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2008**, *866* (1-2), 64-76.
5. Griffiths, W. J.; Wang, Y., Mass spectrometry: from proteomics to metabolomics and lipidomics. *Chemical Society Reviews* **2009**, *38* (7), 1882-1896.
6. Qian, W. J.; Jacobs, J. M.; Liu, T.; Camp, D. G.; Smith, R. D., Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Molecular & Cellular Proteomics* **2006**, *5* (10), 1727-1744.
7. Rontgen, W. C., On a new kind of rays. *Science* **1896**, *3* (59), 227-31.
8. Thomson, J. J.; Rutherford, E., On the passage of electricity through gases exposed to Rontgen rays. *Philos. Mag. (1798-1977)* **1896**, *42* (5), 392.
9. Zeleny, J., VI. On the ratio of the velocities of the two ions produced in gases by Röntgen radiation; and on some related phenomena. *Philosophical Magazine Series 5* **1898**, *46* (278), 120-154.
10. Kingdon, K. H., A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. *Physical Review* **1923**, *21* (4), 408-418.
11. Makarov, A., Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **2000**, *72* (6), 1156-1162.
12. Stephens, W. E., A pulsed mass spectrometer with time dispersion. *Phys. Rev.* **1946**, *69*, 691.
13. Mamyrin, B. A., Russian Patent No. 198034. 1966.

14. Karataev, V. I.; Mamyrin, B. A.; Shmikk, D. A., New principle of the focussing of ion packets in time-of-flight mass spectrometers. *Zh. Tekh. Fiz.* **1971**, *41* (7), 1498-501.
15. Mamyrin, B. A.; Karataev, V. I.; Shmikk, D. V.; Zagulin, V. A., Mass reflectron. New nonmagnetic time-of-flight high-resolution mass spectrometer. *Zh. Eksp. Teor. Fiz.* **1973**, *64* (1), 82-9.
16. Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988**, *60* (20), 2299-301.
17. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yohida, T., Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **1988**, *2* (8), 151-3.
18. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass-spectrometry of large biomolecules. *Science* **1989**, *246* (4926), 64-71.
19. Tyndall, A. M.; Grindley, G. C., Mobility of ions in air. I. Negative ions in moist air. *Proc. R. Soc. London, Ser. A* **1926**, *110*, 341-58.
20. Tyndall, A. M.; Grindley, G. C., Mobility of ions in air. II. Positive ions of short age. *Proc. R. Soc. London, Ser. A* **1926**, *110*, 358-64.
21. Tyndall, A. M.; Phillips, L. R., Mobility of ions in air. III. Air containing organic vapors. *Proc. R. Soc. London, Ser. A* **1926**, *111*, 577-91.
22. Tyndall, A. M.; Starr, L. H.; Powell, C. F., Mobility of ions in air. IV. Investigations by two new methods. *Proc. R. Soc. London, Ser. A* **1928**, *121*, 172-84.
23. Tyndall, A. M.; Grindley, G. C.; Sheppard, P. A., Mobility of ions in air.V. Transformation of positive ions at short ages. *Proc. R. Soc. London, Ser. A* **1928**, *121*, 185-94.
24. Barnes, W. S.; Martin, D. W.; McDaniel, E. W., Mass spectrographic identification of the ion observed in hydrogen mobility experiments. *Phys. Rev. Lett.* **1961**, *6*, 110-11.
25. McAfee, K. B., Jr.; Edelson, D., Identification and mobility of ions in a Townsend discharge by time-resolved mass spectrometry. *Proc. Phys. Soc., London* **1963**, *81* (520), 382-4.
26. Cohen, M. J.; Karasek, F. W., Plasma chromatography TM--new dimension for gas chromatography and mass spectrometry. *J. Chromatogr. Sci.* **1970**, *8* (6), 330-7.

27. Karasek, F. W., Plasma chromatography of the polychlorinated biphenyls. *Anal Chem* **1971**, *43* (14), 1982-6.
28. Karasek, F. W., Trace analysis and fundamental studies by plasma chromatography. *Int J Environ Anal Chem* **1972**, *2* (2), 157-66.
29. Karasek, F. W.; Denney, D. W., Evaluation of the plasma chromatograph as a qualitative detector for liquid chromatography. *Anal. Lett.* **1973**, *6* (11), 993-1004.
30. Gieniec, J.; Mack, L. L.; Nakamae, K.; Gupta, C.; Kumar, V.; Dole, M., Electrospray mass spectroscopy of macromolecules: application of an ion-drift spectrometer. *Biomed. Mass Spectrom.* **1984**, *11* (6), 259-68.
31. Vonhelden, G.; Wytenbach, T.; Bowers, M. T., Conformation of macromolecules in the gas-phase - use of matrix-assisted laser-desorption methods in ion chromatography. *Science* **1995**, *267* (5203), 1483-1485.
32. Wytenbach, T.; vonHelden, G.; Bowers, M. T., Gas-phase conformation of biological molecules: Bradykinin. *Journal of the American Chemical Society* **1996**, *118* (35), 8355-8364.
33. Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F., Naked protein conformations - cytochrome-c in the gas-phase. *Journal of the American Chemical Society* **1995**, *117* (40), 10141-10142.
34. Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F., Protein structure in vacuo: Gas-phase confirmations of BPTI and cytochrome c. *Journal of the American Chemical Society* **1997**, *119* (9), 2240-2248.
35. McLean, J. A.; Ridenour, W. B.; Caprioli, R. M., Profiling and imaging of tissues by imaging ion mobility-mass spectrometry. *Journal of Mass Spectrometry* **2007**, *42* (8), 1099-1105.
36. 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* (5754), 1658-1661.
37. Giddings, J. C., Two-dimensional separations - concept and promise. *Analytical Chemistry* **1984**, *56* (12), 1258-&.
38. McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H., Ion mobility-mass spectrometry: a new paradigm for proteomics. *International Journal of Mass Spectrometry* **2005**, *240* (3), 301-315.
39. Ruotolo, B. T.; Gillig, K. J.; Stone, E. G.; Russell, D. H., Peak capacity of ion mobility mass spectrometry: Separation of peptides in helium buffer gas. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002**, *782* (1-2), 385-392.

40. Ruotolo, B. T.; McLean, J. A.; Gillig, K. J.; Russell, D. H., Peak capacity of ion mobility mass spectrometry: the utility of varying drift gas polarizability for the separation of tryptic peptides. *Journal of Mass Spectrometry* **2004**, *39* (4), 361-367.
41. Shen, Y. F.; Tolic, N.; Zhao, R.; Pasa-Tolic, L.; Li, L. J.; Berger, S. J.; Harkewicz, R.; Anderson, G. A.; Belov, M. E.; Smith, R. D., High-throughput proteomics using high efficiency multiple-capillary liquid chromatography with on-line high-performance ESI FTICR mass spectrometry. *Analytical Chemistry* **2001**, *73* (13), 3011-3021.
42. Hoadlund-Hyzer, C. S.; Li, J. W.; Clemmer, D. E., Mobility labeling for parallel CID of ion mixtures. *Analytical Chemistry* **2000**, *72* (13), 2737-2740.
43. Valentine, S. J.; Kulchania, M.; Barnes, C. A. S.; Clemmer, D. E., Multidimensional separations of complex peptide mixtures: a combined high-performance liquid chromatography/ion mobility/time-of-flight mass spectrometry approach. *International Journal of Mass Spectrometry* **2001**, *212* (1-3), 97-109.
44. Koomen, J. M.; Ruotolo, B. T.; Gillig, K. J.; McLean, J. A.; Russell, D. H.; Kang, M. J.; Dunbar, K. R.; Fuhrer, K.; Gonin, M.; Schultz, J. A., Oligonucleotide analysis with MALDI-ion-mobility-TOFMS. *Analytical and Bioanalytical Chemistry* **2002**, *373* (7), 612-617.
45. Woods, A. S.; Ugarov, M.; Egan, T.; Koomen, J.; Gillig, K. J.; Fuhrer, K.; Gonin, M.; Schultz, J. A., Lipid/peptide/nucleotide separation with MALDI-ion mobility-TOF MS. *Analytical Chemistry* **2004**, *76* (8), 2187-2195.
46. Fenn, L. S.; McLean, J. A., Biomolecular structural separations by ion mobility-mass spectrometry. *Analytical and Bioanalytical Chemistry* **2008**, *391* (3), 905-909.
47. 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. *Analytical and Bioanalytical Chemistry* **2009**, *394* (1), 235-244.
48. McLean, J. A., The Mass-Mobility Correlation Redux: The Conformational Landscape of Anhydrous Biomolecules. *Journal of the American Society for Mass Spectrometry* **2009**, *20* (10), 1775-1781.
49. Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F., Design of helices that are stable in vacuo. *Journal of the American Chemical Society* **1998**, *120* (49), 12974-12975.
50. Ruotolo, B. T.; Verbeck, G. F.; Thomson, L. M.; Gillig, K. J.; Russell, D. H., Observation of conserved solution-phase secondary structure in gas-phase tryptic peptides. *Journal of the American Chemical Society* **2002**, *124* (16), 4214-4215.

51. Ruotolo, B. T.; Russell, D. H., Gas-phase conformations of proteolytically derived protein fragments: Influence of solvent on peptide conformation. *Journal of Physical Chemistry B* **2004**, *108* (39), 15321-15331.
52. Counterman, A. E.; Clemmer, D. E., Cis-trans signatures of proline-containing tryptic peptides in the gas phase. *Analytical Chemistry* **2002**, *74* (9), 1946-1951.
53. Shvartsburg, A. A.; Jarrold, M. F., An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chemical Physics Letters* **1996**, *261* (1-2), 86-91.
54. 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. *Journal of Physical Chemistry* **1996**, *100* (40), 16082-16086.
55. Wyttenbach, T.; vonHelden, G.; Batka, J. J.; Carlat, D.; Bowers, M. T., Effect of the long-range potential on ion mobility measurements. *Journal of the American Society for Mass Spectrometry* **1997**, *8* (3), 275-282.
56. Wyttenbach, T.; Witt, M.; Bowers, M. T., On the stability of amino acid zwitterions in the gas phase: The influence of derivatization, proton affinity, and alkali ion addition. *Journal of the American Chemical Society* **2000**, *122* (14), 3458-3464.
57. Hudgins, R. R.; Jarrold, M. F., Conformations of unsolvated glycine-based peptides. *Journal of Physical Chemistry B* **2000**, *104* (9), 2154-2158.
58. Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F., Mobilities of carbon cluster ions: Critical importance of the molecular attractive potential. *Journal of Chemical Physics* **1998**, *108* (6), 2416-2423.
59. Bernstein, S. L.; Wyttenbach, T.; Baumketner, A.; Shea, J. E.; Bitan, G.; Teplow, D. B.; Bowers, M. T., Amyloid beta-protein: Monomer structure and early aggregation states of A beta 42 and its Pro(19) alloform. *Journal of the American Chemical Society* **2005**, *127* (7), 2075-2084.
60. Jarrold, M. F. MOBCAL - A Program to Calculate Mobilities. <http://www.indiana.edu/~nano/software.html> (accessed April 30).
61. Bowers, M. T. Theory/Analysis. http://bowers.chem.ucsb.edu/theory_analysis/ (accessed April 30).
62. Karasek, F. W.; Hill, H. H.; Kim, S. H., Plasma chromatography of heroin and cocaine with mass-identified mobility spectra. *Journal of Chromatography* **1976**, *117* (2), 327-336.
63. Hudgins, R. R.; Woenckhaus, J.; Jarrold, M. F., High resolution ion mobility measurements for gas phase proteins: correlation between solution phase and gas

- phase conformations. *International Journal of Mass Spectrometry* **1997**, *165*, 497-507.
64. Liu, Y. S.; Valentine, S. J.; Counterman, A. E.; Hoaglund, C. S.; Clemmer, D. E., Injected-ion mobility analysis of biomolecules. *Analytical Chemistry* **1997**, *69* (23), A728-A735.
 65. Wu, C.; Siems, W. F.; Asbury, G. R.; Hill, H. H., Electrospray ionization high-resolution ion mobility spectrometry - Mass spectrometry. *Analytical Chemistry* **1998**, *70* (23), 4929-4938.
 66. Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E., Three-dimensional ion mobility TOFMS analysis of electrosprayed biomolecules. *Analytical Chemistry* **1998**, *70* (11), 2236-2242.
 67. Gillig, K. J.; Ruotolo, B.; Stone, E. G.; Russell, D. H.; Fuhrer, K.; Gonin, M.; Schultz, A. J., Coupling high-pressure MALDI with ion mobility/orthogonal time-of flight mass spectrometry. *Analytical Chemistry* **2000**, *72* (17), 3965-3971.
 68. Vonhelden, G.; Wyttenbach, T.; Bowers, M. T., Inclusion of a MALDI ion source in the ion chromatography technique - conformational information on polymer and biomolecular ions. *International Journal of Mass Spectrometry and Ion Processes* **1995**, *146*, 349-364.
 69. Hoaglund, C. S.; Valentine, S. J.; Clemmer, D. E., An ion trap interface for ESI-ion mobility experiments. *Analytical Chemistry* **1997**, *69* (20), 4156-4161.
 70. Creaser, C. S.; Benyazzar, M.; Griffiths, J. R.; Stygall, J. W., A tandem ion trap/ion mobility spectrometer. *Analytical Chemistry* **2000**, *72* (13), 2724-2729.
 71. Bluhm, B. K.; Gillig, K. J.; Russell, D. H., Development of a Fourier-transform ion cyclotron resonance mass spectrometer-ion mobility spectrometer. *Review of Scientific Instruments* **2000**, *71* (11), 4078-4086.
 72. Javahery, G.; Thomson, B., A segmented radiofrequency-only quadrupole collision cell for measurements of ion collision cross section on a triple quadrupole mass spectrometer. *Journal of the American Society for Mass Spectrometry* **1997**, *8* (7), 697-702.
 73. Guo, Y. Z.; Ling, Y.; Thomson, B. A.; Siu, K. W. M., Combined ion-mobility and mass-spectrometry investigations of metallothionein complexes using a tandem mass spectrometer with a segmented second quadrupole. *Journal of the American Society for Mass Spectrometry* **2005**, *16* (11), 1787-1794.
 74. Guo, Y. Z.; Wang, J. X.; Javahery, G.; Thomson, B. A.; Siu, K. W. M., Ion mobility spectrometer with radial collisional focusing. *Analytical Chemistry* **2005**, *77* (1), 266-275.

75. Guevremont, R., High-field asymmetric waveform ion mobility spectrometry (FAIMS). *Canadian Journal of Analytical Sciences and Spectroscopy* **2004**, *49* (3), 105-113.
76. Shvartsburg, A. A.; Li, F. M.; Tang, K. Q.; Smith, R. D., High-resolution field asymmetric waveform ion mobility spectrometry using new planar geometry analyzers. *Analytical Chemistry* **2006**, *78* (11), 3706-3714.
77. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H., Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Communications in Mass Spectrometry* **2004**, *18* (20), 2401-2414.
78. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinou, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H., An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry* **2007**, *261* (1), 1-12.
79. Shumate, C. B.; Hill, H. H., Coronaspray nebulization and ionization of liquid samples for ion mobility spectrometry. *Analytical Chemistry* **1989**, *61* (6), 601-606.
80. Wytenbach, T.; Kemper, P. R.; Bowers, M. T., Design of a new electrospray ion mobility mass spectrometer. *International Journal of Mass Spectrometry* **2001**, *212* (1-3), 13-23.
81. Hoaglund-Hyzer, C. S.; Lee, Y. J.; Counterman, A. E.; Clemmer, D. E., Coupling ion mobility separations, collisional activation techniques, and multiple stages of MS for analysis of complex peptide mixtures. *Analytical Chemistry* **2002**, *74* (5), 992-1006.
82. Valentine, S. J.; Koeniger, S. L.; Clemmer, D. E., A split-field drift tube for separation and efficient fragmentation of biomolecular ions. *Analytical Chemistry* **2003**, *75* (22), 6202-6208.
83. Clowers, B. H.; Hill, H. H., Mass analysis of mobility-selected ion populations using dual gate, ion mobility, quadrupole ion trap mass spectrometry. *Analytical Chemistry* **2005**, *77* (18), 5877-5885.
84. Steiner, W. E.; Clowers, B. H.; English, W. A.; Hill, H. H., Atmospheric pressure matrix-assisted laser desorption/ionization with analysis by ion mobility time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **2004**, *18* (8), 882-888.
85. Sundarapandian, S.; May, J. C.; McLean, J. A., Dual Source Ion Mobility-Mass Spectrometer for Direct Comparison of Electrospray Ionization and MALDI

- Collision Cross Section Measurements. *Analytical Chemistry* **2010**, 82 (8), 3247-3254.
86. Shelimov, K. B.; Jarrold, M. F., Conformations, unfolding, and refolding of apomyoglobin in vacuum: An activation barrier for gas-phase protein folding. *Journal of the American Chemical Society* **1997**, 119 (13), 2987-2994.
 87. Myung, S.; Badman, E. R.; Lee, Y. J.; Clemmer, D. E., Structural transitions of electrosprayed ubiquitin ions stored in an ion trap over similar to 10 ms to 30 s. *Journal of Physical Chemistry A* **2002**, 106 (42), 9976-9982.
 88. Scarff, C. A.; Thalassinou, K.; Hilton, G. R.; Scrivens, J. H., Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Rapid Communications in Mass Spectrometry* **2008**, 22 (20), 3297-3304.
 89. Kaddis, C. S.; Lomeli, S. H.; Yin, S.; Berhane, B.; Apostol, M. I.; Kickhoefer, V. A.; Rome, L. H.; Loo, J. A., Sizing large proteins and protein complexes by electrospray ionization mass spectrometry and ion mobility. *Journal of the American Society for Mass Spectrometry* **2007**, 18 (7), 1206-1216.
 90. Winston, R. L.; Fitzgerald, M. C., Mass spectrometry as a readout of protein structure and function. *Mass Spectrometry Reviews* **1997**, 16 (4), 165-179.
 91. Loo, J. A., Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *International Journal of Mass Spectrometry* **2000**, 200 (1-3), 175-186.
 92. Asbury, G. R.; Hill, H. H., Using different drift cases to change separation factors (α) in ion mobility spectrometry. *Analytical Chemistry* **2000**, 72 (3), 580-584.
 93. Verbeck, G. F.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H., Resolution equations for high-field ion mobility. *Journal of the American Society for Mass Spectrometry* **2004**, 15 (9), 1320-1324.
 94. Rokushika, S.; Hatano, H.; Baim, M. A.; Hill, H. H., Resolution measurement for ion mobility spectrometry. *Analytical Chemistry* **1985**, 57 (9), 1902-1907.
 95. Watts, P.; Wilders, A., On the resolution obtainable in practical ion mobility systems. *International Journal of Mass Spectrometry and Ion Processes* **1992**, 112 (2-3), 179-190.
 96. McLean, J. A.; Russell, D. H., Sub-femtomole peptide detection in ion mobility-time-of-flight mass spectrometry measurements. *Journal of Proteome Research* **2003**, 2 (4), 427-430.

97. Gillig, K. J.; Ruotolo, B. T.; Stone, E. G.; Russell, D. H., An electrostatic focusing ion guide for ion mobility-mass spectrometry. *International Journal of Mass Spectrometry* **2004**, 239 (1), 43-49.
98. McLean, J. A.; Russell, D. H.; Egan, T. F.; Ugarov, M. V.; Schultz, J. A. Multiplex data acquisition modes for ion mobility-mass spectrometry. US20060024720A1, 2006.
99. Tang, K.; Shvartsburg, A. A.; Lee, H. N.; Prior, D. C.; Buschbach, M. A.; Li, F. M.; Tolmachev, A. V.; Anderson, G. A.; Smith, R. D., High-sensitivity ion mobility spectrometry/mass spectrometry using electrodynamic ion funnel interfaces. *Analytical Chemistry* **2005**, 77 (10), 3330-3339.
100. Koeniger, S. L.; Merenbloom, S. I.; Valentine, S. J.; Jarrold, M. F.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E., An IMS-IMS analogue of MS-MS. *Analytical Chemistry* **2006**, 78 (12), 4161-4174.
101. Lee, Y. J.; Hoaglund-Hyzer, C. S.; Barnes, C. A. S.; Hilderbrand, A. E.; Valentine, S. J.; Clemmer, D. E., Development of high-throughput liquid chromatography injected ion mobility quadrupole time-of-flight techniques for analysis of complex peptide mixtures. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002**, 782 (1-2), 343-351.
102. Liu, X.; Plasencia, M.; Ragg, S.; Valentine, S. J.; Clemmer, D. E., Development of high throughput dispersive LC-ion mobility-TOFMS techniques for analysing the human plasma proteome. *Briefings in Functional Genomics & Proteomics* **2004**, 3 (2), 177-186.
103. Moon, M. H.; Myung, S.; Plasencia, M.; Hilderbrand, A. E.; Clemmer, D. E., Nanoflow LC/ion mobility/CID/TOF for proteomics: analysis of a human urinary Proteome. *Journal of Proteome Research* **2003**, 2 (6), 589-597.
104. Myung, S.; Lee, Y. J.; Moon, M. H.; Taraszka, J.; Sowell, R.; Koeniger, S.; Hilderbrand, A. E.; Valentine, S. J.; Cherbas, L.; Cherbas, P.; Kaufmann, T. C.; Miller, D. F.; Mechref, Y.; Novotny, M. V.; Ewing, M. A.; Sporleder, C. R.; Clemmer, D. E., Development of high-sensitivity ion trap ion mobility spectrometry time-of-flight techniques: A high-throughput nano-LC-IMS-TOF separation of peptides arising from a *Drosophila* protein extract. *Analytical Chemistry* **2003**, 75 (19), 5137-5145.
105. Taraszka, J. A.; Gao, X. F.; Valentine, S. J.; Sowell, R. A.; Koeniger, S. L.; Miller, D. F.; Kaufman, T. C.; Clemmer, D. E., Proteome profiling for assessing diversity: Analysis of individual heads of *Drosophila melanogaster* using LC-ion mobility-MS. *Journal of Proteome Research* **2005**, 4 (4), 1238-1247.
106. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V., Ion mobility-mass spectrometry analysis of large protein complexes. *Nature Protocols* **2008**, 3 (7), 1139-1152.

107. Williams, J. P.; Scrivens, J. H., Coupling desorption electrospray ionisation and neutral desorption/extractive electrospray ionisation with a travelling-wave based ion mobility mass spectrometer for the analysis of drugs. *Rapid Communications in Mass Spectrometry* **2008**, 22 (2), 187-196.
108. 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. *Journal of Proteome Research* **2002**, 1 (4), 303-306.
109. Kliman, M.; May, J. C.; McLean, J. A., Lipid analysis and lipidomics by structurally selective ion mobility-mass spectrometry. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **2011**, 1811 (11), 935-945.
110. Julian, R. R.; Beauchamp, J. L., Site specific sequestering and stabilization of charge in peptides by supramolecular adduct formation with 18-crown-6 ether by way of electrospray ionization. *International Journal of Mass Spectrometry* **2001**, 210 (1-3), 613-623.
111. Julian, R. R.; Akin, M.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L., Molecular recognition of arginine in small peptides by supramolecular complexation with dibenzo-30-crown-10 ether. *International Journal of Mass Spectrometry* **2002**, 220 (1), 87-96.
112. Julian, R. R.; Beauchamp, J. L., The unusually high proton affinity of aza-18-crown-6 ether: Implications for the molecular recognition of lysine in peptides by lariat crown ethers. *Journal of the American Society for Mass Spectrometry* **2002**, 13 (5), 493-498.
113. Julian, R. R.; Beauchamp, J. L., Selective molecular recognition of arginine by anionic salt bridge formation with bis-phosphate crown ethers: Implications for gas phase peptide acidity from adduct dissociation. *Journal of the American Society for Mass Spectrometry* **2004**, 15 (4), 616-624.
114. Lee, S.; Wyttenbach, T.; Vonhelden, G.; Bowers, M. T., Gas-phase conformations of Li⁺, Na⁺, K⁺, and Cs⁺ complexed with 18-crown-6. *Journal of the American Chemical Society* **1995**, 117 (40), 10159-10160.
115. Crowe, M. C.; Brodbelt, J. S., Evaluation of noncovalent interactions between peptides and polyether compounds via energy-variable collisionally activated dissociation. *Journal of the American Society for Mass Spectrometry* **2003**, 14 (10), 1148-1157.
116. Colgrave, M. L.; Bramwell, C. J.; Creaser, C. S., Nanoelectrospray ion mobility spectrometry and ion trap mass spectrometry studies of the non-covalent complexes of amino acids and peptides with polyethers. *International Journal of Mass Spectrometry* **2003**, 229 (3), 209-216.

117. Creaser, C. S.; Griffiths, J. R.; Stockton, B. M., Gas-phase ion mobility studies of amines and polyether/amine complexes using tandem quadrupole ion trap/ion mobility spectrometry. *European Journal of Mass Spectrometry* **2000**, *6* (2), 213-218.
118. Hilderbrand, A. E.; Myung, S.; Clemmer, D. E., Exploring crown ethers as shift reagents for ion mobility spectrometry. *Analytical Chemistry* **2006**, *78* (19), 6792-6800.
119. Whetstone, P. A.; Butlin, N. G.; Corneillie, T. M.; Meares, C. F., Element-coded affinity tags for peptides and proteins. *Bioconjugate Chemistry* **2004**, *15* (1), 3-6.
120. Ahrends, R.; Pieper, S.; Kuehn, 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. *Molecular & Cellular Proteomics* **2007**, *6* (11), 1907-1916.
121. Kerr, T. J.; McLean, J. A., Peptide quantitation using primary amine selective metal chelation labels for mass spectrometry. *Chemical Communications* **2010**, *46* (30), 5479-5481.
122. Gant-Branum, R. L.; Kerr, T. J.; McLean, J. A., Labeling strategies in mass spectrometry-based protein quantitation. *Analyst* **2009**, *134* (8), 1525-1530.
123. 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**, *307* (1-3), 28-32.
124. Fridriksson, E. K.; Beavil, A.; Holowka, D.; Gould, H. J.; Baird, B.; McLafferty, F. W., Heterogeneous glycosylation of immunoglobulin E constructs characterized by top-down high-resolution 2-D mass spectrometry. *Biochemistry* **2000**, *39* (12), 3369-3376.
125. Henzel, W. J.; Watanabe, C.; Stults, J. T., Protein identification: The origins of peptide mass fingerprinting. *Journal of the American Society for Mass Spectrometry* **2003**, *14* (9), 931-942.
126. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003**, *422* (6928), 198-207.
127. Cravatt, B. F.; Simon, G. M.; Yates, J. R., III, The biological impact of mass-spectrometry-based proteomics. *Nature* **2007**, *450* (7172), 991-1000.
128. Service, R. F., Proteomics - Public projects gear up to chart the protein landscape. *Science* **2003**, *302* (5649), 1316-1318.
129. Bischoff, R., Recent developments in proteomics. *Analytical and Bioanalytical Chemistry* **2003**, *376* (3), 289-291.

130. Manabe, T., Analysis of complex protein-polypeptide systems for proteomic studies. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2003**, 787 (1), 29-41.
131. Wang, H.; Hanash, S., Multi-dimensional liquid phase based separations in proteomics. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2003**, 787 (1), 11-18.
132. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology* **1999**, 17 (7), 676-682.
133. Wolters, D. A.; Washburn, M. P.; Yates, J. R., An automated multidimensional protein identification technology for shotgun proteomics. *Analytical Chemistry* **2001**, 73 (23), 5683-5690.
134. Ruotolo, B. T.; Gillig, K. J.; Woods, A. S.; Egan, T. F.; Ugarov, M. V.; Schultz, J. A.; Russell, D. H., Analysis of phosphorylated peptides by ion mobility-mass spectrometry. *Analytical Chemistry* **2004**, 76 (22), 6727-6733.
135. Thalassinos, K.; Grabenauer, M.; Slade, S. E.; Hilton, G. R.; Bowers, M. T.; Scrivens, J. H., Characterization of Phosphorylated Peptides Using Traveling Wave-Based and Drift Cell Ion Mobility Mass Spectrometry. *Analytical Chemistry* **2009**, 81 (1), 248-254.
136. Barnes, C. A. S.; Hilderbrand, A. E.; Valentine, S. J.; Clemmer, D. E., Resolving isomeric peptide mixtures: A combined HPLC/ion mobility-TOFMS analysis of a 4000-component combinatorial library. *Analytical Chemistry* **2002**, 74 (1), 26-36.
137. Barnes, C. A. S.; Clemmer, D. E., Assessing intrinsic side chain interactions between i and i+4 residues in solvent-free peptides: A combinatorial gas-phase approach. *Journal of Physical Chemistry A* **2003**, 107 (49), 10566-10579.
138. Henderson, S. C.; Li, J. W.; Counterman, A. E.; Clemmer, D. E., Intrinsic size parameters for Val, Ile, Leu, Gln, Thr, Phe, and Trp residues from ion mobility measurements of polyamino acid ions. *Journal of Physical Chemistry B* **1999**, 103 (41), 8780-8785.
139. Hilderbrand, A. E.; Clemmer, D. E., Determination of sequence-specific intrinsic size parameters from cross sections for 162 tripeptides. *Journal of Physical Chemistry B* **2005**, 109 (23), 11802-11809.
140. Shvartsburg, A. A.; Siu, K. W. M.; Clemmer, D. E., Prediction of peptide ion mobilities via a priori calculations from intrinsic size parameters of amino acid residues. *Journal of the American Society for Mass Spectrometry* **2001**, 12 (8), 885-888.

141. Valentine, S. J.; Counterman, A. E.; Clemmer, D. E., A database of 660 peptide ion cross sections: Use of intrinsic size parameters for bona fide predictions of cross sections. *Journal of the American Society for Mass Spectrometry* **1999**, *10* (11), 1188-1211.
142. Valentine, S. J.; Counterman, A. E.; Hoaglund-Hyzer, C. S.; Clemmer, D. E., Intrinsic amino acid size parameters from a series of 113 lysine-terminated tryptic digest peptide ions. *Journal of Physical Chemistry B* **1999**, *103* (8), 1203-1207.
143. Valentine, S. J.; Ewing, M. A.; Dilger, J. M.; Glover, M. S.; Geromanos, S.; Hughes, C.; Clemmer, D. E., Using Ion Mobility Data to Improve Peptide Identification: Intrinsic Amino Acid Size Parameters. *Journal of Proteome Research* **2011**, *10* (5), 2318-2329.
144. Counterman, A. E.; Clemmer, D. E., Volumes of individual amino acid residues in gas-phase peptide ions. *Journal of the American Chemical Society* **1999**, *121* (16), 4031-4039.
145. Stone, E.; Gillig, K. J.; Ruotolo, B.; Fuhrer, K.; Gonin, M.; Schultz, A.; Russell, D. H., Surface-induced dissociation on a MALDI-ion mobility-orthogonal time-of-flight mass spectrometer: Sequencing peptides from an "in-solution" protein digest. *Analytical Chemistry* **2001**, *73* (10), 2233-2238.
146. Fernandez-Lima, F. A.; Becker, C.; Gillig, K. J.; Russell, W. K.; Tichy, S. E.; Russell, D. H., Ion Mobility-Mass Spectrometer Interface for Collisional Activation of Mobility Separated Ions. *Analytical Chemistry* **2009**, *81* (2), 618-624.
147. Badman, E. R.; Myung, S.; Clemmer, D. E., Gas-phase separations of protein and peptide ion fragments generated by collision-induced dissociation in an ion trap. *Analytical Chemistry* **2002**, *74* (19), 4889-4894.
148. Hoaglund-Hyzer, C. S.; Clemmer, D. E., Ion trap/ion mobility/quadrupole/time of flight mass spectrometry for peptide mixture analysis. *Analytical Chemistry* **2001**, *73* (2), 177-184.
149. Lee, Y. J.; Hoaglund-Hyzer, C. S.; Taraszka, J. A.; Zientara, G. A.; Counterman, A. E.; Clemmer, D. E., Collision-induced dissociation of mobility-separated ions using an orifice-skimmer cone at the back of a drift tube. *Analytical Chemistry* **2001**, *73* (15), 3549-3555.
150. Taraszka, J. A.; Kurulugama, R.; Sowell, R. A.; Valentine, S. J.; Koeniger, S. L.; Arnold, R. J.; Miller, D. F.; Kaufman, T. C.; Clemmer, D. E., Mapping the proteome of *Drosophila melanogaster*: Analysis of embryos and adult heads by LC-IMS-MS methods. *Journal of Proteome Research* **2005**, *4* (4), 1223-1237.
151. Valentine, S. J.; Liu, X. Y.; Plasencia, M. D.; Hilderbrand, A. E.; Kurulugama, R. T.; Koeniger, S. L.; Clemmer, D. E., Developing liquid chromatography ion

- mobility mass spectrometry techniques. *Expert Review of Proteomics* **2005**, 2 (4), 553-565.
152. Valentine, S. J.; Plasencia, M. D.; Liu, X.; Krishnan, M.; Naylor, S.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E., Toward plasma proteome profiling with ion mobility-mass spectrometry. *Journal of Proteome Research* **2006**, 5 (11), 2977-2984.
 153. Sowell, R. A.; Hersberger, K. E.; Kaufman, T. C.; Clemmer, D. E., Examining the proteome of Drosophila across organism lifespan. *Journal of Proteome Research* **2007**, 6 (9), 3637-3647.
 154. Liu, X.; Valentine, S. J.; Plasencia, M. D.; Trimpin, S.; Naylor, S.; Clemmer, D. E., Mapping the human plasma proteome by SCX-LC-IMS-MS. *Journal of the American Society for Mass Spectrometry* **2007**, 18 (7), 1249-1264.
 155. Shen, Y. F.; Jacobs, J. M.; Camp, D. G.; Fang, R. H.; Moore, R. J.; Smith, R. D.; Xiao, W. Z.; Davis, R. W.; Tompkins, R. G., Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome. *Analytical Chemistry* **2004**, 76 (4), 1134-1144.
 156. Echan, L. A.; Tang, H. Y.; Ali-Khan, N.; Lee, K.; Speicher, D. W., Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* **2005**, 5 (13), 3292-3303.
 157. Varki, A., Biological roles of oligosaccharides - all of the theories are correct. *Glycobiology* **1993**, 3 (2), 97-130.
 158. Fenn, L. S.; McLean, J. A., Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry. *Physical Chemistry Chemical Physics* **2011**, 13 (6), 2196-2205.
 159. Gorelik, E.; Galili, U.; Raz, A., On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis. *Cancer and Metastasis Reviews* **2001**, 20 (3-4), 245-277.
 160. Bahl, O. P., *Glycoconjugates: Composition, Structure, and Function*. Marcel Dekker: New York, NY, 1992.
 161. Harvey, D. J., Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics* **2001**, 1 (2), 311-328.
 162. Harvey, D. J., Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *International Journal of Mass Spectrometry* **2003**, 226 (1), 1-35.
 163. Zaia, J., Mass spectrometry of oligosaccharides. *Mass Spectrometry Reviews* **2004**, 23 (3), 161-227.

164. Harvey, D. J., Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: An update covering the period 1999-2000. *Mass Spectrometry Reviews* **2006**, 25 (4), 595-662.
165. Harvey, D. J., Structural determination of N-linked glycans by matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry. *Proteomics* **2005**, 5 (7), 1774-1786.
166. Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J.-C., The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* **2006**, 6 (14), 3993-4015.
167. Hardy, M. R.; Townsend, R. R., Separation of positional isomers of oligosaccharides and glycopeptides by high-performance anion-exchange chromatography with pulsed amperometric detection. *Proceedings of the National Academy of Sciences of the United States of America* **1988**, 85 (10), 3289-3293.
168. Ashline, D. J.; Lapadula, A. J.; Liu, Y.-H.; Lin, M.; Grace, M.; Pramanik, B.; Reinhold, V. N., Carbohydrate structural isomers analyzed by sequential mass spectrometry. *Analytical Chemistry* **2007**, 79 (10), 3830-3842.
169. Liu, Y. S.; Clemmer, D. E., Characterizing oligosaccharides using injected-ion mobility mass spectrometry. *Analytical Chemistry* **1997**, 69 (13), 2504-2509.
170. Lee, S.; Wyttenbach, T.; Bowers, M. T., Gas phase structures of sodiated oligosaccharides by ion mobility ion chromatography methods. *International Journal of Mass Spectrometry* **1997**, 167, 605-614.
171. Lee, D. S.; Wu, C.; Hill, H. H., Detection of carbohydrates by electrospray ionization ion mobility spectrometry following microbore high-performance liquid chromatography. *Journal of Chromatography A* **1998**, 822 (1), 1-9.
172. Leavell, M. D.; Gaucher, S. P.; Leary, J. A.; Taraszka, J. A.; Clemmer, D. E., Conformational studies of Zn-ligand-hexose diastereomers using ion mobility measurements and density functional theory calculations. *Journal of the American Society for Mass Spectrometry* **2002**, 13 (3), 284-293.
173. Dwivedi, P.; Bendiak, B.; Clowers, B. H.; Hill, H. H., Jr., Rapid resolution of carbohydrate isomers by electrospray ionization ambient pressure ion mobility spectrometry-time-of-flight mass spectrometry (ESI-APIMS-TOFMS). *Journal of the American Society for Mass Spectrometry* **2007**, 18 (7), 1163-1175.
174. Clowers, B. H.; Dwivedi, P.; Steiner, W. E.; Hill, H. H.; Bendiak, B., Separation of sodiated isobaric disaccharides and trisaccharides using electrospray ionization-atmospheric pressure ion mobility-time of flight mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2005**, 16 (5), 660-669.

175. Isailovic, D.; Kurulugama, R. T.; Plasencia, M. D.; Stokes, S. T.; Kyselova, Z.; Goldman, R.; Mechref, Y.; Novotny, M. V.; Clemmer, D. E., Profiling of human serum glycans associated with liver cancer and cirrhosis by IMS-MS. *Journal of Proteome Research* **2008**, *7* (3), 1109-1117.
176. Vakhrushev, S. Y.; Langridge, J.; Campuzano, I.; Hughes, C.; Peter-Katlinic, J., Ion mobility mass spectrometry analysis of human glycourinome. *Analytical Chemistry* **2008**, *80* (7), 2506-2513.
177. Williams, J. P.; Grabenauer, M.; Holland, R. J.; Carpenter, C. J.; Wormald, M. R.; Giles, K.; Harvey, D. J.; Bateman, R. H.; Scrivens, J. H.; Bowers, M. T., Characterization of simple isomeric oligosaccharides and the rapid separation of glycan mixtures by ion mobility mass spectrometry. *International Journal of Mass Spectrometry* **2010**, *298* (1-3), 119-127.
178. Fenn, L. S.; McLean, J. A., Simultaneous glycoproteomics on the basis of structure using ion mobility-mass spectrometry. *Molecular Biosystems* **2009**, *5* (11), 1298-1302.
179. Gabryelski, W.; Froese, K. L., Rapid and sensitive differentiation of anomers, linkage, and position isomers of disaccharides using High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS). *Journal of the American Society for Mass Spectrometry* **2003**, *14* (3), 265-277.
180. Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Clemmer, D. E., Resolving and Assigning N-Linked Glycan Structural Isomers from Ovalbumin by IMS-MS. *Journal of the American Society for Mass Spectrometry* **2008**, *19* (11), 1706-1715.
181. Wymann, M. P.; Schneider, R., Lipid signalling in disease. *Nature Reviews Molecular Cell Biology* **2008**, *9* (2), 162-176.
182. Janmey, P. A.; Kinnunen, P. K. J., Biophysical properties of lipids and dynamic membranes. *Trends in Cell Biology* **2006**, *16* (10), 538-546.
183. Han, X.; Yang, J.; Yang, K.; Zhao, Z.; Abendschein, D. R.; Gross, R. W., Alterations in myocardial cardiolipin content and composition occur at the very earliest stages of diabetes: A shotgun lipidomics study. *Biochemistry* **2007**, *46* (21), 6417-6428.
184. Thomas, A.; Deglon, J.; Lenglet, S.; Mach, F.; Mangin, P.; Wolfender, J.-L.; Steffens, S.; Staub, C., High-Throughput Phospholipidic Fingerprinting by Online Desorption of Dried Spots and Quadrupole-Linear Ion Trap Mass Spectrometry: Evaluation of Atherosclerosis Biomarkers in Mouse Plasma. *Analytical Chemistry* **2010**, *82* (15), 6687-6694.
185. Han, X. L.; Holtzman, D. M.; McKeel, D. W., Plasmalogen deficiency in early Alzheimer's disease subjects and in animal models: molecular characterization

- using electrospray ionization mass spectrometry. *Journal of Neurochemistry* **2001**, 77 (4), 1168-1180.
186. Watson, A. D., Lipidomics: a global approach to lipid analysis in biological systems. *Journal of Lipid Research* **2006**, 47 (10), 2101-2111.
187. Wenk, M. R., The emerging field of lipidomics. *Nature Reviews Drug Discovery* **2005**, 4 (7), 594-610.
188. Petkovic, M.; Schiller, J.; Muller, M.; Benard, S.; Reichl, S.; Arnold, K.; Arnhold, J., Detection of individual phospholipids in lipid mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: Phosphatidylcholine prevents the detection of further species. *Analytical Biochemistry* **2001**, 289 (2), 202-216.
189. 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. *Analytical Chemistry* **2010**, 82 (5), 1881-1889.
190. Jackson, S. N.; Ugarov, M.; Post, J. D.; Egan, T.; Langlais, D.; Schultz, J. A.; Woods, A. S., A Study of Phospholipids by Ion Mobility TOFMS. *Journal of the American Society for Mass Spectrometry* **2008**, 19 (11), 1655-1662.
191. Kim, H. I.; Kim, H.; Pang, E. S.; Ryu, E. K.; Beegle, L. W.; Loo, J. A.; Goddard, W. A.; Kanik, I., Structural Characterization of Unsaturated Phosphatidylcholines Using Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry* **2009**, 81 (20), 8289-8297.
192. Kliman, M.; Vijayakrishnan, N.; Wang, L.; Tapp, J. T.; Broadie, K.; McLean, J. A., Structural mass spectrometry analysis of lipid changes in a Drosophila epilepsy model brain. *Molecular Biosystems* **2010**, 6 (6), 958-966.
193. Jackson, S. N.; Ugarov, M.; Egan, T.; Post, J. D.; Langlais, D.; Schultz, J. A.; Woods, A. S., MALDI-ion mobility-TOFMS imaging of lipids in rat brain tissue. *Journal of Mass Spectrometry* **2007**, 42 (8), 1093-1098.
194. Alberts, B., The cell as a collection of protein machines: Preparing the next generation of molecular biologists. *Cell* **1998**, 92 (3), 291-294.
195. Sali, A.; Glaeser, R.; Earnest, T.; Baumeister, W., From words to literature in structural proteomics. *Nature* **2003**, 422 (6928), 216-225.
196. Loo, J. A., Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews* **1997**, 16 (1), 1-23.

197. Sharon, M.; Robinson, C. V., The role of mass Spectrometry in structure elucidation of dynamic protein complexes. In *Annual Review of Biochemistry*, 2007; Vol. 76, pp 167-193.
198. Benesch, J. L.; Robinson, C. V., Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Current Opinion in Structural Biology* **2006**, *16* (2), 245-251.
199. Loo, J. A.; Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y. M.; Kaufman, S. L.; Chernushevich, I. V., Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *Journal of the American Society for Mass Spectrometry* **2005**, *16* (7), 998-1008.
200. Leary, J. A.; Schenauer, M. R.; Stefanescu, R.; Andaya, A.; Ruotolo, B. T.; Robinson, C. V.; Thalassinos, K.; Scrivens, J. H.; Sokabe, M.; Hershey, J. W. B., Methodology for Measuring Conformation of Solvent-Disrupted Protein Subunits using T-WAVE Ion Mobility MS: An Investigation into Eukaryotic Initiation Factors. *Journal of the American Society for Mass Spectrometry* **2009**, *20* (9), 1699-1706.
201. Uetrecht, C.; Versluis, C.; Watts, N. R.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. R., Stability and shape of hepatitis B virus capsids in vacuo. *Angewandte Chemie-International Edition* **2008**, *47* (33), 6247-6251.
202. van Duijn, E.; Barendregt, A.; Synowsky, S.; Versluis, C.; Heck, A. J. R., Chaperonin Complexes Monitored by Ion Mobility Mass Spectrometry. *Journal of the American Chemical Society* **2009**, *131* (4), 1452-1459.
203. McKay, A. R.; Ruotolo, B. T.; Ilag, L. L.; Robinson, C. V., Mass measurements of increased accuracy resolve heterogeneous populations of intact ribosomes. *Journal of the American Chemical Society* **2006**, *128* (35), 11433-11442.
204. Sobott, F.; Robinson, C. V., Characterising electrosprayed biomolecules using tandem-MS - the noncovalent GroEL chaperonin assembly. *International Journal of Mass Spectrometry* **2004**, *236* (1-3), 25-32.
205. Ruotolo, B. T.; Hyung, S.-j.; Robinson, P. M.; Giles, K.; Bateman, R. H.; Robinson, C. V., Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes. *Angewandte Chemie, International Edition* **2007**, *46* (42), 8001-8004.
206. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V., Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics. *Journal of the American Chemical Society* **2012**, *134* (7), 3429-3438.
207. Erba, E. B.; Ruotolo, B. T.; Barsky, D.; Robinson, C. V., Ion Mobility-Mass Spectrometry Reveals the Influence of Subunit Packing and Charge on the

- Dissociation of Multiprotein Complexes. *Analytical Chemistry* **2010**, 82 (23), 9702-9710.
208. Pagel, K.; Hyung, S.-J.; Ruotolo, B. T.; Robinson, C. V., Alternate Dissociation Pathways Identified in Charge-Reduced Protein Complex Ions. *Analytical Chemistry* **2010**, 82 (12), 5363-5372.
209. Hyung, S.-J.; Robinson, C. V.; Ruotolo, B. T., Gas-Phase Unfolding and Disassembly Reveals Stability Differences in Ligand-Bound Multiprotein Complexes. *Chemistry & Biology* **2009**, 16 (4), 382-390.
210. Benesch, J. L. P., Collisional Activation of Protein Complexes: Picking Up the Pieces. *Journal of the American Society for Mass Spectrometry* **2009**, 20 (3), 341-348.
211. Bornschein, R. E.; Hyung, S.-J.; Ruotolo, B. T., Ion Mobility-Mass Spectrometry Reveals Conformational Changes in Charge Reduced Multiprotein Complexes. *Journal of the American Society for Mass Spectrometry* **2011**, 22 (10), 1690-1698.
212. Hood, L.; Heath, J. R.; Phelps, M. E.; Lin, B. Y., Systems biology and new technologies enable predictive and preventative medicine. *Science* **2004**, 306 (5696), 640-643.
213. Peterson, R. T., Chemical biology and the limits of reductionism. *Nature Chemical Biology* **2008**, 4 (11), 635-638.
214. Alm, E.; Arkin, A. P., Biological networks. *Current Opinion in Structural Biology* **2003**, 13 (2), 193-202.
215. Van Regenmortel, M. H. V., Reductionism and complexity in molecular biology. *Embo Reports* **2004**, 5 (11), 1016-1020.

CHAPTER II

BIOMOLECULAR SIGNATURES OF DIABETIC WOUND HEALING BY STRUCTURAL MASS SPECTROMETRY

2.1. Introduction

Diabetes is a metabolic disease characterized by abnormally high blood glucose levels resulting from the body's inability to produce or use insulin. The Centers for Disease Control and Prevention (CDC) has reported that diabetes affects 8.3% of the U.S. population.¹ Patients with diabetes are at significantly greater risk of complications such as blindness, kidney failure, and heart disease, making it the seventh leading cause of death in the U.S.¹ Additional complications of diabetes occur in the extremities and include the loss or reduction of nerve sensation and decreased blood flow.^{2,3} The lack of nerve sensation is experienced by nearly 70% of diabetics and is particularly serious when it occurs in the lower extremities.¹ Due to decreased blood flow and microcirculation in the lower limbs, signaling defects in the cytokine response, and the potential for infection, injuries incurred on the feet or lower leg can suffer from a delayed healing process, resulting in the formation of a chronic ulcer at the wound site.^{2,3} These chronic diabetic ulcers contribute significantly to the high number of lower-limb amputations performed per year on diabetics in the U.S. (65,700 in 2010).¹

The increased risk of chronic ulcer formation stems from disruption of the complex process of wound healing by the pathophysiological abnormalities associated with diabetes.^{3,4} Analysis of human diabetic ulcers has revealed the differential

expression of growth factors, chemokines, cytokines and their receptors, which are crucial to several phases of the normal wound healing process.^{5,6} Matrix metalloproteinases (MMPs), a family of endoproteinases involved in tissue remodeling, are also differentially expressed in chronic wounds, causing the dysfunctional breakdown of the extracellular matrix.^{7,8} Macrophages isolated from the wounds of diabetic mice have exhibited decreased ability to remove dead cells, resulting in a prolonged inflammatory response.⁹ Although it is evident that diabetes and hyperglycemia cause widespread disruption of the wound healing process, studies of diabetic wound healing continue to provide only a narrow view of a large and complex process.

The protein detection methods of immunoblotting, microbead (Luminex), and enzyme immunoassays commonly utilized in wound healing studies rely on costly antibodies that are protein specific, requiring multiple antibodies for the analysis of multiple proteins. Proteomic analyses of wound healing have emerged recently with the broader focus of developing prognostic and diagnostic tools and potential therapies for chronic wounds.^{8,10,11} However, both proteomic and antibody-based analyses require rigorous and time-consuming sample preparation procedures to isolate the desired proteins that can alter the original state of the biological sample. To capture the full complexity of diabetic wound healing, it is desirable to utilize a more inclusive analysis requiring minimal sample manipulation.

Ion mobility-mass spectrometry (IM-MS) is a rapid method of analysis which requires minimal sample preparation and offers the flexibility to include pre-ionization separations, making it well suited for holistic studies of complex biological systems. IM-MS is a two-dimensional separation combining gas-phase ion mobility (IM) structural

separations with the mass-to-charge (m/z) separations of mass spectrometry (MS). In gas-phase IM separations, ions travel under the influence of an electric field through a drift cell filled with neutral background gas. The number of collisions that occur between the ion and the neutral gas molecules results in a characteristic drift time, measured in micro- to milliseconds, which is dependent on the collision cross section (effective ion surface area). Within a class of biomolecules, such as lipids, proteins, or carbohydrates, a strong correlation between collision cross sections and m/z ratios is observed, resulting in the separation of each biomolecular class along unique mobility-mass correlation lines according to gas-phase packing efficiencies.^{12, 13,14}

Integrating ion mobility with MS measurements provides four primary advantages over MS-only techniques in the analysis of complex biological samples for the fields of imaging,^{15,16} proteomics,^{17,18} glycomics,^{19,20} lipidomics,^{21,22} metabolomics,^{23,24} and systems biology.²⁵ Firstly, the IM-MS integration provides rapid separations, increased peak capacity, and enhanced peak capacity production rate as a 2D separation over 1D separation dimensions. Secondly, IM-MS is inherently gas-phase and the timescales of separation are well suited for further integration with separations such as liquid chromatography (LC) to further increase peak capacity.^{17, 18, 26-28} Thirdly, different classes of biomolecules are readily distinguished from one another on the basis of their gas-phase structure. This provides higher signal-to-noise for low abundant species of one class (*e.g.* metabolites) that are separated from a species of a different class (*e.g.* peptides) having nearly, or the same, mass, which are unresolvable in the MS alone. Finally, ion activation following the IM separation, but prior to the MS separation

provides the ability to obtain fragment ion spectra for all of the molecules nearly simultaneously.

This study demonstrates the use of IM-MS in identifying biomolecular signatures of diabetic wound healing. Wound exudate, a non-invasive sampling approach for human diabetic wounds, was collected from the wounds of diabetic and non-diabetic rats. The wound fluid was analyzed by IM-MS using electrospray ionization (ESI) after minimal sample preparation. Biomolecular signatures distinguishing diabetic and non-diabetic wound fluid and between wound fluids collected at two time points were revealed by direct comparison of IM-MS spectra and with the assistance of statistical analyses. Ultra-performance liquid chromatography (UPLC), tandem MS, and database searching were used in a targeted manner to identify and validate the biomolecular signatures of diabetic wound healing.

2.2. *Experimental*

Rat Model

Sixteen Sprague Dawley rats (300 g, male, Harlan Laboratories) were divided into treatment ($n=10$) and control ($n=6$) groups. The treatment group was dosed twice with streptozotocin (55 mg/kg), which induces diabetes by selective destruction of the insulin-producing pancreatic β -islet cells. A blood glucose concentration greater than 300 mg/dL confirmed the diabetic state in the treatment group, and all ten treated rats converted to the diabetic phenotype. Four sterilized 10 mm polyvinyl alcohol (PVA) sponges (Merocel, Medtronic, Mystic, CT) were inserted subcutaneously on the dorsal side of each animal. The PVA sponges were used to elicit a geometrically defined granulation

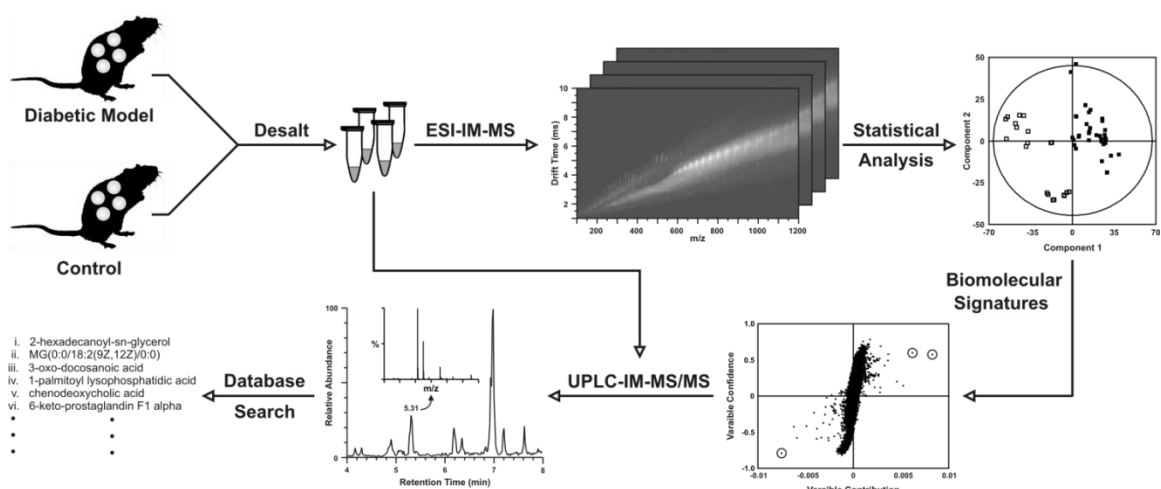
tissue response as a model of wound healing and to retain wound fluid. Sponges were harvested two days post-implantation from five diabetic and three control rats. Sponges were removed from the remaining diabetic ($n=5$) and control ($n=3$) rats five days post-implantation. Approximately 40-50 μL of wound fluid was extracted from each PVA sponge by centrifugation through spin columns (Pierce Spin Cups, cellulose acetate filter, Thermo Fisher Scientific, Waltham, MA). Protein concentrations in the wound fluids ranged from 36 – 114 $\mu\text{g}/\text{mL}$. The remaining wound fluid was flash frozen and stored at -80°C .

Sample Preparation for ESI-IM-MS Direct Infusion Experiments

Scheme 2.1 illustrates a general workflow for the preparation and IM-MS analysis of wound fluid. The 64 different sponge wound fluid samples to be analyzed by ESI-IM-MS were prepared by diluting 20 μL of wound fluid to 1 mL with 4% formic acid (Thermo Fisher Scientific, Waltham, MA) prepared in distilled deionized (DDI) water (18 $\text{M}\Omega$ cm; Millipore, Billerica, MA). The diluted wound fluid was desalted with 1 cc/30 mg Oasis HLB desalting cartridges (Waters, Milford, MA), using CHROMASOLV methanol (Sigma-Aldrich, St. Louis, MO) with 0.1% formic acid as the elution solvent. After desalting, samples were dried and reconstituted with 1 mL of 50% methanol containing 0.1% formic acid. Each of the 64 different samples were run in duplicate for technical replicates, resulting in 128 independent datasets.

Sample Preparation for UPLC-ESI-IM-MS Experiments

50 μL each from two of the previously prepared diabetic day 2 wound fluid



Scheme 2.1. Workflow for the preparation and IM-MS analysis of wound fluid for biomolecular signatures of diabetic wound healing. Briefly, wound fluid collected from PVA sponges was desalted and analyzed by ESI-IM-MS. An orthogonal partial least-squares discriminant analysis (OPLS-DA) statistical analysis of the ESI-IM-MS data revealed biomolecular signatures of diabetic wound healing. UPLC-IM-MS/MS was performed in a targeted fashion to obtain accurate mass and fragmentation data for select biomolecular signatures. These data were then used for database searching, from which tentative identifications of the biomolecular signatures were determined.

samples were combined and diluted with 100 μL H_2O (CHROMASOLV-grade; Sigma-Aldrich, St. Louis, MO) for a total volume of 200 μL . The sample was centrifuged for 10 minutes at 14,500 rpm and transferred to an auto-sampler vial for UPLC separation.

Sample Preparation for UPLC-ESI-IM-MS Experiments

50 μL each from two of the previously prepared diabetic day 2 wound fluid samples were combined and diluted with 100 μL H_2O (CHROMASOLV-grade; Sigma-Aldrich, St. Louis, MO) for a total volume of 200 μL . The sample was centrifuged for 10 minutes at 14,500 rpm and transferred to an auto-sampler vial for UPLC separation.

Instrumentation

IM-MS was performed on a Synapt G2 (Waters, Milford, MA), which utilizes a traveling wave ion mobility cell and an orthogonal time-of-flight (TOF) mass spectrometer. The instrument was operated with electrospray ionization (ESI) and the TOFMS was operated in the single-stage reflectron, or “V,” configuration.^{29, 30}

ESI-IM-MS

An external syringe pump (Harvard Apparatus, Holliston, MA) was used for direct infusion of the samples into the ESI source at a flow rate of 6 $\mu\text{L}/\text{min}$. The ESI source was operated in positive (+) mode. The electrospray voltage was 4.14 kV with sampling and extraction cone voltages of 20 V and 8 V, respectively. The desolvation temperature and gas flow were 150°C and 10 L min^{-1} , respectively. Ion mobility separation was performed through nitrogen gas with a traveling wave velocity of 550 m s^{-1} .

¹ and an amplitude of 40 V. Mass calibration was performed with sodium iodide (2 µg/µL in 50% 2-propanol) in the range of m/z 100 - 3000. Data were acquired over two minutes at a rate of 1 scan s⁻¹. The transfer region was used to perform post-mobility MS/MS experiments.

UPLC-ESI-IM-MS

An ACQUITY (Waters, Milford, MA) UPLC system with an ACQUITY HSS C18 column (1.8 µm, 1.0 X 100 mm; Waters, Milford, MA) was used for the chromatographic separations prior to IM-MS. The column was coupled to the Synapt G2 by the ESI source. The autosampler and column temperatures were maintained at 4°C and 45°C, respectively. Chromatographic separation was performed with a binary solvent system of 95% water/5% acetonitrile (CHROMOSOLV-grade; Sigma-Aldrich, St. Louis, MO) with 10 mM ammonium acetate (Solvent A) and 5% water/95% acetonitrile with 10 mM ammonium acetate (Solvent B). The mobile phase flow rate was set to 70 µL min⁻¹ and the injection volume was 5 µL. The initial gradient conditions were 99% A: 1% B for 1 min, followed by a linear gradient to 1% A: 99% B over the next 7 min. The solvent gradient was held at 1% A: 99% B for 2 min, and then linearly increased to 99% A: 1% B over the next 0.5 min. The positive mode (+) ESI conditions were as follows: capillary +3.5 kV; sampling cone: 35 V; extraction cone: 2 V; source temperature: 130 °C; desolvation temperature: 150 °C. The negative mode (-) ESI conditions were as follows: capillary -2.5 kV; sampling cone: -25 V; extraction cone: -2 V; source temperature: 100 °C; desolvation temperature: 150 °C. Desolvation and cone gas flow were 600 and 20 L/hr, respectively. A traveling wave with a velocity of 550 m s⁻¹ and height of 40 V was

used for IM separation. Leucine-enkephalin (m/z 556.2771) was used for lock mass correction. MS^E was performed in the transfer region with ramping collision energy from 10 – 30 eV.

Biostatistics

The MarkerLynx XS (Waters, Manchester, U.K.) software package was used to perform mass spectral peak detection, alignment and normalization. Mass spectral peak alignment was performed by the combined scan method, where peaks within a 0.1 Da mass window were combined, and all 120 scans in the 2 min ESI-IM-MS acquisitions were combined. The spectral peak detection threshold was set at 1000 counts. Normalization was performed by the constant sum method, where the intensities are normalized such that the sum of intensities of all peaks in a sample sums to 10,000. The aligned data were then exported to the Extended Statistics component of MarkerLynx XS for partial least-squares discriminant analysis (PLS-DA) to show group differences and orthogonal partial least-squares discriminant analysis (OPLS-DA) S-plots to show the molecular species responsible for the group differences. Model parameters (R²_Y, Q²) for PLS-DA and OPLS-DA can be found in the supporting information documentation. Two sample t-tests (assuming unequal variance) were selectively performed to calculate significance p -values for the species revealed to contribute the most to the group differences in the OPLS-DA S-plot analysis of the ESI-IM-MS data. Peak areas from mobility-extracted mass spectra were used for the calculation of p -values, fold-changes and to prepare box-and-whisker plots.

Bioinformatics

Online database searching was performed by accurate mass through the MarkerLynx XS software package. The metabolite databases utilized included ChemSpider, the Human Metabolite Database (<http://www.hmdb.ca>), KEGG (<http://www.kegg.com>), LipidMAPS (<http://www.lipidmaps.org>), and METLIN (<http://metlin.scripps.edu/>). Searches were performed with a mass tolerance of 0.05 Da. Protein sequence similarity searches were performed with the UniProt BLAST program (<http://www.uniprot.org/?tab=blast>) and the UniProtKB protein database.

Validation

The tentative identifications from the bioinformatics were validated by MS/MS fragmentation experiments. When available, chemical standards of the tentatively identified species were purchased from Sigma Aldrich (St. Louis, MO). To be considered a positive identification, all MS and MS/MS peaks of the standard were required to match those of the analyte.

2.3. Results and Discussion

Two-dimensional projections of IM-MS data representative of control day 2 and diabetic day 2 wound fluid are illustrated in Figure 2.1(a),(b). The mass-to-charge (m/z) ratio is displayed along the x-axis, and drift time, measured in milliseconds (ms), is displayed along the y-axis. Intensity (counts), displayed on the z-axis in three-dimensional IM-MS spectra, is represented by false coloring in the two-dimensional projection. Integrating the drift time dimension of IM-MS data yields a mass spectrum,

where intensity is displayed on the y-axis and m/z on the x-axis (Figure 2.1c). Mass spectra may also be extracted from a user-defined region of mobility-mass space and are referred to as mobility-extracted mass spectra (Figure 2.1d).

Several mobility-mass correlation lines were apparent in the IM-MS spectra of the control day 2 and diabetic day 2 wound fluid (Figure 2.1a,b). These lines were predicted to represent biomolecular species such as lipids, proteins, or carbohydrates. While most of these correlation lines were consistently present in all samples, a line of strong signals was observed between m/z 700-950 and 4.2-5.6 ms in the diabetic day 2 mobility-mass spectra (*e.g.*, white dashed box in Figure 2.1b) that was not distinguishable in control day 2 samples or day 5 samples (day 2 control vs. diabetic: fold-change = 5.0; $p = 9.56 \times 10^{-5}$; see Figure B.1a). Mass spectra extracted from the area outlined by the white rectangle in Figure 2.1(b) revealed several highly charged signals (Figure 2.1d). From the isotopic distributions, the signals were identified as the +11 to +14 charge states of an approximately 10.15 kDa species. The number of charges carried by this species and its corresponding molecular weight suggested it was a small protein.

While these signals were strongly apparent in the ion mobility-mass spectra, identifying these unique features of the diabetic day 2 wound fluid from conventional mass spectra would be more challenging. The region between m/z 600-1200 in the mobility-mass spectrum of diabetic day 2 wound fluid (Figure 2.1b) contained another correlation line of highly charged signals which was separated in drift time from the series of nearly isobaric signals corresponding to the 10.15 kDa protein. Without separation in the drift time dimension, these signals were collapsed onto one another in

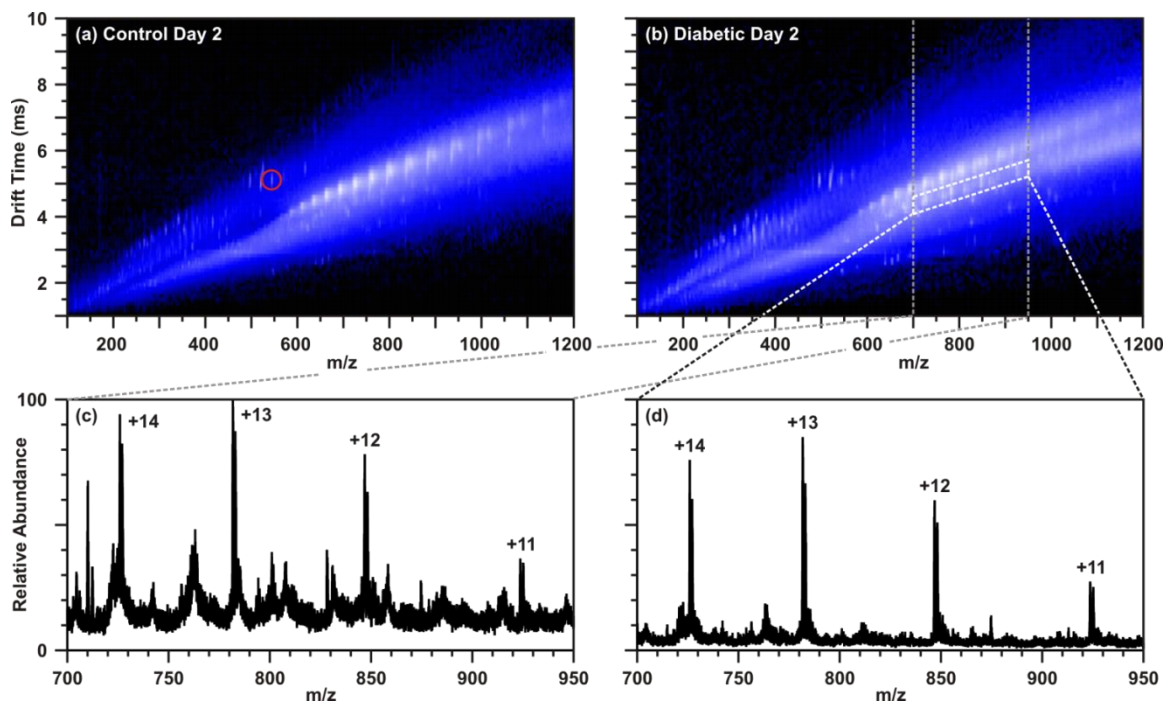


Figure 2.1. Three-dimensional ion mobility-mass spectra (a and b), two-dimensional mass spectra (c and d) representative of the ESI-IM-MS analysis of the 64 wound fluid samples. (a) IM-MS spectrum representative of control day 2 wound fluid. The red circle corresponds to m/z 544.4 as discussed in the text. (b) IM-MS spectrum representative of diabetic day 2 wound fluid. The region outlined by the white dotted rectangle contains the +11 to +14 charge states of protein S100-A8. (c) MS-only spectrum of the region m/z 700-950 (as indicated by the grey dotted lines) of diabetic day 2 wound fluid (b), obtained by collapsing the IM dimension. (d) Mobility-selected mass spectrum of the region outlined by white-dashed lines in (b). A 2.7-fold increase in the chemical signal-to-noise ratio (S/N) of the +13 charge state signal of S100-A8, m/z 781, was observed between the two-dimensional mass spectrum (c) and mobility-extracted mass spectrum (d) of the same m/z region.

the corresponding mass spectrum (Figure 2.1c). A mass spectrum extracted from the region of the mobility-mass spectrum containing the 10.15 kDa species (outlined in Figure 2.1b) showed a significant reduction in the chemical noise baseline (Figure 2.1d). For the +13 peak, m/z 781.7, a 2.7-fold improvement in the signal-to-noise ratio (S/N) was observed between the extracted mass spectrum and the full mass spectrum.

Tandem MS (MS/MS) experiments are another approach widely used in attempts to identify unknown species, but the low S/N due to the overlapping isotopic distributions of several species would provide fragmentation spectra from which it would be challenging to make confident identifications. Two-dimensional IM-MS analysis allows ready discrimination of the chemical noise in a typical MS measurement from isobaric chemical species and overlapping isotopic distributions that are not resolved in one-dimensional separations (for examples, see Figures B.2 and B.3 in Appendix B).^{12,24,30}

The challenges presented to MS/MS experiments by isobaric signals and overlapping isotopic distributions can be overcome with the addition of ion mobility.³¹ The present instrumentation enables MS/MS experiments to be performed before or after IM separation, termed pre-mobility or post-mobility MS/MS, respectively, or both in series. While pre-mobility MS/MS (also known as ion activation) provides the unique drift times of precursor and fragment ions, post-mobility MS/MS yields fragment ions with the same drift time as the precursor ion. Fragmentation spectra are then mobility-organized, separating the precursor of interest and its corresponding fragment ions from isobaric signals in drift time. In contrast to conventional MS/MS, the current IM-MS/MS instrumentation does not require preselecting particular ion species for fragmentation

analysis, since the current IM-MS/MS instrument supports multiplexed MS/MS measurements during the initial sample run, with a wealth of fragmentation data readily available for subsequent analysis without rerunning the sample for targeted MS/MS based upon an initial analysis.

Post-mobility MS/MS experiments were performed to identify the differentially expressed protein observed in the diabetic day 2 wound fluid. The MS/MS spectra taken for several charge states of this species yielded the amino acid sequence NFEEFLVLV from a series of triply charged γ -ions (Figure 2.2). A UniProt BLAST sequence search of “NFEEFLVLV” yielded a match with 100% identity, a score of 70, and an E-value of 0.29 to residues 67-75 of protein S100-A8 (89 residues), also known as Calgranulin-A. The experimental monoisotopic molecular weight was calculated from the raw data to be $10,144.35 \pm 0.04$ Da, which corresponds to the calculated monoisotopic molecular weight of S100-A8 with 26 ppm mass accuracy. Protein S100-A8 is a member of the S100 family of proteins, characterized by the presence of two calcium binding domains which regulate cell functions such as cell stress, signal transduction, chemotaxis and inflammation.^{32,33} The heterodimer S100-A8/9 has been found at increased levels in the serum of type I diabetics.³⁴ S100-A8 and S100-A9 were also found to be differentially expressed in wounds caused by scalpel incisions, and the role of S100-A8 in wound healing is suspected to be the differentiation of fibroblasts and accumulation of monocytes in areas of inflammation.^{35,36}

To identify additional signatures of diabetic wound healing, statistical analysis was performed using MarkerLynx XS software. Data were aligned by mass-to-charge

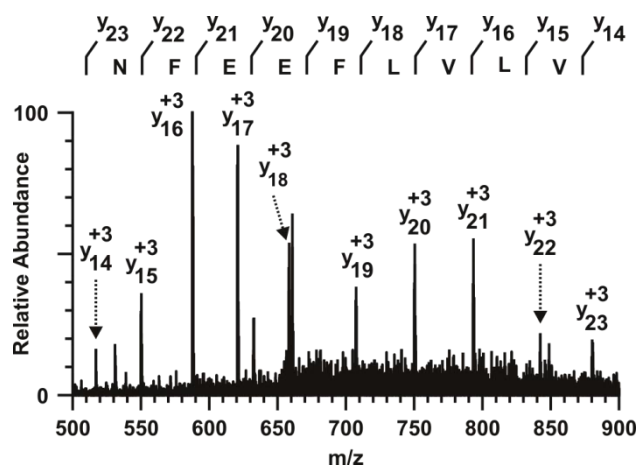


Figure 2.2. Mobility-extracted fragmentation spectrum for collision-induced dissociation (CID) of m/z 677.4, present only in diabetic day 2 wound fluid. The series of triply charged ions were identified as y-ions 14-23 (residues 67-75) of the protein S100-A8, corresponding to the partial sequence “NFEEFLVLV.” The hit to the UniProt protein database had an identity score of 100% and an E-value of 0.29. The experimental monoisotopic molecular weight from the raw data corresponded to the theoretical monoisotopic molecular weight for S100-A8 with 26 ppm mass accuracy.

only, as MarkerLynx XS did not have the capability of aligning by drift time. Partial least-squares discriminant analysis (PLS-DA) and orthogonal partial least-squares discriminant analysis (OPLS-DA) were used to identify signals which contributed most significantly to the differences between control and diabetic samples. The results of these statistical tests were visualized in score plots and S-Plots (Figure 2.3). PLS-DA scores depict the separation of control and diabetic wound fluid by group (Figure 2.3a,c,e), while OPLS-DA S-Plots (Figure 2.3b,d,f) depict which molecular signals are contributing to the group separation. As shown by the arrows in Figure 2.3(b), movement of a variable away from the origin in the x-direction correlates to its contribution to, or abundance in, a specific condition. Movement away from the origin in the y-direction is correlated with increased confidence in that variable's uniqueness to a specific condition, *i.e.*, the variable displays a substantial change between the diabetic and control groups. Features in the lower left quadrant of the S-Plots (Figure 2.3b, d,f) relate to the respective control samples, while features in the upper right quadrant relate to the respective diabetic samples.

Three signals strongly contributed to the differences between the diabetic and control groups: m/z 544.4, 373.3, and 355.3. The maximum group separation was observed in the PLS-DA score plot for control day 2 and diabetic day 2 wound fluids (Figure 2.3a). In the corresponding S-Plot (Figure 2.3b), this separation was attributed to the signals m/z 544.4, 355.3 and 373.3, which were at their maximal separations from the other molecular signals. The signal m/z 544.4 contributed to the separation of control day 5 and diabetic day 5 (Figure 2.3c,d) as well as control day 2 and diabetic day 5 (Figure 2.3 e, f) wound fluid; however, overall group separation decreased due to the lower

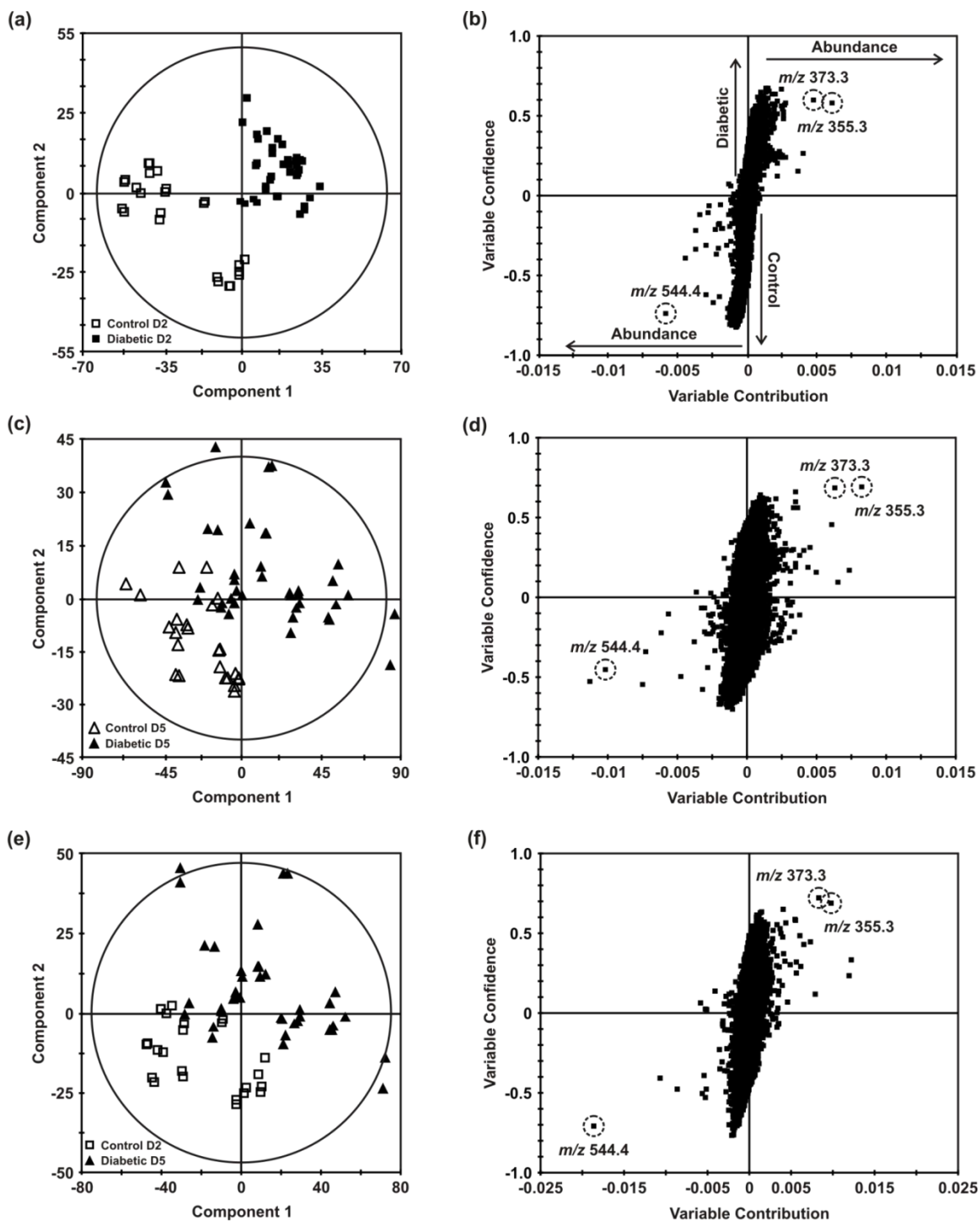


Figure 2.3. Partial least-squares discriminant analysis (PLS-DA) scores (a,c,e) and orthogonal partial least-squares discriminant analysis (OPLS-DA) S-Plots (b,d,f) for ESI-IM-MS data of control day 2 and diabetic day 2 wound fluids (a,b), control day 5 and diabetic day 5 wound fluids (c,d), and control day 2 and diabetic day 5 wound fluids (e,f). Model parameters (R^2Y , Q^2) can be found in Appendix B.

contribution of m/z 355.3 and 373.3 to the diabetic groups. This result inferred that the diabetic day 5 wounds became more similar to the control wounds. The absence of S100-A8 in the control day 2, control day 5, and diabetic day 5 wound fluids (Figure 2.1a,b) supported this inference.

A four-fold increase in the intensity ($p = 3.66 \times 10^{-11}$, Figure B.1b) of m/z 544.4 occurred in the control day 2 group relative to the diabetic day 2 group, and a 2.7-fold increase was observed in the control day 5 group relative to the diabetic day 5 group ($p = 8.37 \times 10^{-6}$, Figure B.1b). In the IM-MS spectra of control day 2 wound fluid, this species was located above the region of the mobility-mass spectra containing the protein S100-A8, with a drift time of 5.03 *ms* (red circle in Figure 2.1a). The area of mobility-mass spectra above the protein region (*i.e.*, with longer drift times) has been demonstrated previously as the region which lipids occupy, and localization of m/z 544.4 in this general region indicated it was a lipid species.^{13,14}

Extracted post-mobility MS/MS spectra of m/z 544.4 yielded fragments m/z 104 and m/z 184 (Figure 2.4). This fragmentation pattern is diagnostic of a glycerophosphocholine (GPCCho) lipid species, as the fragment m/z 184 is indicative of the loss of the lipid's phosphocholine head group. A mass-based database search in MarkerLynx XS suggested that m/z 544.4 was lysophosphatidylcholine (20:4) (LysoPC 20:4, LPC 20:4). Interrogation of the mobility-mass spectra had previously revealed this species to be a lipid due to its location above the protein mobility-mass correlation region, and the fragmentation and database results provided additional confirmation. Lipidomic analyses of human plasma have revealed decreased levels of lysophosphatidylcholines (LPCs), in general, for individuals with pre-diabetic conditions,

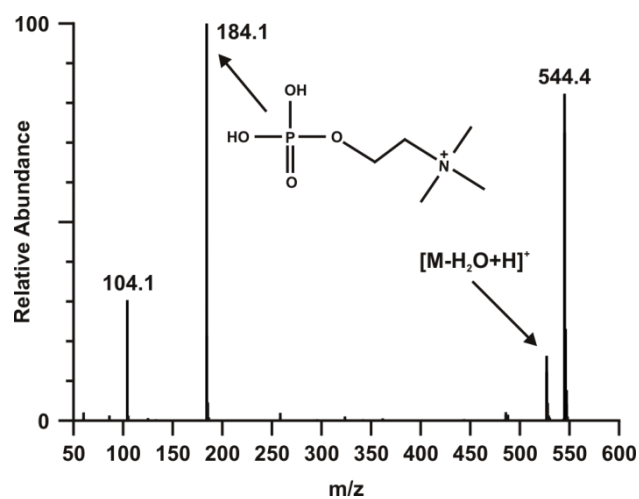


Figure 2.4. Post-mobility MS/MS of m/z 544.4 revealed a base peak fragment at m/z 184.1. This fragmentation pattern is indicative of the loss of the phosphocholine head group from a glycerophosphocholine lipid species.

such as insulin resistance and glucose intolerance.³⁷⁻⁴⁰ LPCs have also been studied in regards to their involvement in inflammation,⁴¹ and recent work has indicated potential anti-inflammatory properties of polyunsaturated LPCs such as LPC (20:4).^{42, 43}

Nine- ($p = 2.3 \times 10^{-10}$; Figure B.1d) and 2.6-fold ($p = 2.5 \times 10^{-8}$; Figure B.1c) increases in intensity were observed for m/z 373.3 and 355.3, respectively, in the diabetic day 2 group relative to the control day 2 group. Furthermore, m/z 355.3 and m/z 373.3 presented similar trends in intensity across the four sample groups, in which intensity was most abundant in the diabetic day 2 group, but also had increased intensity in the diabetic day 5 group relative to its control (Figure B.1c,d). The species m/z 355.3 and 373.3 were located within the same mobility-mass correlation line in the IM-MS spectra of diabetic day 2 wound fluids, with respective drift times of 3.28 and 3.50 *ms* (Figure 2.5), and generally located between the lipid and protein regions of the mobility-mass spectra. The proximity of these two signals in drift time and in m/z , as well as the similarity in expression across the sample groups, indicated that these signals shared a common structure and perhaps a similar function. The structural similarities of these species, as indicated by their adjacent drift times, strongly guided further analysis to identify the nature of these signals.

The mass difference of 18 Da between m/z 355.3 and 373.3 was suspected to correspond to a neutral water loss and to occur via in-source fragmentation. A pre-ionization separation by UPLC was paired with IM-MS in a targeted analysis to determine a common precursor from which these species were generated in the ionization process. In the positive-mode targeted UPLC-IM-MS analysis of diabetic day 2 wound fluid, m/z 355.26 and 373.27 were observed to co-elute from the column with a signal at

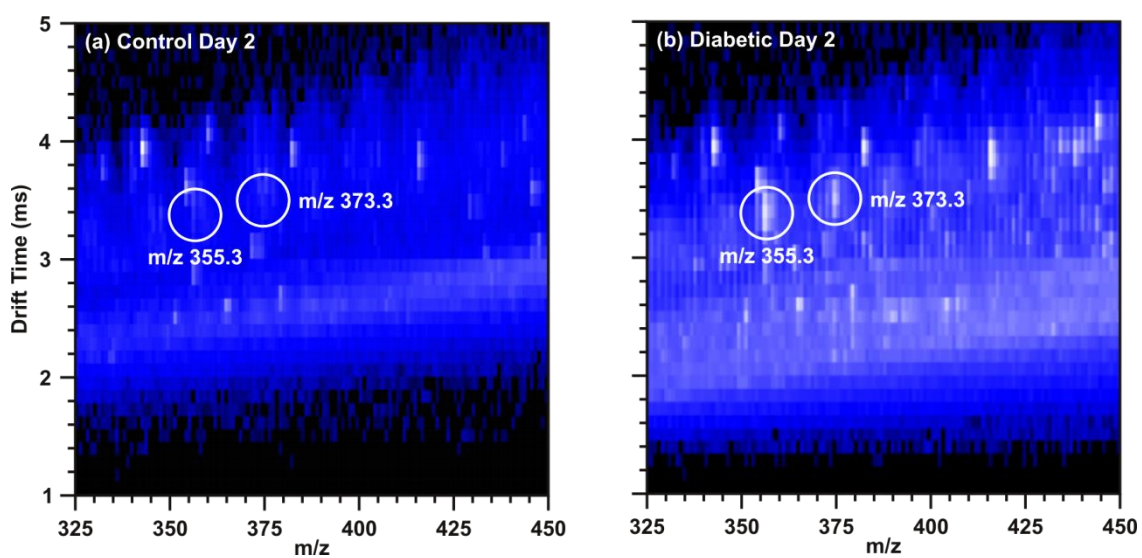


Figure 2.5. Three-dimensional ion mobility-mass spectrum of control day 2 (a) and diabetic day 2 (b) wound fluid. The signals m/z 355.3 and m/z 373.3, encircled in white, are located at 3.28 and 3.50 ms, respectively. These signals are present at decreased intensities in the control day 2 (a) wound fluids relative to the diabetic day 2 (b) wound fluids.

426.32 at 5.31 min (Figure 2.6a,b), confirming the two signals were chemically, as well as structurally, similar. Mobility-extracted fragmentation spectra from post-mobility MS/MS of m/z 373.27 were dominated by a fragment at m/z 355.26. The loss of 18 Da from m/z 373.26 to form the fragment at m/z 355.26 was consistent with a neutral water loss. Both the chromatographic and fragmentation data supported the hypothesis generated from the mobility data that m/z 355.26 and 373.27 are homologous species.

The targeted UPLC-IM-MS analysis was repeated in negative ionization mode and a peak at 5.31 min was observed in the chromatogram (Figure 2.6c), similar to that observed in the positive mode UPLC chromatogram (Figure 2.6a). The mass spectra at 5.31 min contained a signal at m/z 407.2809 (Figure 2.6d), on which post-mobility fragmentation was performed (Figure 2.7a). An accurate mass-based search of the METLIN metabolomics database (<http://www.metlin.scripps.edu>) for m/z 407.2809 suggested that this species was the $[M-H]^-$ signal of cholic acid, with an experimental mass accuracy of 1.47 ppm.

Tandem MS data provided by METLIN and Lipid Maps (<http://www.lipidmaps.org>) matched that obtained from the analysis of diabetic day 2 wound fluids (Figure 2.7a), strongly suggesting that m/z 407.2809 was cholic acid. Fragmentation of a cholic acid standard (Figure 2.7b) was performed to validate the assignment of m/z 407.2809 as cholic acid.

Database searches of the signals present in the ESI+ chromatographic peak at 5.31 min were also performed (Figure 2.6a,b). A mass-based search for m/z 426.3216 strongly suggested that this species was the $[M+NH_4]^+$ adduct of cholic acid, with an experimental mass accuracy of 0.47 ppm. The signals m/z 355.2627 and 373.2735 corresponded to

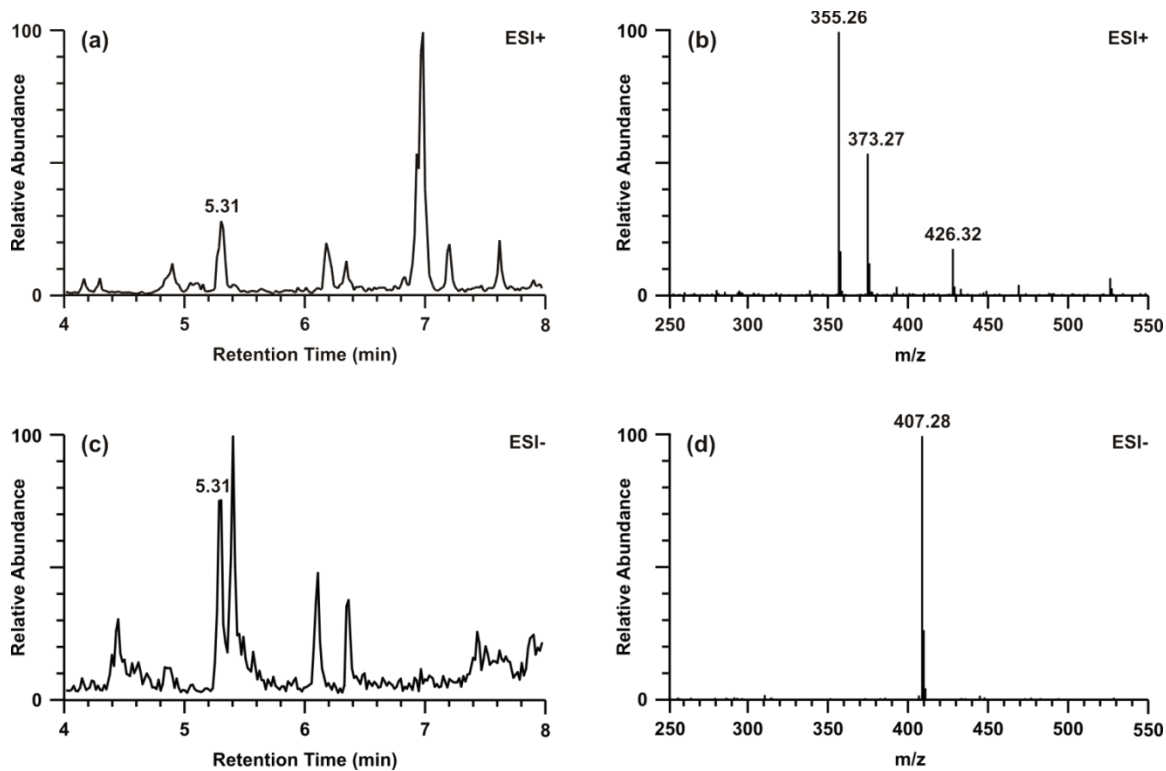


Figure 2.6. Chromatograms (a and c) and mass spectra (b and d) from the positive (a and b) and negative (c and d) mode UPLC-ESI-IM-MS analyses of diabetic day 2 wound fluid. The species m/z 355.26 and 373.27 (b) were observed at 5.31 min in the ESI+ chromatogram (a), with an additional signal at m/z 426.32. A peak at 5.31 min was also observed in the ESI- chromatogram (c), resulting from the species m/z 407.28 (d).

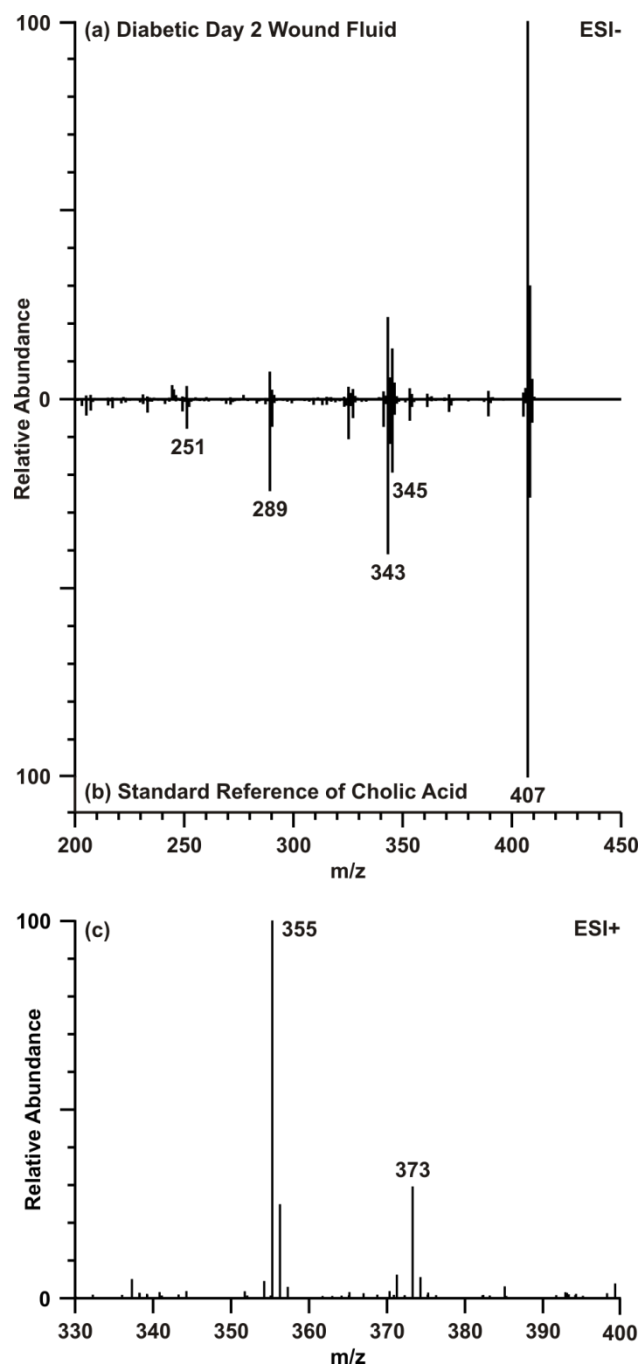


Figure 2.7. (a) Post-mobility MS/MS of m/z 407 (ESI-) in diabetic day 2 wound fluid revealed fragments characteristic of cholic acid. (b) Fragmentation of a standard reference of cholic acid (1 $\mu\text{g/mL}$) was performed to validate the assignment of m/z 407 as cholic acid. (c) The primary species observed from the positive ionization (ESI+) MS analysis of the cholic acid standard were m/z 355 and 373. These signals corresponded to three and two neutral waters losses from cholic acid, respectively.

corresponded to three and two neutral water losses from cholic acid, indicating that these species may be the result of in-source fragmentation of cholic acid in positive ionization mode. An ESI+ analysis of the cholic acid standard confirmed that fragments corresponding to two and three neutral water losses are the primary species observed from cholic acid in positive ionization mode (Figure 2.7c), while the $[M-H]^-$ of cholic acid is the primary species generated by ESI-. Bile acids, such as cholic acid and its derivatives, and their receptors are emerging as important factors in the regulation of glucose homeostasis and diabetes.⁴⁴⁻⁴⁶

2.4. Conclusions

We have illustrated the utility of IM-MS in the analysis of complex biological samples for the study of diabetic wound healing. After minimal sample preparation, IM-MS spectra were rapidly collected for 64 samples of diabetic and non-diabetic wound fluid and four biomolecules distinguishing diabetic and control wound fluid were identified. IM-MS spectra revealed a 10.15 kDa protein, in the form of a unique mobility-mass correlation line, was highly enriched in diabetic day 2 wound fluid. Post-mobility MS/MS aided in the identification of this protein as S100-A8. The separation of S100-A8 in ion mobility drift time from another highly charged, high mass protein provided a 2.7-fold increase in S/N over the MS-only spectrum of the +13 charge state of S100-A8. An OPLS-DA statistical analysis revealed three additional species distinguishing diabetic from control wound fluid, and subsequent UPLC separations and post-mobility MS/MS were performed in a targeted manner to identify and validate these species. The locations of these species in mobility-mass space greatly aided in their identification, as the drift

times indicated that m/z 544.4 was a lipid and the proximity of m/z 355.3 and 373.3 in drift time revealed they were structurally homologous species. A combination of fragmentation experiments and accurate mass database searching led to the identifications of lysophosphatidylcholine (20:4) and cholic acid, which were found at increased intensities in control and diabetic wound fluid, respectively.

The methodology demonstrated in this study can be applied to the analysis of a variety of complex biological systems to rapidly identify biomolecular signatures of diseases and biological processes with IM-MS. In a disease such as the diabetic model examined herein, there can be both increased and decreased (or absent) abundance of chemical species relative to the control. After an intervention, such as the wounding of the tissue, the kinetics of the intervention-induced changes may also differ between the disease model and control. In this study, we focused on the three most abundant and most significantly different species in the first and third quadrants of the S-plot, but with a sufficient investment in effort, an IM-MS/MS workflow as outlined in this paper could be applied to every species in Figure 2.3(b),(d),(f). We expect that such an examination of just the subsets of species that lie within the three we have studied provides, in an efficient manner, an even deeper physiological investigation of the processes that differ between normal and diabetic wound healing. We believe that such an analysis lies midway between narrowly targeted and totally untargeted searches for important contributors to the wound healing process in health and disease.

The growing reliance on structural mass spectrometry stems in part from the flexibility of IM-MS systems, which span smoothly the range between fully targeted and totally untargeted analyses. The workflow that we present demonstrates that a single

instrument can both identify and validate chemical species whose concentrations differ statistically between two experimental conditions. One of the contemporary challenges in metabolomic measurements on complex samples is to transition from untargeted to targeted analysis, not because of instrumentation limitations, but because of the need for both standard samples for target validation and new bioinformatic and biostatistical tools optimized for identification of the salient features in differential and time-series experiments.

2.5. Acknowledgements

This chapter contains the research article: Kelly M. Hines, Samir Ashfaq, Jeffrey M. Davidson, Susan R. Opalenik, John P. Wikswo, and John A. McLean, Biomolecular Signatures of Diabetic Wound Healing by Structural Mass Spectrometry. *Analytical Chemistry* **2013**, 85 (7), 3651-3659.

Financial support for this research was provided by the Vanderbilt University College of Arts and Science, the Vanderbilt Institute for Chemical Biology, the Vanderbilt Institute for Integrative Biosystems Research and Education, funds to JAM from the Defense Threat Reduction Agency (HDTRA1-09-1-0013) and the National Institutes of Health (NIH/NIDA RC2DA028981), Waters Corp., funds provided to SA from the Vanderbilt Diabetes Research and Training Center Summer Diabetes Research Program, and funds provided to JMD from the NIH (AG006528; AR056138) and the Department of Veterans Affairs. The authors thank Dr. Michal Kliman for his discussions and Allison Price for her editorial assistance.

2.6. References

1. National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. Prevention, C. f. D. C. a., Ed. Department of Health and Human Services, Centers for Disease Control and Prevention: Atlanta, GA: U.S., 2011.
2. Jeffcoate, W. J.; Harding, K. G., Diabetic foot ulcers. *Lancet* **2003**, *361* (9368), 1545-1551.
3. Falanga, V., Wound healing and its impairment in the diabetic foot. *Lancet* **2005**, *366* (9498), 1736-1743.
4. Sibbald, R. G.; Woo, K. Y., The biology of chronic foot ulcers in persons with diabetes. *Diabetes-Metabolism Research and Reviews* **2008**, *24*, S25-S30.
5. Werner, S.; Grose, R., Regulation of wound healing by growth factors and cytokines. *Physiological Reviews* **2003**, *83* (3), 835-870.
6. Galkowska, H.; Wojewodzka, U.; Olszewski, W. L., Chemokines, cytokines, and growth factors in keratinocytes and dermal endothelial cells in the margin of chronic diabetic foot ulcers. *Wound Repair and Regeneration* **2006**, *14* (5), 558-565.
7. Lobmann, R.; Ambrosch, A.; Schultz, G.; Waldmann, K.; Schiweck, S.; Lehnert, H., Expression of matrix-metalloproteinases and their inhibitors in the wounds of diabetic and non-diabetic patients. *Diabetologia* **2002**, *45* (7), 1011-1016.
8. Eming, S. A.; Koch, M.; Krieger, A.; Brachvogel, B.; Kreft, S.; Bruckner-Tuderman, L.; Krieg, T.; Shannon, J. D.; Fox, J. W., Differential Proteomic Analysis Distinguishes Tissue Repair Biomarker Signatures in Wound Exudates Obtained from Normal Healing and Chronic Wounds. *Journal of Proteome Research* **2010**, *9* (9), 4758-4766.
9. Khanna, S.; Biswas, S.; Shang, Y.; Collard, E.; Azad, A.; Kauh, C.; Bhasker, V.; Gordillo, G. M.; Sen, C. K.; Roy, S., Macrophage Dysfunction Impairs Resolution of Inflammation in the Wounds of Diabetic Mice. *Plos One* **2010**, *5* (3).
10. Fernandez, M. L.; Broadbent, J. A.; Shooter, G. K.; Malda, J.; Upton, Z., Development of an enhanced proteomic method to detect prognostic and diagnostic markers of healing in chronic wound fluid. *British Journal of Dermatology* **2008**, *158* (2), 281-290.
11. Broadbent, J.; Walsh, T.; Upton, Z., Proteomics in chronic wound research: Potentials in healing and health. *Proteomics Clinical Applications* **2010**, *4* (2), 204-214.

12. Fenn, L. S.; McLean, J. A., Biomolecular structural separations by ion mobility-mass spectrometry. *Analytical and Bioanalytical Chemistry* **2008**, *391* (3), 905-909.
13. McLean, J. A., The Mass-Mobility Correlation Redux: The Conformational Landscape of Anhydrous Biomolecules. *Journal of the American Society for Mass Spectrometry* **2009**, *20* (10), 1775-1781.
14. 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. *Analytical and Bioanalytical Chemistry* **2009**, *394* (1), 235-244.
15. Jackson, S. N.; Ugarov, M.; Egan, T.; Post, J. D.; Langlais, D.; Schultz, J. A.; Woods, A. S., MALDI-ion mobility-TOFMS imaging of lipids in rat brain tissue. *Journal of Mass Spectrometry* **2007**, *42* (8), 1093-1098.
16. McLean, J. A.; Ridenour, W. B.; Caprioli, R. M., Profiling and imaging of tissues by imaging ion mobility-mass spectrometry. *Journal of Mass Spectrometry* **2007**, *42* (8), 1099-1105.
17. McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H., Ion mobility-mass spectrometry: a new paradigm for proteomics. *International Journal of Mass Spectrometry* **2005**, *240* (3), 301-315.
18. Valentine, S. J.; Plasencia, M. D.; Liu, X.; Krishnan, M.; Naylor, S.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E., Toward plasma proteome profiling with ion mobility-mass spectrometry. *Journal of Proteome Research* **2006**, *5* (11), 2977-2984.
19. Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Clemmer, D. E., Resolving and Assigning N-Linked Glycan Structural Isomers from Ovalbumin by IMS-MS. *Journal of the American Society for Mass Spectrometry* **2008**, *19* (11), 1706-1715.
20. Fenn, L. S.; McLean, J. A., Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry. *Physical Chemistry Chemical Physics* **2011**, *13* (6), 2196-2205.
21. Kim, H. I.; Kim, H.; Pang, E. S.; Ryu, E. K.; Beegle, L. W.; Loo, J. A.; Goddard, W. A.; Kanik, I., Structural Characterization of Unsaturated Phosphatidylcholines Using Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry* **2009**, *81* (20), 8289-8297.
22. Kliman, M.; Vijayakrishnan, N.; Wang, L.; Tapp, J. T.; Broadie, K.; McLean, J. A., Structural mass spectrometry analysis of lipid changes in a *Drosophila* epilepsy model brain. *Molecular Biosystems* **2010**, *6* (6), 958-966.

23. Dwivedi, P.; Wu, P.; Klopsch, S. J.; Puzon, G. J.; Xun, L.; Hill, H. H., Metabolic profiling by ion mobility mass spectrometry (IMMS). *Metabolomics* **2008**, *4* (1), 63-80.
24. Dwivedi, P.; Schultz, A. J.; Hill, H. H., Metabolic profiling of human blood by high-resolution ion mobility mass spectrometry (IM-MS). *International Journal of Mass Spectrometry* **2010**, *298* (1-3), 78-90.
25. Enders, J. R.; Marasco, C. C.; Kole, A.; Nguyen, B.; Sevugarajan, S.; Seale, K. T.; Wikswo, J. P.; McLean, J. A., Towards monitoring real-time cellular response using an integrated microfluidics-matrix assisted laser desorption ionisation/nanoelectrospray ionisation-ion mobility-mass spectrometry platform. *Iet Systems Biology* **2010**, *4* (6), 416-427.
26. Ruotolo, B. T.; Gillig, K. J.; Stone, E. G.; Russell, D. H., Peak capacity of ion mobility mass spectrometry: Separation of peptides in helium buffer gas. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002**, *782* (1-2), 385-392.
27. Ruotolo, B. T.; McLean, J. A.; Gillig, K. J.; Russell, D. H., Peak capacity of ion mobility mass spectrometry: the utility of varying drift gas polarizability for the separation of tryptic peptides. *Journal of Mass Spectrometry* **2004**, *39* (4), 361-367.
28. Liu, X.; Plasencia, M.; Ragg, S.; Valentine, S. J.; Clemmer, D. E., Development of high throughput dispersive LC-ion mobility-TOFMS techniques for analysing the human plasma proteome. *Briefings in Functional Genomics & Proteomics* **2004**, *3* (2), 177-186.
29. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H., Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Communications in Mass Spectrometry* **2004**, *18* (20), 2401-2414.
30. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H., An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry* **2007**, *261* (1), 1-12.
31. Hoadlund-Hyzer, C. S.; Li, J. W.; Clemmer, D. E., Mobility labeling for parallel CID of ion mixtures. *Analytical Chemistry* **2000**, *72* (13), 2737-2740.
32. Raftery, M. J.; Geczy, C. L., Identification of posttranslational modifications and cDNA sequencing errors in the rat S100 proteins MRP8 and 14 using electrospray ionization mass spectrometry. *Analytical Biochemistry* **1998**, *258* (2), 285-292.

33. Eckert, R. L.; Broome, A. M.; Ruse, M.; Robinson, N.; Ryan, D.; Lee, K., S100 proteins in the epidermis. *Journal of Investigative Dermatology* **2004**, *123* (1), 23-33.
34. Bouma, G.; Lam-Tse, W. K.; Wierenga-Wolf, A. F.; Drexhage, H. A.; Versnel, M. A., Increased serum levels of MRP-8/14 in type 1 diabetes induce an increased expression of CD11b and an enhanced adhesion of circulating monocytes to fibronectin. *Diabetes* **2004**, *53* (8), 1979-1986.
35. Hessian, P. A.; Edgeworth, J.; Hogg, N., MRP-8 AND MRP-14, 2 ABUNDANT CA-2+-BINDING PROTEINS OF NEUTROPHILS AND MONOCYTES. *Journal of Leukocyte Biology* **1993**, *53* (2), 197-204.
36. Wu, N. J.; Davidson, J. M., Migration inhibitory factor-related protein (MRP)8 and MRP14 are differentially expressed in free-electron laser and scalpel incisions. *Wound Repair and Regeneration* **2004**, *12* (3), 327-336.
37. Graessler, J.; Schwudke, D.; Schwarz, P. E. H.; Herzog, R.; Shevchenko, A.; Bornstein, S. R., Top-Down Lipidomics Reveals Ether Lipid Deficiency in Blood Plasma of Hypertensive Patients. *Plos One* **2009**, *4* (7).
38. Zhao, X.; Fritsche, J.; Wang, J.; Chen, J.; Rittig, K.; Schmitt-Kopplin, P.; Fritsche, A.; Haering, H.-U.; Schleicher, E. D.; Xu, G.; Lehmann, R., Metabonomic fingerprints of fasting plasma and spot urine reveal human pre-diabetic metabolic traits. *Metabolomics* **2010**, *6* (3).
39. Rhee, E. P.; Cheng, S.; Larson, M. G.; Walford, G. A.; Lewis, G. D.; McCabe, E.; Yang, E.; Farrell, L.; Fox, C. S.; O'Donnell, C. J.; Carr, S. A.; Vasani, R. S.; Florez, J. C.; Clish, C. B.; Wang, T. J.; Gerszten, R. E., Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *Journal of Clinical Investigation* **2011**, *121* (4).
40. Barber, M. N.; Risis, S.; Yang, C.; Meikle, P. J.; Staples, M.; Febbraio, M. A.; Bruce, C. R., Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes. *PLoS One* **2012**, *7* (7), e41456.
41. Sevastou, I.; Kaffe, E.; Mouratis, M.-A.; Aidinis, V., Lysoglycerophospholipids in chronic inflammatory disorders: The PLA2/LPC and ATX/LPA axes. *Biochimica et Biophysica Acta, Molecular and Cell Biology of Lipids*, Ahead of Print.
42. Hung, N. D.; Kim, M. R.; Sok, D.-E., Anti-inflammatory action of arachidonoyl lysophosphatidylcholine or 15-hydroperoxy derivative in zymosan A-induced peritonitis. *Prostaglandins & Other Lipid Mediators* **2009**, *90* (3-4).
43. Nguyen Dang, H.; Sok, D.-E.; Kim, M. R., Prevention of 1-palmitoyl lysophosphatidylcholine-induced inflammation by polyunsaturated acyl lysophosphatidylcholine. *Inflammation Research* **2012**, *61* (5).

44. Duran-Sandoval, D.; Mautino, G.; Martin, G. V.; Percevault, F.; Barbier, O.; Fruchart, J. C.; Kuipers, F.; Staels, B., Glucose regulates the expression of the farnesoid X receptor in liver. *Diabetes* **2004**, *53* (4), 890-898.
45. Lefebvre, P.; Cariou, B.; Lien, F.; Kuipers, F.; Staels, B., Role of Bile Acids and Bile Acid Receptors in Metabolic Regulation. *Physiological Reviews* **2009**, *89* (1), 147-191.
46. Cai, S.; Huo, T. G.; Xu, J. H.; Lu, X. M.; Zheng, S. N.; Li, F. M., Effect of mitiglinide on Streptozotocin-induced experimental type 2 diabetic rats: A urinary metabonomics study based on ultra-performance liquid chromatography-tandem mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2009**, *877* (29), 3619-3624.

CHAPTER III

STRUCTURAL MASS SPECTROMETRY OF TISSUE EXTRACTS TO DISTINGUISH CANCEROUS AND NON-CANCEROUS BREAST DISEASES

3.1. Introduction

Among women in the United States, breast cancer is the most prevalent type of invasive cancer, affecting 118.7 women per 100,000 in 2010, the most recent year for which statistics are available.¹ Although it is currently the second leading cause of cancer deaths, the mortality due to breast cancer has steadily decreased from 26 women per 100,000 in 2001 to 21.9 in 2010.¹ The general decrease in breast cancer-related mortalities is in part due to the discovery of novel diagnostic and prognostic markers, and the subsequent development of targeted cancer therapies born of extensive genomic and proteomic profiling endeavors.^{2,3} Despite these advances, there still remain breast cancer subtypes, such as triple negative breast cancer (named so due to the absence of genetic markers HER2, PR, and ER), which do not respond to the currently available targeted therapies.^{4,5}

Metabolomics approaches have also been applied in the efforts to discover molecular differences between tumor and healthy cells. As the downstream endpoint of changes in the genome or proteome, metabolites best represent the cellular phenotype while also reflecting environmental influences.^{6,7} Cellular metabolism is significantly altered in the transformation of healthy cells into malignant cells, likely due to the rapid cellular proliferation in cancer.⁶ Affected pathways include glycolysis, oxidative

phosphorylation, choline metabolism, and protection against reactive oxygen species, giving rise to a metabolic phenotype common to cancers in general.^{6,8,9} Among the most notable hypotheses of aberrant metabolism in cancer is the Warburg effect, which describes cancer cells' preferential use of glycolysis, typically an anaerobic process, to generate energy despite aerobic conditions amenable to oxidative phosphorylation.^{6,8-10} Thus, interrogation of the metabolic phenotype of cancer may provide targets for therapeutics designed to decrease tumor viability by disrupting the metabolic drivers specific to tumors.⁶

Among the tools typically used to perform cancer metabolomics studies are nuclear magnetic resonance (NMR) and mass spectrometry (MS).⁷⁻⁹ The former offers a number of tailored approaches, from proton (¹H) and carbon (¹³C) NMR to phosphate (³¹P) NMR which may be used to monitor pathway-specific high-energy phosphate metabolites.^{11,12} Analysis of solid tissues is feasible with magnetic resonance spectroscopy (MRS) approaches such as high resolution magic angle spinning (HR-MAS).^{13,14} These techniques are nondestructive, which is an advantage when working with limited volumes of sample or tissue; however, they are generally less sensitive than MS-based approaches.^{14,15} Chromatographic separations, such as gas chromatography (GC) and ultra-performance liquid chromatography (UPLC), are frequently combined with MS to increase peak capacity, reduce ionization suppression effects, and reduce mass spectral complexity. However, the derivatization necessitated for GC-MS analyses can be undesirable. MS-based cancer metabolomics are amenable to a variety of sample types including serum, plasma, urine and tissues.¹⁵ These studies may be targeted or untargeted, where targeted approaches aim to measure a predefined subset of molecules

based on their class or pathway and untargeted approaches aim to observe as many metabolites as possible without bias.¹⁶

Ion mobility-mass spectrometry (IM-MS) is a hybrid two-dimensional technique which combines the gas phase structural separation of IM with mass-to-charge (m/z) separation of mass spectrometry. Briefly, IM separation occurs as ions travel through a drift cell containing a neutral buffer gas, such as helium or nitrogen, under the influence of a static or dynamic electric field. As the ions traverse the drift cell, they experience a number of collisions with the neutral buffer gas dependent on the collision cross section (CCS; effective ion surface area) of the ion. This process results in a characteristic drift time in the range of micro- to milliseconds, which can be used to calculate CCSs and determine coarse structural information.¹⁷

IM-MS-based analyses have been demonstrated for fields ranging from proteomics^{18, 19} to systems biology.^{20, 21} For metabolomics analyses of complex biological samples,²²⁻²⁵ IM-MS offers unique advantages relative to MS-only approaches. IM-MS platforms are highly flexible, allowing a range of pre-ionization separations, such as UPLC, to be combined in tandem with the IM-MS experiment. Relative to UPLC-MS approaches, UPLC-IM-MS provides even greater peak capacity without increasing the time of analysis due to the complementary time scales of minutes, milliseconds and microseconds for the UPLC, IM, and MS dimensions, respectively.^{18, 19, 26-28} For classes of biomolecules, such as lipids, proteins, and carbohydrates, a correlation is observed between mass and size, or m/z and CCS, based on their gas phase packing efficiencies. In an IM-MS experiment, this correlation results in the separation of each biomolecular class into unique regions, or trendlines, in IM-MS conformation space.²⁹⁻³¹ The

separation of biomolecular classes in IM-MS conformation space enables a more holistic approach with minimal sample preparation for the analysis of complex biological samples, where multiple species of biomolecules may be present.³² Rather than performing sample purification strategies to isolate a particular class of biomolecules, the molecules of interest (*e.g.* metabolites) can be separated from other biomolecular classes in the IM dimension, effectively increasing the signal-to-noise (S/N). For metabolomics, this is particularly advantageous as isobaric species may be resolved in the IM dimension based on the differences in their gas phase structures. Lastly, data independent acquisition of MS/MS spectra post-ion mobility enables multiplexed fragmentation experiments to be performed nearly simultaneously, and minimizing the need to perform additional experiments to obtain fragmentation data.

Here, we describe an UPLC-IM-MS/MS approach to the characterization of the metabolites differentially expressed in breast cancer. Based on a previously demonstrated workflow for the analysis of complex biological samples,³² disease (n=3) and control (n=3) breast tissues were homogenized, polar metabolites were extracted and characterized by UPLC-IM-MS/MS.³³ Although it was known which tissues represented healthy and disease, researchers were blind to the exact pathologies of the tissues. Principal components analysis (PCA) revealed an unexpected grouping of the breast tissues. Features giving rise to this separation in the PCA were interrogated to determine tentative molecular identifications based on chromatographic retention time, accurate mass, drift time, and fragmentation analysis. The merits of IM-MS for untargeted metabolomics analyses are demonstrated in the transition from a list of features to

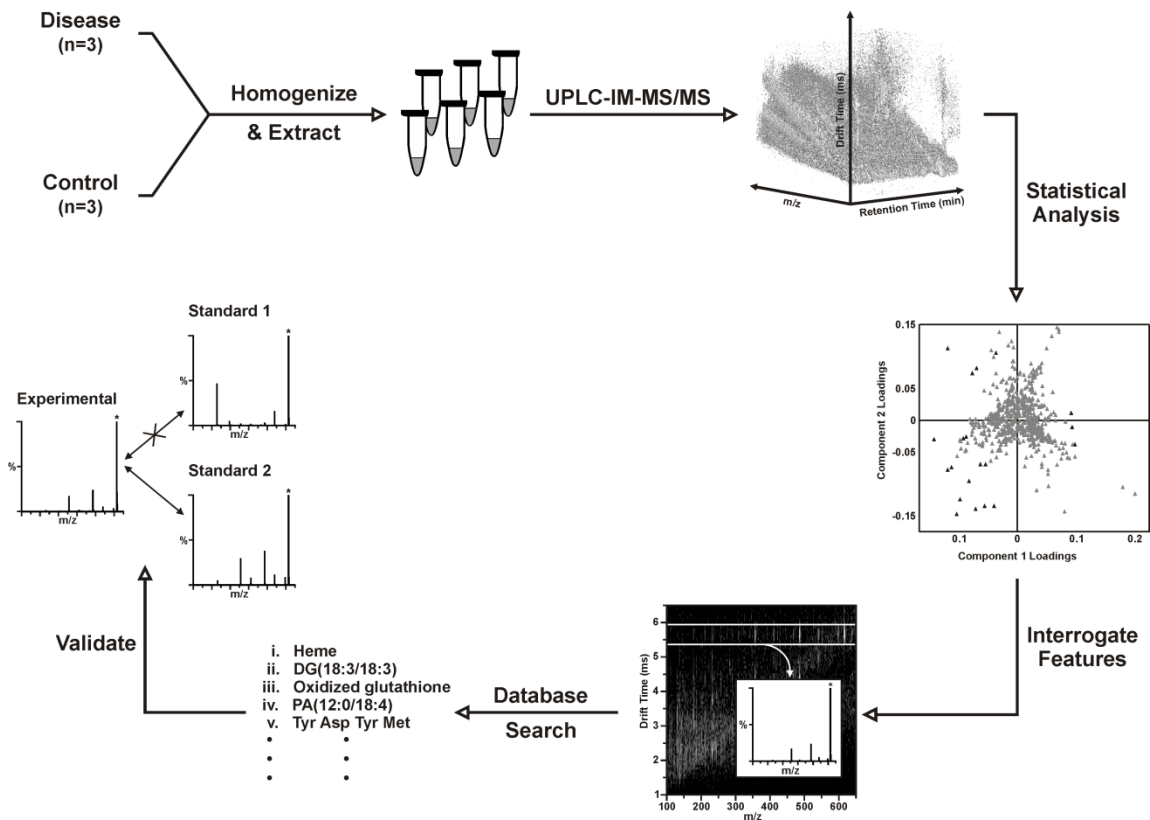
validated identifications of the biomolecular species distinguishing cancerous from non-cancerous tissues.

3.2. *Experimental*

The workflow demonstrated here for the UPLC-IM-MS/MS analysis of human tissues, as described below, is based on a previously developed workflow to transition from wholly untargeted to targeted analyses of complex biological samples in the pursuit of identifying key biomolecular features distinguishing disease and healthy conditions.³² The general workflow demonstrated here is illustrated in Scheme 3.1, where this methodology has been adapted to include the additional online UPLC separation, methods for chromatographic peak picking and alignment, post-mobility data-independent acquisition of MS/MS spectra, and inclusion of MS/MS spectra in addition to accurate mass as a parameter for generating tentative identifications.

Tissues

Six surgically resected fresh-frozen tissues were selected from the Meharry Medical College Translational Pathology Shared Resource Core Facility to control for gender, age and ethnicity. Matched controls were used when available, such that the control would be grossly normal peritumoral tissue from the same patient. The specific pathology of the disease tissues was withheld from the researchers to create a partially blinded experiment, however it was known which three of the six tissues represented a form of breast disease (C, control; D, disease; matched pairs indicated by consecutive numbering (*i.e.* 1C and 2D are a matched pair)). Tissues were stored at -80°C.



Scheme 3.1. Illustration of the workflow for the preparation and extraction of breast tissues, including the steps necessary to transition from a multidimensional dataset to the identification of statistically-prioritized molecular features.

Sample Preparation

The procedure for homogenization and extraction of the human breast tissues was adapted from the methods described by Want *et al.* for the extraction of polar metabolites from tissues for UPLC-MS analysis.^{33,34} Intact tissues were initially coarsely homogenized on ice in a dounce homogenizer, and 50 ± 4 mg (wet) of each tissue was transferred to an eppendorf tube. To each tissue, 1 mL of cold 1:1 methanol/water (v/v) (Chromasolv, Sigma-Aldrich, St. Louis, MO) was added and the samples were further homogenized on ice using a hand-help homogenizer with disposable plastic probes (Omni International, Kennesaw, GA). A fresh disposable probe was used for each sample. An additional 500 μ L of cold 1:1 methanol/water was added to each sample for a total volume of 1.5 mL, and extraction was allowed to proceed overnight at -20°C .

Samples were centrifuged (16,500 g for 5 min at 2°C ; Heraeus Fresco 21, Thermo Scientific) and the supernatants were transferred to fresh eppendorf tubes. For the UPLC-IM-MS/MS analysis, 750 μ L of each supernatant was dried on a speed-vac and reconstituted with 200 μ L of H_2O . Protein precipitation was performed by adding cold (-20°C) methanol in a 3:1 ratio, or 600 μ L, to the samples on ice. Samples were then transferred to dry ice for 5 min, after which they were vortexed and centrifuged. Supernatants were transferred to fresh tubes, dried on a speed-vac, and then stored at -80°C .

Samples were reconstituted with 500 μ L of H_2O with 0.1% formic acid (HPLC-grade, Fisher Scientific), vortexed, and centrifuged briefly. From each sample, 250 μ L was transferred to autosampler vials. A pooled quality control (QC) sample was prepared with 30 μ L of each sample (taken from the 250 μ L), for a total volume of 180 μ L.

UPLC-IM-MS/MS

A nanoACQUITY (Waters, Milford, MA.) UPLC system, modified to provide 60 $\mu\text{L min}^{-1}$ flow rates, was used to perform chromatographic separations on an ACQUITY HSS C18 column (1.8 μm , 1.0 x 100mm; Waters, Milford, MA.). The autosampler and column temperatures were maintained at 4 and 40°C, respectively, and an injection volume of 6 μL was used to overfill the 5 μL loop. Chromatographic separation was performed with a binary solvent system, where solvent A was 0.1% formic acid in water (Fisher Scientific) and solvent B was 0.1% formic acid in acetonitrile (Chromasolv, Sigma-Aldrich). Gradient conditions for the 25 min run with 60 $\mu\text{L min}^{-1}$ flow rate were as follows: initial, 99% A; 1 min, 99% A; 6 min, 40% A; 16 min, 1% A; 18 min, 1% A; 19 min, 99% A; 21 min, 99% A.

IM-MS/MS was performed on a Synapt G2 ion mobility-mass spectrometer (Waters, Milford, MA.), which utilizes a traveling wave IM cell pressurized with nitrogen gas and an orthogonal time-of-flight mass spectrometer (TOFMS) operated in the single stage reflectron configuration.^{35, 36} The outflow from the chromatographic system was couple to the instrument through the electrospray ionization (ESI) source. Conditions for positive mode ESI were as follows: capillary, 3 kV; sampling cone, 40 V; extraction cone, 7 V; source temperature, 80°C; desolvation temperature, 150°C; cone gas flow, 20 L/hr; desolvation gas flow, 300 L/hr. Mass calibration was performed with sodium formate in the range of m/z 50-1400, and a leucine enkephalin lockmass signal was continuously acquired throughout the MS acquisition for external mass correction of the data. IM separation was achieved with a traveling wave velocity of 650 m s^{-1} and height

of 40 V. Data independent MS/MS by collision induced dissociation (CID) was performed in the transfer region with collision energies alternating between 10 and 30 V.

The sample queue was prepared to run one set of technical replicates (injections of a sample) in a randomized order, bracketed by QCs samples. This was repeated two more times to provide three technical replicates for each sample and four QC injections through the queue to monitor performance.

Biostatistics

Data was first mass corrected using the continuously acquired lockmass signal, and the data was centroided during this process. The corrected data were then processed with ProteoWizard (version 3.0.4243) MSConvert to convert the .raw files to .mzXML files.³⁷ The .mzXML files were processed with XCMS (Scripps, La Jolla, CA.) to perform peak picking and peak alignment in the chromatographic domain.^{38,39} Briefly, the “matchedFilter” algorithm was used for peak picking and peak alignment and retention time correction was performed with the “obiwarp” algorithm. Details of the XCMS processing are provided in Appendix C. The output from XCMS was normalized such that the sum of all intensities within a sample equaled 10,000. This dataset was then imported into Extended Statistics (Umetrics) for visualization of multivariate statistical analyses. PCA was used to determine the quality of the dataset, in terms of the grouping of triplicate technical replicates and grouping of QCs injections near the origin of the PCA plot. Model parameters (R2Y and Q2) for the PCA are provided in Appendix C. Analysis of the corresponding loadings plot was used to identify the features contributing to the score of each sample along PC1 and PC2 of the PCA plot. To determine

significance, p -values were calculated with the Student's t -test for means (two-tailed, equal variance, $\alpha = 0.05$) using the normalized aligned dataset. Means were calculated from the three technical replicates of each sample, and all reported p -values were calculated between the matched disease and control samples.

Bioinformatics

For the statistically prioritized molecular features, values of m/z were retrieved from the lockmass-corrected .raw files to ensure the best mass accuracy. In addition, the IM-MS spectra containing the post-mobility data-independent MS/MS acquisitions (saved as function 2 of the data file) were accessed and mobility-organized fragmentation spectra for the features were extracted. This was performed by extracting a defined window of retention time containing the chromatographic peak in which the feature eluted, followed by extracting a defined window of drift time across all m/z which bracketed the drift time of the feature (see Figures 3.2 and 3.4). When possible, both accurate mass and fragmentation information were used to make tentative identifications from database searches. Databases used for generating tentative identifications by accurate mass included the Human Metabolite Database (HMDB, <http://www.hmdb.ca>), KEGG (<http://www.kegg.com>), and METLIN⁴⁰⁻⁴² (<http://metlin.scripps.edu>). The METLIN MS/MS spectrum match feature or the MetFrag⁴³ (<http://msbi.ipb-halle.de/MetFrag/>) *in silico* fragmentation tool were used to search the experimental fragmentation spectra peak lists, which were filtered by intensity to include only the top 30 peaks.

Validation

Tentative identifications were validated with standards when possible. The experimental MS and MS/MS spectra, as well as the drift times, were compared against those of the standard. Standards of glutathione, oxidized glutathione, and adenosine 5'-monophosphate were purchased from Sigma-Aldrich (St. Louis, MO.). Standards were prepared at concentrations from 1-3 $\mu\text{g/mL}$ in 1:1 methanol/water with 0.1% formic acid. Each standard was directly infused into the ESI source of the Synapt G2 with an external syringe pump (10 $\mu\text{l/min}$ flow rate) with ionization conditions identical to those described above. Post-mobility fragmentation was performed at collision energies between 10 and 30 V to approximate the conditions of the data-independent MS/MS acquisitions.

3.3. Results and Discussion

Data representative of the UPLC-IM-MS analysis of sample 4D is presented in Figure 3.1. The IM-MS plot (Figure 3.1a), shows the dimensions of m/z on the x -axis and drift time (ms) on the y -axis. Intensity, measured as counts, is depicted as a color scale, where white indicates high intensity signals, black indicates the absence of signals, and blue represents low-to-medium intensity signals. A portion of the UPLC chromatogram from the analysis of 4D is presented in panel (b). User-defined regions of the multidimensional dataset may be selectively extracted to isolate species of interest and effectively increase the S/N by separating those species from chemical noise.³² Figure 3.1 demonstrates the extraction of two acetylated polypeptides, thymosins $\beta 4$ and $\beta 10$, in

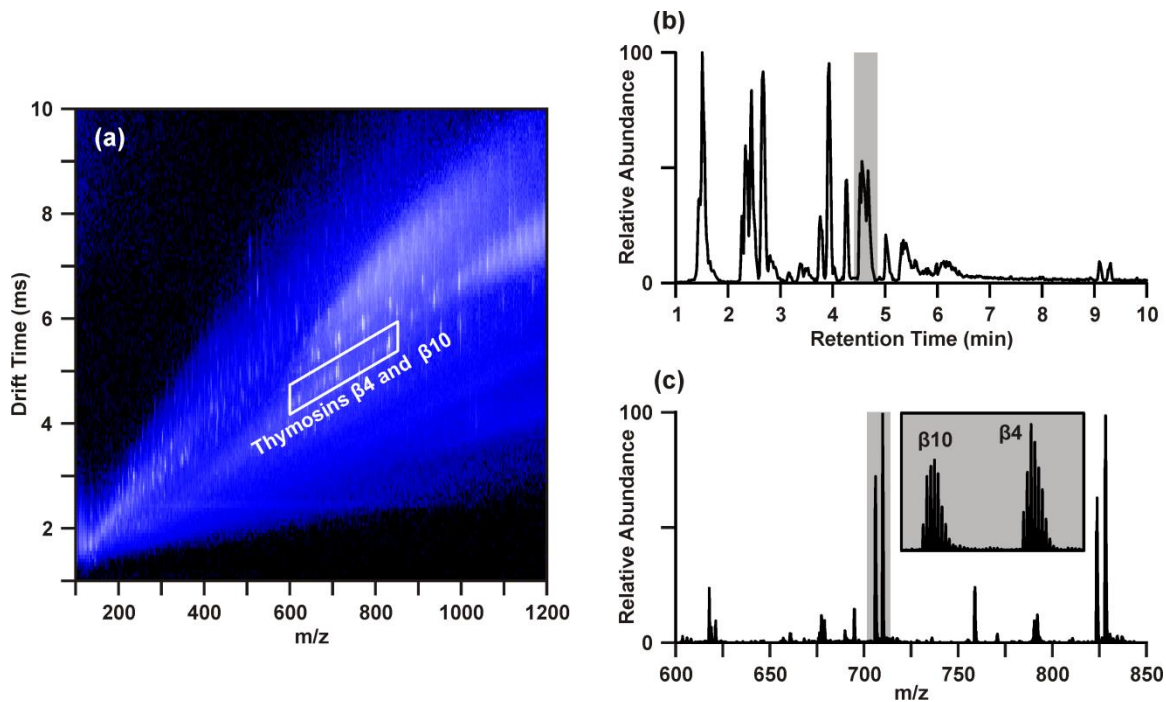


Figure 3.1. Data representative of the UPLC-IM-MS/MS characterization of breast tissue extract 4D. A user-defined region of conformation space, containing the polypeptides Thymosin β 4 and Thymosin β 10, is indicated on the IM-MS spectrum (a). Panel (b) represents a portion of the UPLC chromatogram from the same analysis, where the chromatographic peak in which the Thymosin polypeptides eluted is highlighted by the grey bar. Extraction of these signals in both the drift time and chromatographic dimensions provides a mass spectrum (c) with improved S/N in which the isotopic distributions (inset, magnification of the area highlighted by grey bar) of the 7+ species are resolved.

both the chromatographic (b) and drift time dimensions (a). The chromatographic peak containing thymosins $\beta 4$ and $\beta 10$, as indicated in Figure 3.1(b) by the grey bar, may be extracted to yield the IM-MS plot containing the polypeptides and any co-eluting species. Likewise, a user defined area (outlined in white in (a)) of m/z -drift time space, or conformation space, containing the polypeptides may be extracted to provide an extracted ion chromatogram for thymosins $\beta 4$ and $\beta 10$. Extraction in both the chromatographic and ion mobility dimensions yields a mass spectrum (c) of the polypeptides (highlighted in grey), in which the isotopic distributions (inset) of the multi-charged species are well-enough resolved to determine their charge states as $7+$. This strategy may be applied in both the analysis of IM-MS and post-mobility-MS/MS spectra, where extraction in chromatographic and IM dimensions effectively increases confidence in tentative identifications by reducing isobaric interferences.^{31,32,36}

Although the presence or absence of molecular features may be visually observed from the IM-MS plots representing different disease statuses (*i.e.* disease vs. control), multivariate statistical analyses are required to detect more subtle variations in the expression of molecular species as a function of disease status. However, utilization of the drift time information in the initial peak picking and peak alignment remains a challenge for biostatistical tools. Therefore, peak picking and alignment of data resulting from the analysis of the breast tissue extracts was performed at the chromatographic level and the ion mobility dimension was returned to as an aid in generating or filtering putative identifications of statistically significant molecular features.

Results from the statistical analysis of the breast tissue extract dataset are summarized in Figure 3.2. As described previously, samples were injected in three

batches, where each batch contained one replicate injection for each sample a in randomized order bracketed by QC samples (not shown). The PCA score plot (Figure 3.2a) provides an overview of the samples. Grouping of the technical replicates for each sample, as seen in the score plot, indicates good reproducibility throughout the analysis. Samples forming a matched pair of disease and control tissues are identified with the dashed ellipses on the score plot (Fig. 3.2a). Generally, the 4D and 6D separated from the other samples along principal component (PC) 1, but were separated from each other along PC2. Interestingly, 2D grouped with the control tissues. This suggested that although 2D represented a tissue affected by disease, this tissue was generally more metabolically normal than the pathologically abnormal tissues 4D and 6D. The initial hypothesis as to the location of 2D in the PCA score plot was that this tissue represented a benign breast disease, whereas 4D and 6D potentially represented more malignant diseases of the breast.

The corresponding loadings plot (Figure 3.2b) was investigated to determine the specific molecular features giving rise to the separations observed in the PCA score plot. In general, the loadings plot presents the coefficients or weights assigned to each variable in the process of generating the principal components. The values of the PC1 and PC2 loadings for a particular feature are representative of the correlation between the respective PC and the feature. For example, feature 16 in Figure 3.2(b) has a loading of 0.09 for PC1 and a loading near zero (0.01) for PC2. This indicates that PC1 is highly correlated with feature 16 and PC2 is poorly correlated with feature 16. Similar to interpretation of the PCA score plot, features which group together in the loadings plot demonstrate similar behavior.

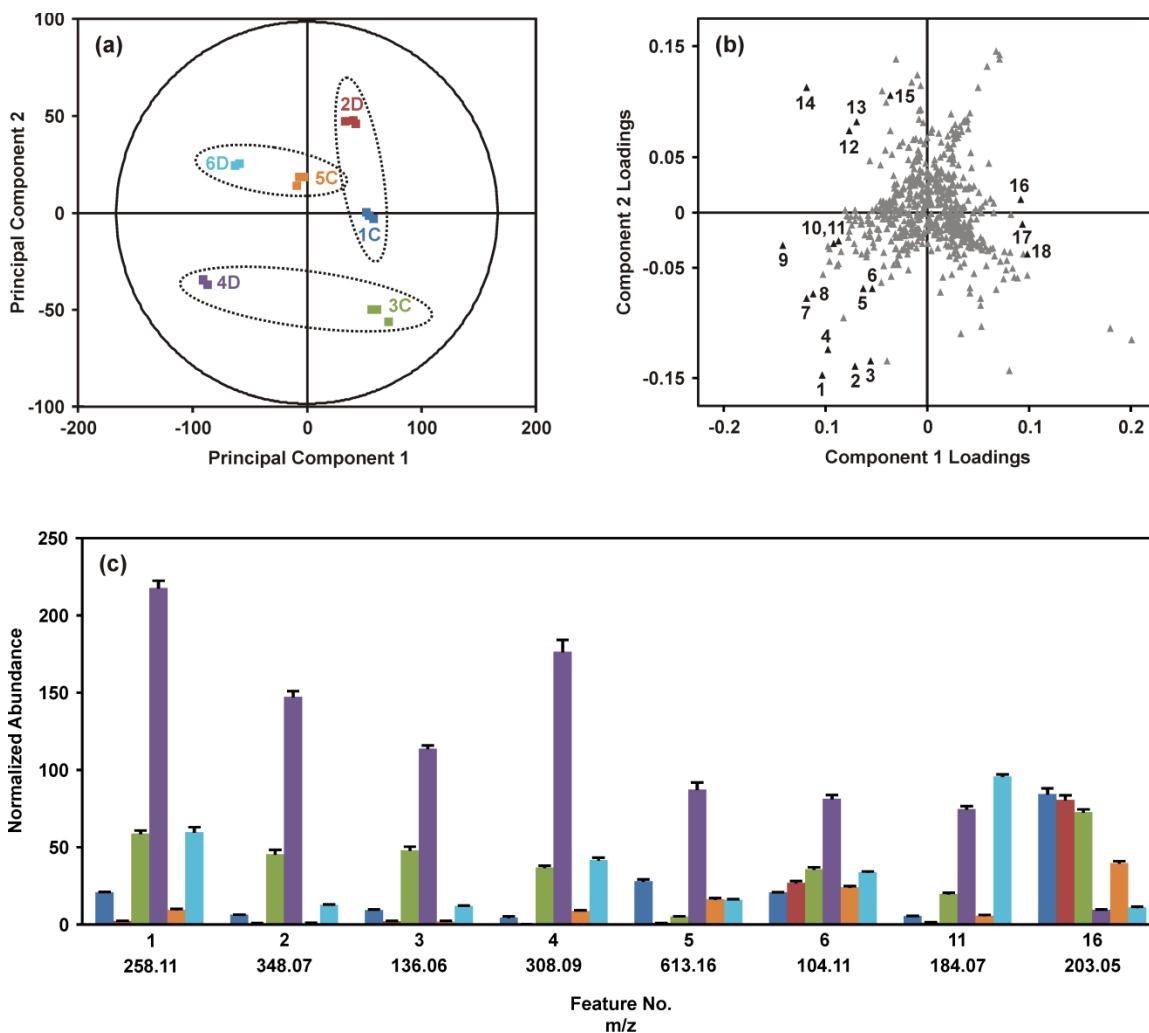


Figure 3.2. Summary of the multivariate statistical analysis of the breast tissue extract dataset. (a) Grouping of technical replicates and separation of biological replicates is shown by PCA of the UPLC-IM-MS/MS data. The plot contains three technical replicates for each sample, and each sample is represented by a different color. Matched control and disease tissues are indicated with ellipses. (b) Analysis of the corresponding loadings plot indicates the molecular features contributing to the separation of the samples along PC1 and PC2. Select features (in black) showing statistical significance are numbered. Average normalized abundances from three technical replicates for select features are shown in (c). The coloring scheme is the same as (a). Error bars indicate the standard deviations of the means.

The loadings (Fig. 3.2b) and score plots (Fig. 3.2a) may be compared to understand the relationship between the features, shown in the loadings plot, and the samples, shown in the score plot. The chart shown in Figure 3.2(c) provides the average normalized abundances of several of the features labeled in Fig. 3.2(b) across the six tissues. Looking at feature 16, a clear pattern emerges where the samples to the right of PC1 (2D, 1C and 3C) all have high abundances of feature 16 (m/z 203.05), while samples to the left of PC1 (4D and 6D) have significantly lower abundances. Sample 5C, which has a score near zero on PC1, has an intermediate abundance of m/z 203.05. This feature contributes significantly to the variability extracted by PC1, which generally separates the disease and control samples. Likewise, features 1-4 in the lower left quadrant of the loadings plot are all substantially more abundant in tissue 4D than the other tissues, giving rise to the location of 4D in the lower left quadrant of the PCA score plot. Fold-changes for the data shown in Figure 3.2(c) may be found in Table 3.1, along with the m/z and retention time associated with the feature.

The raw data was then revisited to determine tentative identities of the differentially abundant molecular features. Upon investigating feature 4, m/z 308.09, the extracted ion chromatogram indicated this species was found in two chromatographic peaks (retention times 2.47 and 3.78 min). While the peak at 2.47 min contained m/z 308.09 as the base ion, the chromatographic peak at 3.78 min contained a base ion of m/z 613.16. Closer inspection of the mass spectrum around m/z 308.1 revealed that there were overlapping isotopic envelopes present (Figure 3.3b). Analysis of the IM-MS spectrum using the same approach described above revealed that there were in fact two species contributing to the spectrum in Fig. 3.3(b). In contrast to one dimensional MS analyses,

Table 3.1. Tentative identifications, mass accuracy, and fold-changes of the features highlighted in loadings plot (Figure 3.2b).

No.	m/z	RT	Tentative ID	ppm	2D ^{a, b}	4D ^{a, c}	6D ^{a, d}
1	258.11	1.51	Glycerophosphocholine	9.7	↓ 8.1	↑ 3.7	↑ 6.0
2	348.07	2.33	Adenosine monophosphate (V)	3.2	↓ 5.7	↑ 3.2	↑ 10.5
3	136.06	2.33	Adenine ^e	6.6	↓ 3.9	↑ 3.4	↑ 5.2
4	308.09	2.45	Reduced glutathione (V)	5.2	↓ 7.8	↑ 4.8	↑ 4.6
5	613.16	3.76	Oxidized glutathione (V)	5.5	↓ 20.1	↑ 15.6	N.S. ^f
6	104.11	1.51	Choline	7.7	↑ 1.3	↑ 2.2	↑ 1.4
7	705.94	4.56	Thymosin β10, acetylated (7+)	6.5	↓ 7.4	↑ 22.1	↑ 2.3
8	823.44	4.56	Thymosin β10, acetylated (6+)	9.6	↓ 6.8	↑ 29.9	↑ 2.5
9	827.76	4.56	Thymosin β4, acetylated (6+)	5.9	↓ 11.8	↑ 10.4	↑ 2.0
10	152.06	2.26	Guanine	6.6	↓ 4.4	↑ 17.1	↑ 3.5
11	184.07	1.51	Phosphocholine	6.0	↓ 3.5	↑ 3.7	↑ 5.8
12	137.05	3.93	Hypoxanthine	9.5	↑ 1.3	↑ 2.0	↑ 1.5
13	246.17	5.01	2-Methylbutyrylcarnitine Pivaloylcarnitine	3.7	↑ 1.9	N.S. ^f	↑ 67.8
14	137.05	2.67	Allopurinol	10.9	↑ 3.3	↑ 3.5	↑ 2.6
15	152.06	3.93	2- or 8-Hydroxyadenine	5.9	↑ 2.1	↑ 5.7	↑ 1.7
16	203.05	1.45	Monosaccharide [M+Na] ⁺	5.9	N.S. ^f	↓ 7.3	↓ 3.5
17	437.19	9.31	Unknown	-	N.S. ^f	↓ 3.2	↓ 1.9
18	387.19	9.14	Unknown	-	N.S. ^f	↓ 7.0	↓ 2.1

^a Arrows indicate directionality of fold-change, where ↑ represents increase in tumor and ↓ represents decrease in tumor relative to the respective matched control. The *p*-values (shown in supporting information document) were calculated with the two-tailed student's *t*-test with equal variance. *p* ≤ 0.05 was used to determine significance; ^b Fold change and *p*-value calculated against the matched control, 1C; ^c Fold change and *p*-value calculated against the matched control, 3C; ^d Fold change and *p*-value calculated against the matched control, 5C; ^e MS/MS of adenosine 5'-monophosphate standard (Appendix C) revealed presence of *m/z* 136.06 with no or little collision energy applied. Feature #4 may be adenine fragment of adenosine monophosphate (feature #3). ^f N.S. indicates no significance, as determined by *p* > 0.05. "V" indicates features whose identifications were validated with standards.

these overlapping isotopic distributions can be easily resolved in the IM dimension as the separation occurs on the basis of size-to-charge, where the general trend is that multiply ($\geq 2+$) charged species have short drift times owing to greater charge repulsion effects.^{17,31,36} This can be seen in Figure 3.3(a-d), where (b) represents the mass spectrum obtained from summing across all drift time as demonstrated with the white dashed lines in (a), mimicking a typical one dimensional MS analysis. Creating a user-define area of drift time- m/z space by combining the vertical white dashed lines and the horizontal blue dashed lines allows the peak inside the blue box to be extracted away from the isobaric inference. This yields the spectrum presented in (c), from which it is evident that this species, m/z 308.09 is singly charged. Performing the same steps for the region outline in (a) with green dashed lines provides a mass spectrum (d) of a doubly charged species, m/z 307.09. This corresponds to a molecular weight of 612.16, or $[M+H]^+$ of 613.16, which corresponds to feature 5 (Table 3.1). Database searches of the m/z and molecular weight values suggested that m/z 308.09 (feature 4) and m/z 307.09 (*i.e.* $[M+H]^+$ 613.16, feature 5) were reduced and oxidized glutathione, respectively.

The post-mobility fragmentation spectra for m/z 308.09 and m/z 613.16 were studied to assess the plausibility of the potential identifications of reduced and oxidized glutathione as features 4 and 5, respectively. The chromatographic peak containing m/z 613.16 at 3.78 min was extracted to provide the IM-MS spectrum of the data independent MS/MS acquired post-mobility separation, shown in Figure 3.4(a). Multiple regions of fragmentation were evident in the IM-MS spectrum, as outlined by the purple, green and blue lines. As fragmentation occurs after mobility separation, the fragment ion should

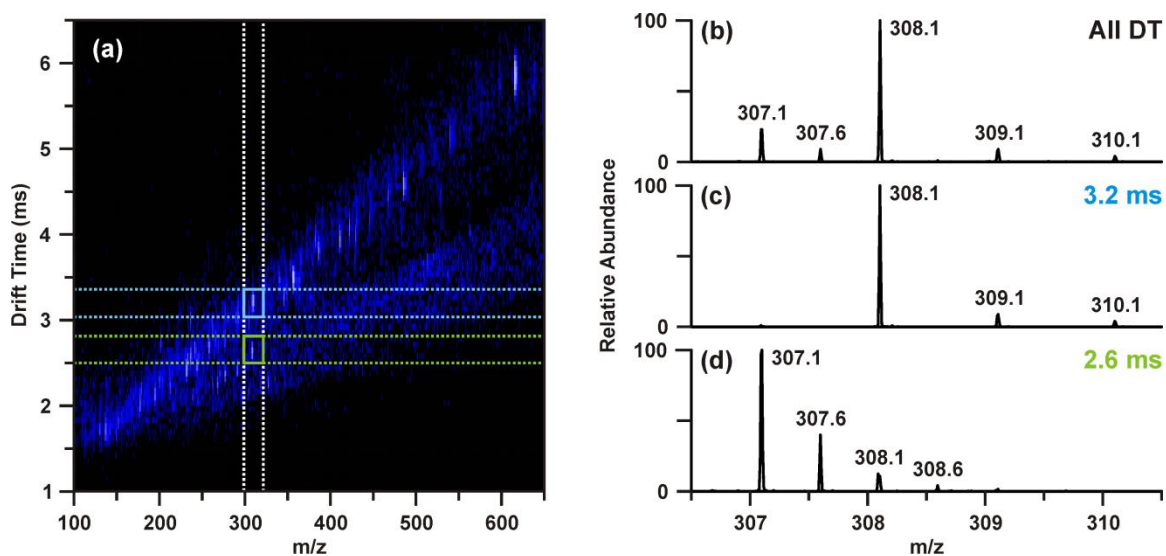


Figure 3.3. Demonstration of IM separation of isobaric overlapping isotopic distributions convoluting the mass spectrum around m/z 308.1 at retention time 3.78 min. (a) IM-MS plot with regions correlating to extracting the mobility separated isobaric ion. Extracting a defined window around m/z 308 across all drift times (outlined by the white dashed lines) yields the extracted mass spectrum (b). Combining the defined window about m/z 308 with a discrete window of drift times (blue dashed lines) allows the peak in the blue box to be extracted (c) away from the isobaric interference (green dashed lines, c). The end result is separate mobility-separated mass spectra resolving the isotopic distributions to that of a singly-charged m/z 308.1 (feature 4) and doubly charged m/z 307.1 (feature 5).

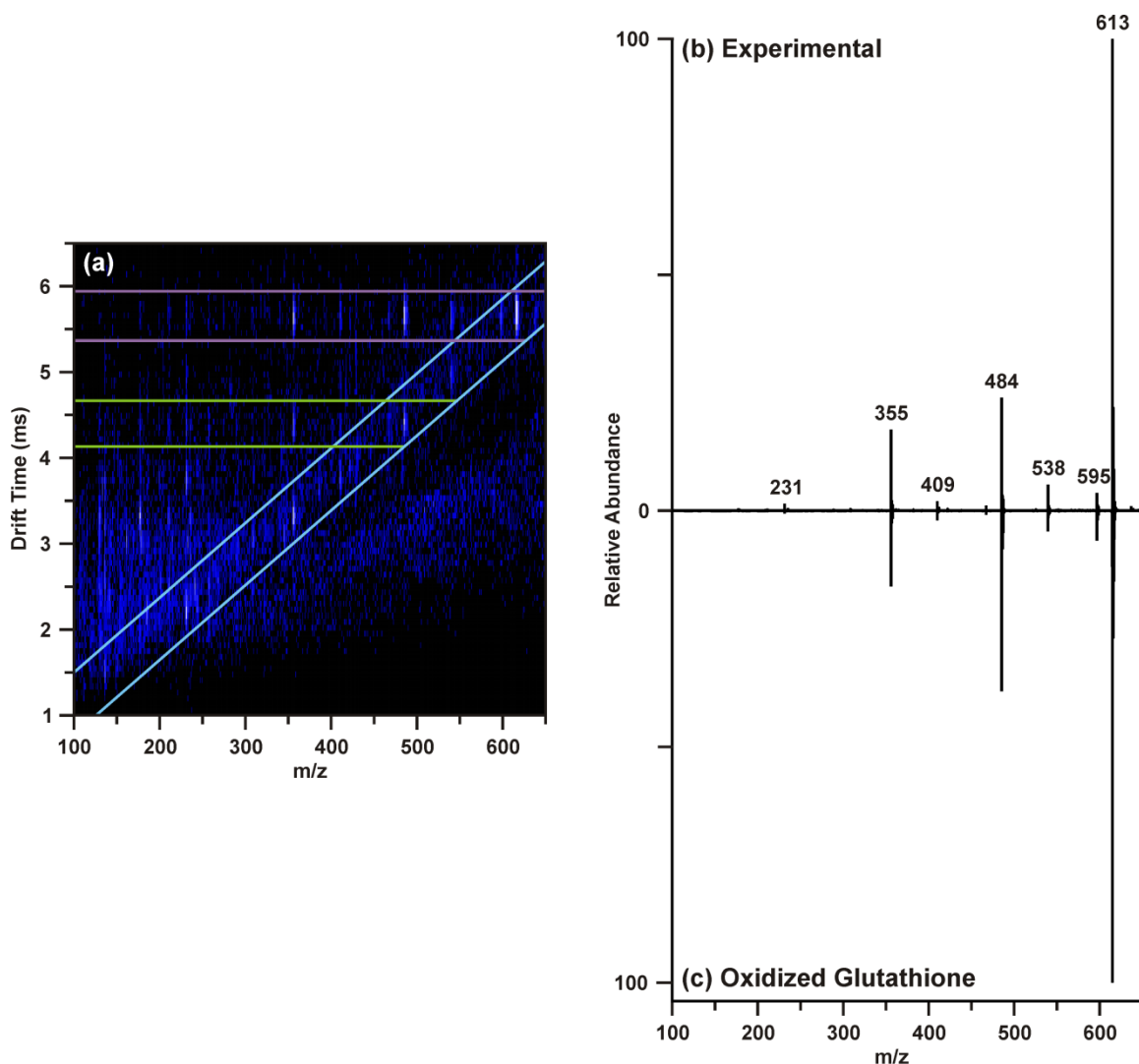


Figure 3.4. Interpretation of post-mobility MS/MS of m/z 613.16 at 3.78 min acquired by data-independent acquisition. (a) The IM-MS spectrum resulting from extraction of the chromatographic peak at 3.78 min reveals multiple regions of fragmentation. The region outlined in blue is indicative of fragmentation which occurred prior to the mobility separation (*e.g.* in-source fragmentation), while the purple region corresponds to post-mobility fragmentation of the intact molecular ion (m/z 613.16), and the green region corresponds to post-mobility fragmentation of an in-source fragment ion (m/z 484) of m/z 613.16. Extracting the purple region yields a true MS/MS spectrum of m/z 613.16 (b). MS/MS was performed on a standard of oxidized glutathione (c) to validate the tentative identification.

retain the drift time of the corresponding precursor ion, thus generating a horizontal line of precursor and fragment ions (regions outlined by purple and green lines).⁴⁴ Effectively, this mobility organization of fragments circumvents complications which may arise from the presence isobaric species and eliminates the requirement of mass selection prior to MS/MS experiments. Similarly, fragmentation which occurred prior to IM can be discerned from the post-mobility fragmentation as each ion, regardless of whether it is a precursor or fragment, will have a unique drift time. This may arise in situations where in-source fragmentation has occurred, and such a phenomenon is indicated by the blue lines in Figure 3.4(a). In order to obtain a true fragmentation spectrum of the intact molecular ion m/z 613.16, only a discrete window of drift time centered at the drift time of m/z 613.16 is extracted. As shown in (a) with the purple lines, this area contains the mobility-organized fragments of m/z 613.16. Extracting this region provides the MS/MS spectrum shown in Figure 3.4(b). Database searches and *in silico* fragmentation analysis of this spectrum supported the tentative identification of m/z 613.16 as oxidized glutathione. To validate this identification, a standard of oxidized glutathione was characterized by ESI-IM-MS/MS. The resulting extracted post-mobility MS/MS spectrum is shown in (c), and confirms the proposed identification as there is a peak-to-peak match between the experimental and standard spectra. A similar process was performed to obtain the identifications shown in Table 3.1, and standards were used to validate IDs where possible (noted in Table 3.1).

Although tissue 4D is the only sample to reside in the lower left quadrant of the PCA (Figure 3.2a), a number of the more significant molecular features (No. 1-11) are located in the corresponding quadrant of the loadings plot (Figure 3.2b). Interestingly,

many of these features were identified as biomolecular species previously demonstrated to be differentially expressed in the tumor environment. For example, thymosins β 4 and β 10 (features #7-9) are highly conserved, highly abundant polar polypeptides which are overexpressed in a number of tumor types, including breast cancer.⁴⁵ A primary function of thymosin peptides is to bind G-actin, the primary component of the cellular cytoskeleton, and inhibit G-actin polymerization.⁴⁵ As actin sequestration increases the motility of the cell, thymosin polypeptides have been suspected to play a key role in the processes of cell migration and tumor metastasis.⁴⁵⁻⁴⁷ Characterization of the effects of thymosin β 4 overexpression on lung tumor metastasis revealed increases in tumor sizes, number of metastatic nodules, neoangiogenesis, and cell migration, strongly suggesting that thymosin β 4 stimulates tumor metastasis.⁴⁶ In the analysis described here, thymosin β 4 and β 10 were found to be increased greater than 10 and 20-fold (based on single peak in isotopic distribution of multiple charged species), respectively, between tissue 4D and its corresponding control 3C (p -values: β 10 (#7), $p = 7.3 \times 10^{-7}$; β 10 (#8), $p = 2.9 \times 10^{-7}$; β 4 (#9), $p = 1.5 \times 10^{-5}$). A bar chart demonstrating the differential expression of thymosins β 4 and β 10 across the disease and control breast tissues using summed peak areas for the whole isotopic envelope of the 6+ charge states can be found in Appendix C. Using the peak area results, fold-changes of 38 (p -value = 7.0×10^{-7}) and 75 (p -value = 1.4×10^{-6}) were observed between samples 3C and 4D for thymosins β 4 and β 10, respectively

Three features contributing significantly to location of 4D within the PCA were identified as choline-containing metabolites: (#1) glycerophosphocholine, (#6) choline, and (#11) phosphocholine. Aberrant metabolism of choline phospholipids is a hallmark of cancer cells and tumors, including breast cancer, typically presenting in the form of

increased levels of phosphocholine (PC) and total choline-containing species (tCho; free choline (Cho) + PC + glycerophosphocholine (GPC)) due to increased activity and/or expression of choline kinase, choline transporters and phospholipases.⁴⁸⁻⁵² Generally, the extent to which levels of PC and tCho are increased is reflective of the tumor aggressiveness or malignancy.^{12, 50} From Figure 3.2(c) and Table 3.1, it is evident that both 4D and 6D showed significant increases in GPC (4D: 3.7-fold, p -value = 1.4×10^{-6} ; 6D: 6.0-fold, p -value = 3.0×10^{-5}) and PC (4D: 3.7-fold, p -value = 3.3×10^{-6} ; 6D: 5.8-fold, p -value = 1.3×10^{-7}) relative to the respective controls, suggesting both may represent types of breast tumors. Sample 2D, suspected to represent a non-cancerous breast disease, had significantly less GPC (8.1-fold decrease, p = 6.5×10^{-7}) and PC (3.5-fold decrease, p = 6.2×10^{-5}) relative to the matched control.

Additionally, samples 4D and 6D presented higher levels of glutathione (#4; 4D: 4.8-fold, p -value = 1.3×10^{-5} ; 6D: 4.6-fold, p -value = 9.9×10^{-6}) relative to the controls, while only 4D showed increased levels of oxidized glutathione (#5; 15.6-fold, p -value = 1.5×10^{-5} ; see Table 3.1 and Figure 3.2c). Glutathione (GSH) is the primary intracellular thiol responsible for protection against free radicals.⁵³ This detoxification may occur directly or in conjunction with the enzyme glutathione s-transferase, which conjugates electrophilic compounds to reducing sulfhydryl (-SH) group of glutathione's cysteine residue. Oxidized glutathione, or glutathione disulfide, is composed to two glutathione units linked through a disulfide bond between the cysteine residues. The relative abundances of reduced and oxidized glutathione (GSSG) in tissues have been examined as an indicator of the redox status of the tissue given the potential to detoxify to cancer drugs which work via oxidative damage.⁵³⁻⁵⁶ Previous studies of glutathione levels in

breast tumors have found significantly elevated levels of reduced and oxidized glutathione in tumor tissues relative to matched peritumoral (*i.e.* control) tissues.^{53,54,56} However, it was observed in the tumor tissues that reduced glutathione levels were significantly greater than oxidized glutathione levels, representing an increase in the detoxification capacity of the tumors.⁵⁶ Our results were consistent with these findings, where abundances of GSH were approximately 2-fold greater than GSSG in both 4D and 6D (Figure 3.2c and Appendix C).

For many of the tentatively identified features selectively highlighted in Figure 3.2(c) and Table 3.1, a consistent pattern has been observed in which abundances are increased in 4D and 6D relative to their controls, while 2D demonstrates the opposite correlations despite also representing a tissue affected by breast disease. Our general hypothesis has thus been that 4D and 6D represented cancerous breast diseases, perhaps differing in malignancies, given the identities of the most significant molecular features and their known involvement in breast cancer. On the other hand, 2D has been suspected to represent a benign cancer or a non-cancerous breast disease. Examination of the pathology reports for 2D, 4D and 6D revealed our initial data-driven hypotheses to be generally accurate. Sample 2D was diagnosed as a fibroadenoma, a benign breast tumor most commonly diagnosed in patients in their early 20s.⁵⁷ However, this sample did not represent a typical fibroadenoma diagnosis as the patient from whom this biopsy was taken was 47 years old. Sample 4D, as suspected, was diagnosed as an infiltrating duct carcinoma of grade 3 and pathological state IIA which was found to be ER, HER2/NEU, and PR negative, often referred to as triple negative cancer. This particular type of breast cancer is challenging to treat as it does not respond to targeted therapies and is often

associated with a shorter time between relapse and death.⁴ Lastly, the diagnosis of sample 6D was a pseudoangiomatous stromal hyperplasia, a benign breast tumor.⁵⁸ Similar to triple negative tumors, psuedoangiomatous stromal hyperplasias are highly metabolic and perhaps this was the primary director for the separation of 4D and 6D away from the other tissues.

3.4. Conclusions

Metabolic profiling of breast tissues using the UPLC-IM-MS/MS-based platform described here was demonstrated to be a highly sensitive technique for the differentiation of breast tissues representing a range of benign to cancerous breast diseases. The ion mobility aspect of the analysis provided an additional dimension of separation orthogonal to that provided in the chromatographic dimension. This enabled simultaneous isolation of features of interest in both IM and LC dimensions, improving confidence in locating features of interest while also increasing their signal-to-noise ratios. In the instance of co-eluting isobaric species, it was demonstrated that IM could effectively separate thses species and eliminate the overlapping isotopic peaks which may confound accurate mass determination and subsequent identification. Data-independent tandem MS acquired post-mobility separation provided a means to distinguish fragmentation occurring prior to the mobility and collision cells from that of true collision induced dissociation, providing MS/MS spectra from which feature identifications could be made with high confidence.

The molecular features giving rise to the distinction of cancerous and benign tissues from peritumoral control tissues included species known to be affected by the aberrant metabolism observed in breast cancer. These included choline, phosphocholine,

glycerophosphocholine, glutathione, and oxidized glutathione. Similarly, our analysis revealed that the actin-sequestering polypeptides thymosin β 4 and β 10 were differentially expressed between disease and control tissues. Thus, the UPLC-IM-MS/MS platform provided a truly untargeted approach in which features from multiple classes of biomolecules could be utilized to differentiate tissues representing an array of breast diseases from cancerous to benign. In general, this approach would be straightforward to apply to both larger datasets and different biological matrices.

3.5. Acknowledgements

The authors would like to thank Purmina Ghose of the Meharry Medical College Translational Pathology Shared Resource Core Facility for assistance in selection of the tissues and in locating information regarding the pathologies of the tissues. The Translational Pathology Shared Resource Core Facility is funded by the U54/National Cancer Institute Grant (NIH/NCI 5U54CA163069-02). Financial support for this research was provided by funds to J.A.M. from the National Institutes of Health (UH2TR000491) and the Defense Threat Reduction Agency (HDTRA-09-1-00-13 and DTRA100271 A-5196), the Vanderbilt University College of Arts and Science, the Vanderbilt Institute for Chemical Biology, and the Vanderbilt Institute for Integrative Biosystems Research and Education.

3.6. References

1. U.S. Cancer Statistics Working Group. *United States Cancer Statistics: 1999–2010 Incidence and Mortality Web-based Report*; U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute: Atlanta, **2013**; Available at: www.cdc.gov/uscs.
2. Etzioni, R.; Urban, N.; Ramsey, S.; McIntosh, M.; Schwartz, S.; Reid, B.; Radich, J.; Anderson, G.; Hartwell, L., The case for early detection. *Nature Reviews Cancer* **2003**, *3* (4), 243-252.
3. Vucic, E. A.; Thu, K. L.; Robison, K.; Rybaczyk, L. A.; Chari, R.; Alvarez, C. E.; Lam, W. L., Translating cancer 'omics' to improved outcomes. *Genome Research* **2012**, *22* (2), 188-195.
4. Hudis, C. A.; Gianni, L., Triple-Negative Breast Cancer: An Unmet Medical Need. *Oncologist* **2011**, *16*, 1-11.
5. Senkus, E.; Cardoso, F.; Pagani, O., Time for more optimism in metastatic breast cancer? *Cancer Treatment Reviews* **2014**, *40* (2), 220-228.
6. Coller, H. A., Is Cancer a Metabolic Disease? *American Journal of Pathology* **2014**, *184* (1), 4-17.
7. Spratlin, J. L.; Serkova, N. J.; Eckhardt, S. G., Clinical Applications of Metabolomics in Oncology: A Review. *Clinical Cancer Research* **2009**, *15* (2), 431-440.
8. Oakman, C.; Tenori, L.; Biganzoli, L.; Santarpia, L.; Cappadona, S.; Luchinat, C.; Di Leo, A., Uncovering the metabolomic fingerprint of breast cancer. *International Journal of Biochemistry & Cell Biology* **2011**, *43* (7), 1010-1020.
9. Armitage, E. G.; Barbas, C., Metabolomics in cancer biomarker discovery: Current trends and future perspectives. *Journal of Pharmaceutical and Biomedical Analysis* **2014**, *87*, 1-11.
10. Heiden, M. G. V.; Cantley, L. C.; Thompson, C. B., Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **2009**, *324* (5930), 1029-1033.
11. Griffin, J. L.; Shockcor, J. P., Metabolic profiles of cancer cells. *Nature Reviews Cancer* **2004**, *4* (7), 551-561.
12. Chen, J. H.; Mehta, R. S.; Baek, H. M.; Nie, K.; Liu, H.; Lin, M. Q.; Yu, H. J.; Nalcioglu, O.; Su, M. Y., Clinical characteristics and biomarkers of breast cancer associated with choline concentration measured by H-1 MRS. *Nmr in Biomedicine* **2011**, *24* (3), 316-324.

13. Bathen, T. F.; Sitter, B.; Sjobakk, T. E.; Tessem, M.-B.; Gribbestad, I. S., Magnetic Resonance Metabolomics of Intact Tissue: A Biotechnological Tool in Cancer Diagnostics and Treatment Evaluation. *Cancer Research* **2010**, *70* (17), 6692-6696.
14. Sitter, B.; Bathen, T. F.; Tessem, M.-B.; Gribbestad, I. S., High-resolution magic angle spinning (HR MAS) MR spectroscopy in metabolic characterization of human cancer. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2009**, *54* (3-4), 239-254.
15. Liesenfeld, D. B.; Habermann, N.; Owen, R. W.; Scalbert, A.; Ulrich, C. M., Review of Mass Spectrometry-Based Metabolomics in Cancer Research. *Cancer Epidemiology Biomarkers & Prevention* **2013**, *22* (12), 2182-2201.
16. Patti, G. J.; Yanes, O.; Siuzdak, G., Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* **2012**, *13* (4), 263-269.
17. Hines, K. M.; Enders, J. R.; McLean, J. A., Multidimensional Separations by Ion Mobility-Mass Spectrometry. In *Encyclopedia of Analytical Chemistry*, Myers, R. A., Ed. John Wiley & Sons: 2012.
18. McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H., Ion mobility-mass spectrometry: a new paradigm for proteomics. *International Journal of Mass Spectrometry* **2005**, *240* (3), 301-315.
19. Valentine, S. J.; Plasencia, M. D.; Liu, X.; Krishnan, M.; Naylor, S.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E., Toward plasma proteome profiling with ion mobility-mass spectrometry. *Journal of Proteome Research* **2006**, *5* (11), 2977-2984.
20. Enders, J. R.; Marasco, C. C.; Kole, A.; Nguyen, B.; Sevugarajan, S.; Seale, K. T.; Wikswo, J. P.; McLean, J. A., Towards monitoring real-time cellular response using an integrated microfluidics-matrix assisted laser desorption ionisation/nanoelectrospray ionisation-ion mobility-mass spectrometry platform. *IET Systems Biology* **2010**, *4* (6), 416-427.
21. Enders, J. R.; Goodwin, C. R.; Marasco, C. C.; Seale, K. T.; Wikswo, J. P.; McLean, J. A., Advanced structural mass spectrometry for systems biology: Pulling the needles from haystacks. *Current Trends in Mass Spectrometry* **2011**, 18-23.
22. Dwivedi, P.; Wu, P.; Klopsch, S. J.; Puzon, G. J.; Xun, L.; Hill, H. H., Jr., Metabolic profiling by ion mobility mass spectrometry (IMMS). *Metabolomics* **2008**, *4* (1), 63-80.
23. May, J. C.; Goodwin, C. R.; McLean, J. A., Gas-Phase Ion Mobility-Mass Spectrometry (IM-MS) and Tandem IM-MS/MS Strategies for Metabolism

Studies and Metabolomics. *Encyclopedia of Drug Metabolism and Interactions* **2012**.

24. Kaplan, K. A.; Chiu, V. M.; Lukus, P. A.; Zhang, X.; Siems, W. F.; Schenk, J. O.; Hill, H. H., Jr., Neuronal metabolomics by ion mobility mass spectrometry: cocaine effects on glucose and selected biogenic amine metabolites in the frontal cortex, striatum, and thalamus of the rat. *Anal Bioanal Chem* **2013**.
25. Kaplan, K.; Dwivedi, P.; Davidson, S.; Yang, Q.; Tso, P.; Siems, W.; Hill, H. H., Monitoring Dynamic Changes in Lymph Metabolome of Fasting and Fed Rats by Electrospray Ionization-Ion Mobility Mass Spectrometry (ESI-IMMS). *Analytical Chemistry* **2009**, *81* (19), 7944-7953.
26. Ruotolo, B. T.; Gillig, K. J.; Stone, E. G.; Russell, D. H., Peak capacity of ion mobility mass spectrometry: Separation of peptides in helium buffer gas. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2002**, *782* (1-2), 385-392.
27. Ruotolo, B. T.; McLean, J. A.; Gillig, K. J.; Russell, D. H., Peak capacity of ion mobility mass spectrometry: the utility of varying drift gas polarizability for the separation of tryptic peptides. *Journal of Mass Spectrometry* **2004**, *39* (4), 361-367.
28. Liu, X.; Plasencia, M.; Ragg, S.; Valentine, S. J.; Clemmer, D. E., Development of high throughput dispersive LC-ion mobility-TOFMS techniques for analysing the human plasma proteome. *Briefings in Functional Genomics & Proteomics* **2004**, *3* (2), 177-186.
29. 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. *Analytical and Bioanalytical Chemistry* **2009**, *394* (1), 235-244.
30. McLean, J. A., The Mass-Mobility Correlation Redux: The Conformational Landscape of Anhydrous Biomolecules. *Journal of the American Society for Mass Spectrometry* **2009**, *20* (10), 1775-1781.
31. Fenn, L. S.; McLean, J. A., Biomolecular structural separations by ion mobility-mass spectrometry. *Analytical and Bioanalytical Chemistry* **2008**, *391* (3), 905-909.
32. Hines, K. M.; Ashfaq, S.; Davidson, J. M.; Opalenik, S. R.; Wikswo, J. P.; McLean, J. A., Biomolecular Signatures of Diabetic Wound Healing by Structural Mass Spectrometry. *Analytical Chemistry* **2013**, *85* (7), 3651-3659.
33. Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K., Global metabolic

- profiling of animal and human tissues via UPLC-MS. *Nature Protocols* **2013**, *8* (1), 17-32.
34. Masson, P.; Alves, A. C.; Ebbels, T. M. D.; Nicholson, J. K.; Want, E. J., Optimization and Evaluation of Metabolite Extraction Protocols for Untargeted Metabolic Profiling of Liver Samples by UPLC-MS. *Analytical Chemistry* **2010**, *82* (18), 7779-7786.
 35. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H., Applications of a travelling wave-based radio-frequency only stacked ring ion guide. *Rapid Communications in Mass Spectrometry* **2004**, *18* (20), 2401-2414.
 36. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinou, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H., An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry* **2007**, *261* (1), 1-12.
 37. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P., A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology* **2012**, *30* (10), 918-920.
 38. Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: Processing mass spectrometry data for metabolite profiling using Nonlinear peak alignment, matching, and identification. *Analytical Chemistry* **2006**, *78* (3), 779-787.
 39. Tautenhahn, R.; Bottcher, C.; Neumann, S., Highly sensitive feature detection for high resolution LC/MS. *Bmc Bioinformatics* **2008**, *9* (504).
 40. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN - A metabolite mass spectral database. *Therapeutic Drug Monitoring* **2005**, *27* (6), 747-751.
 41. Tautenhahn, R.; Cho, K.; Uritboonthai, W.; Zhu, Z.; Patti, G. J.; Siuzdak, G., An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnology* **2012**, *30* (9), 826-828.
 42. Zhu, Z.-J.; Schultz, A. W.; Wang, J.; Johnson, C. H.; Yannone, S. M.; Patti, G. J.; Siuzdak, G., Liquid chromatography quadrupole time-of-flight mass spectrometry

- characterization of metabolites guided by the METLIN database. *Nature Protocols* **2013**, 8 (3), 451-460.
43. Wolf, S.; Schmidt, S.; Mueller-Hannemann, M.; Neumann, S., In silico fragmentation for computer assisted identification of metabolite mass spectra. *Bmc Bioinformatics* **2010**, 11.
 44. Hoaglund-Hyzer, C. S.; Li, J. W.; Clemmer, D. E., Mobility labeling for parallel CID of ion mixtures. *Analytical Chemistry* **2000**, 72 (13), 2737-2740.
 45. Huff, T.; Muller, C. S. G.; Otto, A. M.; Netzker, R.; Hannappel, E., beta-thymosins, small acidic peptides with multiple functions. *International Journal of Biochemistry & Cell Biology* **2001**, 33 (3), 205-220.
 46. Cha, H. J.; Jeong, M. J.; Kleinman, H. K., Role of thymosin beta(4) in tumor metastasis and angiogenesis. *Journal of the National Cancer Institute* **2003**, 95 (22), 1674-1680.
 47. Maelan, A. E.; Rasmussen, T. K.; Larsson, L.-I., Localization of thymosin beta 10 in breast cancer cells: relationship to actin cytoskeletal remodeling and cell motility. *Histochemistry and Cell Biology* **2007**, 127 (1), 109-113.
 48. Katz-Brull, R.; Seger, D.; Rivenson-Segal, D.; Rushkin, E.; Degani, H., Metabolic markers of breast cancer: Enhanced choline metabolism and reduced choline-ether-phospholipid synthesis. *Cancer Research* **2002**, 62 (7), 1966-1970.
 49. Glunde, K.; Jie, C.; Bhujwala, Z. M., Molecular causes of the aberrant choline phospholipid metabolism in breast cancer. *Cancer Research* **2004**, 64 (12), 4270-4276.
 50. Glunde, K.; Jacobs, M. A.; Bhujwala, Z. M., Choline metabolism in cancer: implications for diagnosis and therapy. *Expert Review of Molecular Diagnostics* **2006**, 6 (6), 821-829.
 51. van Hove, E. R. A.; Blackwell, T. R.; Klinkert, I.; Eijkel, G. B.; Heeren, R. M. A.; Glunde, K., Multimodal Mass Spectrometric Imaging of Small Molecules Reveals Distinct Spatio-Molecular Signatures in Differentially Metastatic Breast Tumor Models. *Cancer Research* **2010**, 70 (22), 9012-9021.
 52. Glunde, K.; Bhujwala, Z. M.; Ronen, S. M., Choline metabolism in malignant transformation. *Nature Reviews Cancer* **2011**, 11 (12), 835-848.
 53. Gamcsik, M. P.; Kasibhatla, M. S.; Teeter, S. D.; Colvin, O. M., Glutathione levels in human tumors. *Biomarkers* **2012**, 17 (8), 671-691.
 54. Perquin, M.; Oster, T.; Maul, A.; Froment, N.; Untereiner, M.; Bagrel, D., The glutathione-related detoxification pathway in the human breast: a highly

- coordinated system disrupted in the tumour tissues. *Cancer Letters* **2000**, *158* (1), 7-16.
55. Schafer, F. Q.; Buettner, G. R., Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biology and Medicine* **2001**, *30* (11), 1191-1212.
 56. Perquin, M.; Oster, T.; Maul, A.; Froment, N.; Untereiner, M.; Bagrel, D., The glutathione-related detoxification system is increased in human breast cancer in correlation with clinical and histopathological features. *Journal of Cancer Research and Clinical Oncology* **2001**, *127* (6), 368-374.
 57. Dupont, W. D.; Page, D. L.; Parl, F. F.; Vnencakjones, C. L.; Plummer, W. D.; Rados, M. S.; Schuyler, P. A., Long-term risk of breast cancer in women with fibroadenoma. *New England Journal of Medicine* **1994**, *331* (1), 10-15.
 58. Ferreira, M.; Albarracin, C. T.; Resetkova, E., Pseudoangiomatous stromal hyperplasia tumor: a clinical, radiologic and pathologic study of 26 cases. *Modern Pathology* **2008**, *21* (2), 201-207.

CHAPTER IV

PHOSPHORYLATION OF SERINE 106 IN ASEF2 REGULATES CELL MIGRATION AND ADHESION TURNOVER

4.1. Introduction

Cell migration is a complex, actin-dependent process that plays a central role in embryonic development and wound healing.¹ The tightly controlled signaling pathways that mediate cell migration can be altered in pathological states, such as tumor metastasis and atherosclerosis.^{2,3} Cell migration involves several canonical steps: the extension of actin-rich protrusions, the assembly of nascent adhesions at the leading edge, the translocation of the cell body, and the retraction of the rear of the cell.⁴ The assembly of leading edge adhesions, which are sites of contact between cells and the extracellular matrix, stabilizes protrusions and provides traction to propel the forward movement of cells.⁴⁻⁶ Once formed, nascent adhesions can disassemble, or they can continue to grow into larger, more stable adhesions.^{7,8} The constant assembly and disassembly of leading edge adhesions, termed adhesion turnover, is crucial for efficient cell migration^{7,9} but not well understood on a molecular level.

Small GTPases that comprise the Rho family, including Rho, Rac, and Cdc42, are key modulators of cell migration through their ability to regulate processes underlying migration, such as adhesion assembly, disassembly, and maturation.^{1,10-12} Rho family GTPases, like other small GTPases, function by cycling between a GTP-bound active form and a GDP-bound inactive form.¹³ This cycling is dependent on GEFs that catalyze

the exchange of GDP for GTP and GTPase activating proteins (GAPs), which promote the hydrolysis of GTP.¹⁴⁻¹⁷ Upon activation by GEFs, the Rho GTPases, in turn, activate a series of downstream effector proteins that regulate adhesion and actin dynamics.^{12,18,19} While the role of the Rho GTPases in regulating cell migration has been studied, less is known about the function of the various GEFs and GAPs in modulating migration and its underlying processes.

Asef2 is a recently discovered GEF that has been implicated in the regulation of cell migration.²⁰⁻²² This 652-amino acid protein is composed of several functional domains: an APC-binding region (ABR), a Src homology 3 (SH3) domain, a Dbl homology (DH) domain, and a pleckstrin homology (PH) domain.²⁰ The DH domain mediates GTP exchange for Rac and Cdc42, while the PH domain is most likely involved in membrane localization.²⁰⁻²² The ABR and SH3 domains work in concert to regulate Asef2 activity.²⁰ Asef2 exists in an autoinhibited conformation that prevents nucleotide exchange by the DH domain; once the tumor suppressor APC binds to the tandem ABR and SH3 domains, Asef2 undergoes a conformational change that stimulates its GEF activity.^{20,23,24} While the mechanism of APC binding to Asef2 and relieving autoinhibition has been studied,²⁰ little is currently known about other potential modes of Asef2 regulation. For example, post-translational modification of Asef2 is one possible avenue of modulating its activity and function. The addition of chemical moieties, such as acetyl, phosphate, or glycosyl groups, to a protein is a common mechanism for altering its conformation, localization, and activity.²⁵ Indeed, it has previously been shown that phosphorylation of GEFs is necessary for proper function.²⁶⁻²⁸ These data point to a possible role for phosphorylation in regulating Asef2 activity and function.

Here, we describe the identification of phosphorylation sites in Asef2 using a liquid chromatography-mass spectrometry (LC-MS) approach consisting of high-mass resolution Orbitrap MS, data-dependent tandem MS (MS/MS), multiple protease and denaturing strategies, and bioinformatics-based peptide and protein assignments.²⁹ This methodology yielded 94.6% sequence coverage and identified six sites of phosphorylation. The 5.4% of the sequence does not contain serine, threonine or tyrosine residues; therefore, 100% coverage of possible phosphorylatable sites was achieved. The majority of these sites are clustered in the N-terminus of Asef2; one site (S106) is located in the ABR domain, suggesting that it could regulate Asef2 activity. Indeed, we show that S106 phosphorylation is crucial for Asef2-promoted Rac activation, cell migration, and adhesion turnover, pointing to a new regulatory mechanism for Asef2 activity and function.

4.2. Experimental Methods

Reagents and Plasmids

Mouse IgG agarose, FLAG M2-agarose affinity gel, FLAG peptide (DYKDDDDK), FLAG monoclonal antibody (clone M2), and fibronectin were obtained from Sigma (St. Louis, MO). Sodium vanadate was purchased from Fisher Scientific (Fairlawn, NJ), and calyculin A was obtained from EMD Millipore (Billerica, MA). Peroxovanadate was prepared as previously described.²⁹ Glutathione sepharose beads were purchased from GE Healthcare Life Sciences (Piscataway, NJ). Phosphoserine polyclonal antibody (catalog number 61-8100) was obtained from Life Technologies (Grand Island, NY). GFP-Asef2 was generated by cloning human Asef2 (accession

number: NM_153023.2) into EGFP-C3 vector (Clontech, Mountain View, CA) at *EcoRI* sites as previously described.²² FLAG-GFP was prepared as previously described,²⁹ and FLAG-GFP-Asef2 was generated by inserting human Asef2 into the FLAG-GFP vector at *EcoRI* sites. Asef2 serine 106 mutants were created via site-directed mutagenesis using the following primers: serine 106 to alanine (Asef2-S106A), forward (5'-GGTACTGAGCCCGCTGCCTTAGTGGAT-3') and reverse (5'-ATCCACTAAGGCAGCGGGCTCAGTACC-3'); serine 106 to aspartic acid (Asef2-S106D), forward (5'-GGTACTGAGCCCGATGCCTTAGTGGAT-3') and reverse (5'-ATCCACTAAGGCATCGGGCTCAGTACC-3'). mCherry-paxillin was a generous gift from Steve Hanks (Vanderbilt University, Nashville, TN).

Cell Culture and Transfection

HT1080 fibrosarcoma cells and human embryonic kidney 293 (HEK293) cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies, Grand Island, NY), which was supplemented with 10% fetal bovine serum (FBS; Thermo Scientific, Waltham, MA) and penicillin/streptomycin (Life Technologies, Grand Island, NY). These cells were maintained in an incubator with 5% carbon dioxide (CO₂) at 37 °C. Cells were transiently transfected with appropriate cDNAs using LipofectamineTM 2000 (Life Technologies, Grand Island, NY) according to the manufacturer's instructions.

Protein Purification

HEK293 cells were cultured in eight 150 mm dishes (Corning, Tewksbury, MA)

for 24 h and then transfected with FLAG-GFP-Asef2 cDNA (12 μ g per dish). After approximately 40 h, cells were treated with 1 mM peroxovanadate and 50 nM calyculin A for 10 min and then extracted with 25 mM Tris, 137 mM NaCl, 1% NP-40, 10% glycerol, and 2 mM EDTA (pH 7.4) containing a phosphatase inhibitor cocktail (Sigma, St. Louis, MO; catalogue number P2714) for 30 min on ice. The lysate was precleared with mouse IgG agarose for 1 h at 4 °C with end-over-end mixing; the lysate was then precleared a second time by overnight incubation with IgG agarose. After preclearing, the lysate was incubated with FLAG-agarose for 2 h at 4 °C with end-over-end mixing, and the beads were washed three times (15 min each, 4 °C) with 25 mM Tris and 100 mM NaCl (pH 7.4). FLAG-GFP-Asef2 protein was eluted from the beads by incubation with 0.2 mg/mL FLAG peptide suspended in 25 mM Tris and 100 mM NaCl (pH 7.4) for 30 min at 4 °C; this elution was repeated, and the eluates were pooled. The eluate was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 10% slab, followed by staining with Coomassie® Brilliant Blue R-250 (EMD Millipore, Billerica, MA) to determine the protein purity.

To examine serine phosphorylation, HEK293 cells from a single 150 mm dish per condition were transfected with 8 μ g of either FLAG-GFP or FLAG-GFP-Asef2 cDNAs and were immunoprecipitated using the aforementioned protocol. The eluates were subjected to SDS-PAGE and transferred to nitrocellulose membranes. Membranes were incubated with either phosphoserine polyclonal antibody or M2 FLAG monoclonal antibody, followed by incubation with AlexaFluor 680 anti-rabbit IgG (Life Technologies, Grand Island, NY) or IRDye 800 anti-mouse IgG (Rockland

Immunochemicals Inc., Gilbertsville, PA). Membranes were scanned with a LI-COR[®] Odyssey[®] Infrared Imaging System (LI-COR Biosciences, Lincoln, NE).

Enzymatic Proteolysis

Purified Asef2 was separated into three aliquots containing equal amounts of protein and was subjected to enzymatic digestion using trypsin, chymotrypsin and Glu-C proteases (Promega, Madison, WI), as described previously.²⁹ Briefly, approximately 5 µg of purified Asef2 was resuspended in 75 µL of 25 mM ammonium bicarbonate and were aliquoted into three 25 µL samples containing approximately 1.7 µg of protein. Reduction and alkylation of cysteine sulfhydryl groups were performed by the addition of 1.5 µL of 45 mM dithiothreitol (DTT) and incubation for 30 min at 55 °C, followed by the addition of 2.5 µL of 100 mM iodoacetamide (IAM) and incubation in darkness for 30 min at room temperature. Digestion was performed by adding 42 ng of trypsin, chymotrypsin, or Glu-C at a ratio of 1:40 protease:protein (w/w), followed by incubation at 37 °C for 16, 4, and 6 h, respectively. To quench proteolysis, 1 µL of 88% formic acid was added. The digested material was lyophilized and reconstituted in 25 µL of 0.1% formic acid for MS analysis.

Two additional trypsin digestions were performed at strongly denaturing conditions using heat and organic solvent. Both samples contained approximately 2 µg of purified and aliquoted protein. For denaturation by heat, the sample was reconstituted with 25 µL of 25 mM ammonium bicarbonate and denatured for 15 min at 90 °C. For denaturation by high organic solvent, the respective sample was reconstituted with 20 µL of acetonitrile (HPLC Grade) and 5 µL of 25 mM ammonium bicarbonate to achieve a

solution of 80% acetonitrile.^{30,31} Both samples were treated with DTT and IAM as described above to reduce and alkylate cysteine sulfhydryl groups. Digestion was performed by adding 52 ng (1:40 protease:protein, w/w) of trypsin (Promega, Madison, WI) to each sample. The high organic digestion (referred to as Trypsin^{Org}) was stopped after 1 h of incubation at 37 °C, while the high temperature denatured digestion (referred to as Trypsin^{Temp}) was allowed to proceed for 16 h at 37 °C. The digestions were quenched, dried, and reconstituted as described above.

LC-MS/MS

Initial digestions of Asef2 were loaded onto a reverse-phase capillary trap column using a helium-pressurized cell (pressure bomb). The trap column (360 µm OD x 150 µm ID) was fitted with a filter end-fitting (IDEX Health & Science, Oak Harbor, WA) and packed with 4 cm of C18 reverse phase material (Jupiter, 5 µm beads, 300 Å; Phenomenex, Torrance, CA). After sample loading, an M-520 microfilter union (IDEX Health & Science) was used to connect the trap column to a capillary analytical column (360 µm OD x 100 µm ID) equipped with a laser-pulled emitter tip and was packed with 20 cm of C18 material (Jupiter, 3µm beads, 300Å; Phenomenex, Torrance, CA). Peptides were gradient-eluted at a flow rate of 500 nL/min using an Eksigent NanoLC Ultra HPLC, and the mobile phase solvents consisted of 0.1% formic acid, 99.9% water (solvent A) and 0.1% formic acid, 99.9% acetonitrile (solvent B). The gradient consisted of the following: 2-45% B in 40 min, 45-90% B in 10 min, 90% B for 5 min, 90-2% B in 10 min. Subsequent trypsin digestions of Asef2 were loaded directly onto the capillary analytical column using the Eksigent NanoLC Ultra HPLC and autosampler, and the

same reverse phase gradient was performed. Upon gradient-elution, peptides were mass analyzed on a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, San Jose, CA) equipped with a nanoelectrospray ionization source. The instrument was operated using a data-dependent method with dynamic exclusion enabled. Full scan (m/z 300-2000) spectra were acquired with the Orbitrap as the mass analyzer (resolution 60,000), and the top ten most abundant ions in each MS scan were selected for fragmentation in the LTQ. An isolation width of 2 m/z , activation time of 10 ms, and normalized collision energy of 35% were used to generate MS/MS spectra. The MSⁿ AGC target value was set to 1×10^4 , and the maximum injection time was 100 ms.

Bioinformatics

For peptide identification, tandem mass spectra were converted into DTA files using Scansifter and searched using a custom version of Sequest (Thermo Fisher Scientific, San Jose, CA) on the Vanderbilt ACCRE Linux cluster against a concatenated forward and reversed (decoy) database containing the *Homo sapiens* subset of the UniProtKB (www.uniprot.org) protein database. The protease was identified for samples digested with trypsin, while non-specific protease was indicated for the chymotrypsin and Glu-C digests. A maximum of 10 missed cleavages was allowed for trypsin digests, and zero missed cleavages were allowed for chymotrypsin and Glu-C digests. Spectra were searched using a 2.5 Da mass tolerance for the precursor peptide mass, and parameters were set to search for monoisotopic masses of the product ions. Allowable variable modifications were limited to carbamidomethyl derivatization of cysteine, oxidation of methionine, and phosphorylation of serine, tyrosine, and threonine. Scaffold version

3.3.3 (Proteome Software Inc., Portland, OR) was used to visualize and validate peptide and protein identifications based on MS/MS data. A threshold of 95% probability was required for peptide identifications, and a minimum of two identified peptides and a probability threshold of 99% were required for protein identifications.^{32,33} Identifications made to Asef2 were based on the sequence associated with accession number A2VEA_HUMAN. All potential phosphopeptides with probabilities of 95% or greater were manually validated, and labeled MS/MS spectra from which these assignments were based are available in Appendix D. Additionally, several peptides not found by the bioinformatics were manually identified and validated.

Migration Analysis and Microscopy

HT1080 cells were transfected with 1.5 µg of GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D cDNAs and were incubated for 24 h at 37 °C. Then, the cells were plated on tissue culture dishes that were coated with 5 µg/ml fibronectin (diluted in Dulbecco's Phosphate Buffered Saline (DPBS, Life Technologies, Grand Island, NY)) and allowed to adhere for 1 h at 37 °C. Prior to imaging, the culture medium was replaced with SFM4MABTM medium (Hyclone, Logan, UT) supplemented with 2% FBS. Images were obtained every 5 min for 6 h using an inverted Olympus IX71 microscope (Melville, NY) with a Retiga EXi CCD camera (QImaging, Surrey, BC), a 10X objective (NA 0.3), and MetaMorph software (Molecular Devices, Sunnyvale, CA) connected to a Lambda 10-2 automated controller (Sutter Instruments, Novato, CA). GFP-expressing cells were visualized with an Endow GFP Bandpass filter cube (excitation HQ470/40, emission HQ525/50, Q495LP dichroic mirror) (Chroma,

Brattleboro, VT). MetaMorph software was used to track cell movement, and the migration speed was calculated by dividing the net distance traveled (μm) by the migration time (h). Wind-Rose plots were generated as previously described.³⁴ For migration, adhesion turnover, and Rac activity assays, statistical analyses were performed using a Student's t-test.

Adhesion Turnover Assay

HT1080 cells were co-transfected with 1.5 μg mCherry-paxillin cDNA and 1.5 μg of either GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D cDNAs and were incubated for 24 h. Cells were then plated on glass-bottom dishes coated with fibronectin (5 $\mu\text{g}/\text{mL}$) and were allowed to adhere for 1 h at 37 °C. Fluorescent time-lapse images were acquired at 15 s intervals for 20 min using the Olympus IX71 microscope setup described above with a PlanApo 60X OTIRM objective (NA 1.45) and Metamorph software. mCherry was visualized with a TRITC/Cy3 cube (excitation HQ545/30, emission HQ610/75, Q570LP dichroic mirror). The $t_{1/2}$ values for adhesion assembly and disassembly were measured as previously described.^{7, 22}

Rac Activity Assay

Glutathione-S-transferase (GST)-tagged p21-activated kinase (PAK) binding domain (GST-PBD) was expressed and attached to glutathione sepharose beads as previously described.³⁵ HT1080 cells were cultured on 60 mm tissue culture dishes coated with 5 $\mu\text{g}/\text{mL}$ fibronectin and co-transfected with 2 μg FLAG-Rac1 cDNA and 4 μg of GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D cDNAs. After 24 h,

cells were lysed and assayed for Rac activity as previously described.^{34, 35} Briefly, cells were lysed with 50 mM Tris, 1% NP-40, 10% glycerol, 100 mM NaCl, 2 mM MgCl₂, and a protease inhibitor cocktail (lysis buffer). A small fraction of the each lysate was kept to determine the amount of total Rac. The remaining lysate was incubated with GST-PBD beads for 1 h at 4 °C with end-over-end mixing. The beads were washed three times with lysis buffer. Then, the bound protein was eluted from the beads with Laemmli sample buffer and analyzed via Western blot. The amount of active Rac pulled down was normalized to total Rac for each condition.

4.3. Results and Discussion

Identification of Phosphorylation Sites in Human Asef2

Because Asef2 phosphorylation, which could be an important regulatory mechanism for the activity and function of this protein, had not been previously investigated, we utilized a MS-based approach to uncover potential phosphorylation sites in Asef2.²⁹ To perform MS analyses, FLAG-GFP-Asef2 was expressed in HEK293 cells and then purified according to the immunoprecipitation protocol outlined in Figure 4.1(a).²⁹ A predominant band with a molecular mass corresponding to that of FLAG-GFP-Asef2 was observed when the immunoprecipitated protein sample was subjected to SDS-PAGE followed by Coomassie Blue staining (Figure 4.1b). This band was confirmed to be FLAG-GFP-Asef2 via Western blot analysis (Figure 4.1c). We next examined the phosphorylation state of Asef2 by using a phosphoserine antibody; a distinct band was observed for the Asef2 sample compared to the control sample (FLAG-GFP), demonstrating that Asef2 is phosphorylated on serine residues (Figure 4.1c).

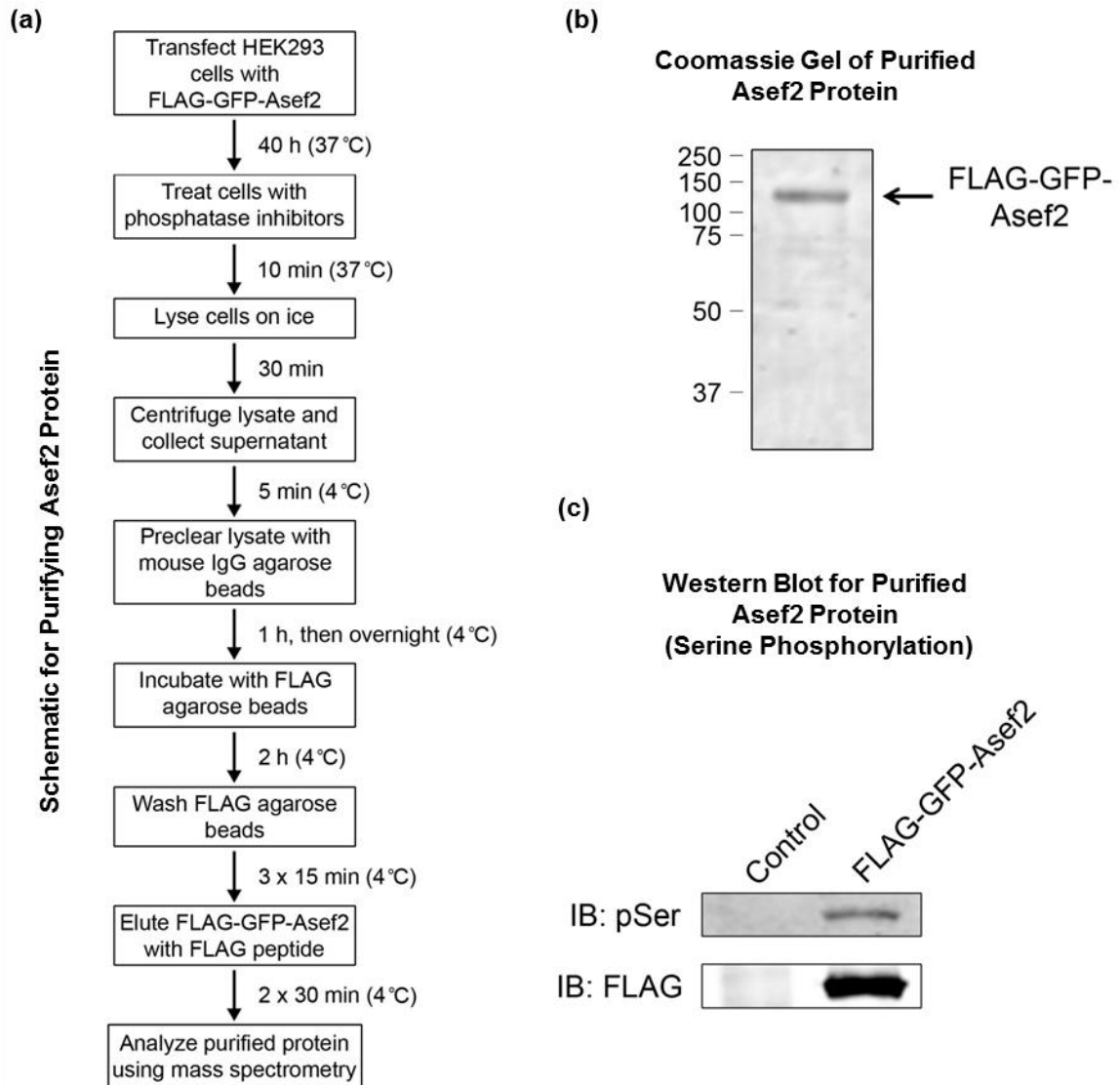


Figure 4.1. Purification of FLAG-GFP-Asef2. (a) Schematic showing the protocol used to purify FLAG-GFP-Asef2 for MS analysis. (b) SDS-PAGE gel of immunoprecipitated FLAG-GFP-Asef2 that was stained with Coomassie Blue. The arrow points to the band representing purified FLAG-GFP-Asef2 in the eluted sample. (c) Purified FLAG-GFP (Control) and FLAG-GFP-Asef2 were subjected to SDS-PAGE, followed by immunoblotting (IB) with phosphoserine (pSer, upper panel) and FLAG (lower panel) antibodies. These panels show that FLAG-GFP-Asef2 is phosphorylated on serine residues.

Collectively, these results indicate that the immunoprecipitated protein sample is suitable for MS analysis to identify specific phosphorylated residues in Asef2.

Multiple proteases were used to obtain complete coverage of the potential sites of phosphorylation in Asef2. Initially, trypsin, chymotrypsin, and Glu-C digestions were used, providing partial (85.6%) sequence coverage, with 93% coverage of serine, threonine, and tyrosine residues. However, two significant stretches of the Asef2 protein sequence from R492-K518 and R561-Y596 (Figure 4.2a) were not covered in the trypsin, chymotrypsin, or Glu-C digestions. These regions have a high abundance of aspartic acid and glutamic acid residues, and therefore may not provide peptides of suitable length for MS analysis upon digestion with Glu-C. While these amino acid sequences have multiple lysine and arginine residues, regions R492-K518 and R561-Y596 were found to be inaccessible or resistant to trypsin and chymotrypsin under standard digestion conditions.

In order to obtain sequence coverage of regions R492-K518 and R561-Y596 of Asef2, additional trypsin digestions were performed using two strongly denaturing approaches: heat and a high organic solvent concentration. Digestions done in mixed aqueous-organic solvent conditions have been demonstrated to increase peptide identifications from proteolysis-resistant proteins, while simultaneously allowing for shorter digestion times.^{30,31} Our results showed that an additional 44 residues were covered in regions R492-K518 and R561-Y596 with the strongly denaturing digestions. Of these 44 amino acids, approximately 64% were observed in the high organic solvent digestion (Trypsin^{Org}). Combined, the two strongly denaturing digestions yielded nearly

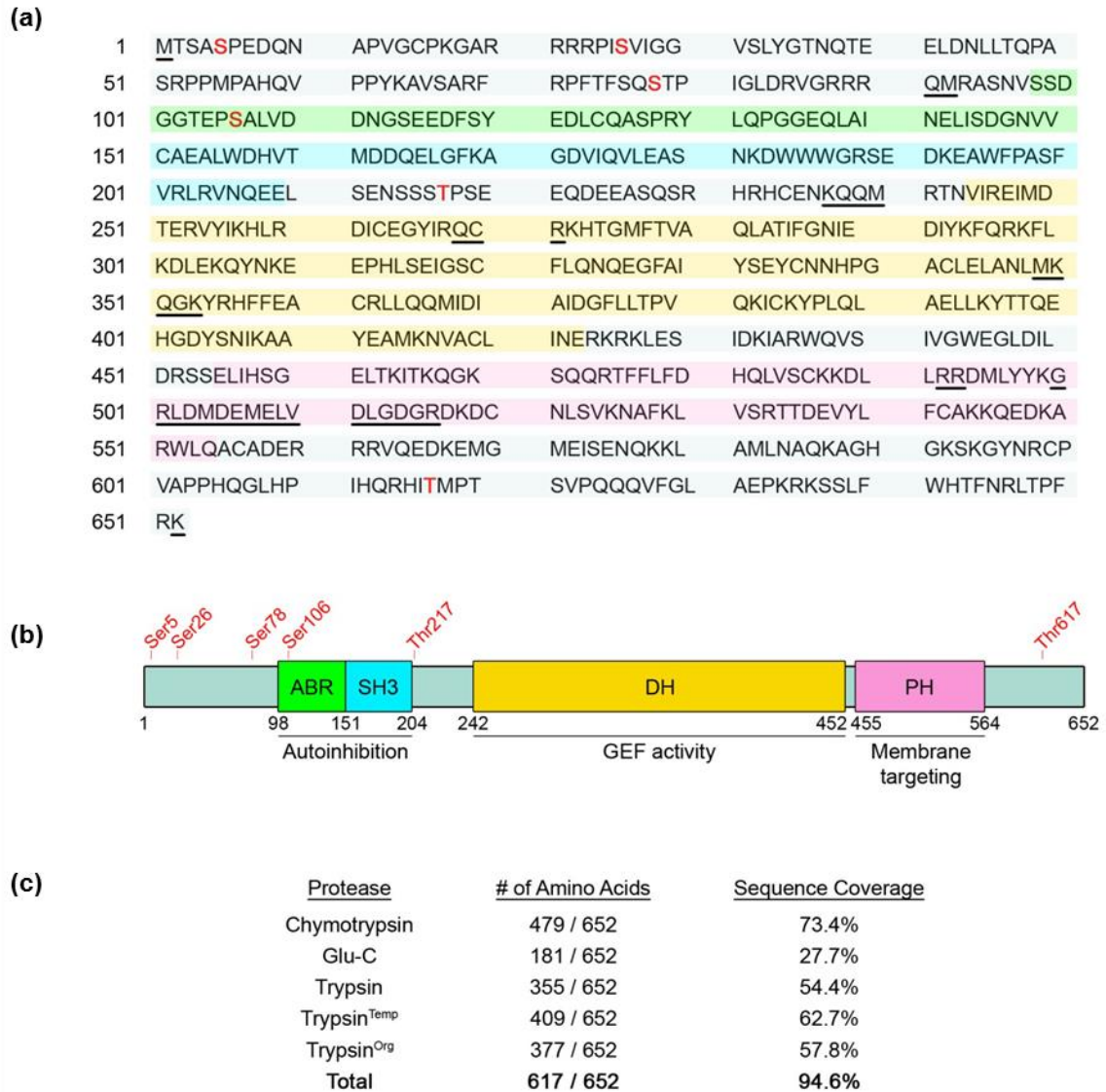


Figure 4.2. Phosphorylation sites identified in Asef2. (a) The protein sequence of Asef2 is shown with the phosphorylation sites that were detected by MS in red. Colored shading represents the conserved domains in Asef2 as shown in panel B. Underlined residues were not detected in the MS analyses. (b) Schematic of Asef2 showing conserved domains and the location of the six identified phosphorylation sites (red). The domain numbering is based on Kawasaki et al.²¹ (c) Summary of Asef2 amino acid sequence coverage by MS analyses. Purified Asef2 samples were treated with multiple proteases (trypsin, chymotrypsin, or Glu-C) to achieve high sequence coverage. Additional trypsin digestions were performed with strongly denaturing conditions, which included high temperature (90 °C, Trypsin^{Temp}) or a high percentage of organic solvent (80% acetonitrile, Trypsin^{Org}), to obtain sequence coverage of regions that were resistant to trypsin under standard digestion conditions.

83% sequence coverage and accessed additional regions of Asef2 compared to the more conventional, aqueous-based digestions. Sequence coverage of 94.6% was achieved with the five different digestions (Figure 4.2c), and all of the serine, threonine, and tyrosine residues were found in the identified peptides.

The identified phosphopeptides are shown in Table 4.1, along with the type of enzymatic digestion used and the associated mass error. Phosphopeptide identities were initially revealed by SEQUEST, but each was manually validated to confirm the location of the phosphorylation. Four sites of phosphorylation were identified in the digestions using standard conditions: pS5, pS78, pS106 and pT217. Two additional sites, pS26 and pT617, were identified from the strongly denatured digestions. Three (pS78, pS106 and pT217) of the six total sites of phosphorylation were observed in multiple digests. An example of MS/MS data for the phosphorylated peptide, $^{94}\text{ASNVSSDGGTEPpSALVDDNGSEEDFSYEDLCQASPR}^{129}$ is shown in Figure 4.3. This peptide was conserved in all three trypsin digests, and the observation of y_{23}^{+3} enabled the exact site of phosphorylation (pS106) to be discerned.

Serine Phosphorylation Stimulates Asef2 GEF Activity

A majority of the confirmed phosphorylation sites in Asef2 are concentrated at the N-terminus (Figure 4.2b). One residue (S106) is in the ABR domain, while four others (S5, S26, S78, and T217) bracket the adjacent ABR-SH3 domains. The sixth phosphorylation site (T617), conversely, resides in the C-terminus of Asef2 (Figure 4.2b). Of these six phosphorylation sites, S106 was particularly intriguing because of its location in the ABR domain, which is a critical region for Asef2 activation. This led us

Table 1. Phosphorylation sites identified within Asef 2 by MS analysis.

Peptide ^a	Sequence Position	Protease ^b	m/z (Charge)	Mass Error (ppm)
² TSA p SPEDQNAPVGC*PK ¹⁷	5S	Chymotrypsin	869.36 (+2)	-1.27
²¹ RRRPI p SVIGGVSLYG TNQTEELDNLLTQPASRPPMPAHQVPPYK ⁶⁴	26S	Trypsin ^{Org}	993.52 (+5)	1.81
²² RRPI p SVIGGVSLYG TNQTEELDNLLTQPASRPPMPAHQVPPYK ⁶⁴	26S	Trypsin ^{Org}	962.30 (+5)	1.77
²² RRPI p SVIGGVSLYG TNQTEELDNLLTQPASRPPMPAHQVPPYKAVSAR ⁶⁹	26S	Trypsin ^{Org}	1059.15 (+5)	2.55
²³ RPI p SVIGGVSLYG TNQTEELDNLLTQPASRPPMPAHQVPPYK ⁶⁴	26S	Trypsin ^{Temp}	1163.59 (+4)	0.52
²³ RPI p SVIGGVSLYG TNQTEELDNLLTQPASRPPMPAHQVPPYK ⁶⁴	26S	Trypsin ^{Org}	1163.59 (+4)	2.49
⁷⁰ FRPFTFSQ p STPIGLDR ⁸⁵	78S	Trypsin	650.32 (+3)	0.92
⁷⁰ FRPFTFSQ p STPIGLDR ⁸⁵	78S	Trypsin ^{Temp}	650.32 (+3)	0.77
⁷⁰ FRPFTFSQ p STPIGLDR ⁸⁵	78S	Trypsin ^{Org}	650.32 (+3)	2.00
⁷⁰ FRPFTFSQ p STPIGLDRVGR ⁸⁸	78S	Trypsin ^{Org}	566.04 (+4)	0.53
⁷⁰ FRPFTFSQ p STPIGLDRVGR ⁸⁹	78S	Trypsin ^{Org}	605.06 (+4)	0.66
⁹⁴ ASNVSSDGGTEP p SALVDDNGSEEDFSYEDLC*QASPR ¹²⁹	106S	Trypsin	1295.86 (+3)	-0.23
⁹⁴ ASNVSSDGGTEP p SALVDDNGSEEDFSYEDLC*QASPR ¹²⁹	106S	Trypsin ^{Temp}	1295.86 (+3)	-0.39
⁹⁴ ASNVSSDGGTEP p SALVDDNGSEEDFSYEDLC*QASPR ¹²⁹	106S	Trypsin ^{Org}	1295.86 (+3)	1.00
²⁰⁵ VNQEELSENSSS p TPSEEQDEEASQSR ²³⁰	217T	Trypsin	992.73 (+3)	-0.40
²⁰⁵ VNQEELSENSSS p TPSEEQDEEASQSR ²³⁰	217T	Trypsin ^{Temp}	992.73 (+3)	0.71
²⁰³ LRVNQEELSENSSS p TPSEEQDEEASQSR ²³⁰	217T	Trypsin ^{Org}	1082.46 (+3)	0.00
²¹⁰ LSSENSSS p TPSEEQDEEASQSRHRHC*E ²³⁵	217T	Glu-C	774.81 (+4)	-0.26
⁶¹⁵ HI p TMPTSVPQQVFLAEPK ⁶³⁴	617T	Trypsin ^{Org}	763.38 (+3)	1.18

^a The “p” denotes a site of phosphorylation; an asterisk “*” denotes carboxyamidomethylation of cysteine. ^b The superscript “Org” denotes sample denatured by high percent organic solvent; the superscript “Temp” denotes sample denatured by heat.

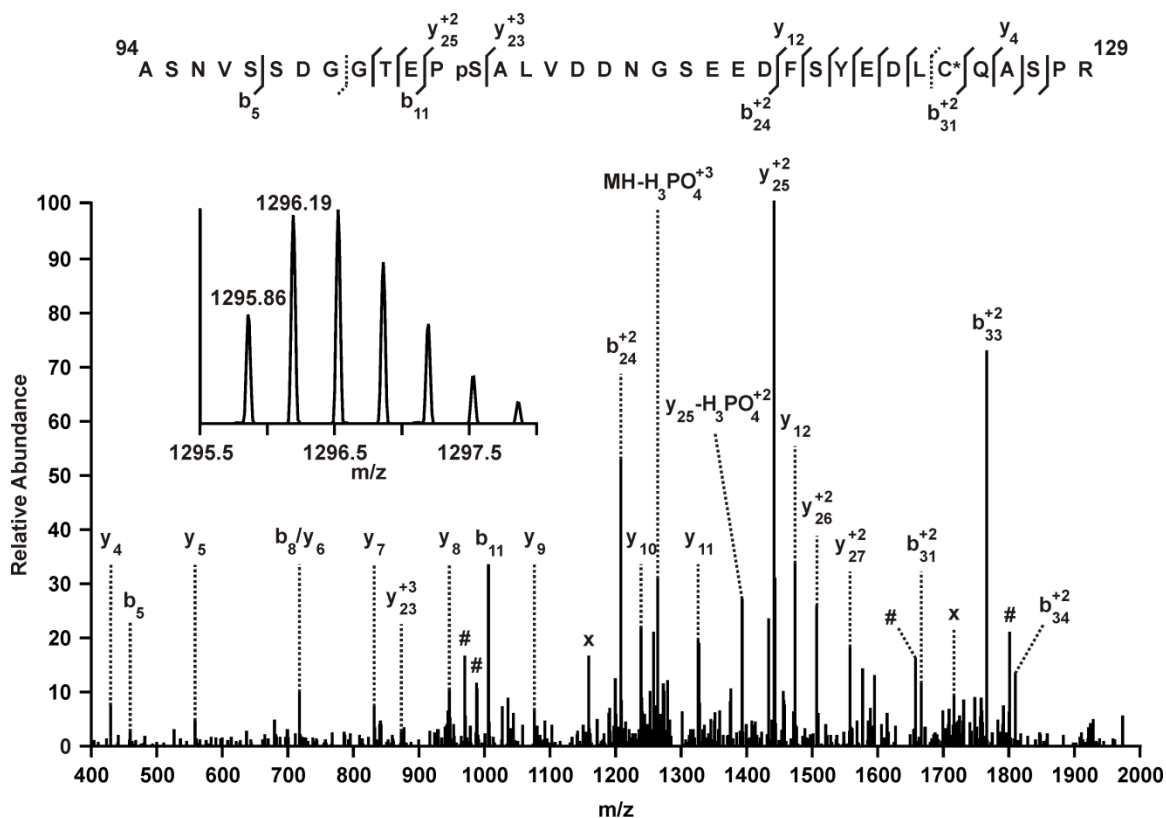


Figure 4.3. MS and MS/MS spectra for an Asef2 phosphorylated peptide. Data-targeted MS/MS scan of m/z 1295.86 in the Trypsin^{Org} sample, corresponding to the triply charged (inset) phosphopeptide $^{94}ASNVS)SDG)G)T)E)P)S)A)LVDDNGSEED)F)S)Y)E)D)L)C^*)Q)A)S)P)R^{129}$. Backbone cleavages (b and y -ions) observed are marked on the sequence (top) and additional ions are labeled in the spectrum. Sequence positions for b_8 and y_6 are shown with dashed markers (top) due to isobaric m/z values of 718.3. “C*” denotes a carbamidomethyl modified cysteine. “#” indicates ions corresponding to $-H_2O$ from b_{11}^{+2} (2), b_{31}^{+2} , and b_{34}^{+2} , respectively. “X” denotes ions corresponding to $-H_3PO_4$ from b_{24}^{+2} and b_{33}^{+2} , respectively.

to hypothesize that phosphorylation of S106 is an important regulatory mechanism for Asef2 GEF activity. To investigate the effect of S106 phosphorylation on Asef2 activity, we mutated this residue to either alanine (S106A) or aspartic acid (S106D) using site-directed mutagenesis; these substitutions represent non-phosphorylatable and phosphomimetic analogues, respectively.³⁶⁻³⁸ Then, we assessed the effect of S106 mutation on the activation of the small GTPase Rac using a GTPase activity assay.^{22, 34} In this assay, the GST-tagged binding domain from the Rac effector PAK (GST-PBD) is used to detect the active form of Rac from lysates of GFP- and GFP-Asef2-expressing cells (Figure 4.4a). As expected, expression of wild-type Asef2 caused a significant increase in the level of active Rac (Figure 4.4b). Quantification showed that the amount of active Rac was increased approximately 8-fold in GFP-Asef2-expressing cells compared with control cells expressing GFP (Figure 4.4c). In contrast, GFP-Asef2-S106A expression caused an approximately 80% decrease in active Rac compared to expression of GFP-Asef2 (Figure 4.4b, c), suggesting that phosphorylation of Asef2 at S106 promotes its GEF activity towards Rac. Expression of GFP-Asef2-S106D resulted in an approximately 2-fold increase in active Rac compared to expression of GFP-Asef2 (Figure 4.4b,c). Therefore, these results point to S106 as an important phosphorylation site in Asef2 that mediates its ability to activate Rac.

S106 Phosphorylation Regulates Cell Migration

We have previously shown that Asef2 promotes the migration of HT1080 cells plated on fibronectin via active Rac.²² In this study, we demonstrate that phosphorylation of S106 is critical for Asef2-mediated activation of Rac. This led us to hypothesize that

(a) Schematic for Rac Activity Assay

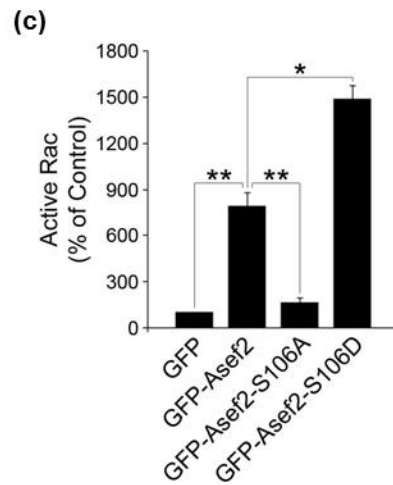
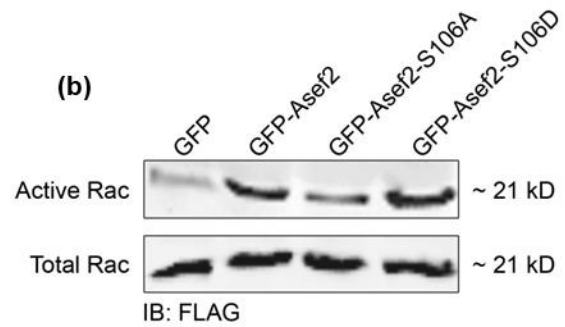
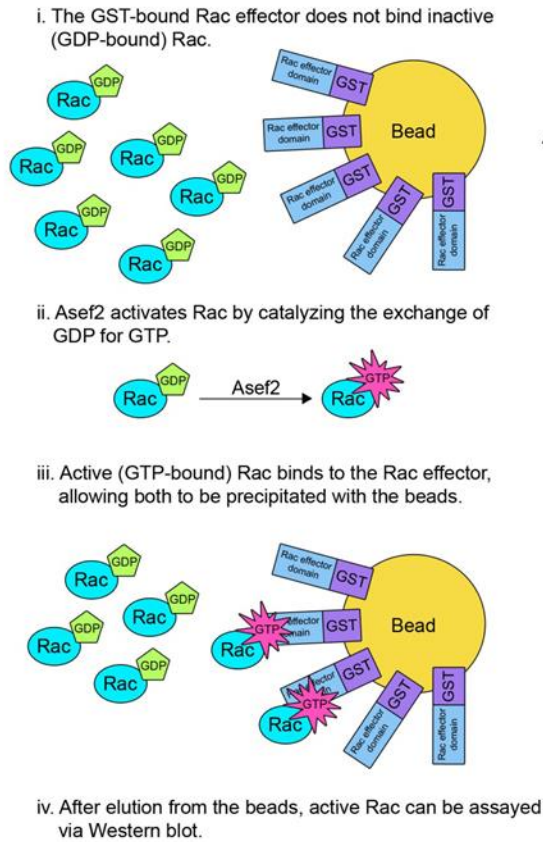


Figure 4.4. Phosphorylation of S106 stimulates Asef2 GEF activity. (a) Schematic depicting the protocol used to detect active Rac (Rac activity assay). (b) HT1080 cells were co-transfected with FLAG-Rac1 cDNA and either GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D, and three days later active Rac was pulled down from lysates from these cells. The amount of total Rac is shown as a control. (c) Quantification of the amount of active Rac from 3-7 separate experiments is shown. Error bars represent s.e.m. *, $p < 0.002$, **, $p < 0.0001$.

S106 phosphorylation of Asef2 plays a role in regulating cell migration. To test this hypothesis, we transfected HT1080 cells with GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D cDNAs, then plated the cells on fibronectin-coated dishes, and assessed cell migration using live-cell imaging. Migration data were generated by tracking individual cells and were used to calculate the migration speed. Figure 4.5(a) shows the individual tracks of GFP-, GFP-Asef2, GFP-Asef2-S106A, and GFP-Asef2-S106D-expressing cells. The migration paths of GFP-Asef2-expressing cells were significantly longer than those of control cells expressing GFP. Quantification showed an approximately 1.3-fold increase in migration speed in GFP-Asef2-expressing cells compared to those expressing GFP (Figure 4.5b). Intriguingly, mutation of serine 106 to alanine abolished this increase in migration (Figure 4.5b). The migration speed of GFP-Asef2-S106A-expressing cells was significantly decreased compared to cells expressing GFP-Asef2, suggesting that S106 phosphorylation is important for Asef2-promoted cell migration. All of the GFP-tagged proteins were expressed at similar levels (Figure 4.5c), indicating that changes in migration speed were not due to differential protein expression. Expression of the phosphomimetic S106D mutant resulted in an increase in migration speed compared to that observed with GFP expression (Figure 4.5b). However, the migration speed of cells expressing GFP-Asef2-S106D was not significantly different than the migration speed of GFP-Asef2-expressing cells. This result is somewhat unexpected given the additional increase in Rac activity detected in the GFP-Asef2-S106D-expressing cells (Figure 4.4c). The expression of wild-type Asef2 may be sufficient to maximally stimulate Asef2 signaling, at least in terms of promoting cell migration. Specifically, the high level of active Rac resulting from wild-type Asef2

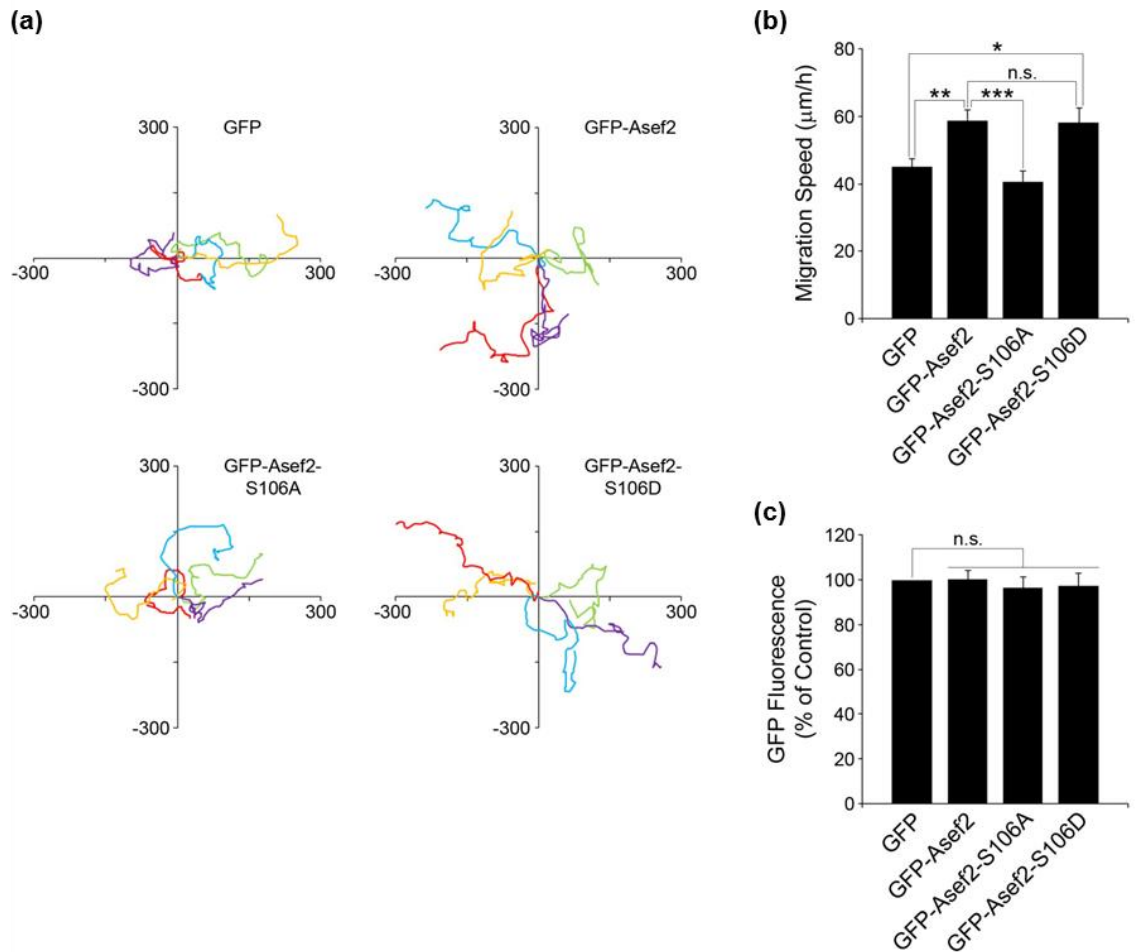


Figure 4.5. Phosphorylation of S106 is critical for Asef2-promoted cell migration. (a) HT1080 cells expressing GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D were plated on fibronectin-coated dishes and imaged using time-lapse microscopy. The migration of individual cells was tracked and analyzed. Wind-Rose plots depicting the migrations tracks for individual cells are shown. (b) Migration speed was quantified for GFP-, GFP-Asef2-, GFP-Asef2-S106A-, and GFP-Asef2-S106D-expressing cells. (c) All GFP-tagged proteins were expressed to similar levels; therefore differential protein expression was not responsible for changes in migration speed. Error bars represent s.e.m. for 4-9 independent experiments (*, $p < 0.02$, **, $p < 0.004$, ***, $p < 0.002$). For (b) and (c), “n.s.” denotes no statistically significant difference.

expression could be adequate to saturate downstream signaling; thus, a further increase in active Rac, such as that caused by GFP-Asef2-S106D expression, would not yield higher migration speed. Consistent with this, a previous study showed that expression of constitutively-active Rac did not further increase in migration speed compared to that observed with wild-type Rac expression.³⁹ Nevertheless, these results underscore the importance of S106 phosphorylation in regulating Asef2-mediated cell migration.

Phosphorylation of S106 Modulates Adhesion Turnover

Because the ability of cells to migrate efficiently is dependent on the proper assembly and disassembly of their adhesions (adhesion turnover), and because Asef2-Rac signaling plays an important role in regulating adhesion dynamics,^{7,11,22} S106 phosphorylation may affect migration by modulating adhesion turnover. Therefore, we co-expressed mCherry-paxillin, a well-characterized adhesion marker, with GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D in HT1080 cells and analyzed adhesion turnover using an assay that we previously developed.^{7,22} In this assay, mCherry-paxillin-containing adhesions from these cells were imaged using time-lapse microscopy (Figure 4.6a), and the change in fluorescence intensity in individual adhesions was used to calculate $t_{1/2}$ values for adhesion assembly and disassembly. Expression of GFP-Asef2 resulted in an approximately 50% decrease in the $t_{1/2}$ values for adhesion assembly and disassembly (Figure 4.6b); this suggests that adhesions in GFP-Asef2-expressing cells turn over more quickly, resulting in faster cell migration speeds.²² Conversely, the $t_{1/2}$ values for both adhesion assembly and disassembly were significantly larger in GFP-

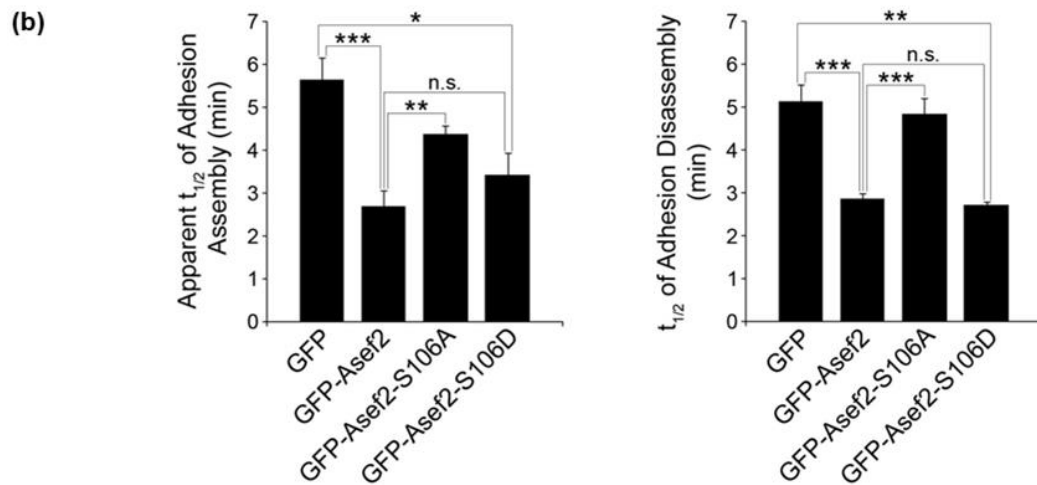
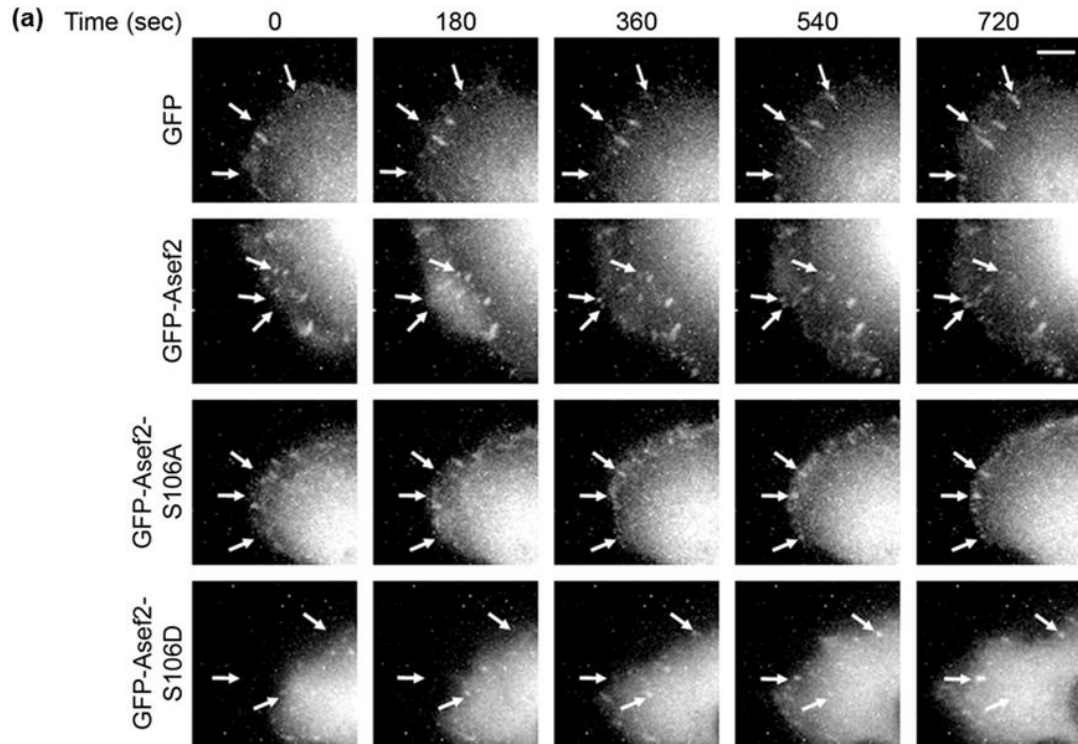


Figure 4.6. S106 phosphorylation regulates adhesion turnover. (a) HT1080 cells were co-transfected with mCherry-paxillin cDNA and GFP, GFP-Asef2, GFP-Asef2-S106A, or GFP-Asef2-S106D cDNAs and were subsequently used in adhesion turnover assays. Time-lapse images show adhesions that are assembling and disassembling at the leading edge of migrating cells (arrows). Bar=5 μ m. (b) Quantification of the apparent $t_{1/2}$ of adhesion assembly and the $t_{1/2}$ of adhesion disassembly for transfected cells is shown. For each condition, 34-70 adhesions from 11-22 cells were analyzed. Error bars represent s.e.m. for 3-6 independent experiments (*, $p < 0.04$, **, $p < 0.006$, ***, $p < 0.004$). “n.s.” denotes no statistically significant difference.

Asef2-S106A-expressing cells compared to those cells expressing GFP-Asef2 (Figure 4.6b). These data are consistent with the slower cell migration speed that was observed in GFP-Asef2-S106A-expressing cells (Figure 4.5b), further emphasizing the importance of phosphorylation of this residue for efficient cell migration. Expression of the S106D mutant, on the other hand, resulted in $t_{1/2}$ values that were comparable to those observed with GFP-Asef2 expression. Collectively, these results suggest that the phosphorylation of Asef2 at S106 promotes faster adhesion turnover, which is critical for proficient cell migration.

4.4. Conclusions

Asef2 is emerging as an important GEF in modulating cellular processes, such as migration and adhesion dynamics; however, the mechanisms that regulate the activity and function of Asef2 are currently not well understood. In this study, we identified six phosphorylation sites in Asef2 by MS analysis. We demonstrate that phosphorylation of one of these sites, S106, which is located in the ABR domain, is important for modulating Asef2 GEF activity as well as for Asef2 function in cell migration and adhesion turnover. Four of the other detected phosphorylation sites (S5, S26, S78, and T217) are congregated toward the N-terminus of Asef2. This region of the protein contains the ABR-SH3 module, which maintains Asef2 in an autoinhibited, inactive state.²⁰ Therefore, the location of these phosphorylation sites puts them in a potential position to regulate the autoinhibitory state of Asef2, even though these sites are not found within a known domain of the protein. The sixth phosphorylation site (T617) is located in the C-terminus of Asef2. The C-terminus associates with the ABR-SH3 module to maintain

Asef2 in an autoinhibitory state.²⁰ Thus, phosphorylation within this region could also affect Asef2 GEF activity. Furthermore, the C-terminus of Asef2 is involved in mediating protein-protein interactions; for example, Asef2 interacts with the actin-binding protein spinophilin, via this region.⁴⁰ Phosphorylation of T617 could be involved in regulating this association or other protein-protein interactions. Future studies are needed to determine the significance of these phosphorylation sites on Asef2 activity and function.

4.5. Acknowledgements

This chapter contains the research article submitted for publication in February 2014: J. Corey Evans*, Kelly M. Hines*, Jay G. Forsythe*, Begum Erdogan, Mingjian Shi, Salisha Hill, Kristie L. Rose, John A. McLean, and Donna J. Webb, Phosphorylation of Serine 106 in Asef2 Regulates Cell Migration and Adhesion Turnover. *Journal of Proteome Research* **2014**, Submitted. (*These authors contributed to an equal extent).

The authors thank the Vanderbilt Mass Spectrometry Research Center Proteomics Core Facility for data acquisition on a LTQ Orbitrap Velos mass spectrometer funded by National Institutes of Health (S10RR027714). We are grateful to Lan Hu for assistance in preparing cDNA constructs. This work was supported by National Institutes of Health Grant (NIH) GM092914 (D.J.W) and in part by National Institutes of Health Grant UH2TR000491 (K.M.H., J.G.F. and J.A.M), the Defense Threat Reduction Agency HDTRA-09-1-00-13 and DTRA100271 A-5196 (K.M.H., J.G.F., and J.A.M.), the Vanderbilt Institute of Chemical Biology (J.A.M.), and the Vanderbilt Institute for Integrative Biosystems Research and Education (D.J.W. and J.A.M).

4.6. *References*

1. Vicente-Manzanares, M.; Horwitz, A. R., Cell migration: an overview. *Methods Mol Biol* **2011**, *769*, 1-24.
2. Yamaguchi, H.; Wyckoff, J.; Condeelis, J., Cell migration in tumors. *Curr Opin Cell Biol* **2005**, *17* (5), 559-64.
3. Hopkins, P. N., Molecular biology of atherosclerosis. *Physiol Rev* **2013**, *93* (3), 1317-542.
4. Lauffenburger, D. A.; Horwitz, A. F., Cell migration: a physically integrated molecular process. *Cell* **1996**, *84* (3), 359-69.
5. Beningo, K. A.; Dembo, M.; Kaverina, I.; Small, J. V.; Wang, Y. L., Nascent focal adhesions are responsible for the generation of strong propulsive forces in migrating fibroblasts. *J Cell Biol* **2001**, *153* (4), 881-8.
6. Laukaitis, C. M.; Webb, D. J.; Donais, K.; Horwitz, A. F., Differential dynamics of alpha 5 integrin, paxillin, and alpha-actinin during formation and disassembly of adhesions in migrating cells. *J. Cell Biol.* **2001**, *153* (7), 1427-40.
7. Webb, D. J.; Donais, K.; Whitmore, L. A.; Thomas, S. M.; Turner, C. E.; Parsons, J. T.; Horwitz, A. F., FAK-Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly. *Nat Cell Biol* **2004**, *6* (2), 154-61.
8. Choi, C. K.; Vicente-Manzanares, M.; Zareno, J.; Whitmore, L. A.; Mogilner, A.; Horwitz, A. R., Actin and alpha-actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nat Cell Biol* **2008**, *10* (9), 1039-50.
9. Webb, D. J.; Parsons, J. T.; Horwitz, A. F., Adhesion assembly, disassembly and turnover in migrating cells -- over and over and over again. *Nat Cell Biol* **2002**, *4* (4), E97-100.
10. Chrzanowska-Wodnicka, M.; Burridge, K., Rho-stimulated contractility drives the formation of stress fibers and focal adhesions. *J Cell Biol* **1996**, *133* (6), 1403-15.
11. Rottner, K.; Hall, A.; Small, J. V., Interplay between Rac and Rho in the control of substrate contact dynamics. *Curr Biol* **1999**, *9* (12), 640-8.
12. Nobes, C. D.; Hall, A., Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **1995**, *81* (1), 53-62.
13. Jaffe, A. B.; Hall, A., Rho GTPases: biochemistry and biology. *Annu Rev Cell Dev Biol* **2005**, *21*, 247-69.

14. West, M.; Kung, H. F.; Kamata, T., A novel membrane factor stimulates guanine nucleotide exchange reaction of ras proteins. *FEBS Lett* **1990**, *259* (2), 245-8.
15. Cherfils, J.; Zeghouf, M., Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* **2013**, *93* (1), 269-309.
16. Trahey, M.; McCormick, F., A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants. *Science* **1987**, *238* (4826), 542-5.
17. Gibbs, J. B.; Schaber, M. D.; Allard, W. J.; Sigal, I. S.; Scolnick, E. M., Purification of ras GTPase activating protein from bovine brain. *Proc Natl Acad Sci U S A* **1988**, *85* (14), 5026-30.
18. Ridley, A. J.; Hall, A., The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* **1992**, *70* (3), 389-99.
19. Ridley, A. J.; Paterson, H. F.; Johnston, C. L.; Diekmann, D.; Hall, A., The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* **1992**, *70* (3), 401-10.
20. Hamann, M. J.; Lubking, C. M.; Luchini, D. N.; Billadeau, D. D., Asef2 functions as a Cdc42 exchange factor and is stimulated by the release of an autoinhibitory module from a concealed C-terminal activation element. *Mol Cell Biol* **2007**, *27* (4), 1380-93.
21. Kawasaki, Y.; Sagara, M.; Shibata, Y.; Shirouzu, M.; Yokoyama, S.; Akiyama, T., Identification and characterization of Asef2, a guanine-nucleotide exchange factor specific for Rac1 and Cdc42. *Oncogene* **2007**, *26* (55), 7620-267.
22. Bristow, J. M.; Sellers, M. H.; Majumdar, D.; Anderson, B.; Hu, L.; Webb, D. J., The Rho-family GEF Asef2 activates Rac to modulate adhesion and actin dynamics and thereby regulate cell migration. *J Cell Sci* **2009**, *122* (Pt 24), 4535-46.
23. Murayama, K.; Shirouzu, M.; Kawasaki, Y.; Kato-Murayama, M.; Hanawa-Suetsugu, K.; Sakamoto, A.; Katsura, Y.; Suenaga, A.; Toyama, M.; Terada, T.; Taiji, M.; Akiyama, T.; Yokoyama, S., Crystal structure of the rac activator, Asef, reveals its autoinhibitory mechanism. *J Biol Chem* **2007**, *282* (7), 4238-42.
24. Zhang, Z.; Chen, L.; Gao, L.; Lin, K.; Zhu, L.; Lu, Y.; Shi, X.; Gao, Y.; Zhou, J.; Xu, P.; Zhang, J.; Wu, G., Structural basis for the recognition of Asef by adenomatous polyposis coli. *Cell Res* **2012**, *22* (2), 372-86.
25. Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr., Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl* **2005**, *44* (45), 7342-72.

26. Kato, J.; Kaziro, Y.; Satoh, T., Activation of the guanine nucleotide exchange factor Dbl following ACK1-dependent tyrosine phosphorylation. *Biochem Biophys Res Commun* **2000**, *268* (1), 141-7.
27. Servitja, J. M.; Marinissen, M. J.; Sodhi, A.; Bustelo, X. R.; Gutkind, J. S., Rac1 function is required for Src-induced transformation. Evidence of a role for Tiam1 and Vav2 in Rac activation by Src. *J Biol Chem* **2003**, *278* (36), 34339-46.
28. Miyamoto, Y.; Torii, T.; Yamamori, N.; Ogata, T.; Tanoue, A.; Yamauchi, J., Akt and PP2A reciprocally regulate the guanine nucleotide exchange factor Dock6 to control axon growth of sensory neurons. *Sci Signal* **2013**, *6* (265), ra15.
29. Gant-Branum, R. L.; Broussard, J. A.; Mahsut, A.; Webb, D. J.; McLean, J. A., Identification of phosphorylation sites within the signaling adaptor APPL1 by mass spectrometry. *J Proteome Res* **2010**, *9* (3), 1541-8.
30. Russell, W. K.; Park, Z. Y.; Russell, D. H., Proteolysis in mixed organic-aqueous solvent systems: applications for peptide mass mapping using mass spectrometry. *Anal Chem* **2001**, *73* (11), 2682-5.
31. Strader, M. B.; Tabb, D. L.; Hervey, W. J.; Pan, C.; Hurst, G. B., Efficient and specific trypsin digestion of microgram to nanogram quantities of proteins in organic-aqueous solvent systems. *Anal Chem* **2006**, *78* (1), 125-34.
32. Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R., Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **2002**, *74* (20), 5383-92.
33. Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R., A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **2003**, *75* (17), 4646-58.
34. Jean, L.; Majumdar, D.; Shi, M.; Hinkle, L. E.; Diggins, N. L.; Ao, M.; Broussard, J. A.; Evans, J. C.; Choma, D. P.; Webb, D. J., Activation of Rac by Asef2 promotes myosin II-dependent contractility to inhibit cell migration on type I collagen. *J Cell Sci* **2013**, *126* (Pt 24), 5585-97.
35. Knaus, U. G.; Bamberg, A.; Bokoch, G. M., Rac and Rap GTPase activation assays. *Methods Mol Biol* **2007**, *412*, 59-67.
36. Langan, T. A.; Rall, S. C.; Cole, R. D., Variation in primary structure at a phosphorylation site in lysine-rich histones. *J Biol Chem* **1971**, *246* (6), 1942-4.
37. Thorsness, P. E.; Koshland, D. E., Jr., Inactivation of isocitrate dehydrogenase by phosphorylation is mediated by the negative charge of the phosphate. *J Biol Chem* **1987**, *262* (22), 10422-5.

38. Tarrant, M. K.; Cole, P. A., The chemical biology of protein phosphorylation. *Annu Rev Biochem* **2009**, *78*, 797-825.
39. Pankov, R.; Endo, Y.; Even-Ram, S.; Araki, M.; Clark, K.; Cukierman, E.; Matsumoto, K.; Yamada, K. M., A Rac switch regulates random versus directionally persistent cell migration. *J Cell Biol* **2005**, *170* (5), 793-802.
40. Sagara, M.; Kawasaki, Y.; Iemura, S. I.; Natsume, T.; Takai, Y.; Akiyama, T., Asef2 and Neurabin2 cooperatively regulate actin cytoskeletal organization and are involved in HGF-induced cell migration. *Oncogene* **2009**, *28* (10), 1357-65.

CHAPTER V

PERSPECTIVES ON EMERGING AND FUTURE DIRECTIONS

5.1. Summary

The biomolecular signatures of diseases have been investigated using advanced IM and MS techniques and bioinformatics strategies. A workflow for the untargeted ESI-IM-MS analysis of 64 samples of diabetic wound fluid was developed, which included statistically-driven prioritization of differentially expressed biomolecules, bioinformatics-based tentative identifications, targeted UPLC-IM-MS analyses, MS/MS experiments, and validation of tentative identifications with standards. Cholic acid, lysophosphatidylcholine 20:4, and the protein S100-A8 were found to be differentially expressed between wound fluids from diabetic and non-diabetic animals across two time points after the initial wound. The utility of IM-MS for complex biological sample analysis was demonstrated in the identification of biomolecular signatures of disease representing protein, lipid and metabolite classes of biomolecular species.

The workflow developed for the wound fluid analysis was expanded upon for the study of breast cancer metabolomics. Breast tissues representing disease and matched controls were homogenized and polar species extracted. An untargeted analysis by UPLC-IM-MS/MS, combined with biostatistics and bioinformatics tools, revealed an unexpected grouping of the samples. Although researchers were blinded to the particular pathologies of the tissues, grouping in the PCA led to the hypothesis that the disease tissues represented both cancerous and benign breast diseases, and was later confirmed

upon revealing of the tissue pathologies. Key molecular features distinguishing these states represented species previously known to be affected by the aberrant metabolism in cancer, including the polypeptides thymosin β 4 and thymosin β 10, choline-containing metabolites, adenosine monophosphate, and reduced and oxidized glutathione. The transition from unknown molecular features to validated identifications of biomolecular signatures using an IM-based approach was demonstrated for the glutathione species, where structural separation in IM dimension was critical for resolving isobaric species and distinguishing multiple sources of fragmentation.

An MS approach was used to study another biomolecular species implicated in cancer. Asef2 is a protein suspected to regulate the cell migration, a process central to the development of metastasis in cancer. Phosphorylation of Asef2 was previously uncharacterized, but was implicated as a possible method of regulation based on results for other proteins in the GEF family. Using an LC-MS/MS proteomics approach incorporating multiple digestion strategies and bioinformatics based peptide and protein identification, the phosphorylation of the protein Asef2 was mapped. While initial digestions with trypsin, chymotrypsin and Glu-C provided moderate sequence coverage (85.6%), additional trypsin digestions under strongly denaturing conditions of heat and high organic solvent content increases the sequence coverage of the 652 amino acid protein to 94.6%. Most notably, 100% of the phosphorylatable residues (serine, threonine, tyrosine) were covered by the combine analyses, and six sites of phosphorylation were identified. One particular site, serine 106, was located in a region of Asef2 critical for relief of autoinhibition. Mutation analysis revealed this site of

phosphorylation regulated adhesion turnover and cell migration, and promoted Rac activation.

5.2. *Future Directions*

5.2.1. *Biomolecular Signatures of Diabetic Wound Healing*

Many of the biomolecular signatures found to distinguish control and diabetic wound fluids have also been characterized from biological matrices such as blood, serum and plasma. As the molecular contents of wound fluid originate in some part from the circulatory system, a dual characterization of wound fluids and serum may provide new insight into the wound healing process. Relative to wound fluid, serum represent a more abundant and more accessible biological sample from which the state of wound healing may be assessed. Of particular interest would be whether these different biological matrices reveal the same biomolecular signatures for distinguishing diabetic versus non-diabetic. Additional time points prior to the earliest time point used in our study would provide a better perspective on the rate at which diabetic and non-diabetic wound diverge in the wound healing process.

5.2.2. *Distinguishing Cancerous and Noncancerous Breast Diseases*

From Figure 3.2(b), it is apparent that there are many more molecular species contributing to the differentiation of cancerous, benign and healthy tissues than those 18 which were focused on in our analysis. While interrogation of all these features is feasible with a sufficient investment in time, a more targeted strategy may be to first investigate the features which are connected to those already identified through known

metabolic pathways. This is possible with use of the KEGG metabolite database and software such as Cytoscape and Metscape, which can be used to map known features to the pathways in which they participate and reveal molecules in those pathways.

The analysis of breast cancer metabolomics could be expanded by adding more samples to each of the disease subsets used here to describe their metabolic differences with more statistical power. Power analysis for the features identified in the current work indicates that a total of three biological replicates per current sample (*i.e.* 18 tissues: 3 triple negative, 3 fibroadenoma, 3 pseudoangiomatous hyperplasia and 9 appropriate matched controls) would provide a statistical power of one for a two-tailed t-test where $\alpha=0.01$ (calculation based on means for choline, m/z 104, for samples 3C and 4D). The study may also be expanded by analyzing the nonpolar extracts of the tissues, which would reflect the changes in species such as lipids, steroids, and fatty acids between cancerous and non-cancerous breast tissues.

5.2.3. Characterizing the Phosphorylation of Asef2

A total of six sites of phosphorylation were identified in our analysis of Asef2, many of which clustered near the N-terminus of the protein. The site S106 was chosen for the initial experiments based on its location within the APC binding region (ABR; Figure 4.2b). As discussed in section 4.1, relief of Asef2's autoinhibition occurs through tandem binding of the ABR and SH3 domains by APC. One of the five unexplored phosphorylation sites, T217, lies just beyond the SH3 domain. Given the significance of the ABR and SH3 domains, mutation analyses similar to those described in Chapter IV could be performed for T217 to determine if this phosphorylation impacts the

autoinhibition of Asef2 in a manner comparable to S106. In addition, dual mutation experiments for both S106 and pT217 simultaneous may provide further insight into the role of phosphorylation in regulating the activity of Asef2. Another promising site is T617, located within the C-terminus of Asef2. As described in Chapter IV, the C-terminus is involved in maintaining the autoinhibitory state of Asef2 by associating with the ABR-SH3 domains. Thus, T617 presents a third potential avenue for modulating the activity and autoinhibition of Asef2.

5.3. Conclusions

Mapping the phosphorylation of the protein Asef2 with an MS-based proteomics strategy was a substantial undertaking in terms of the data generated and interpreted. The development of robust bioinformatics contributed substantially to the wide-spread acceptance of MS-based proteomics as the standard approach for identifying proteins. Given the enormity of such datasets, manual interpretation would require an extensive investment of time. IM-MS approaches face a similar challenge for the identification of biomolecular signatures of diseases from complex samples. The development of bioinformatics and biostatistics platforms utilizing the data in the IM dimension would vastly expand the utility and approachability of IM for large-scale analyses of highly complex samples. As discussed in Chapters II and III, peak picking and data alignment in the IM dimension is typically not supported by the most widely used bioinformatics platforms. This often necessitates incorporation of chromatographic separations, for which most bioinformatics packages were developed, effectively excluding the IM dimension and reducing the dimensionality of the dataset in order to perform peak

picking and data alignment. In this case, the information in the IM dimension is not being utilized to the maximum but used as supplementary information in support of results from the LC-MS-based bioinformatics.

Despite the current challenges described above, the works described in this dissertation strongly demonstrate the merits of incorporating IM separations into the traditional strategies used for untargeted analyses of complex biological samples. Chapters II and III demonstrated the advantages of incorporating ion mobility into workflows for characterizing complex biological samples. In the analysis of wound fluids, the improvement in S/N gained by extracting signals away from overlapping chemical noise in the IM dimension was demonstrated. Additionally, the structural information provided by IM separation enabled hypothesis driven identification of unknown molecular species based on their locations within conformation space or the proximity between species in conformation space. The flexibility of extracting information from multidimensional UPLC-IM-MS/MS analysis was demonstrated in the context of the breast cancer metabolomics study. The separation in the IM dimension of nearly isobaric species and different sources of fragmentation from data-independent acquisition of MS/MS spectra aided in the identification of a significantly differentially expressed metabolite. Further, the additional information in the form of a drift time provided a third piece of evidence on which experimental unknowns and standards could be compared in the validation of tentative identifications. Thus, the incorporation of IM into contemporary MS techniques greatly benefits the field of biologically driven studies targeting the biomolecular signatures of diseases.

APPENDIX A

REFERENCES OF ADAPTATION FOR CHAPTERS

- Chapter I. **Kelly M. Hines**, Jeffrey R. Enders, and John A. McLean, “Multidimensional Separations by Ion Mobility-Mass Spectrometry,” Invited chapter for inclusion in “Encyclopedia of Analytical Chemistry (online),” Robert Myers and David Muddiman, Eds. *John Wiley & Sons, Ltd.* December 2012.
- Chapter II. **Kelly M. Hines**, Samir Ashfaq, Jeffrey M. Davidson, Susan R. Opalenik, John P. Wikswo and John A. McLean, “Biomolecular Signatures of Diabetic Wound Healing by Structural Mass Spectrometry,” *Analytical Chemistry*, 2013, 85 (7), 3651-3659.
- Chapter III **Kelly M. Hines**, Billy R. Ballard, Dana M. Marshall and John A. McLean, “IM-MS-based Profiling of Human Tissues to Distinguish Cancerous and Non-cancerous Breast Diseases,” In preparation for *Analytical Chemistry*, 2014.
- Chapter IV J. Corey Evans, **Kelly M. Hines**, Jay G. Forsythe, Begum Erdogan, Mingjian Shi, Salisha Hill, Kristie L. Rose, John A. McLean and Donna J. Webb, “Phosphorylation of Serine 106 in Asef2 Regulates Cell Migration and Adhesion Turnover,” Submitted to *Journal of Proteome Research*, February 2014.

APPENDIX B

SUPPORTING INFORMATION FOR CHAPTER II

B.1. Model Parameters for PLS-DA Score Plots and OPLS-DA S-Plots

The following model parameters (Tables A.1-A.3) correspond with the PLS-DA score plots and OPLS-DA S-plots found in the main text of Chapter II as Figure 2.3. The values were reported directly from the Extended Statistic software, and represent the variability extracted from the data by each principal component. Values of R2Y and Q2 for principal components beyond component 1 (Comp. 1) represent cumulative totals of the variability accounted for by addition of each consecutive principal component. Therefore, the variability extracted by one particular component can be determined by subtracting the value of R2Y or Q2 corresponding to the previous component. Components for the OPLS-DA analyses are labeled “Princ.” and “Orthog.” to indicate the principal component which extracts the variability due to the group differences, and the orthogonal components which extract all other variability from the dataset, respectively.

Table B.1. Day 2 Control vs. Diabetic PLS-DA and OPLS-DA Model Parameters.

PLS-DA Model Parameters		
Component No.	R²Y(cum)	Q²(cum)
Comp. 1	0.69	0.61
Comp. 2	0.83	0.71
Comp. 3	0.90	0.79
Comp. 4	0.96	0.83
Comp. 5	0.98	0.89
Comp. 6	0.99	0.93

OPLS-DA Model Parameters		
Component No.	R²Y(cum)	Q²(cum)
Comp. 1 (Princ.)	0.90	0.61
Comp. 2 (Orthog.)	0.90	0.72
Comp. 3 (Orthog.)	0.90	0.82

Table B.2. Day 5 Control vs. Diabetic PLS-DA and OPLS-DA Model Parameters.

PLS-DA Model Parameters		
Component No.	R²Y(cum)	Q²(cum)
Comp. 1	0.33	0.26
Comp. 2	0.66	0.48
Comp. 3	0.86	0.61
Comp. 4	0.92	0.72
Comp. 5	0.97	0.82
Comp. 6	0.98	0.87

OPLS-DA Model Parameters		
Component No.	R²Y(cum)	Q²(cum)
Comp. 1 (Princ.)	0.86	0.26
Comp. 2 (Orthog.)	0.86	0.46
Comp. 3 (Orthog.)	0.86	0.66

Table B.3. Day 5 Control vs. Day 2 Diabetic PLS-DA and OPLS-DA Model Parameters.

<i>PLS-DA: Day 5 Control vs. Day 2 Diabetic</i>		
Component No.	R²Y(cum)	Q²(cum)
Comp. 1	0.31	0.22
Comp. 2	0.63	0.50
Comp. 3	0.92	0.68
<i>OPLS-DA: Day 5 Control vs. Day 2 Diabetic</i>		
Component No.	R²Y(cum)	Q²(cum)
Comp. 1 (Princ.)	0.92	0.22
Comp. 2 (Orthog.)	0.92	0.47
Comp. 3 (Orthog.)	0.92	0.70

B.2. Box-and-Whisker Plots for Biomolecular Signatures

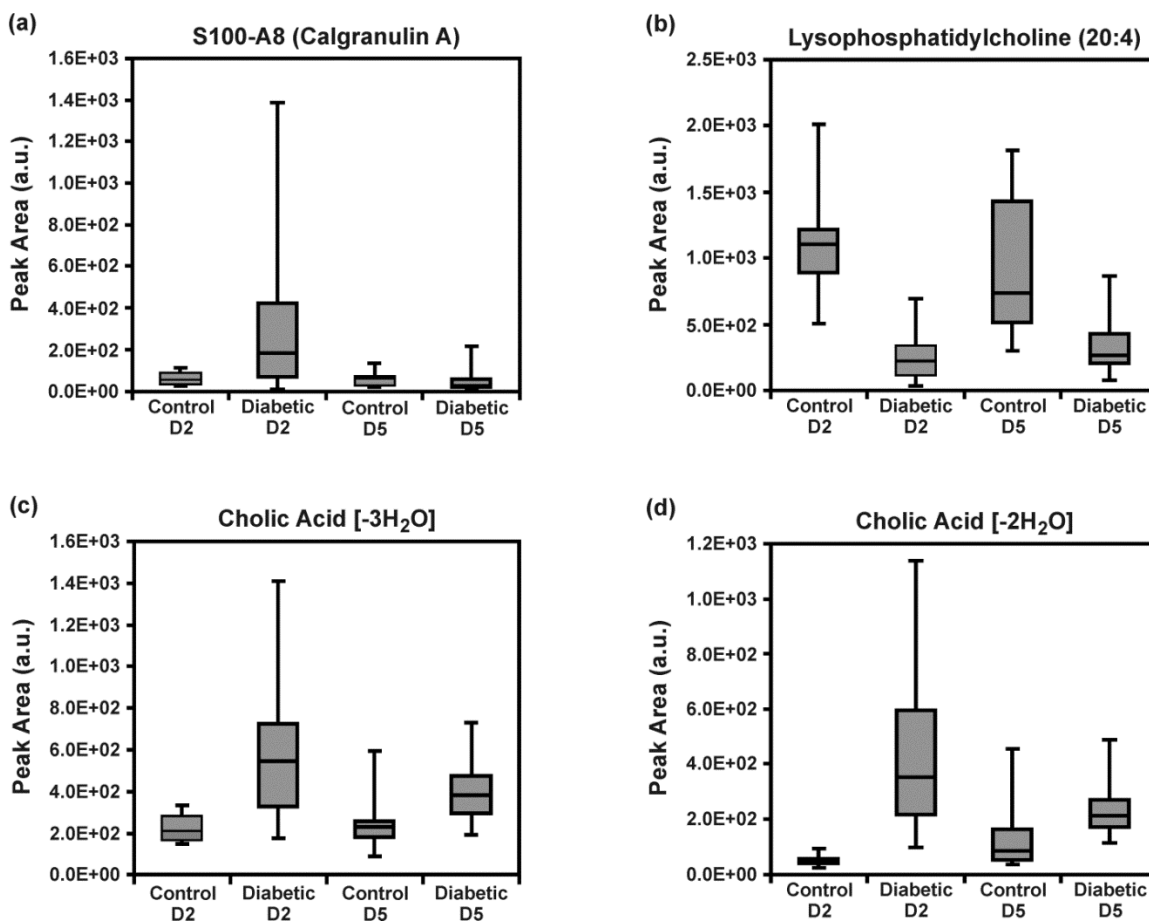


Figure B.1. Box-and-whisker plots prepared from the mobility-extracted integrated m/z peak areas for the species found to be statistically significant. (a) The S100-A8 was observed with a 5-fold increase in the diabetic day 2 group relative to the control ($p = 9.56 \times 10^{-5}$). No significance was found between the day 5 control and diabetic groups. (b) Lysophosphatidylcholine (20:4) (m/z 544.4) was generally more abundant in the control groups relative to the diabetic groups at both time points. A 4.3-fold increase was observed in control day 2 wound fluids relative to diabetic day 2 wound fluids ($p = 3.66 \times 10^{-11}$), and a 2.7-fold increase was observed in control day 5 wound fluids relative to diabetic day 5 wound fluids ($p = 8.37 \times 10^{-6}$). (c) The cholic acid species [-3H₂O] (m/z 355.3) was generally more abundant in the diabetic wound fluids relative to control wound fluids at both time points. A 2.6-fold increase ($p = 2.5 \times 10^{-8}$) was observed in the diabetic day 2 group relative to its control, and a 1.5-fold increase ($p = 1.8 \times 10^{-4}$) was observed in the diabetic day 5 wound fluids relative to their controls. (d) The cholic acid [-2H₂O] species (m/z 373.3) was generally more abundant in the diabetic wound fluids relative to control wound fluids at both time points. An 8.8-fold increase ($p = 2.3 \times 10^{-10}$) was observed in the diabetic day 2 group relative to its control, and a 1.7-fold increase ($p = 1.1 \times 10^{-3}$) was observed in the diabetic day 5 wound fluids relative to their controls.

B.3. Demonstration of IM-MS for Separating Interfering Signals

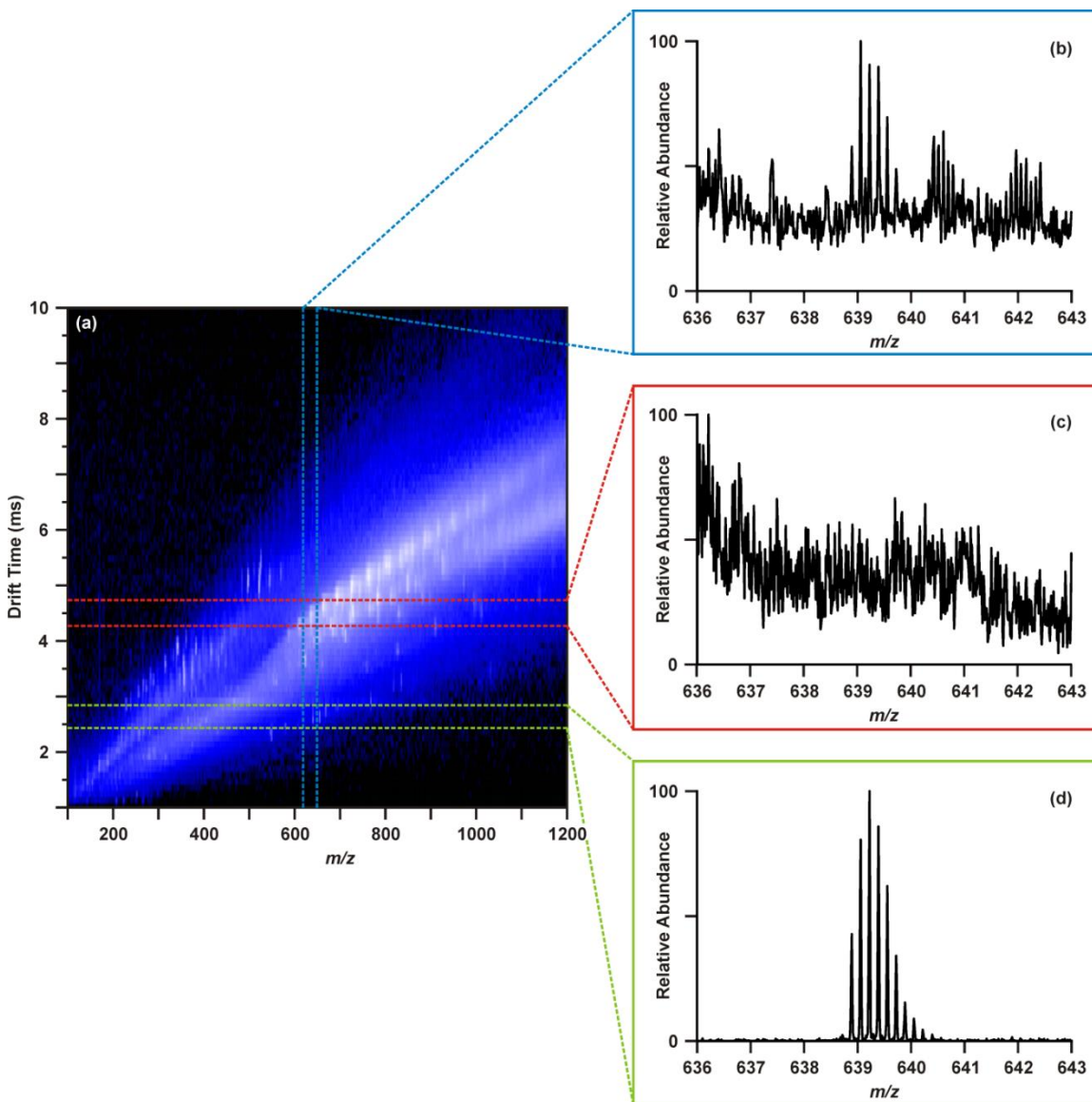


Figure B.2. Demonstration of IM-MS for separation of chemical noise from signals. (a) The area outlined with blue lines corresponds to extracting a defined region of m/z space across all drift times. The corresponding mass spectrum is presented in (b), and represents a typical mass spectrum without mobility separation. Individual isotopic clusters are not baseline resolved due to high background chemical noise in (b). The area outlined with red lines corresponds to extracting the signals contained in the defined region of IM. The combination of the blue lines and the red lines defines a region of drift time- m/z space, yielding the extracted mass spectrum (c). The combination of the blue lines and the green lines defines the region of drift time- m/z space containing the multiply-charged species, shown in (d). Structural separation of the background chemical noise in (c) from the full mass spectrum in (b) yields (d), in which the isotopic distribution of the multiply-charged species is baseline resolved.

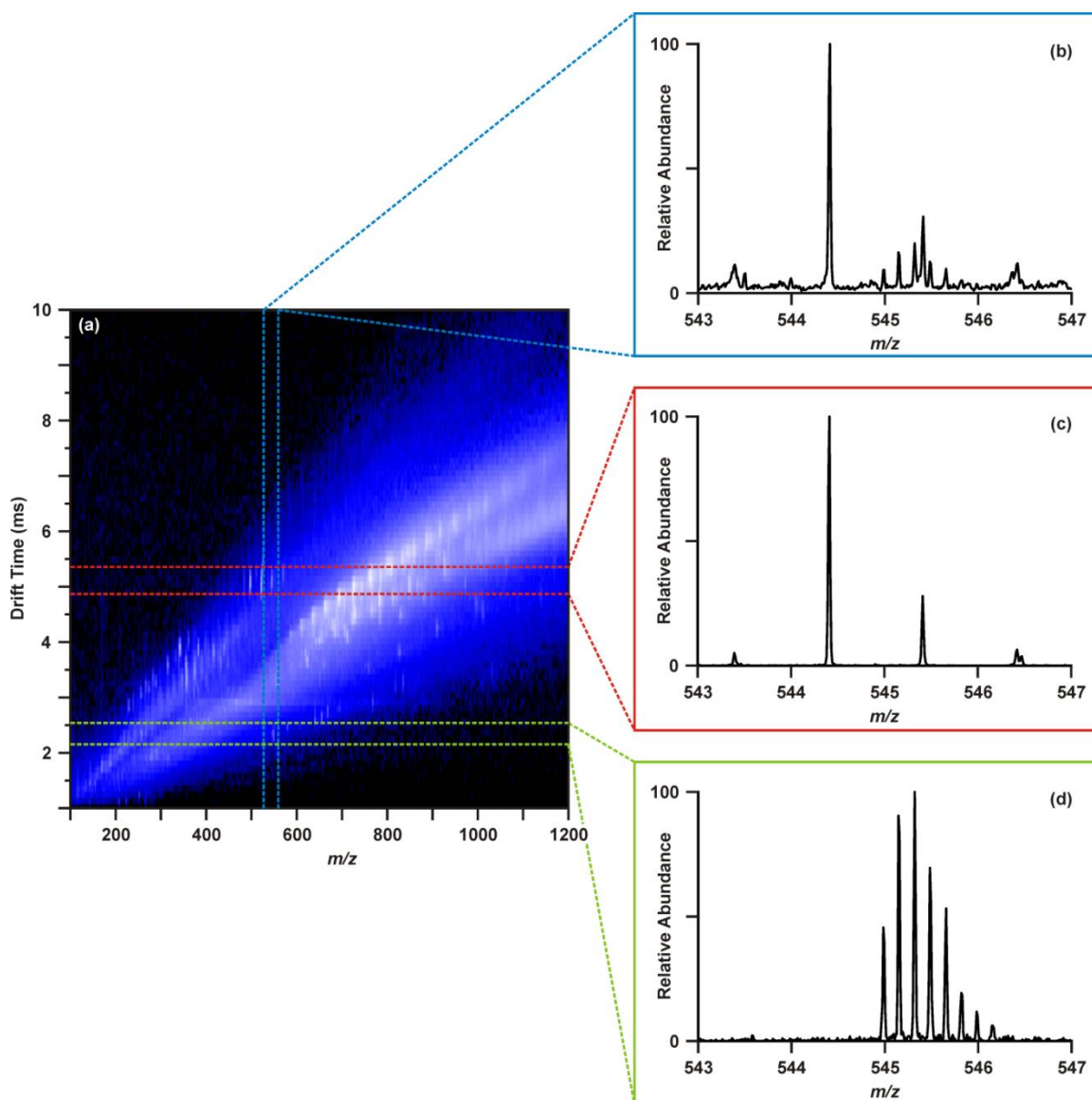


Figure B.3. Demonstration of IM-MS for resolving overlapping isotopic distributions of nearly-isobaric species. (a) The area outlined with blue lines corresponds to extracting a defined region of m/z space across all drift times. The corresponding mass spectrum is presented in (b), and represents a typical mass spectrum without mobility separation. While the monoisotopic peak of m/z 544.4 is resolved, the second isotopic peak m/z 545.5 is not baseline resolved. The area outlined with red lines corresponds to extracting the signals contained in the defined region of IM space. The combination of the blue lines and the red defines a region of drift time- m/z space, yielding the mass spectrum in (c). Likewise, the combination of the blue lines and the green lines defines the region of drift time- m/z space containing the multiply-charged species, shown in (d). Structural separation of these two species yield baseline resolve mass spectra for both, clearly showing a singly-charged species at m/z 544.4 in (c) and a 6^+ species at m/z 544.9.

APPENDIX C

SUPPORTING INFORMATION FOR CHAPTER III

C.1. Tissue Weights

Table C.1. Weights of Tissues used for Polar Metabolite Extraction

Sample Name	Tissue Weight (mg, wet)
1C	44.93
2D	50.57
3C	40.69
4D	48.99
5C	52.09
6D	48.80
Average	47.68
Standard Deviation	± 3.81

C.2. Method for Data Processing with XCMS

The R package can be downloaded from <http://www.r-project.org/>. After installing R, XCMS may be downloaded using the following code in an active R session:

```
source("http://bioconductor.org/biocLite.R")
biocLite("xcms", dep=T)
```

Additional details may be found on the XCMS Bioconductor website (<http://www.bioconductor.org/packages/release/bioc/html/xcms.html>). Prior to processing data with XCMS, .raw files must be converted to .mzXML using ProteoWizard (<http://proteowizard.sourceforge.net/>) MSConvert and the “sortByScanTime” function.

C.2.1. XCMS Method

```
xset <-xcmsSet()
xset
xset <-group(xset)
xset2 <-retcor(xset, method="obiwarp",
              plottype="deviation")
xset2 <-group(xset2, bw=10)
xset3 <-fillPeaks(xset2)
xset3
reporttab <-diffreport(xset3, "Group_1", "Group_2",
                      "File_Name", [Number of EICs, boxplots, to
                      generate])
```

The aligned data in the diffreport was normalized such that all intensities within a sample summed to 10,000. The data was then transposed such that the sample and file names were in a column instead of a row, and this dataset was imported into Umetrics Extended Statistics for statistical analysis.

C.3. Principal Components Analysis with QC Samples

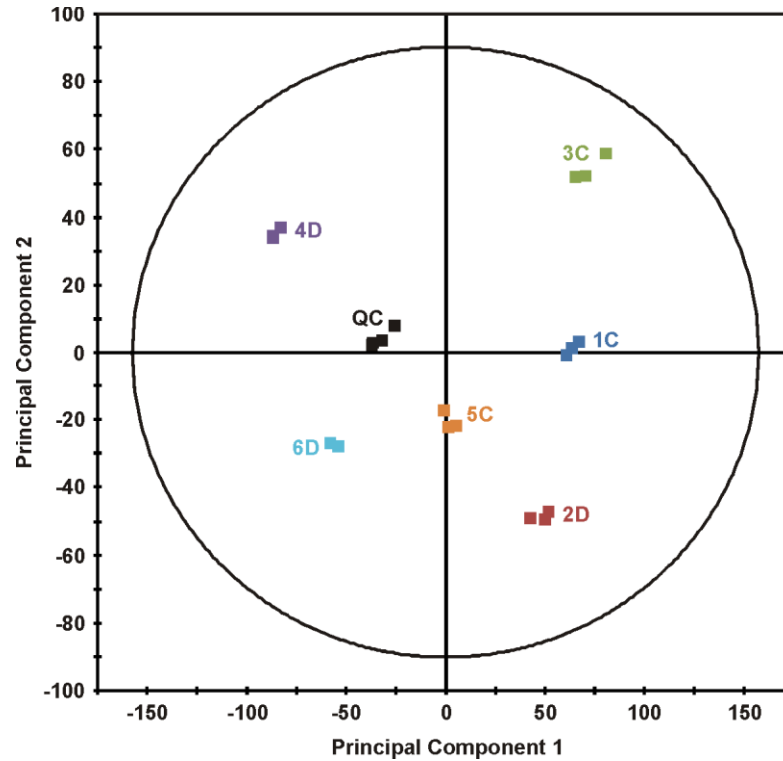


Figure C.1. Principal component analysis score plot for breast cancer tissue dataset including quality control (QC) samples for diagnostic purposes. QC samples group together well, indicating good reproducibility throughout the queue. QC samples cluster near the center of the PCA indicating they are representative of all samples.

C.4. Goodness of Fit for PCA (Figure 3.2(a))

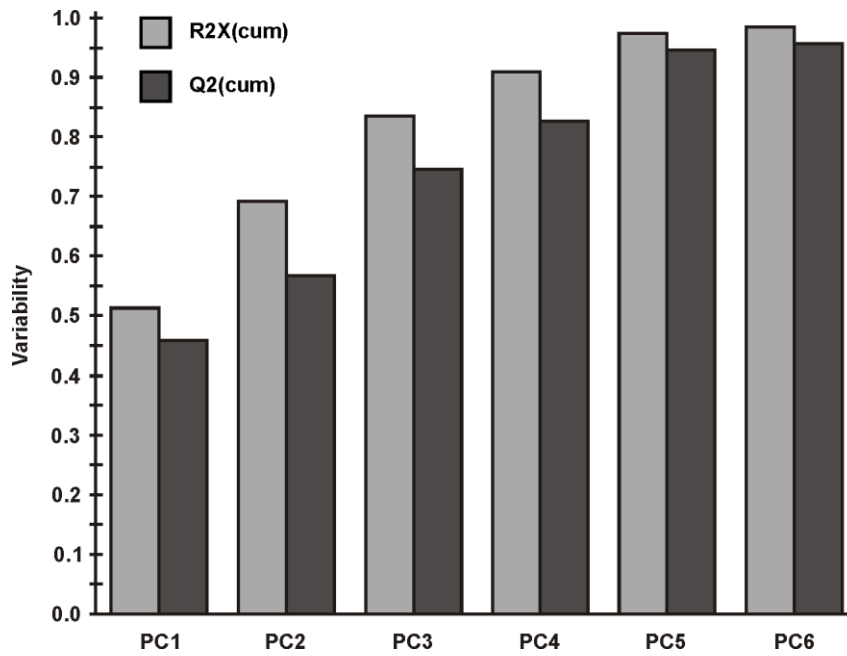


Figure C.2. Goodness of fit for PCA analysis of breast tissue extracts. Model parameters R2X (light grey) and Q2 (dark grey) describe the cumulative variability extracted by including each consecutive principal component (PC1-6). For example, including PC1 and PC2 extracts approximately 70% (R2X) of the variability in the dataset.

C.5. Data Used to Prepare Figure 3.2(c) and Table 3.1

Data from XCMS peak aligned output. All comparisons are made between matched disease and control pairs. Fold-changes were calculated by dividing the larger of the two tissues (control or diabetic) by the smaller to obtain numbers ≥ 1 . The student's t-test with two-tails, equal variance, and $\alpha=0.05$ was used to calculate p -values.

Table C.2. Normalized abundances, fold-changes and *p*-values for Feature #1.

Feature 1 <i>m/z</i> 258.11, 1.53 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	20.62	21.12	0.37	8.1	6.5E-7
1C_02	21.23				
1C_03	21.51				
2D_01	3.01	2.60	0.30		
2D_02	2.44				
2D_03	2.34				
3C_01	57.87	59.18	2.02	3.7	1.4E-6
3C_02	57.65				
3C_03	62.03				
4D_01	222.92	217.99	4.51		
4D_02	219.03				
4D_03	212.03				
5C_01	10.41	9.93	0.49	6.0	3.0E-5
5C_02	10.13				
5C_03	9.26				
6D_01	58.93	60.05	3.33		
6D_02	56.65				
6D_03	64.57				

Table C.3. Normalized abundances, fold-changes and *p*-values for Feature #2.

Feature 2 <i>m/z</i> 348.07, 2.36 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	6.64	6.77	0.09	5.7	3.2E-06
1C_02	6.86				
1C_03	6.80				
2D_01	1.01	1.19	0.19		
2D_02	1.46				
2D_03	1.11				
3C_01	44.55	45.84	2.82	3.2	7.5E-06
3C_02	43.20				
3C_03	49.76				
4D_01	152.84	147.42	3.89		
4D_02	145.58				
4D_03	143.86				
5C_01	1.40	1.25	0.24	10.5	1.1E-06
5C_02	1.43				
5C_03	0.91				
6D_01	13.42	13.16	0.26		
6D_02	12.81				
6D_03	13.25				

Table C.4. Normalized abundances, fold-changes and *p*-values for Feature #3.

Feature 3 <i>m/z</i> 136.06, 2.36 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	10.41	9.70	0.52	3.9	5.3E-05
1C_02	9.18				
1C_03	9.50				
2D_01	2.44	2.48	0.20		
2D_02	2.26				
2D_03	2.74				
3C_01	46.06	48.43	2.33	2.4	7.3E-06
3C_02	47.64				
3C_03	51.60				
4D_01	116.81	114.16	2.04		
4D_02	113.81				
4D_03	111.86				
5C_01	2.71	2.37	0.33	5.2	4.4E-06
5C_02	2.47				
5C_03	1.93				
6D_01	12.69	12.36	0.25		
6D_02	12.09				
6D_03	12.29				

Table C.5. Normalized abundances, fold-changes and *p*-values for Feature #4.

Feature 4 <i>m/z</i> 308.09, 2.45 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	6.01	4.86	0.81	7.8	1.8E-03
1C_02	4.26				
1C_03	4.33				
2D_01	0.65	0.62	0.02		
2D_02	0.60				
2D_03	0.62				
3C_01	35.38	37.10	1.30	4.8	1.3E-05
3C_02	37.38				
3C_03	38.53				
4D_01	186.32	176.75	7.53		
4D_02	176.02				
4D_03	167.92				
5C_01	8.66	9.12	0.58	4.6	9.9E-06
5C_02	9.94				
5C_03	8.77				
6D_01	42.75	42.05	1.57		
6D_02	43.53				
6D_03	39.88				

Table C.6. Normalized abundances, fold-changes and *p*-values for Feature #5.

Feature 5 <i>m/z</i> 613.01, 3.79 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	27.39	28.29	1.24	20.1	6.8E-06
1C_02	27.43				
1C_03	30.04				
2D_01	1.36	1.41	0.04		
2D_02	1.45				
2D_03	1.42				
3C_01	5.66	5.61	0.12	15.6	1.5E-05
3C_02	5.44				
3C_03	5.74				
4D_01	94.05	87.57	4.66		
4D_02	85.30				
4D_03	83.34				
5C_01	17.34	16.74	0.71	1.0	6.4E-01
5C_02	15.74				
5C_03	17.13				
6D_01	16.07	16.45	0.37		
6D_02	16.33				
6D_03	16.96				

Table C.7. Normalized abundances, fold-changes and *p*-values for Feature #6.

Feature 6 <i>m/z</i> 104.17, 1.49 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	21.28	21.11	0.18	1.3	2.0E-03
1C_02	21.19				
1C_03	20.87				
2D_01	27.71	27.26	1.20		
2D_02	28.45				
2D_03	25.63				
3C_01	34.93	36.19	1.29	2.3	2.3E-05
3C_02	35.69				
3C_03	37.96				
4D_01	84.99	81.55	2.53		
4D_02	80.70				
4D_03	78.97				
5C_01	23.26	24.44	0.85	1.4	1.4E-04
5C_02	24.85				
5C_03	25.21				
6D_01	34.34	34.21	0.48		
6D_02	33.57				
6D_03	34.72				

Table C.8. Normalized abundances, fold-changes and *p*-values for Feature #7.

Feature 7 <i>m/z</i> 705.95, 4.64 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	37.77	35.65	2.76	7.4	1.0E-04
1C_02	31.75				
1C_03	37.44				
2D_01	4.54	4.84	0.50		
2D_02	4.44				
2D_03	5.55				
3C_01	12.85	8.67	3.49	22.1	7.3E-07
3C_02	8.88				
3C_03	4.30				
4D_01	187.43	192.01	3.35		
4D_02	193.29				
4D_03	195.32				
5C_01	38.26	34.42	2.75	2.3	3.0E-05
5C_02	31.97				
5C_03	33.03				
6D_01	79.72	79.76	1.29		
6D_02	81.37				
6D_03	78.20				

Table C.9. Normalized abundances, fold-changes and *p*-values for Feature #8.

Feature 8 <i>m/z</i> 823.32, 4.64 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	30.43	29.82	0.46	6.8	9.7E-07
1C_02	29.31				
1C_03	29.73				
2D_01	4.53	4.40	0.55		
2D_02	3.66				
2D_03	5.00				
3C_01	7.15	5.75	1.89	29.8	2.9E-07
3C_02	7.03				
3C_03	3.08				
4D_01	175.74	171.70	2.93		
4D_02	170.48				
4D_03	168.89				
5C_01	30.50	28.01	1.78	2.5	1.7E-05
5C_02	27.13				
5C_03	26.41				
6D_01	69.81	69.05	1.57		
6D_02	70.48				
6D_03	66.87				

Table C.10. Normalized abundances, fold-changes and *p*-values for Feature #9.

Feature 9 <i>m/z</i> 827.82, 4.58 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	159.75	157.54	8.67	11.8	2.3E-05
1C_02	145.99				
1C_03	166.89				
2D_01	12.75	13.37	2.43		
2D_02	10.76				
2D_03	16.60				
3C_01	38.04	25.60	13.36	10.4	1.5E-05
3C_02	31.69				
3C_03	7.06				
4D_01	268.45	266.03	2.25		
4D_02	263.04				
4D_03	266.60				
5C_01	130.72	123.40	6.12	2.0	1.7E-05
5C_02	115.75				
5C_03	123.75				
6D_01	250.24	245.34	3.69		
6D_02	241.32				
6D_03	244.45				

Table C.11. Normalized abundances, fold-changes and *p*-values for Feature #10.

Feature 10 <i>m/z</i> 152.11, 2.30 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	12.37	10.96	1.01	4.4	3.4E-04
1C_02	10.01				
1C_03	10.51				
2D_01	2.08	2.49	0.29		
2D_02	2.67				
2D_03	2.72				
3C_01	4.54	4.92	0.27	17.1	4.9E-06
3C_02	5.05				
3C_03	5.16				
4D_01	88.62	83.88	3.35		
4D_02	81.48				
4D_03	81.53				
5C_01	18.37	18.77	0.36	3.6	3.9E-07
5C_02	19.25				
5C_03	18.70				
6D_01	65.54	66.73	1.02		
6D_02	68.03				
6D_03	66.62				

Table C.12. Normalized abundances, fold-changes and *p*-values for Feature #11.

Feature 11 <i>m/z</i> 184.11, 1.53 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	6.14	5.82	0.23	3.5	6.2E-05
1C_02	5.74				
1C_03	5.59				
2D_01	1.98	1.65	0.25		
2D_02	1.58				
2D_03	1.39				
3C_01	18.82	20.03	0.89	3.7	3.3E-06
3C_02	20.35				
3C_03	20.93				
4D_01	76.97	74.93	1.93		
4D_02	75.48				
4D_03	72.35				
5C_01	5.75	6.08	0.56	15.8	1.3E-07
5C_02	6.87				
5C_03	5.63				
6D_01	98.02	96.06	1.44		
6D_02	95.57				
6D_03	94.59				

Table C.13. Normalized abundances, fold-changes and *p*-values for Feature #12.

Feature 12 <i>m/z</i> 137.10, 3.94 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	160.13	148.58	13.20	1.3	3.3E-02
1C_02	130.10				
1C_03	155.52				
2D_01	206.12	189.00	12.16		
2D_02	181.94				
2D_03	178.94				
3C_01	99.71	106.03	4.81	2.0	2.2E-04
3C_02	106.99				
3C_03	111.38				
4D_01	218.17	206.81	10.10		
4D_02	208.63				
4D_03	193.64				
5C_01	148.78	139.06	8.24	1.5	3.1E-04
5C_02	139.75				
5C_03	128.65				
6D_01	206.67	209.79	2.36		
6D_02	212.36				
6D_03	210.33				

Table C.14. Normalized abundances, fold-changes and *p*-values for Feature #13.

Feature 13 <i>m/z</i> 246.19, 4.99 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	3.69	3.62	0.27	1.9	6.9E-04
1C_02	3.26				
1C_03	3.91				
2D_01	6.63	6.80	0.39		
2D_02	7.34				
2D_03	6.43				
3C_01	5.64	5.05	0.76	1.1	4.2E-01
3C_02	3.97				
3C_03	5.54				
4D_01	4.67	4.55	0.19		
4D_02	4.29				
4D_03	4.70				
5C_01	3.18	3.10	0.22	67.8	4.6E-09
5C_02	3.32				
5C_03	2.79				
6D_01	209.56	209.85	1.52		
6D_02	208.15				
6D_03	211.83				

Table C.15. Normalized abundances, fold-changes and *p*-values for Feature #14.

Feature 14 m/z 137.10, 2.68 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	72.95	77.83	3.72	3.3	1.8E-05
1C_02	81.97				
1C_03	78.58				
2D_01	254.93	255.64	9.81		
2D_02	268.00				
2D_03	243.99				
3C_01	71.24	75.87	5.24	3.5	1.3E-06
3C_02	73.19				
3C_03	83.20				
4D_01	265.46	265.24	2.57		
4D_02	268.27				
4D_03	261.99				
5C_01	114.19	116.44	2.13	2.6	5.3E-06
5C_02	119.30				
5C_03	115.82				
6D_01	308.29	297.68	7.59		
6D_02	293.77				
6D_03	290.96				

Table C.16. Normalized abundances, fold-changes and *p*-values for Feature #15.

Feature 15 <i>m/z</i> 152.11, 3.93 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	34.02	31.56	3.44	2.1	4.66E-04
1C_02	26.69				
1C_03	33.96				
2D_01	70.91	66.60	3.24		
2D_02	65.79				
2D_03	63.11				
3C_01	7.31	7.61	0.57	5.7	4.31E-05
3C_02	7.11				
3C_03	8.41				
4D_01	46.82	43.29	2.56		
4D_02	42.24				
4D_03	40.82				
5C_01	41.50	39.44	1.91	1.7	1.77E-04
5C_02	39.91				
5C_03	36.89				
6D_01	65.51	67.29	2.22		
6D_02	70.43				
6D_03	65.94				

Table C.17. Normalized abundances, fold-changes and *p*-values for Feature #16.

Feature 16 <i>m/z</i> 203.09, 1.48 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	87.57	84.77	3.62	1.0	3.1E-01
1C_02	79.66				
1C_03	87.08				
2D_01	84.01	80.88	3.01		
2D_02	81.81				
2D_03	76.83				
3C_01	71.68	73.02	1.78	7.3	9.8E-07
3C_02	71.84				
3C_03	75.53				
4D_01	10.02	9.97	0.23		
4D_02	9.67				
4D_03	10.22				
5C_01	39.69	40.12	1.25	3.5	7.7E-06
5C_02	41.82				
5C_03	38.84				
6D_01	12.17	11.39	0.56		
6D_02	10.88				
6D_03	11.12				

Table C.18. Normalized abundances, fold-changes and *p*-values for Feature #17.

Feature 17 <i>m/z</i> 437.13, 9.31 min					
	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	119.37	111.95	5.80	1.1	7.1E-02
1C_02	111.26				
1C_03	105.22				
2D_01	96.68	99.06	4.72		
2D_02	105.65				
2D_03	94.86				
3C_01	125.05	119.78	4.68	3.2	5.3E-05
3C_02	113.68				
3C_03	120.62				
4D_01	43.04	36.98	4.40		
4D_02	32.73				
4D_03	35.18				
5C_01	99.13	85.94	9.54	1.9	5.1E-03
5C_02	81.75				
5C_03	76.93				
6D_01	50.84	46.24	3.26		
6D_02	44.27				
6D_03	43.61				

Table C.19. Normalized abundances, fold-changes and *p*-values for Feature #18.

Feature 18 m/z 387.15					
	Norm. Int.	Average	Standard Dev.	Fold- change	<i>p</i>-value
1C_01	86.43	86.41	0.58	1.0	7.2E-01
1C_02	87.11				
1C_03	85.69				
2D_01	82.64	87.65	4.54		
2D_02	93.62				
2D_03	86.69				
3C_01	120.37	130.88	9.03	7.0	6.6E-05
3C_02	129.84				
3C_03	142.42				
4D_01	16.64	18.62	1.65		
4D_02	18.53				
4D_03	20.67				
5C_01	62.81	60.72	1.49	2.1	8.5E-06
5C_02	59.46				
5C_03	59.89				
6D_01	29.30	29.61	0.31		
6D_02	30.03				
6D_03	29.50				

C.6. Expression of Thymosins β 4 and β 10 by Peak Areas for 8+ Species

The region containing thymosins β 4 and β 10 was mobility-extracted from the IM-MS spectra of each sample (18 total) using a defined selection rule similar to that shown in Figure 3.1. Total abundances of the 6+ charge states of thymosins β 4 and β 10 were determined by summing the peak areas in the regions of the mobility-extracted mass spectrum centered about the isotopic distribution for the 8+ species, m/z 827.5-829.5 and m/z 823.0-825.0, respectively. All comparisons are made between matched disease and control pairs. Fold-changes were calculated by dividing the larger of the two tissues (control or diabetic) by the smaller to obtain numbers ≥ 1 . The student's t-test with two-tails, equal variance, and $\alpha=0.05$ was used to calculate p -values.

Table C.20. Abundances of Thymosin β 4 (6+) from IM-Extracted MS Peak Areas.

	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01 1C_02 1C_03	206039 208877 198867	204594	4212	3.0	4.5E-05
2D_01 2D_02 2D_03	60145 61216 80429	67263	9320		
3C_01 3C_02 3C_03	60755 47597 16387	41580	18606	38.1	7.0E-07
4D_01 4D_02 4D_03	1631322 1544283 1579135	1584913	35768		
5C_01 5C_02 5C_03	346773 308806 283154	312911	26134	3.8	3.5E-06
6D_01 6D_02 6D_03	1217249 1184695 1163585	1188510	22074		

Table C.21. Abundances of Thymosin β 10 (6+) from IM-Extracted MS Peak Areas.

	Norm. Int.	Average	Standard Dev.	Fold-change	<i>p</i>-value
1C_01	45840	45745	2470	1.3	3.6E-03
1C_02	48722				
1C_03	42673				
2D_01	34868	35020	159		
2D_02	35239				
2D_03	34954				
3C_01	16016	13416	3401	76.0	1.4E-06
3C_02	15619				
3C_03	8612				
4D_01	1062637	1019036	30959		
4D_02	1000684				
4D_03	993786				
5C_01	96636	80048	12706	4.8	8.2E-06
5C_02	77735				
5C_03	65772				
6D_01	389273	383980	7401		
6D_02	389153				
6D_03	373514				

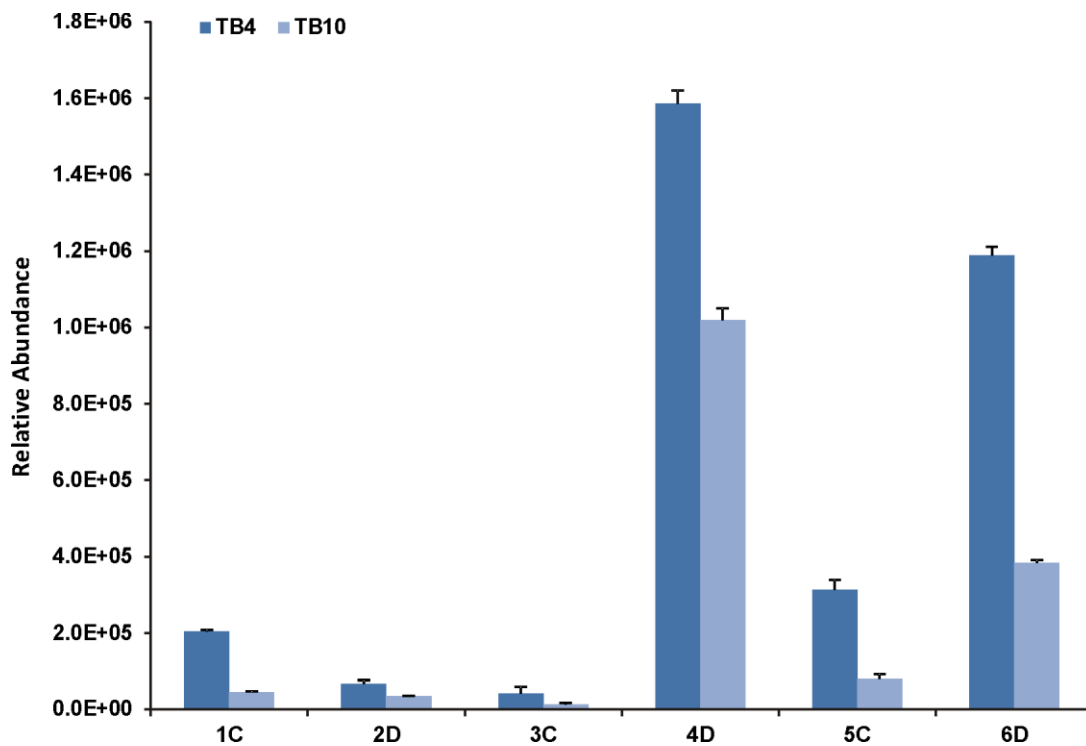


Figure C.3. Expression of thymosins $\beta 4$ and $\beta 10$ across the breast tissue samples. Relative abundances were calculated from the peaks areas of the isotopic distribution for the 6+ charge state of thymosins $\beta 4$ and $\beta 10$.

B.7. MS/MS Spectra for Features in Table 3.1

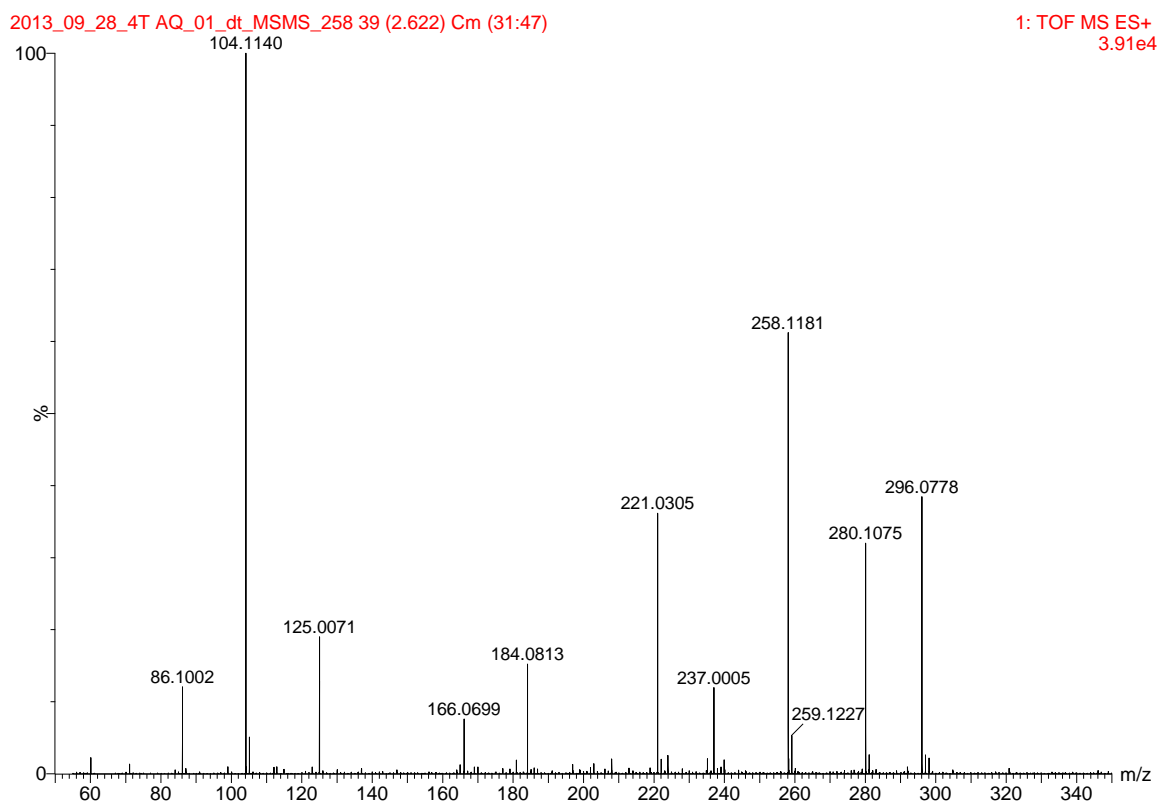


Figure C.4. Mobility- and chromatography-extracted (2.62 *ms*; 1.51 min) DIA MS/MS spectrum of Feature #1, glycerophosphocholine (*m/z* 258.13).

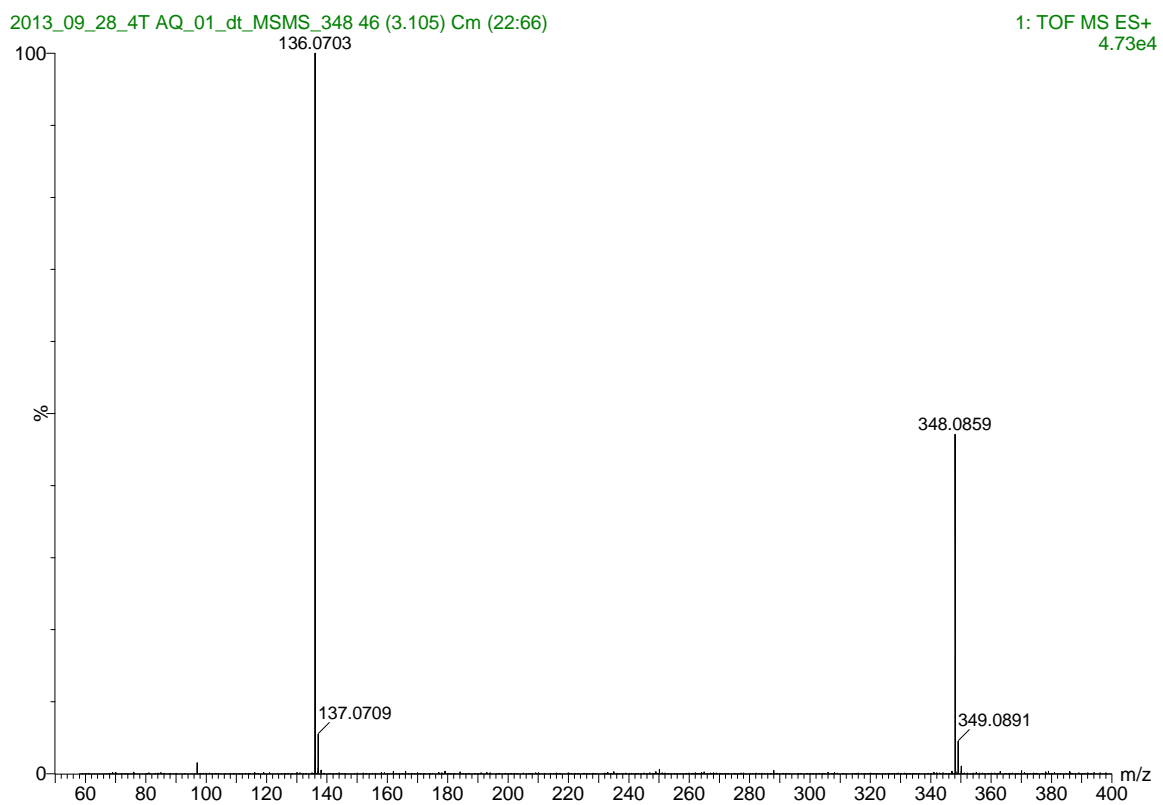


Figure C.5. Mobility- and chromatography-extracted (3.11 *ms*; 2.33 min) DIA MS/MS spectrum of Feature #2, adenosine monophosphate (m/z 348.09).

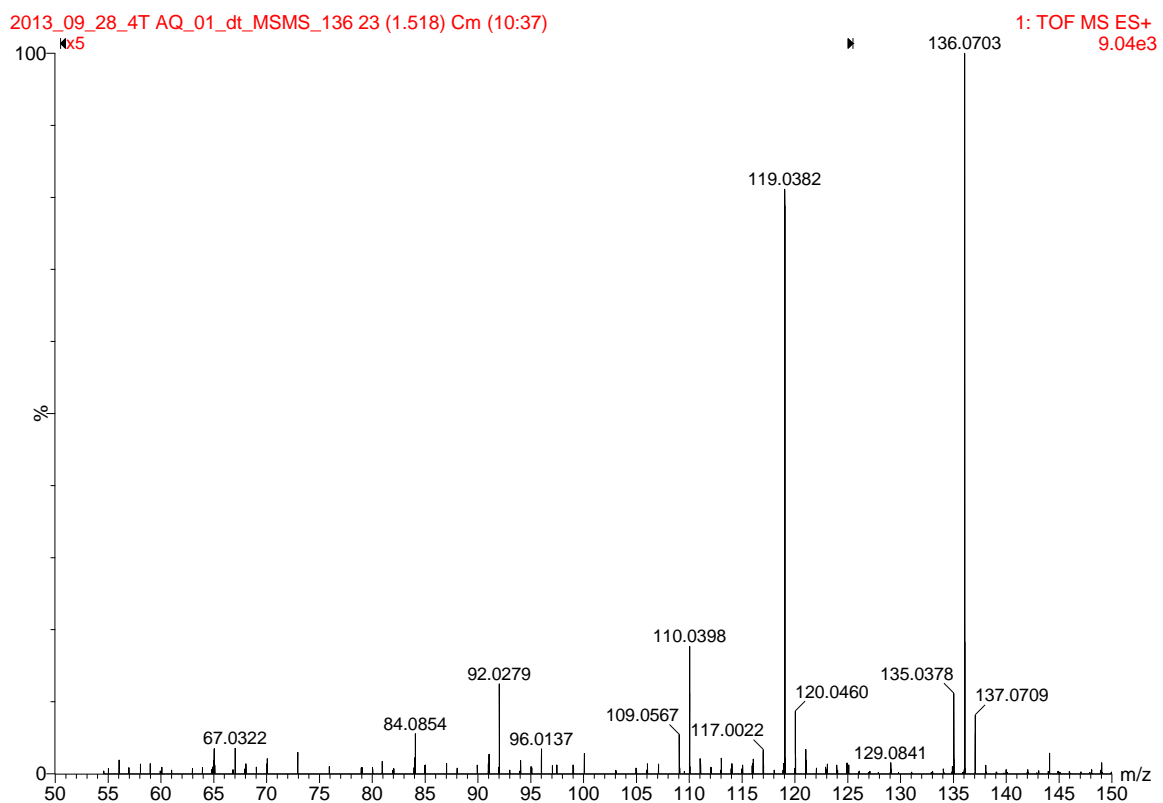


Figure C.6. Mobility- and chromatography-extracted (1.52 *ms*; 2.33) DIA MS/MS spectrum of Feature #3, adenine/fragment of adenosine monophosphate (*m/z* 136.07). The region *m/z* 50-125 is shown at 5X magnification for clarity.

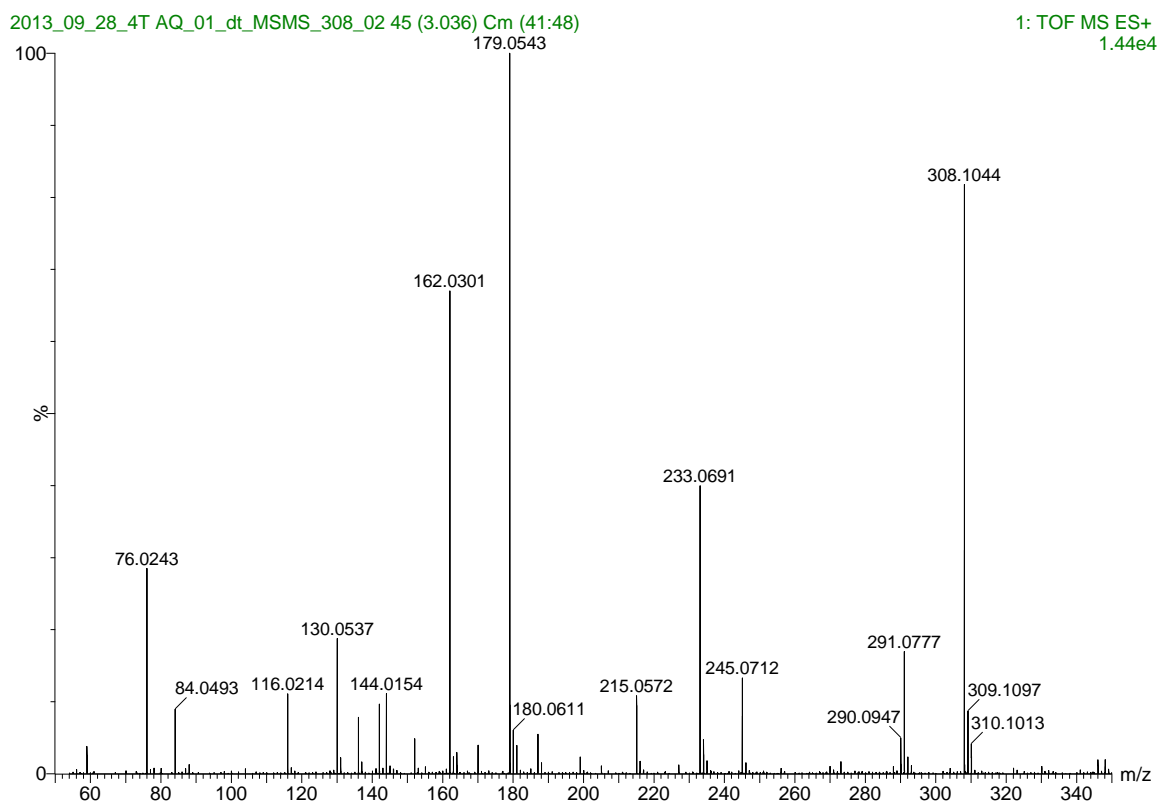


Figure C.7. Mobility- and chromatography-extracted (3.04 *ms*; 2.45 min) DIA MS/MS spectrum of Feature #4, glutathione (*m/z* 308.10).

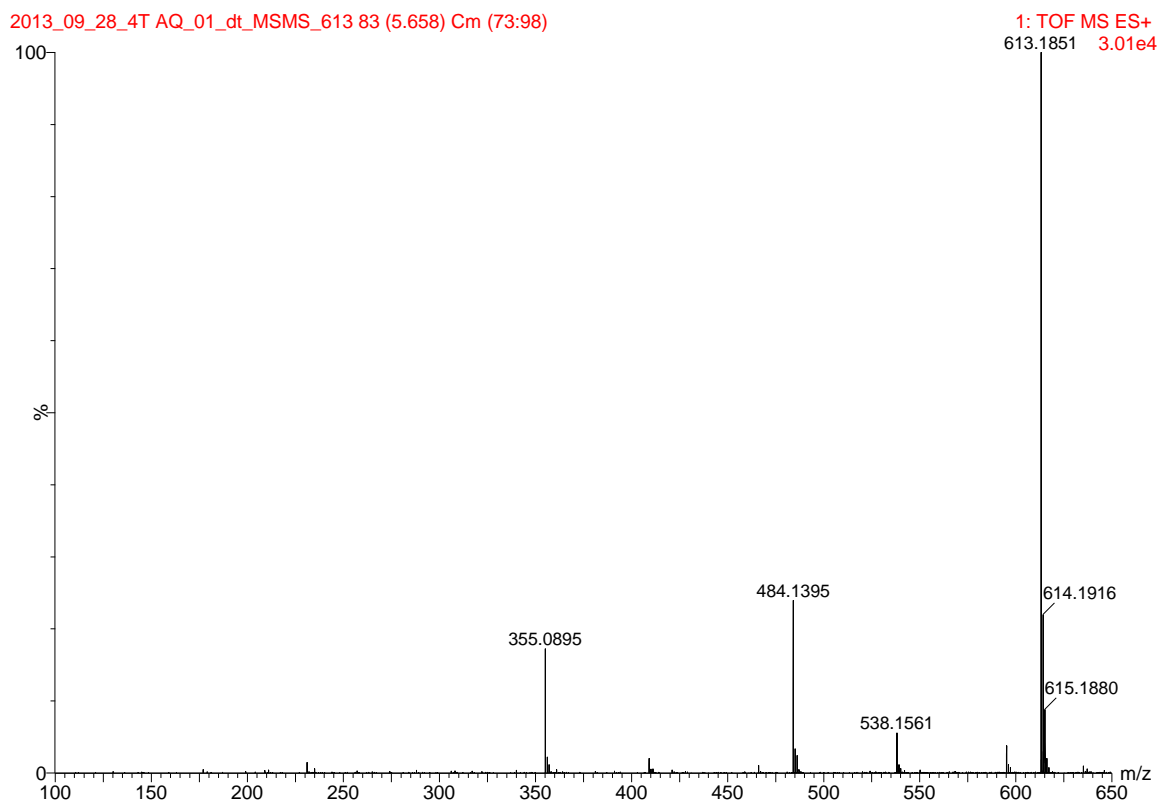


Figure C.8. Mobility- and chromatography-extracted (5.66 ms; 3.76 min) DIA MS/MS spectrum of Feature #5, oxidized glutathione (m/z 613.19).

2013_09_28_4T AQ_01_dt_MSMS_thyosins 705 709 70 (4.761) Cm (58:92)

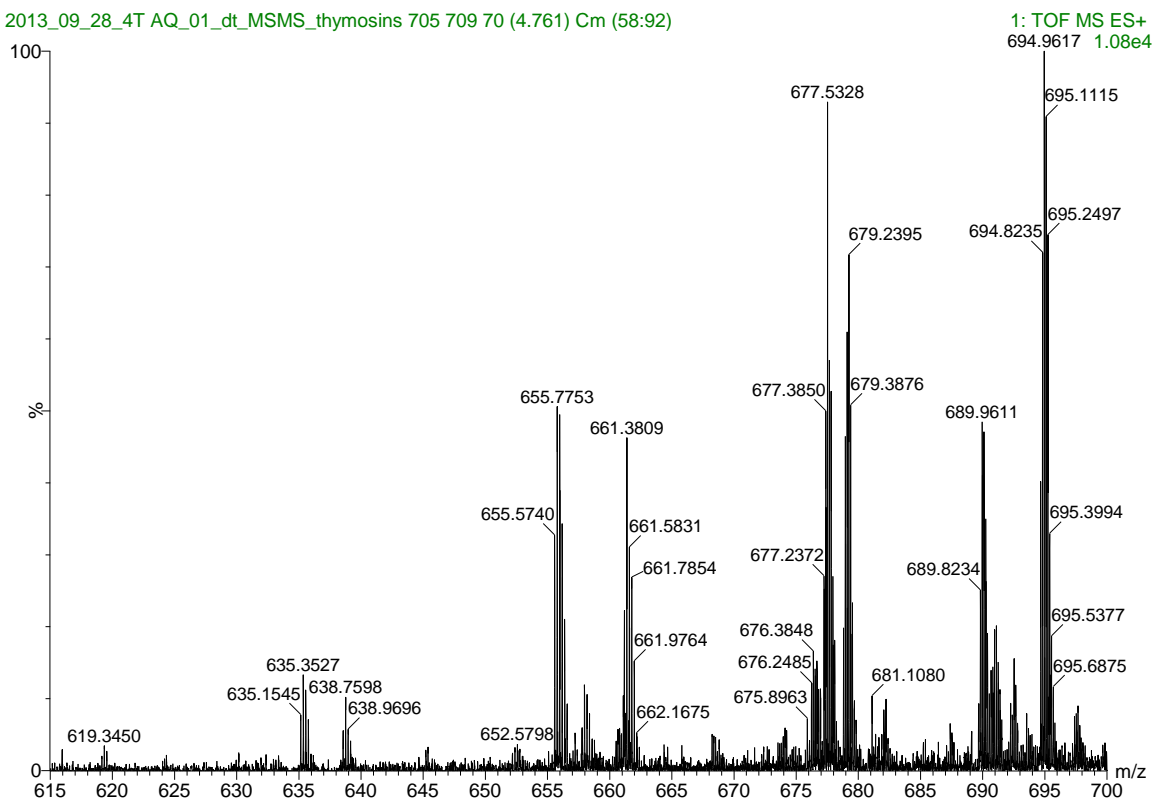


Figure C.9. Mobility- and chromatography-extracted (4.76 ms; 4.56 min) DIA MS/MS spectrum of Feature #7-9, thymosins β 4 and β 10 (m/z 705.94, 823.44, 827.76).

2013_09_28_4T_AQ_01_dt_MSMS_thymosins 705 709 70 (4.761) Cm (58:92)

1: TOF MS ES+
2.34e4

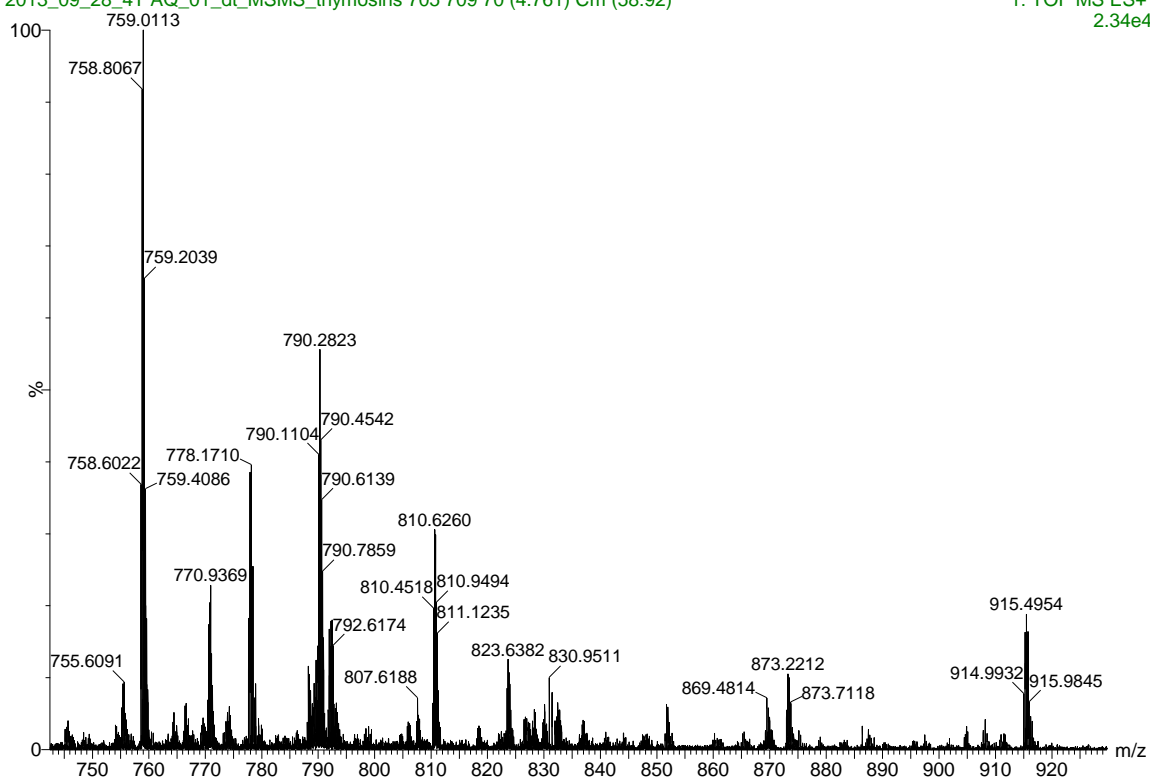


Figure C.10. Mobility- and chromatography-extracted (4.76 ms; 4.56 min) DIA MS/MS spectrum of Features #7-9, thymosins β 4 and β 10 (m/z 705.94, 823.44, 827.76).

Table C.21. Backbone cleavages observed for Thymosin β 4 and β 10

Thymosin β 4 Fragmentation			Thymosin β 10 Fragmentation		
<i>m/z</i>	charge	fragment	<i>m/z</i>	charge	fragment
619.1161	8+	MH-H ₂ O ⁺⁸	635.1545	5+	y ₂₇ -NH ₃ ⁺⁵
635.3738	5+	b ₂₇ -H ₂ O ⁺⁵	638.5611	5+	y ₂₇ ⁺⁵
655.9865	5+	a ₂₈ ⁺⁵	660.7968	5+	y ₂₈ -NH ₃ ⁺⁵
661.5927	5+	b ₂₈ ⁺⁵	677.2372	7+	b ₄₁ +H ₂ O ⁺⁷
676.5419	7+	b ₄₁ ⁺⁷	689.6857	7+	y ₄₂ ⁺⁷
679.2603	7+	b ₄₁ +H ₂ O ⁺⁷			
694.9823	7+	b ₄₂ ⁺⁷			
745.4326	4+	b ₂₅ ⁺⁴			
773.9556	4+	b ₂₆ ⁺⁴			
792.1327	6+	b ₄₁ +H ₂ O ⁺⁶			
810.8069	6+	b ₄₂ ⁺⁶			

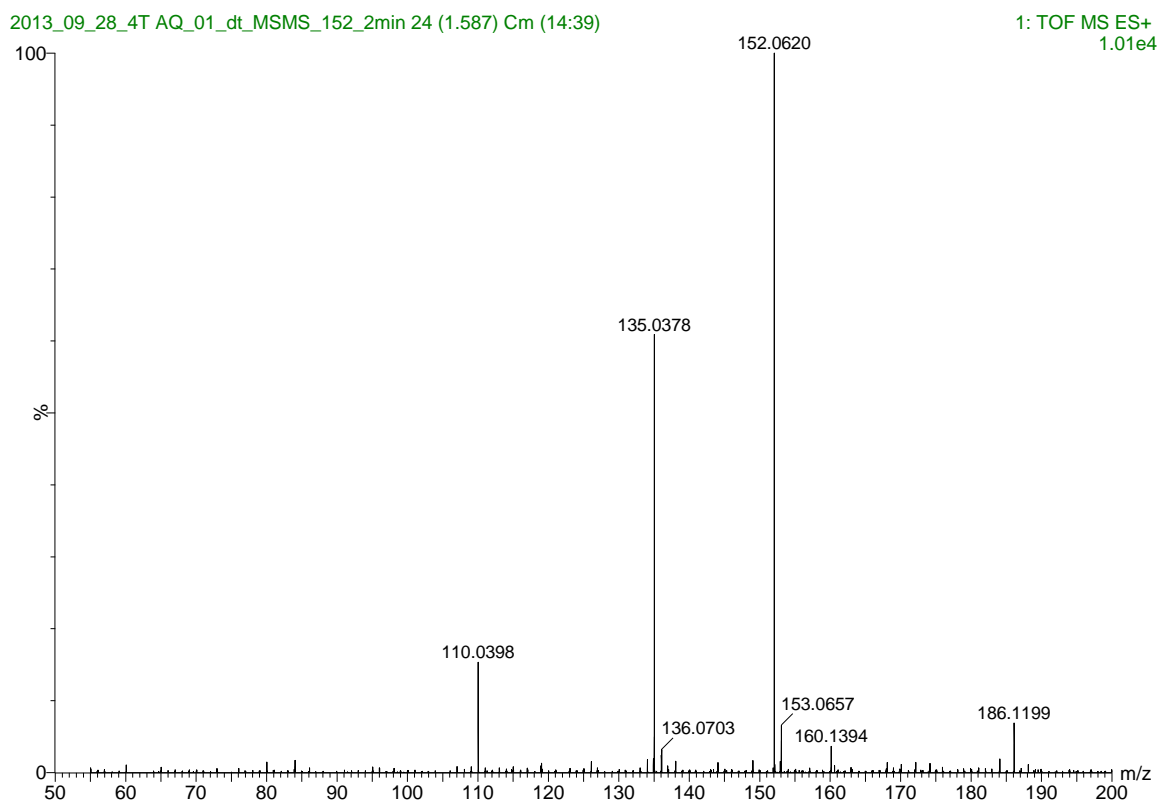


Figure C.11. Mobility- and chromatography-extracted (1.59 *ms*; 2.26 min) DIA MS/MS spectrum of Feature #10 (*m/z* 152.06). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: 2-hydroxyadenine (C₅H₅N₅O; cLogP: 0.56); 8-hydroxyadenine (C₅H₅N₅O; cLogP: 0.48); and guanine (C₅H₅N₅O; cLogP: -1.16). Also observed at 3.93 min in chromatogram, suggesting this particular species is guanine due to CLogP values.

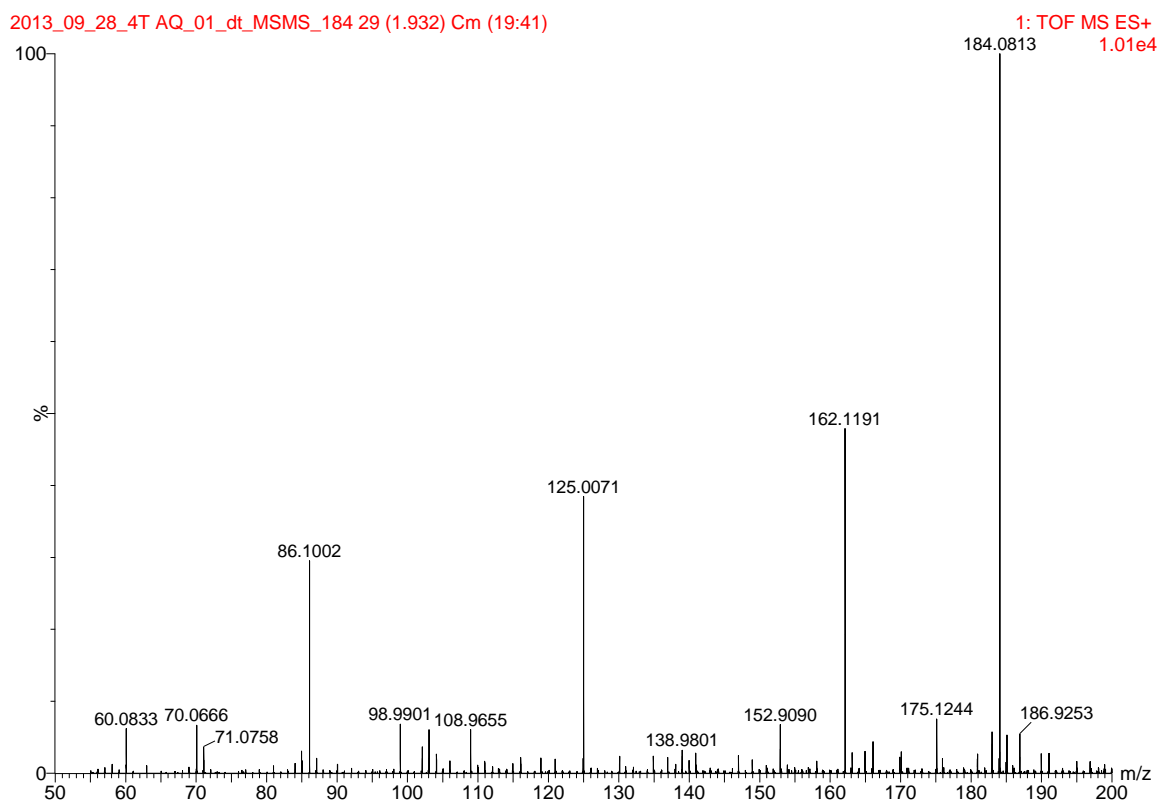


Figure C.11. Mobility- and chromatography-extracted (1.93 ms; 1.51 min) DIA MS/MS spectrum of Feature #11, phosphocholine (m/z 184.08).

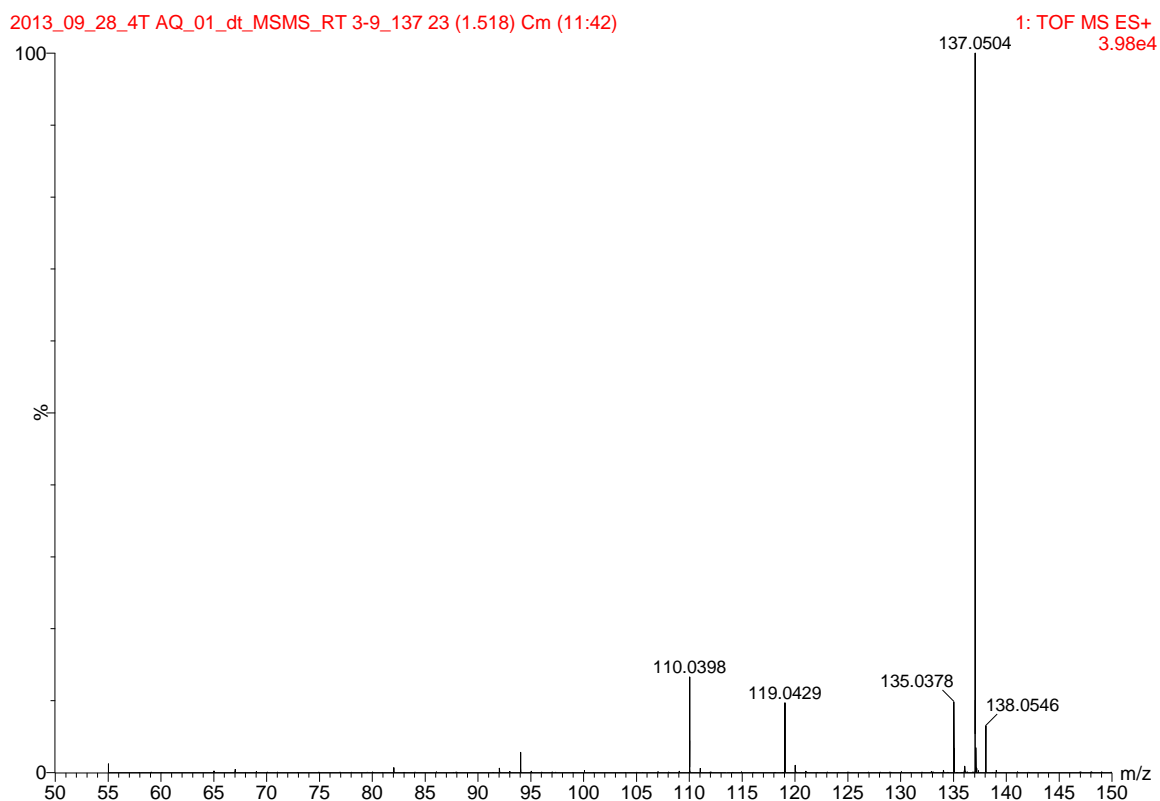


Figure C.12. Mobility- and chromatography-extracted (1.52 *ms*; 3.93 min) DIA MS/MS spectrum of Feature #12 (*m/z* 137.05). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: hypoxanthine (C₅H₄N₄O; CLogP: 0.49); allopurinol (C₅H₄N₄O; CLogP: -0.74). Also observed at 2.67 min in chromatogram, suggesting this particular species is hypoxanthine due to CLogP values. Signals at *m/z* 137, 119, 110 and 94 correspond to Feature #12, while *m/z* 135 and 110 correspond to feature #15 (*m/z* 152).

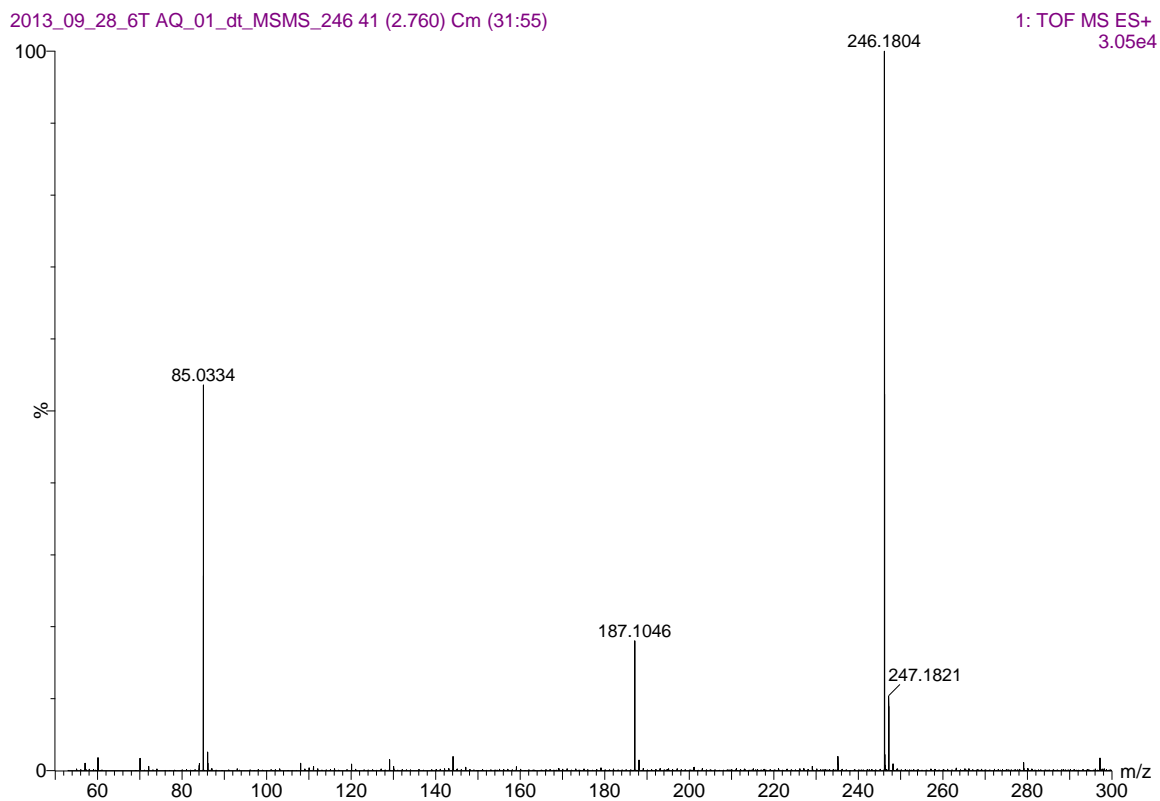


Figure C.13. Mobility- and chromatography-extracted (2.76 *ms*; 5.01 min) DIA MS/MS spectrum of Feature #13 (*m/z* 246.18). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: 2-methylbutyrylcarnitine ($C_{12}H_{23}NO_4$); pivaloylcarnitine ($C_{12}H_{23}NO_4$).

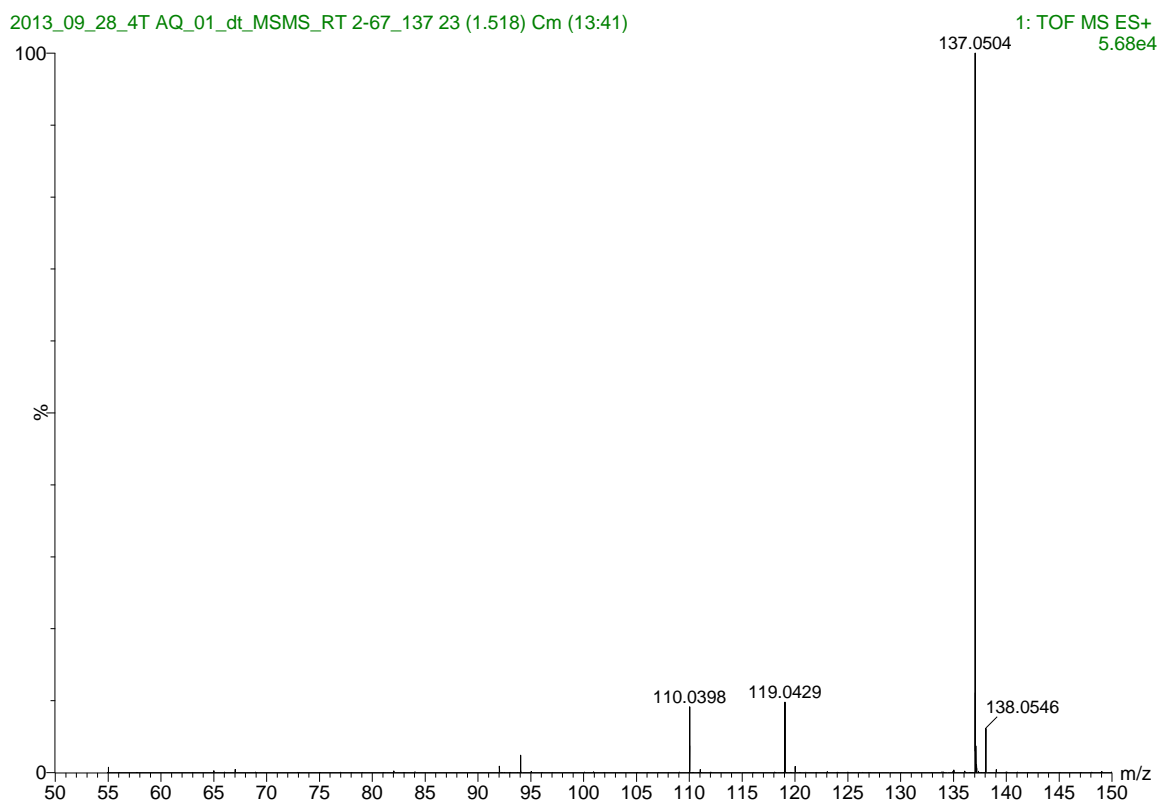


Figure C.14. Mobility- and chromatography-extracted (1.52 *ms*; 2.67 min) DIA MS/MS spectrum of Feature #14 (*m/z* 137.05). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: hypoxanthine (C₅H₄N₄O; CLogP: 0.49); allopurinol (C₅H₄N₄O; CLogP: -0.74). Also observed at 3.93 min in chromatogram, suggesting this particular species is allopurinol due to CLogP values.

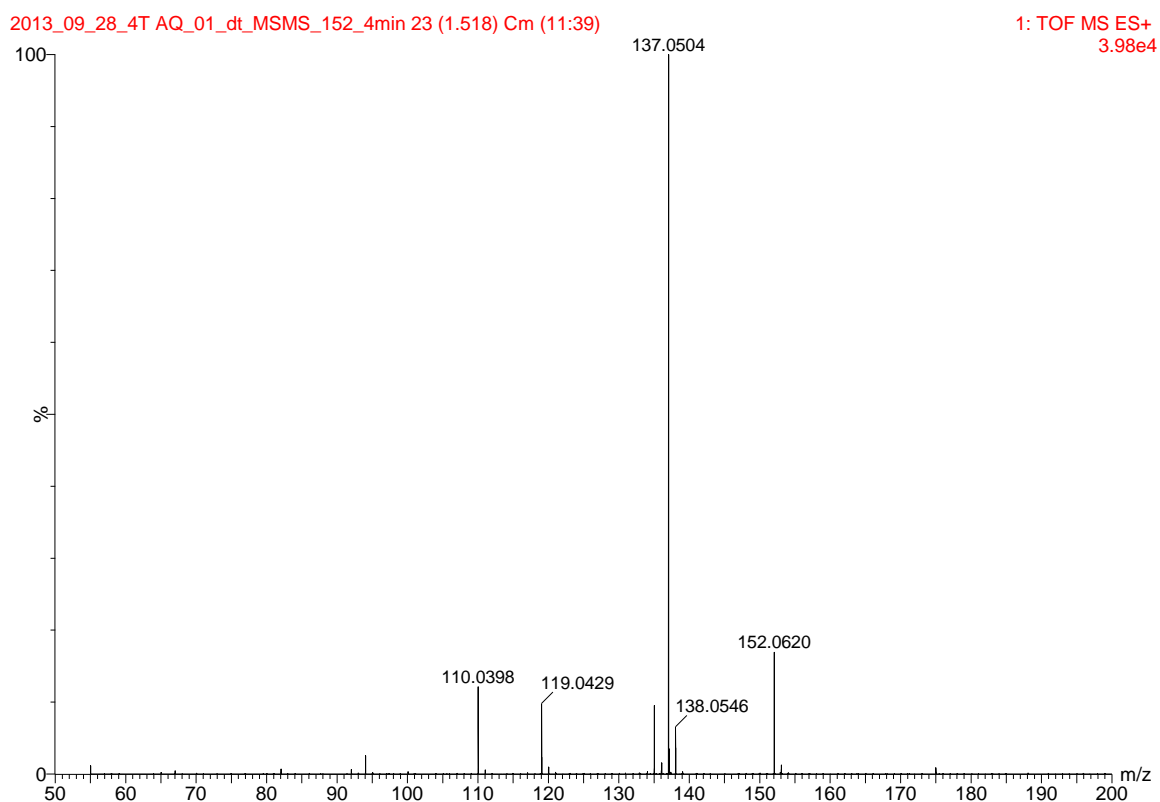


Figure C.15. Mobility- and chromatography-extracted (1.52 *ms*; 3.93 min) DIA MS/MS spectrum of Feature #15 (*m/z* 152.06). Fragmentation pattern and mass accuracy suggest the following molecules as potential identifications: 2-hydroxyadenine ($C_5H_5N_5O$; cLogP: 0.56); 8-hydroxyadenine ($C_5H_5N_5O$; cLogP: 0.48); and guanine ($C_5H_5N_5O$; cLogP: -1.16). Also observed at 2.67 min in chromatogram, suggesting this particular species is hydroxyadenine due to CLogP values. Signals at *m/z* 137, 119, 110 and 94 correspond to Feature #12, while *m/z* 135 and 110 correspond to feature #15.

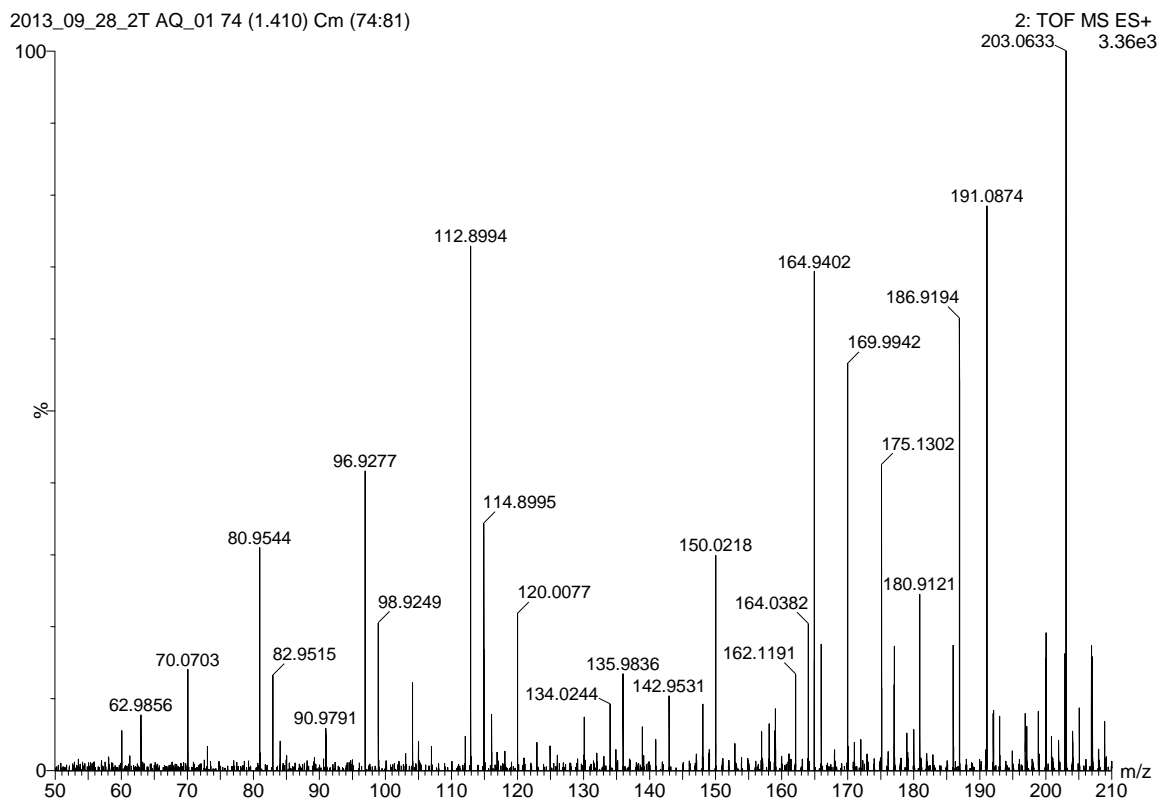


Figure C.16. DIA MS/MS spectrum of Feature #16, a sodiated monosaccharide (m/z 203.05) at 1.5 min.

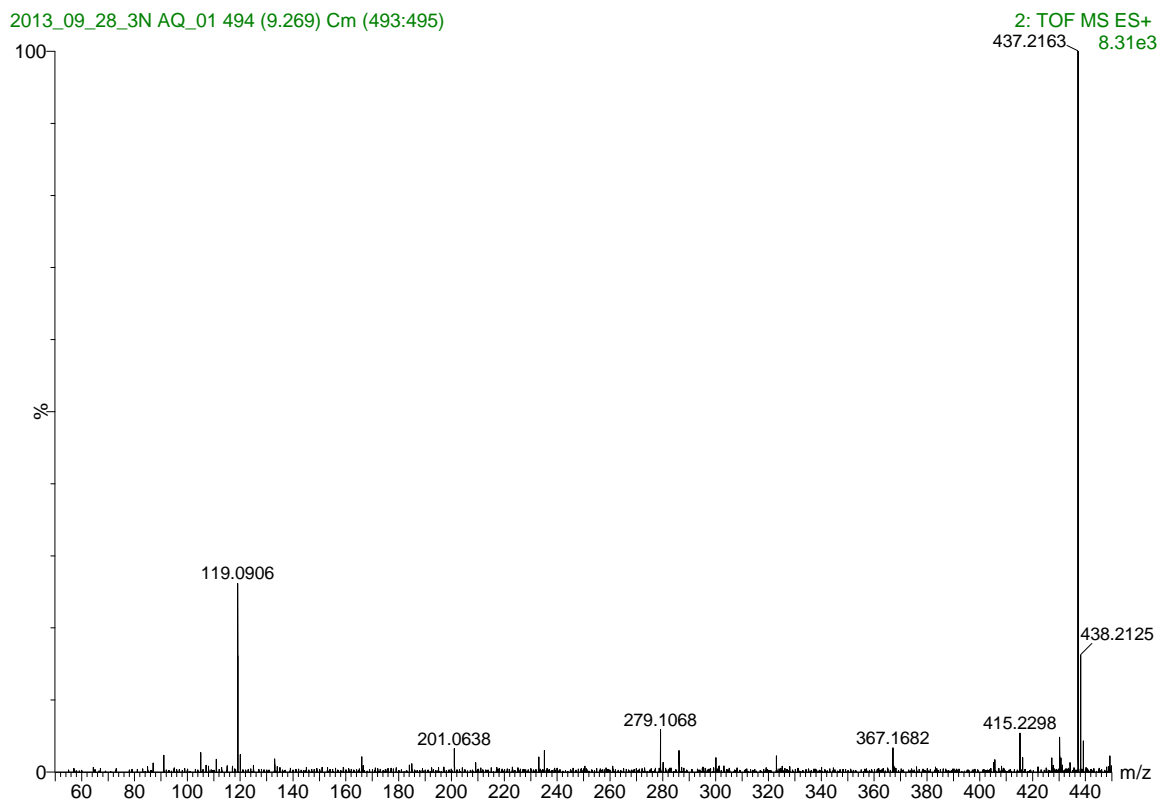


Figure C.17. DIA MS/MS spectrum of Feature #17 (m/z 437.19) at 9.31 min.

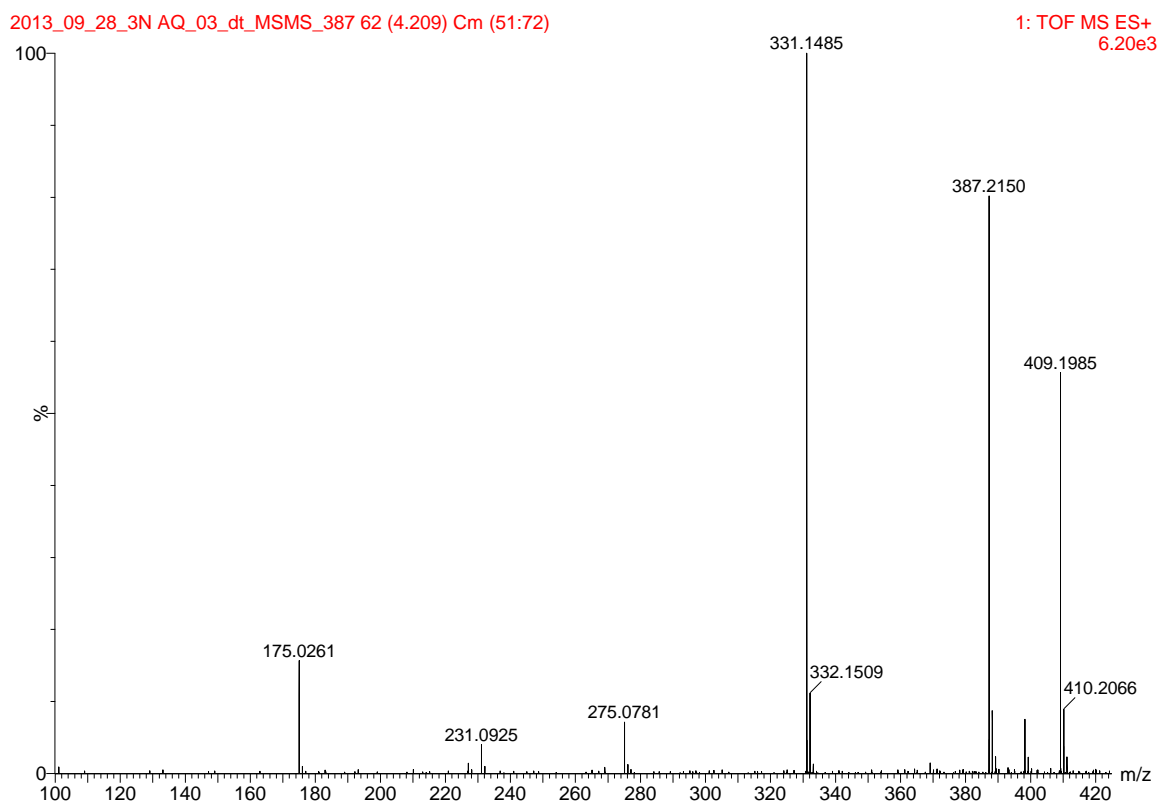


Figure C.18. Mobility- and chromatography-extracted (4.21 *ms*; 9.14 min) DIA MS/MS spectrum of Feature #18, phosphocholine (*m/z* 387.22).

C.8. MS/MS Spectra for Standards

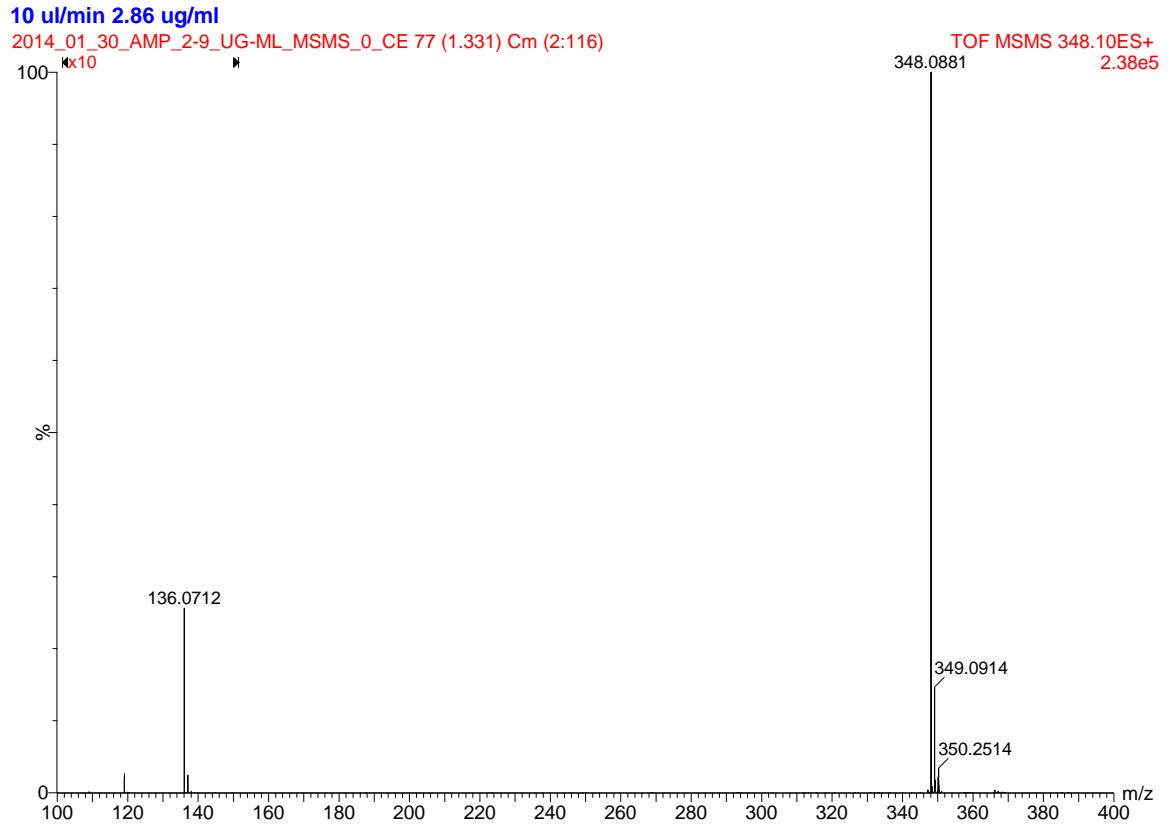


Figure C.19. MS/MS spectrum for 2.86 $\mu\text{g/ml}$ standard of adenosine 5'-monophosphate with 0 V collision energy. The area from m/z 100-150 is shown at 10X magnification for clarity.

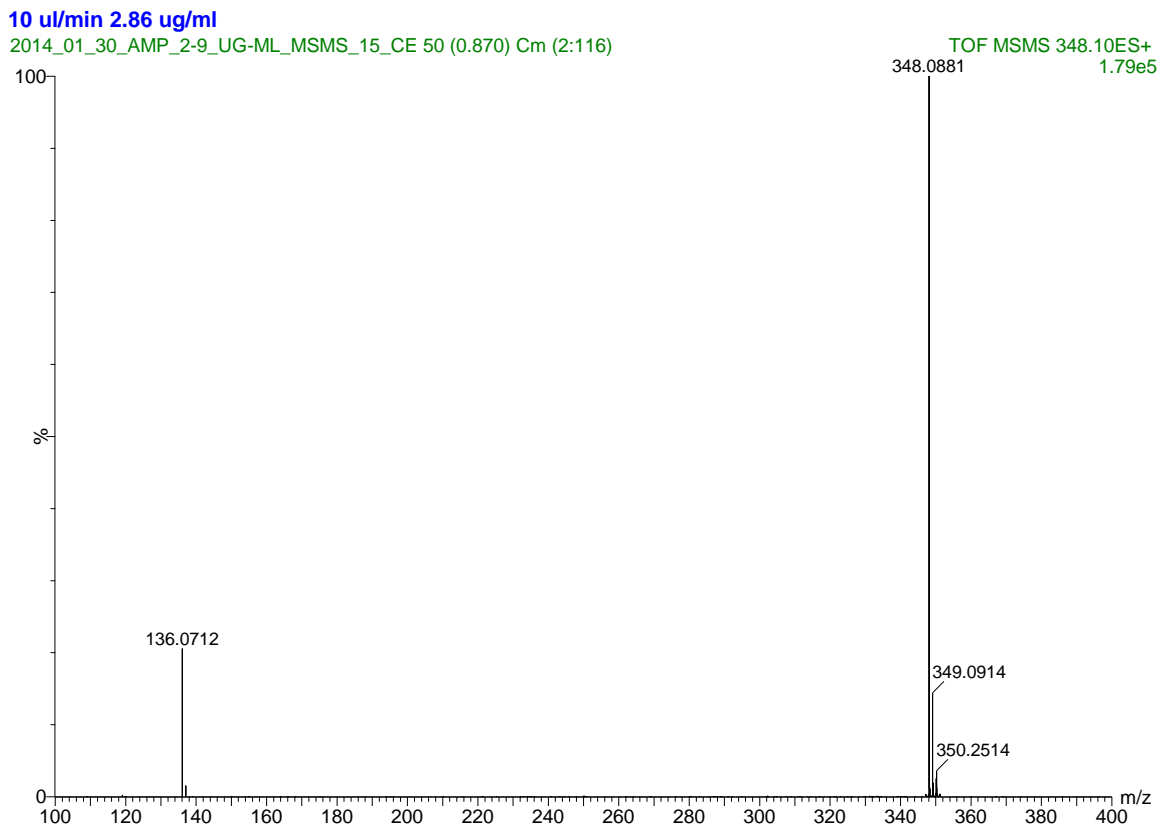


Figure C.20. MS/MS spectrum for 2.86 $\mu\text{g/ml}$ standard of adenosine 5'-monosphosphate (m/z 348.09) with 15 V collision energy.

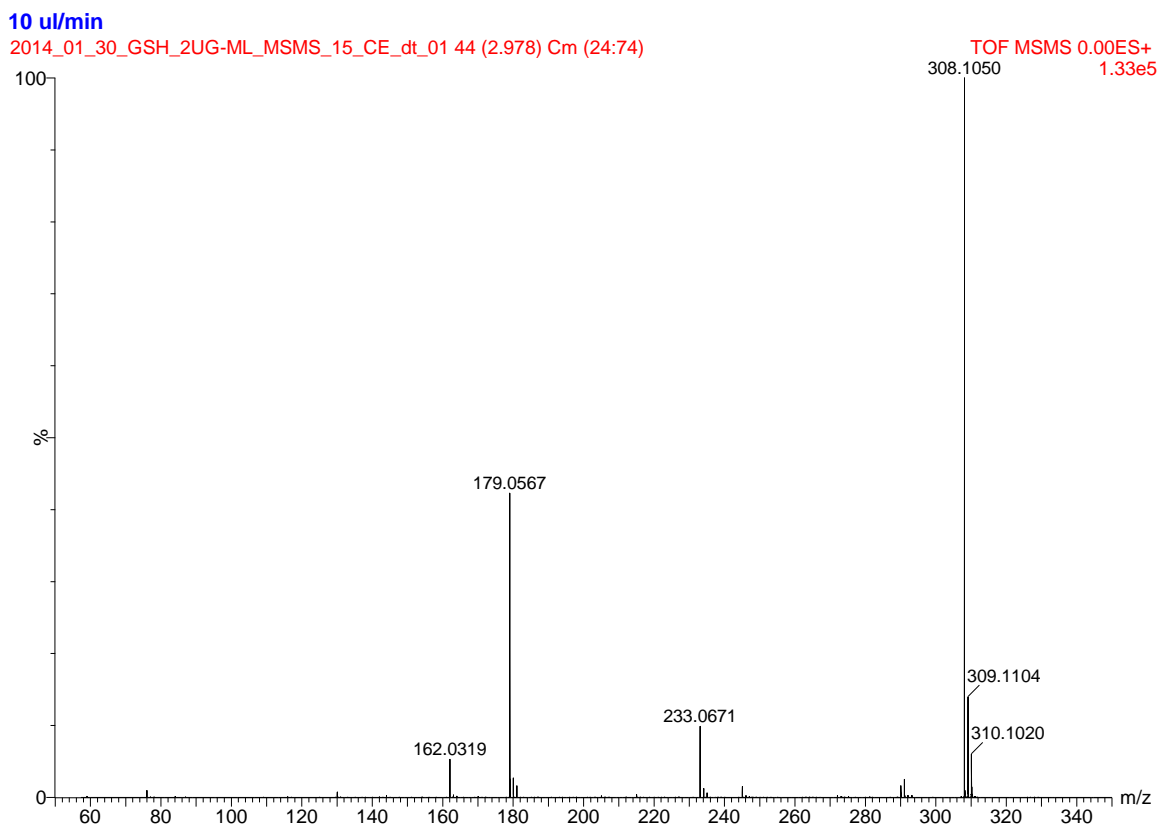


Figure C.21. Mobility-extracted (2.98 *ms*) MS/MS spectrum for 2.12 $\mu\text{g/ml}$ standard of reduced glutathione (m/z 308.11) with 15 V collision energy.

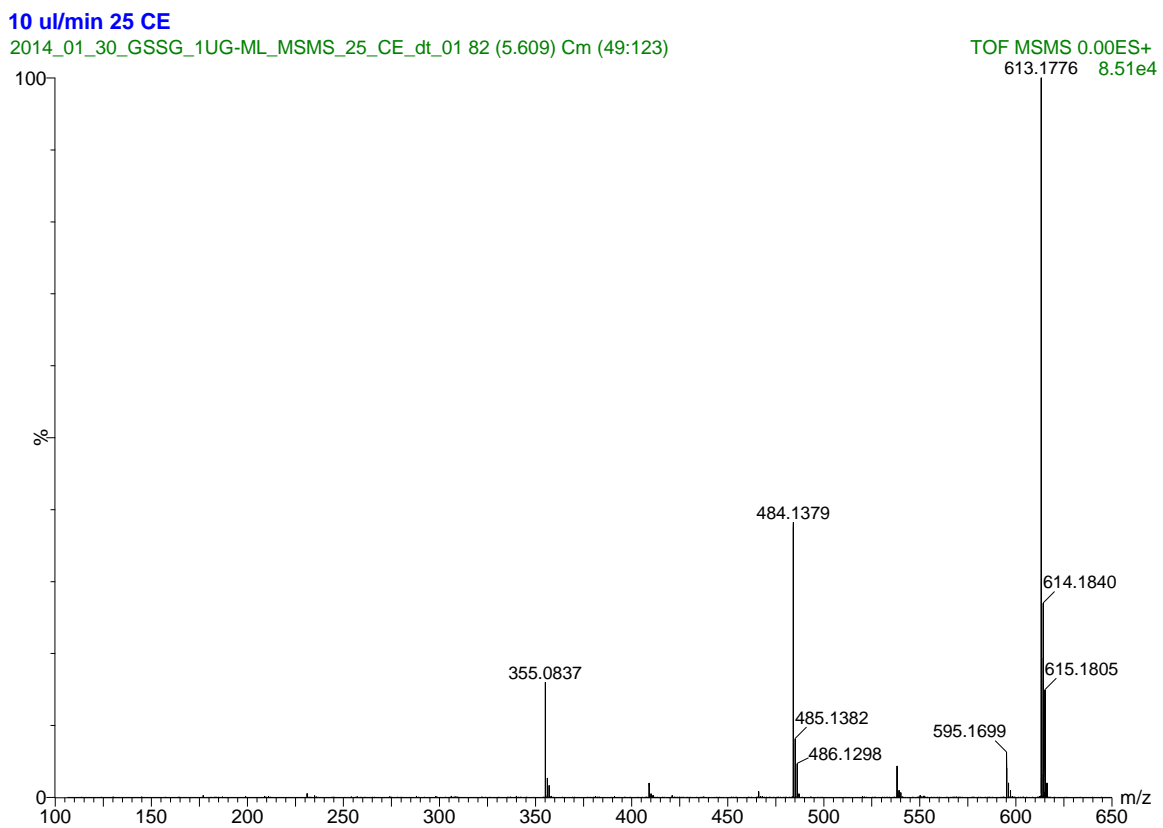


Figure C.22. Mobility-extracted (5.61 ms) MS/MS spectrum for 1.1 $\mu\text{g/ml}$ standard of oxidized glutathione (m/z 613.18) with 25 V collision energy.

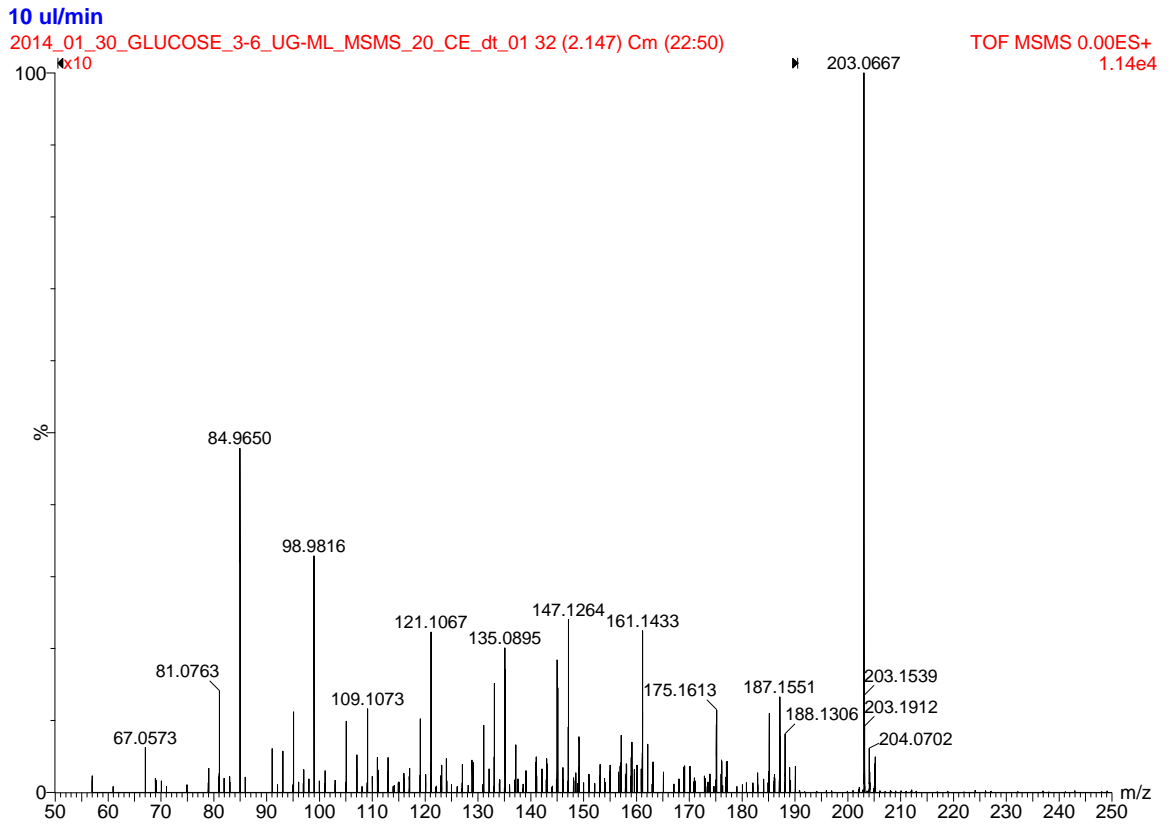


Figure C.23. Mobility-extracted (2.15 ms) MS/MS spectrum for 3.6 $\mu\text{g/ml}$ standard of $\alpha\text{-D-glucose}$ (m/z 203.07) with 20 V collision energy. The region m/z 50-190 is shown at 10X magnification for clarity.

APPENDIX D

SUPPORTING INFORMATION FOR CHAPTER IV

D.1. Supplementary Information for Mass Spectrometry Data Acquisition According to MIAPE-MS Format

1. General Features

1.1. Global Descriptors

- i. Dates on which data was acquired, 01_24_12 and 08_13_12, may be found in the file names in the following format:
“sh_2514_JF_012412_” and “sh_2796_KH_081312_”

- ii. The samples were processed and analyzed by the Vanderbilt Proteomic Core

Contact Info: Kristie L. Rose, Ph.D. - Co-Associate Director

Email: kristie.rose@vanderbilt.edu

Phone: (615) 343-1568

Contact Info: Salisha Hill, M.S. – Sr. Research Specialist

Email: salisha.hill@vanderbilt.edu

Phone: (615) 343-2739

- iii. Instrument Manufacturer and Model

LTQ-Orbitrap: Thermo Scientific LTQ Orbitrap Velos

- iv. Customizations: none

2. Ion sources

2.1. nanoElectrospray ionization (2 kV source voltage, 100 μ A source current, 0V skimmer offset)

Fed by Eksigent NanoLC Ultra HPLC

Laser-pulled emitter tip, fused silica (360 μ m OD x 100 μ m ID)

3.. Post-source component

3.1. Mass Analyzer: MS, Orbitrap, MS/MS, Ion Trap

3.2. Ion trap final MS stage achieved: MS2

MS/MS: CID, Argon, HCD Collision cell

4. Spectrum and peak list generation and annotation

4.1. Data acquisition:

Thermo Xcalibur 2.1.0.1139

Switching criteria: Top ten most intense peaks in MS scan subject to MS/MS by CID in the LTQ

Min. signal required: 500

Isolation width: 2.0

Normalized collision energy: 35%

Default charge state: 4

Activation time: 10 ms

4.2. Analysis Software:

MS/MS File Converter: ScanSifter Desktop, Vanderbilt University

Medical Center, conversion of .raw MS/MS spectra to DTA files

Database Searching: Thermo Fisher Scientific SEQUEST version 27.12 (Custom Version on Vanderbilt ACCRE Linux Cluster, concatenated forward and reverse (decoy) database containing *Homo sapiens* subset of the UniProt KB protein database, 173736 protein sequences for forward and reverse databases)

Parameters for Sequest:

Protease: specified for trypsin digestions, non-specific for Chymotrypsin and Glu-C.

Max. missed cleavages: 10 for trypsin, 0 for chymotrypsin and Glu-C.

Peptide mass tolerance: 2.5 Da, monoisotopic

Variable modifications: carbamidomethyl derivatization of cysteine, oxidation of methionine, phosphorylation of serine, threonine, and tyrosine.

Validation/Visualization: Proteome Software Scaffold 3.3.3

Peptide probability threshold: 95%

Protein probability threshold: 99%, 2 identified peptides required

Algorithm: Protein Prophet, Nesvizhskii et al., Anal. Chem. 2003.

4.3. Resulting data

FDR Statistics for database searches:

Trypsin, Chymotrypsin and Glu-C: 10% Protein FDR, 0.5%

Peptide FDR

Trypsin^{Heat} and Trypsin^{Org}: 0.5% Protein FDR, 4.5% Peptide FDR

Identifications to Asef2 based on sequence associated with accession number A2VEA_HUMAN in custom database described in Section 4.2.

195 Peptides assigned to Asef2, 99% probability for protein identification

Sequence coverage: 617/652 amino acids, 94.6%

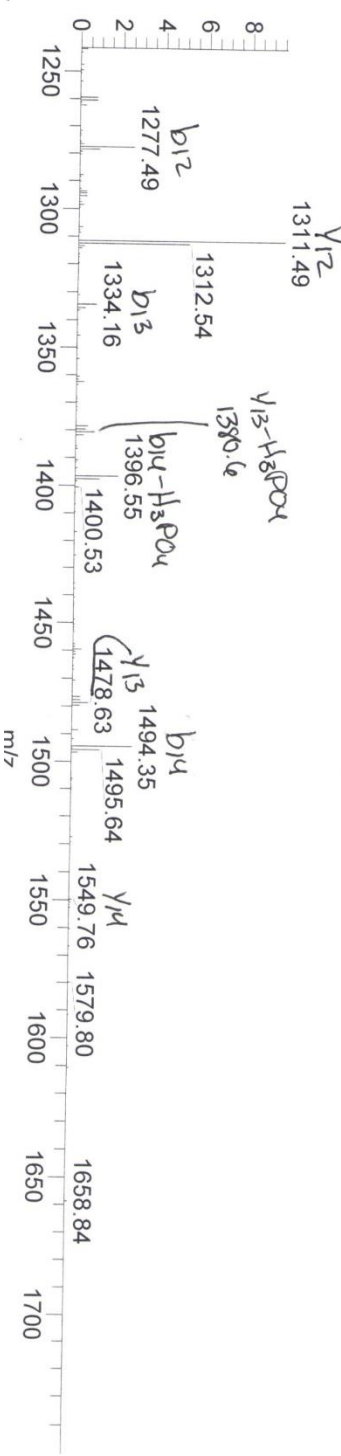
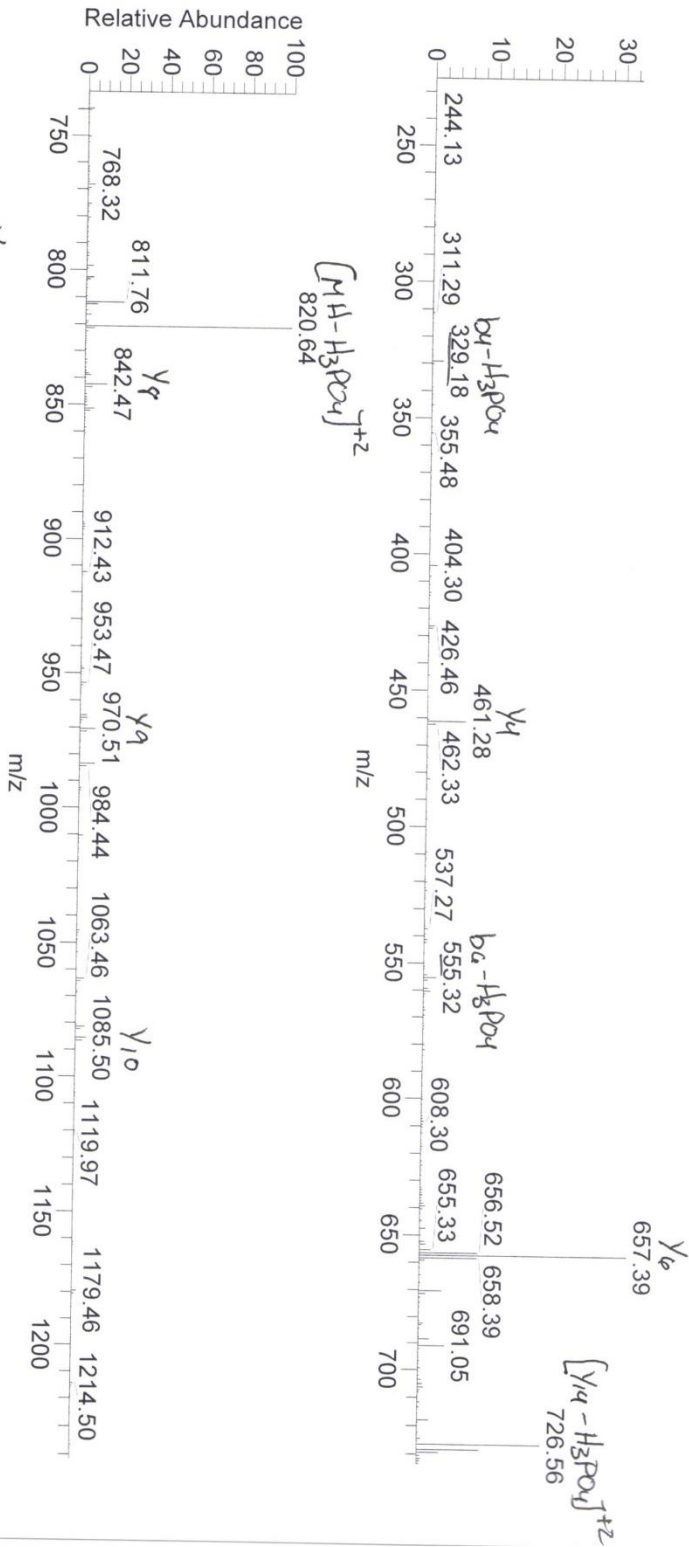
D2. Summary of Database Search Results for Phosphopeptides Identified for Asef2

Table D.1. Sequences, Mass Errors, and Sequest Scores (Xcorr and ΔCn) for Identified Phosphopeptides

Phosphopeptide Sequence	Obs.	z	Theor.	ΔAMU	ΔPPM	Mod.	XCorr	ΔCn	Start	Stop
TRYPsin										
(R)ASNVSSDGGTEP _s ALVDDNGSEEDFSYEDLcQASPR(Y)	1295.8563	3	1295.8566	-0.0003	-0.23	P (+80), C (+57)	4.9310	0.1234	94	129
(R)VNQEELSENSSSiPSEEQDEEASQSR(H)	992.7332	3	992.7336	-0.0004	-0.40	P (+80)	5.5673	0.1236	205	230
(R)FRPFTFSQ _s TPIGLDR(V)	650.3160	3	650.3154	0.0006	0.92	P (+80)	4.0599	0.3762	70	85
CHYMOTRYPsin										
(M)TSA _s PEDQNAPVGCpK(G)	869.3600	2	869.3611	-0.0011	-1.27	P (+80), C (+57)	3.5373	0.1373	2	17
GLU-C										
(E)LSENSSSiPSEEQDEEASQSRHRHcE(N)	774.8111	4	774.8113	-0.0002	-0.26	P (+80), C (+57)	4.1883	0.0768	210	235
TRYPsin ORG										
(R)ASNVSSDGGTEP _s ALVDDNGSEEDFSYEDLcQASPR(Y)	1295.8579	3	1295.8566	0.0013	1.00	P (+80), C (+57)	3.8711	0.3045	94	129
(R)FRPFTFSQ _s TPIGLDR(V)	650.3167	3	650.3154	0.0013	2.00	P (+80)	4.4995	0.4199	70	85
(R)FRPFTFSQ _s TPIGLDRVGR(R)	566.0364	4	566.0361	0.0003	0.53	P (+80)	3.4340	0.1742	70	88
(R)FRPFTFSQ _s TPIGLDRVGR(R)	605.0618	4	605.0614	0.0004	0.66	P (+80)	3.0850	0.1938	70	89
(R)HiMPTSVPOQQVFGLAEPK(R)	763.3772	3	763.3763	0.0009	1.18	P (+80)	3.0999	0.1475	615	634
(R)LRVNQEELSENSSSiPSEEQDEEASQSR(H)	1082.4620	3	1082.4620	0.0000	0.00	P (+80)	5.3512	0.3239	203	230
(R)RPI _s VIGGVSLYGTNQTEELDNLLTQPASRPPMPAHQVPPYK(A)	1163.5928	4	1163.5899	0.0029	2.49	P (+80)	5.0000	0.3509	23	64
(R)RRPI _s VIGGVSLYGTNQTEELDNLLTQPASRPPMPAHQVPPYK(A)	962.2953	5	962.2936	0.0017	1.77	P (+80)	2.9387	0.2696	22	64
(R)RRPI _s VIGGVSLYGTNQTEELDNLLTQPASRPPMPAHQVPPYKAVSAR(F)	1059.1515	5	1059.1488	0.0027	2.55	P (+80)	4.8367	0.2227	22	69
(R)RRRPI _s VIGGVSLYGTNQTEELDNLLTQPASRPPMPAHQVPPYK(A)	993.5156	5	993.5138	0.0018	1.81	P (+80)	4.2402	0.1231	21	64
TRYPsin TEMP										
(R)ASNVSSDGGTEP _s ALVDDNGSEEDFSYEDLcQASPR(Y)	1295.8561	3	1295.8566	-0.0005	-0.39	P (+80), C (+57)	4.6717	0.1855	94	129
(R)FRPFTFSQ _s TPIGLDR(V)	650.3159	3	650.3154	0.0005	0.77	P (+80)	4.9663	0.3846	70	85
(R)RPI _s VIGGVSLYGTNQTEELDNLLTQPASRPPMPAHQVPPYK(A)	1163.5905	4	1163.5899	0.0006	0.52	P (+80)	4.9410	0.3356	23	64
(R)VNQEELSENSSSiPSEEQDEEASQSR(H)	992.7343	3	992.7336	0.0007	0.71	P (+80)	5.3548	0.0555	205	230

D.2. Annotated DDA MS/MS Spectra for Identified Phosphopeptides

sh_2514_JF_012412_Aset2_Chymo
 2A01
 sh_2514_JF_012412_Aset2_Chymo #11666 RT: 39.83 AV: 1 NL: 2.73E5
 F: ITMS + c NSI d Full ms2 869.36@cid35.00 [225.00-1750.00]
 1/25/2012 1:23:42 PM TSAs PEDGNAP V6cPK
 5.000000



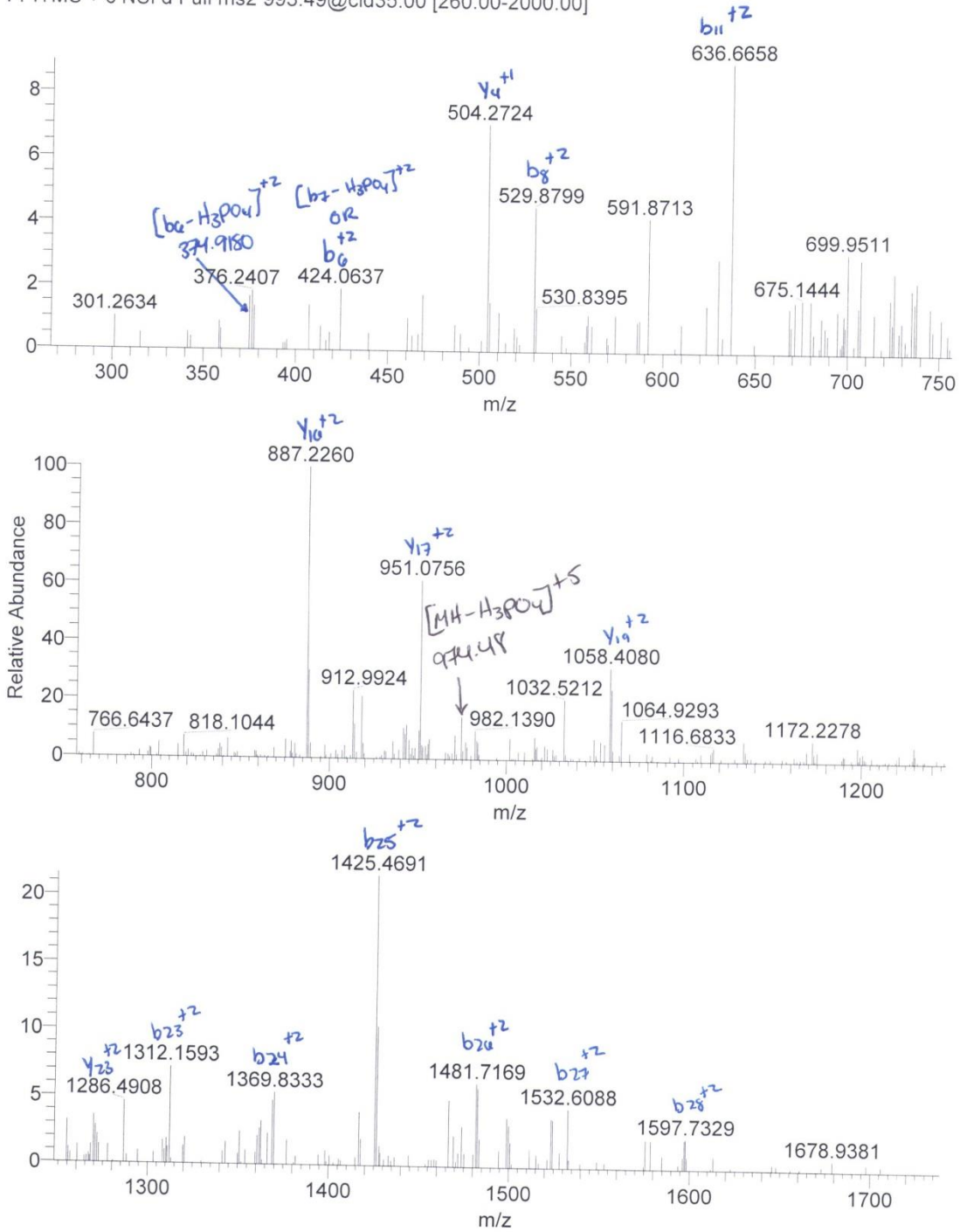
RRRPI₅VIGGVS₁LYE₁TNQTEELD₁LLTQ₁PAS₁RPPM₁PAHQV₁APYK

sh_2796_KH_081312_ASEF_ACN_digest
2E06

8/13/2012 4:42:04 PM
5.000000

sh_2796_KH_081312_ASEF_ACN_digest #20085 RT: 47.71 AV: 1 NL: 2.16E5

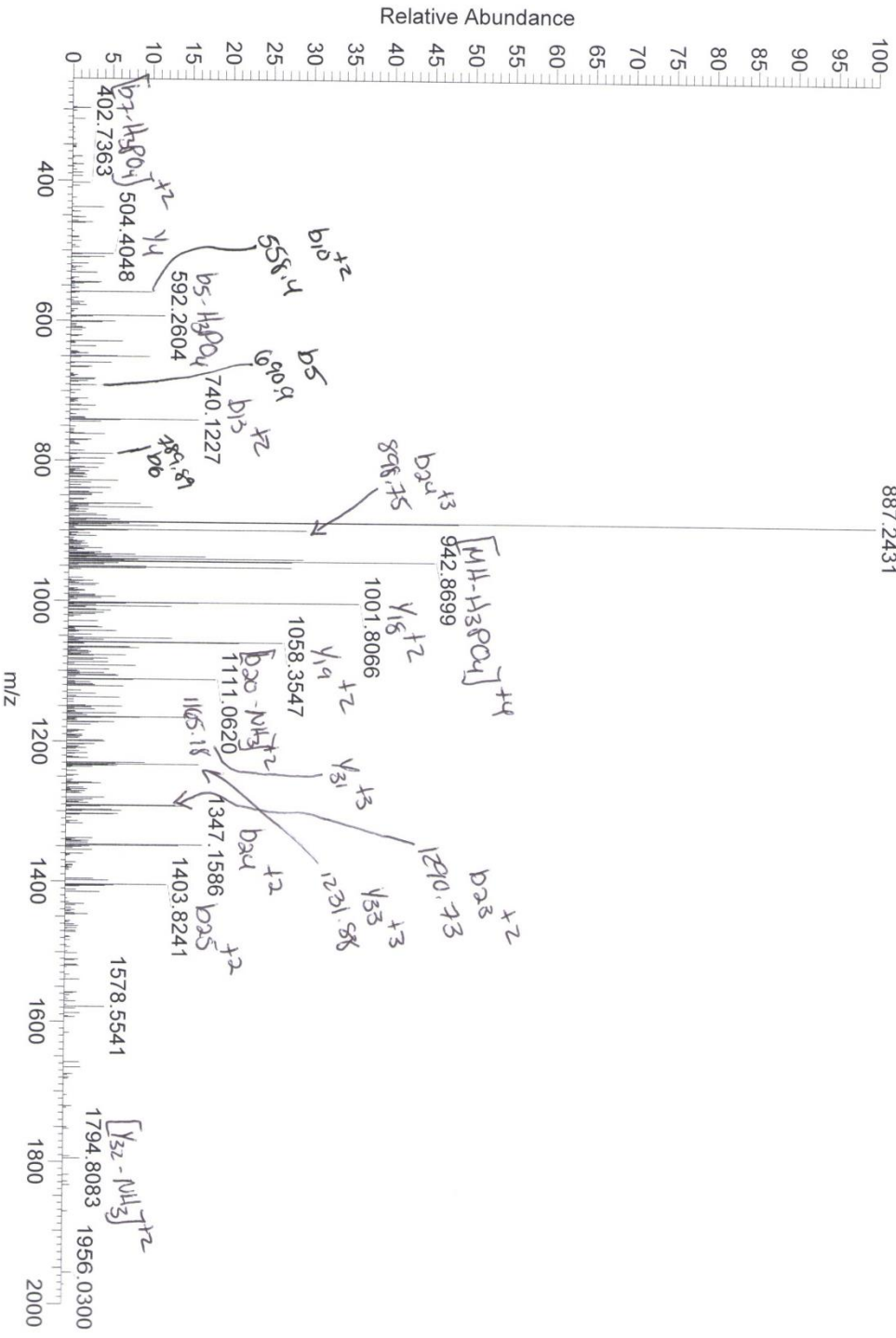
F: ITMS + c NSI d Full ms2 993.49@cid35.00 [260.00-2000.00]



sh_2796_KH_081312_ASEF_ACN_digest
ZE06

sh_2796_KH_081312_ASEF_ACN_digest #20574 RT: 48.93 AV: 1 NL: 6.52E4
F: ITMS + c NSI.d Full.ms2_962.90

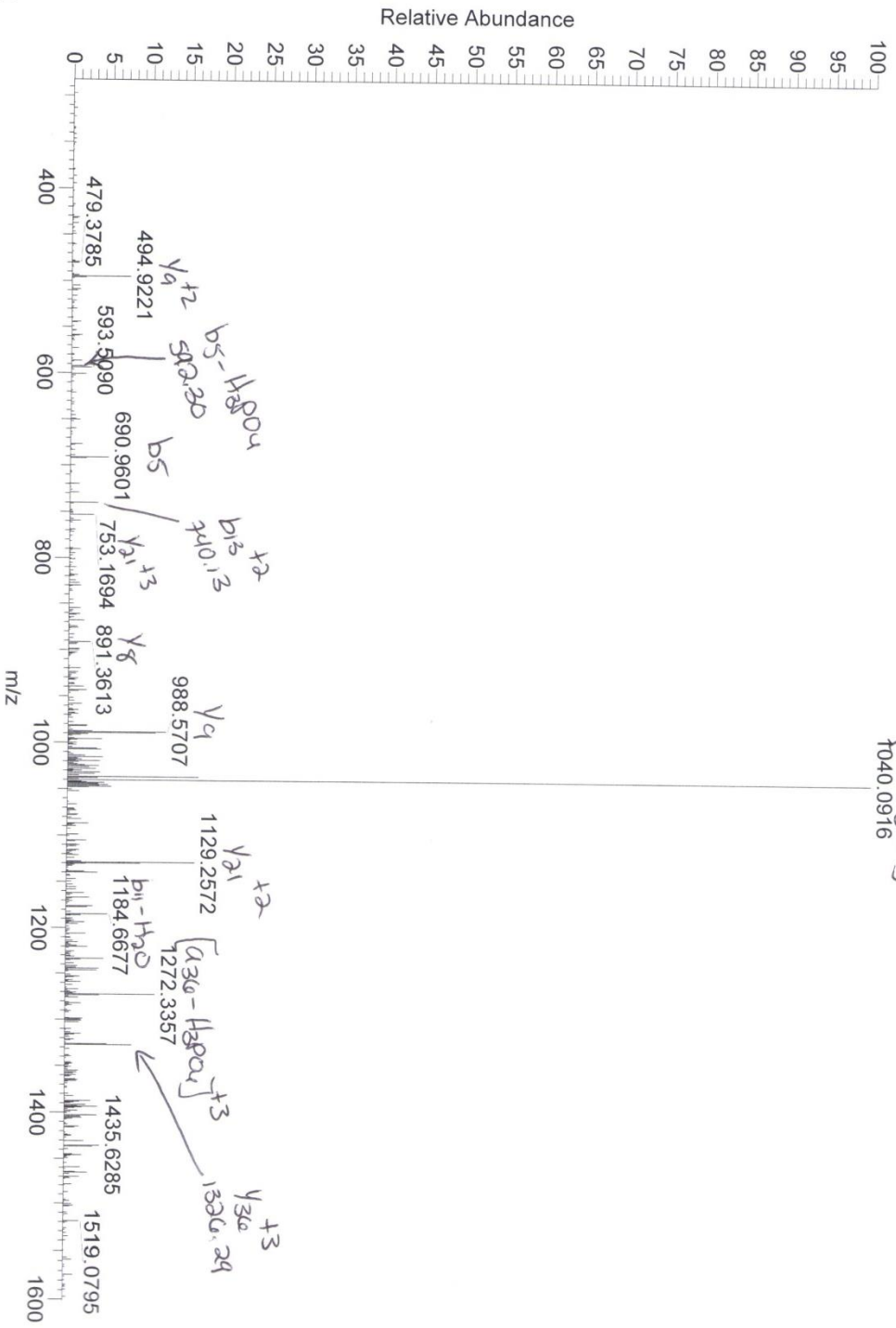
RRPVS YKGVSLVETNQTTEELDILLTQ PASRPPMATHQVPPYK
8/13/2012 4:42:04 PM
5.000000



PEPIS VGENSVLGTIUGTELELDLUL TQPAS RPPM P AH QVPPYKAVBAR

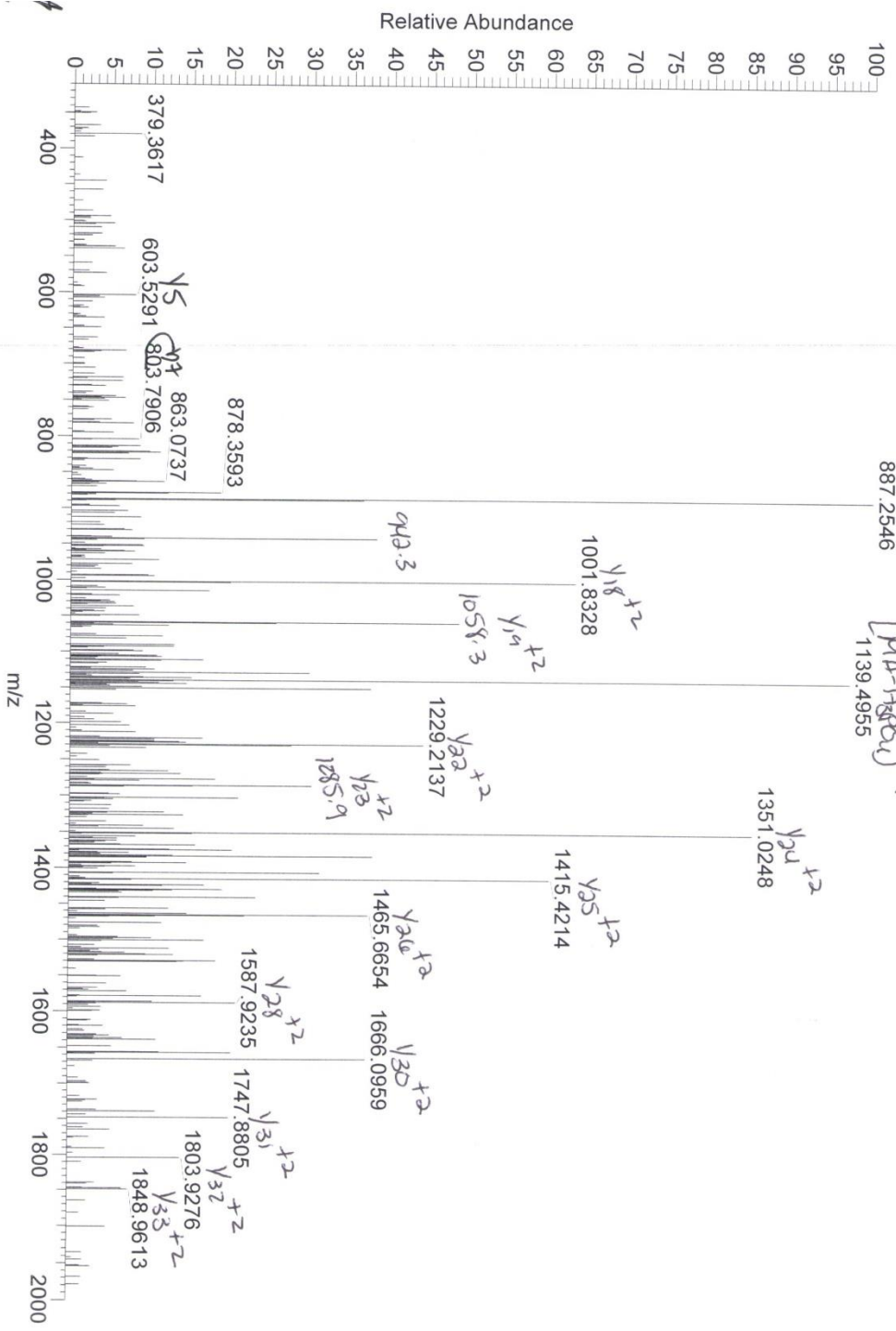
sh_2796_KH_081312_ASEF_ACN_digest
ZE06
sh_2796_KH_081312_ASEF_ACN_digest#20378 RT: 48.44 AV: 1 NL: 7.18E4
F: ITMS + c NSI d Full ms2 1059.75@cid35.00 [280.00-2000.00]
8/13/2012 4:42:04 PM
5.000000

[M+H]⁺
1040.0916



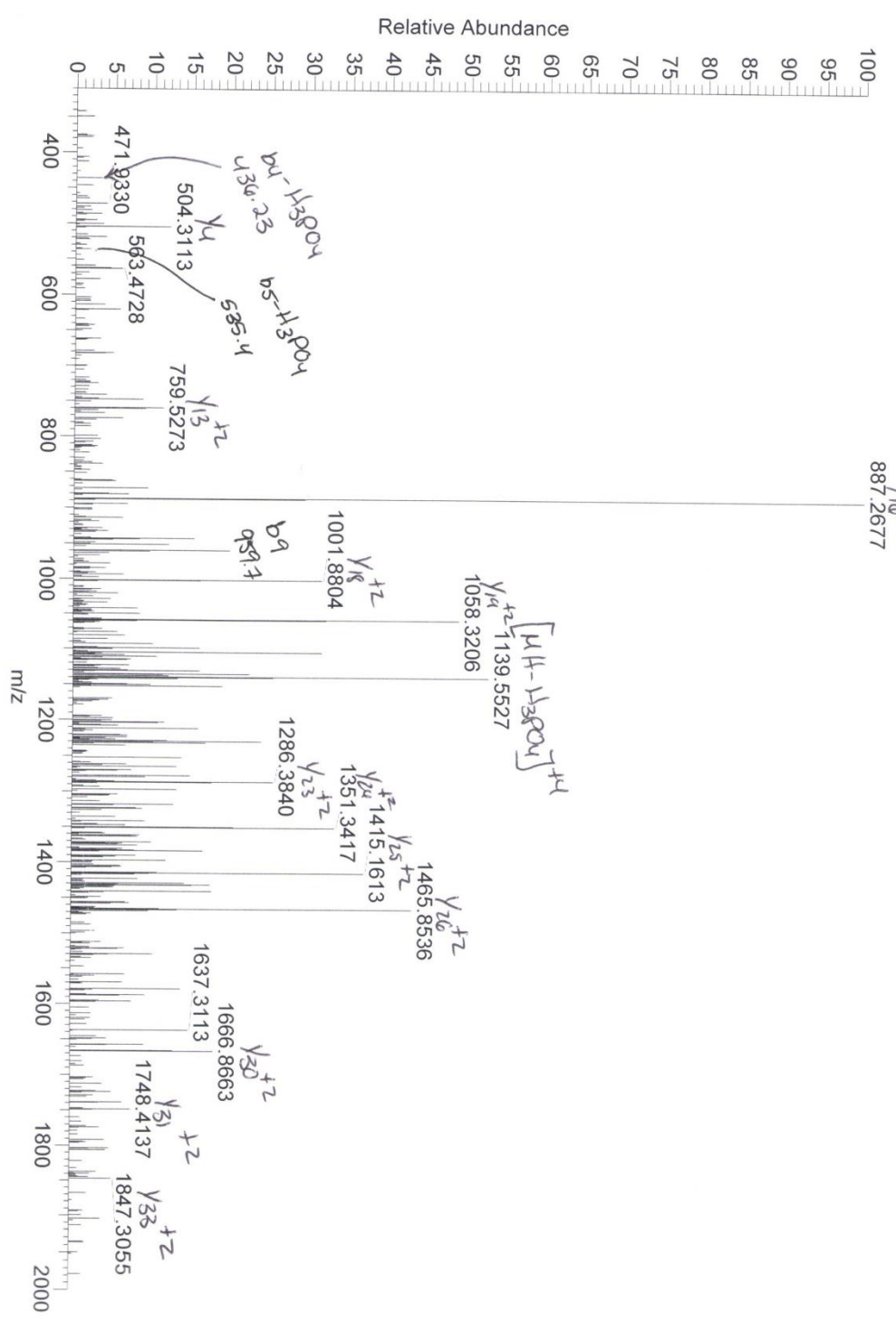
sh_2796_KH_081312_ASEF_amm_bic_digest
 2E05
 sh_2796_KH_081312_ASEF_amm_bic_digest#21712 RT: 49.80 AV: 1 NL: 2.42E4
 F: ITMS + c NSI d Full ms2 1164.09@cids35.00 [310.00-2000.00]

RPLS VIGVSLYGT VQTELELDVLLTQPARSRPPMRAHQVPPYK
 8/13/2012 1:07:47 PM
 5.000000



RPLS VIG6 YSLYGTNQT EELVLLTQ PHSRPPMRAHGVPYK

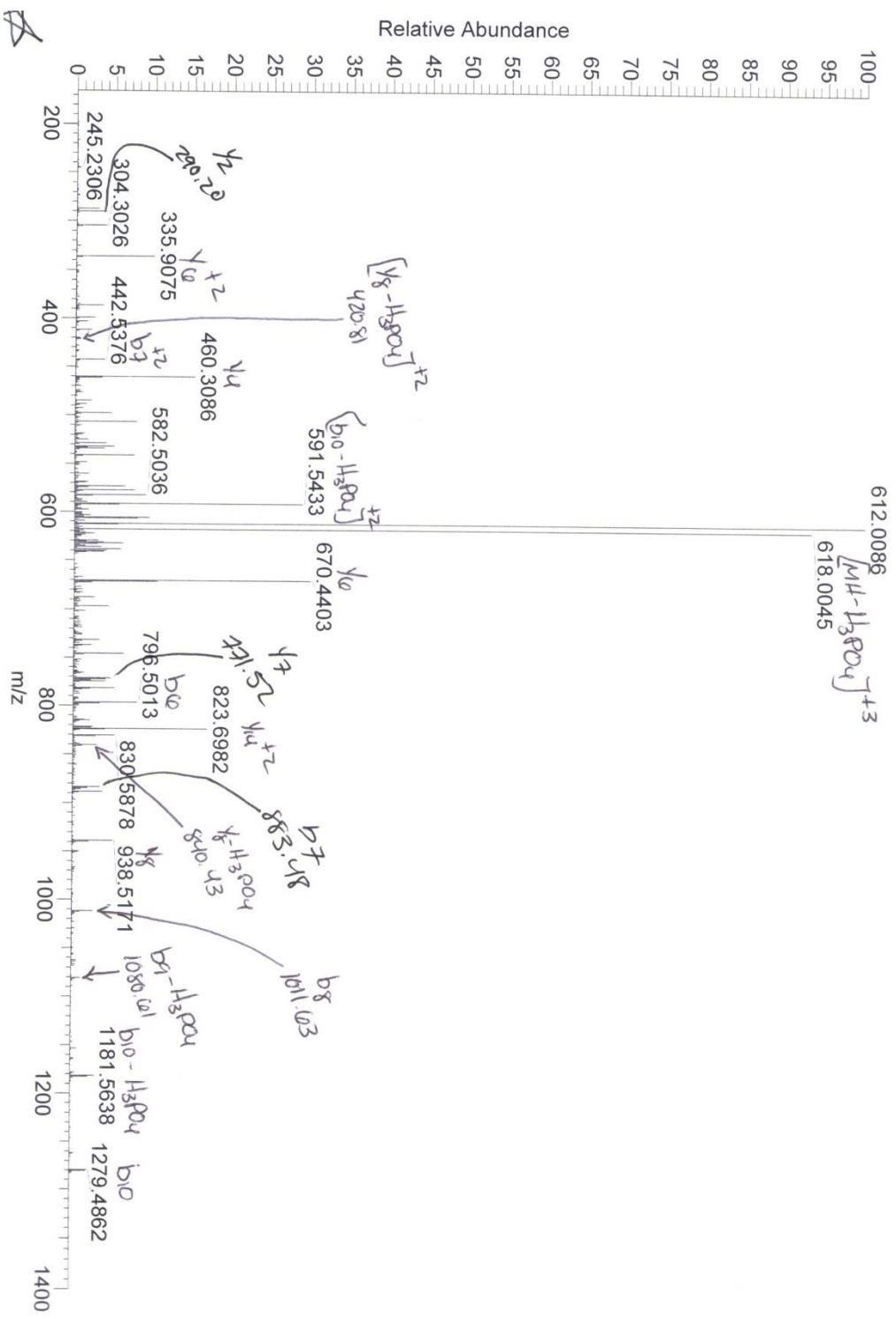
sh_2796_KH_081312_ASEF_ACN_digest 8/13/2012 4:42:04 PM 5.000000
sh_2796_KH_081312_ASEF_ACN_digest#20892 RT: 49.76 AV: 1 NL: 1.32E4
F: TTMS + c NSI d Full ms2 1164.19@cic35.00 [310.00-2000.00]



sh_2514_JF_012412_Asef2_ttyp
 2A01
 sh_2514_JF_012412_Asef2_ttyp #18471 RT: 56.51 AV: 1 NL: 1.85E5
 F: ITMS + c NSI d Full ms2 650.65@cid35.00 [165.00-1965.00]

1/25/2012 8:44:34 AM
 5.000000

FRPPF TFSQS TP16JDR



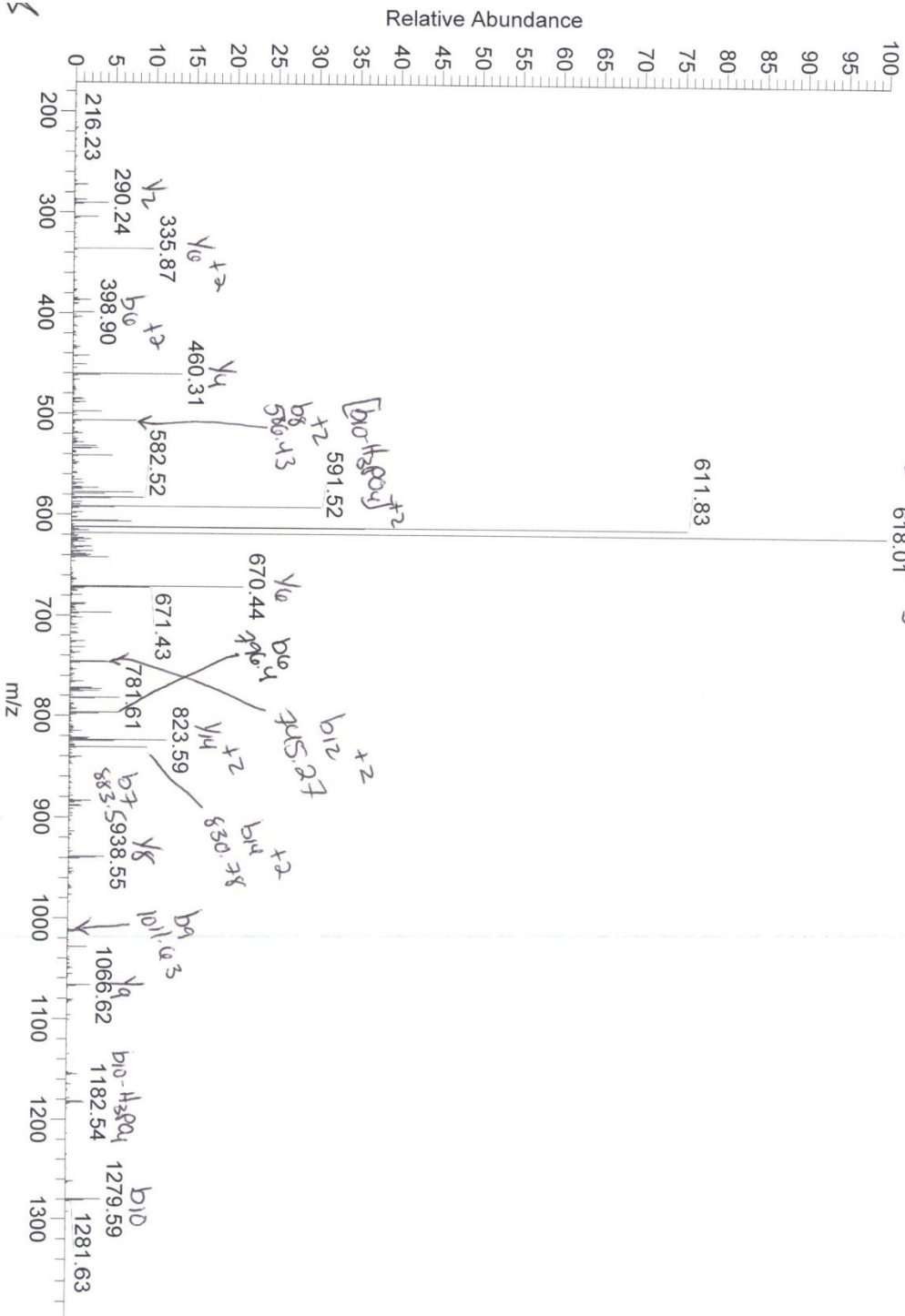
sh_2796_KH_081312_ASEF_amm_bic_digest
2E05

8/13/2012 1:07:47 PM
5.000000

EXPFTS05.TPK6.LDK

sh_2796_KH_081312_ASEF_amm_bic_digest #20665 RT: 47.31 AV: 1 NL: 6.19E5
F: TTMS + c NSI d Full ms2 650.65@cid35.00 [165.00-1965.00]

$[M+H_2O]^+$
618.01



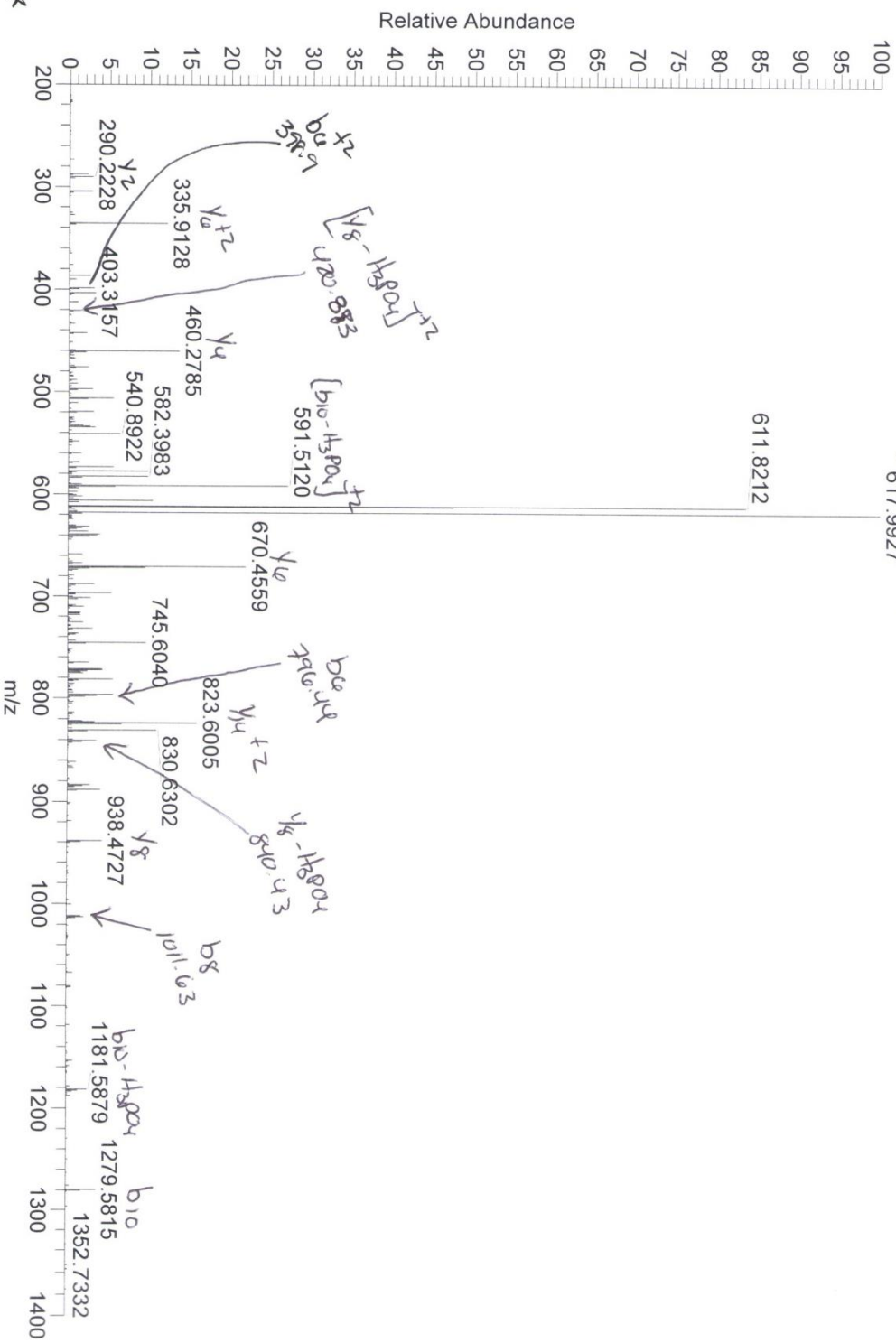
sh_2796_KH_081312_ASEF_ACN_digest
2E06

8/13/2012 4:42:04 PM
5.000000

FRPFTESSQ57P16L02

sh_2796_KH_081312_ASEF_ACN_digest#19934 RT: 47.33 AV: 1 NL: 9.00E5
F: ITMS + c NSI d Full ms2 650.65@cid35.00 [165.00-1965.00]

$[M+H_2PO_4]^+$
617.9927

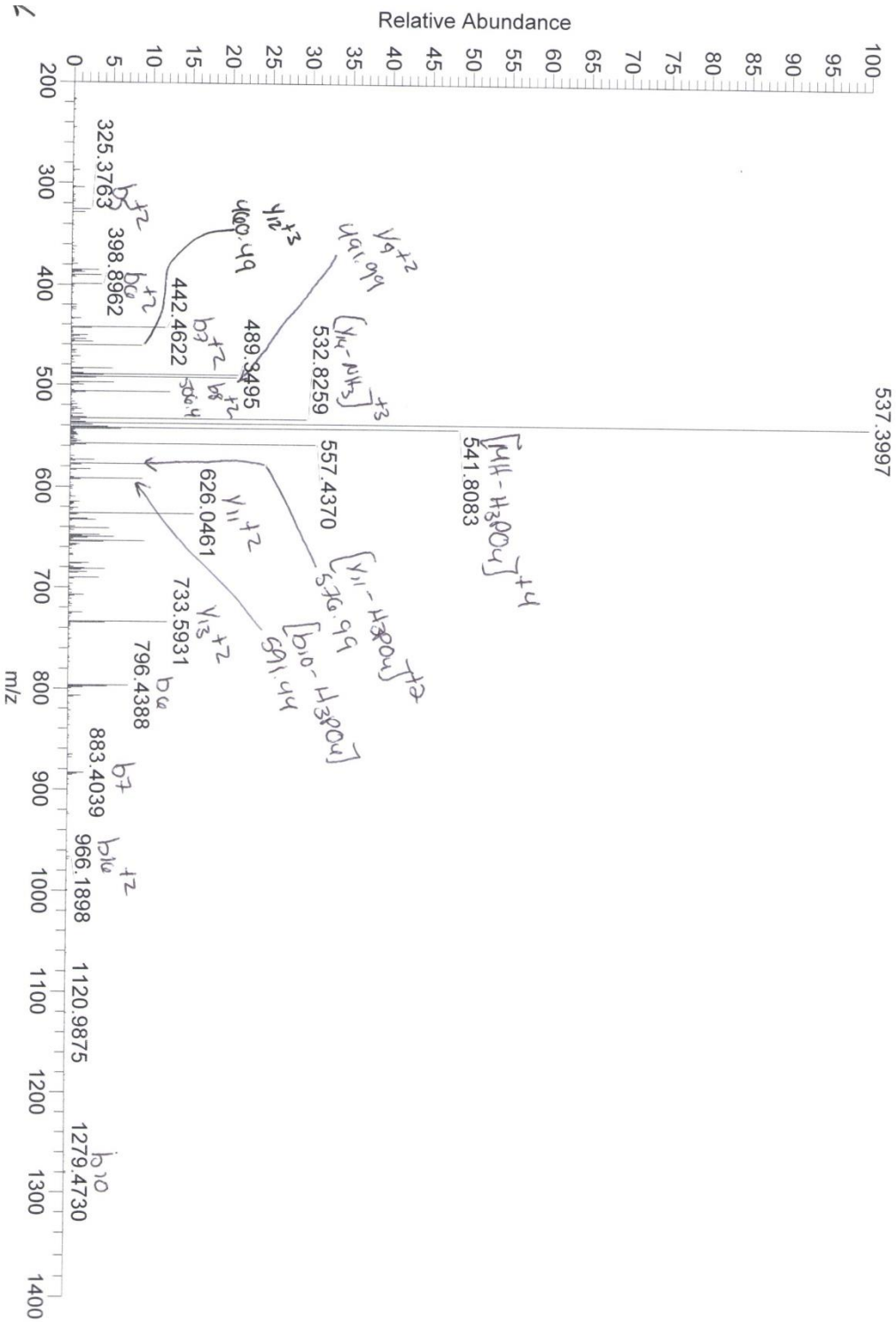


sh_2796_KH_081312_ASEF_ACN_digest
2E06

8/13/2012 4:42:04 PM
5.000000

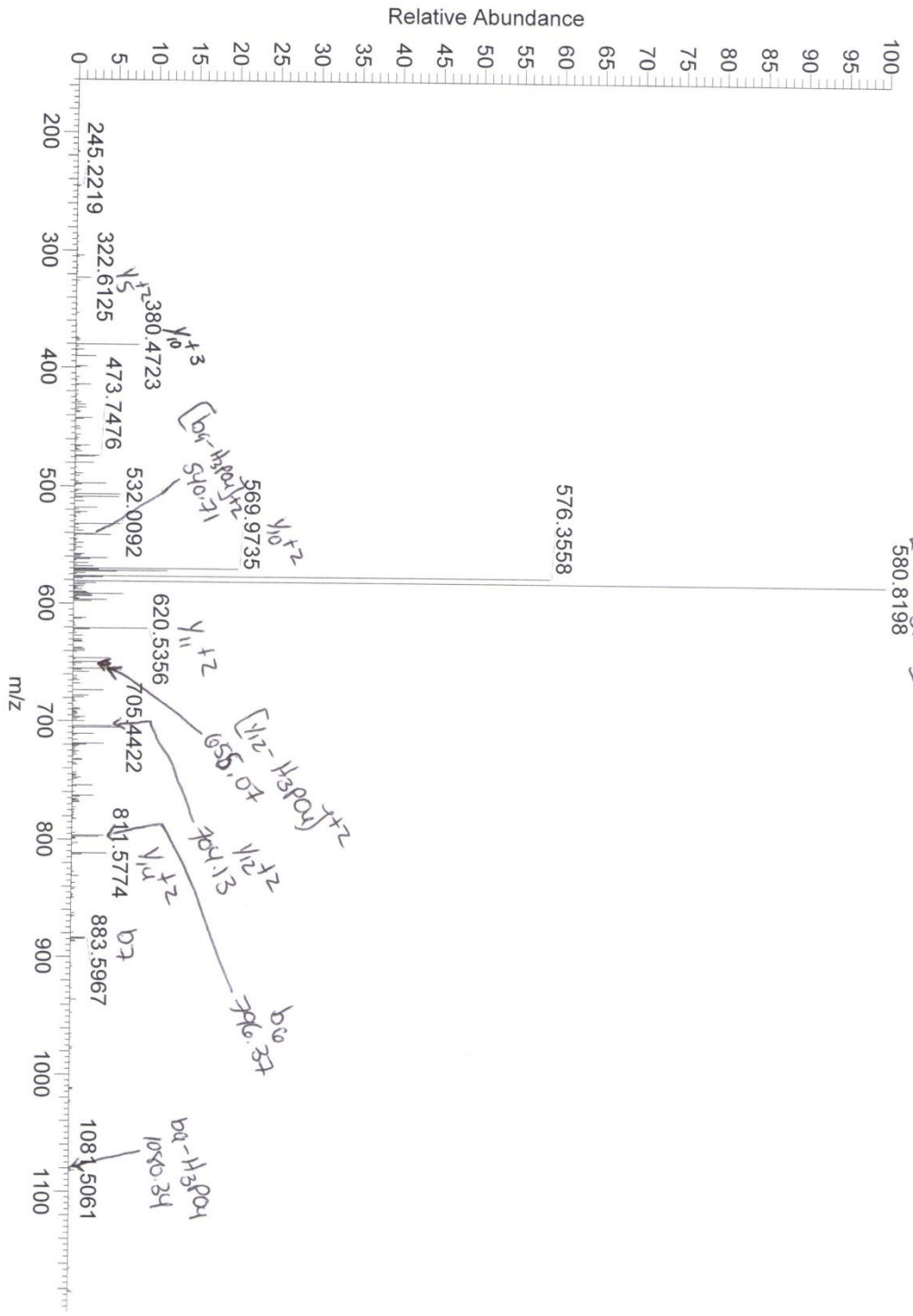
FRPFTFSQs 7P16 LDR V612

sh_2796_KH_081312_ASEF_ACN_digest #19706 RT: 46.76 AV: 1 NL: 6.35E5
F: ITMS + c NSI d Full ms2 566.29@cid35.00 [145.00-2000.00]



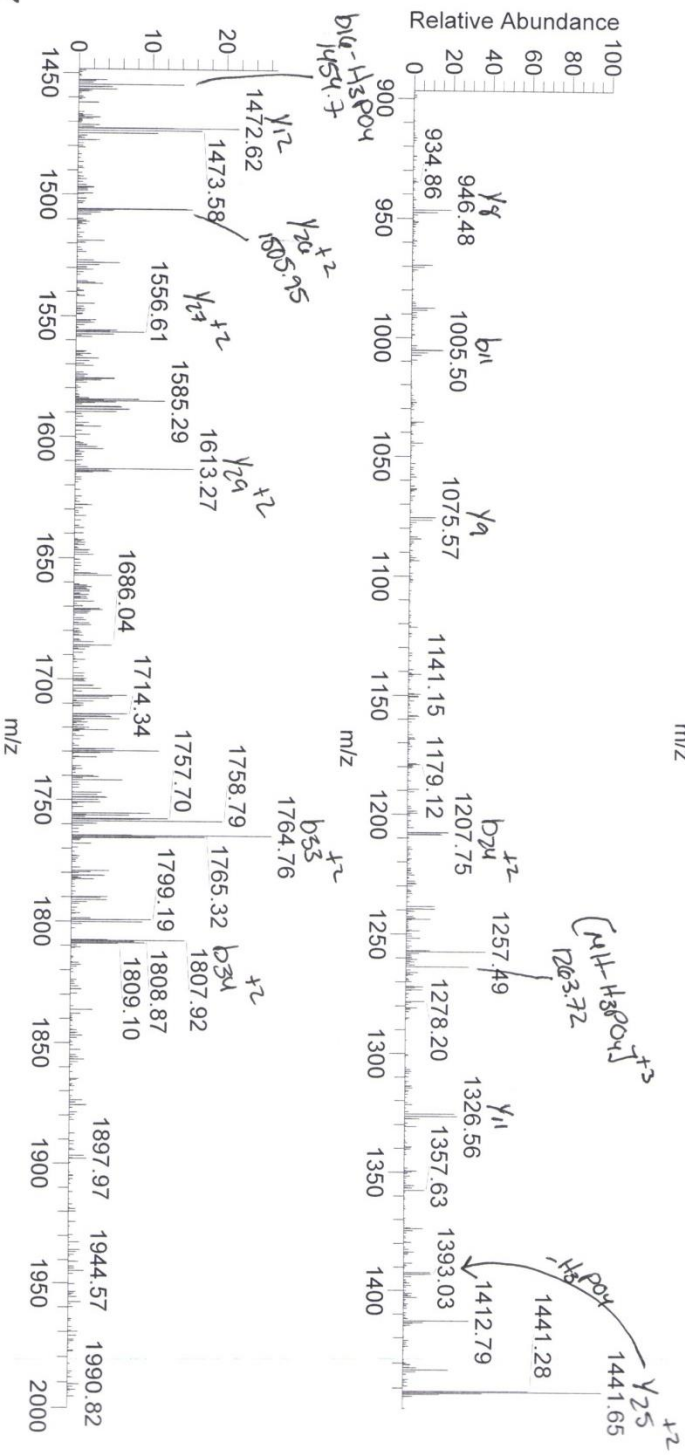
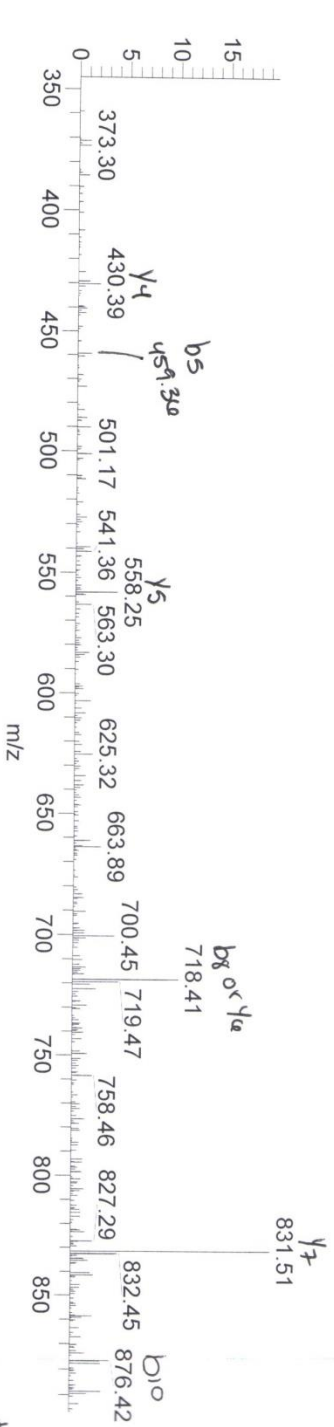
sh_2796_KH_081312_ASEF_ACN_digest 8/13/2012 4:42:04 PM FPPFTFSQS-TP16LDR VERRE

2E06
 sh_2796_KH_081312_ASEF_ACN_digest #19188 RT: 45.46 AV: 1 NL: 5.48E5
 F: ITMS + c NSI d Full ms2 605.31@cid35.00 [155.00-2000.00]
 [M⁺-H₂O]₂⁺
 580.8198



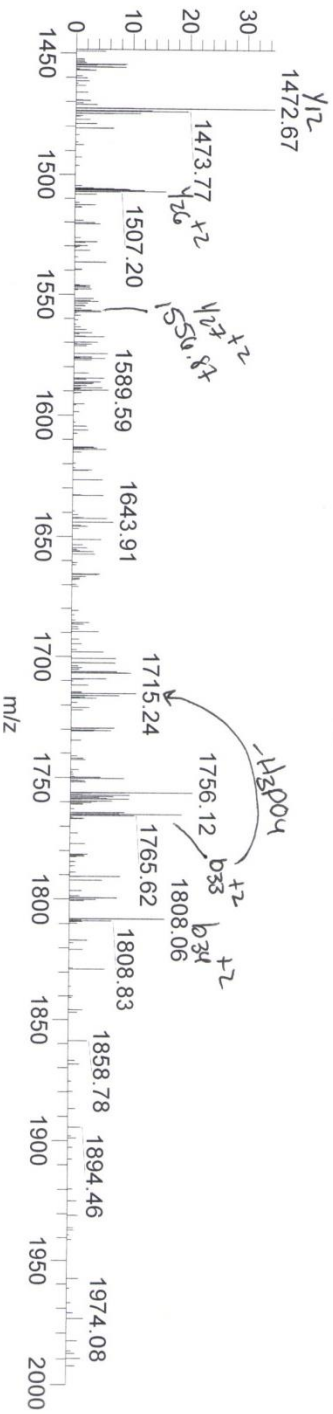
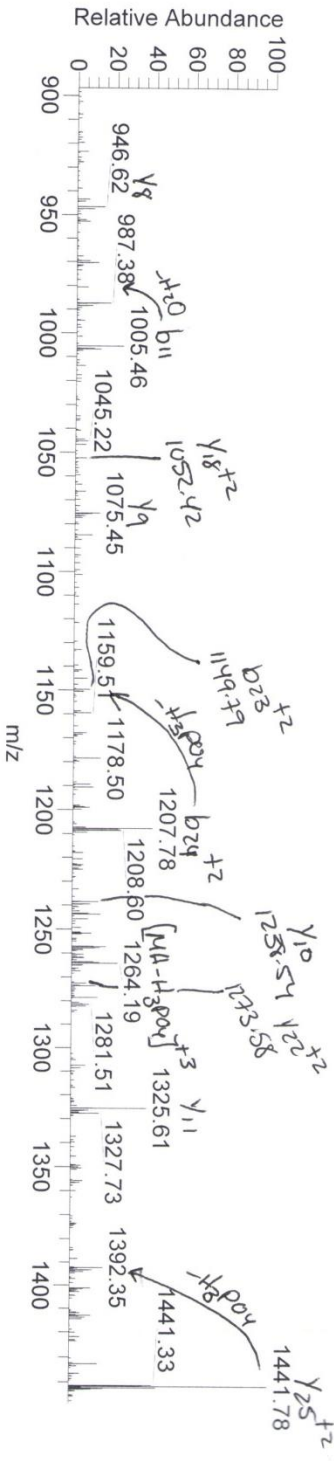
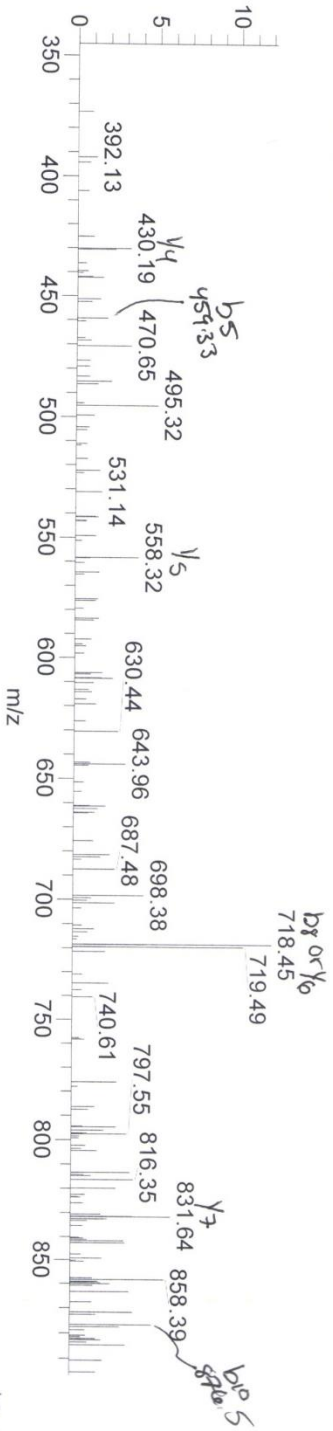
ASNVSSD667EP6 ALVDDW6SSEEDFSYEDL6 QASPR

sh_2514_JF_012412_Asef2_ttyp
 2A01
 1/25/2012 8:44:34 AM
 5.000000
 sh_2514_JF_012412_Asef2_ttyp #16672-16728 RT: 52.06-52.10 AV: 2 NL: 9.46E3
 F: ITMS + c NSI d Full ms2 1296.50



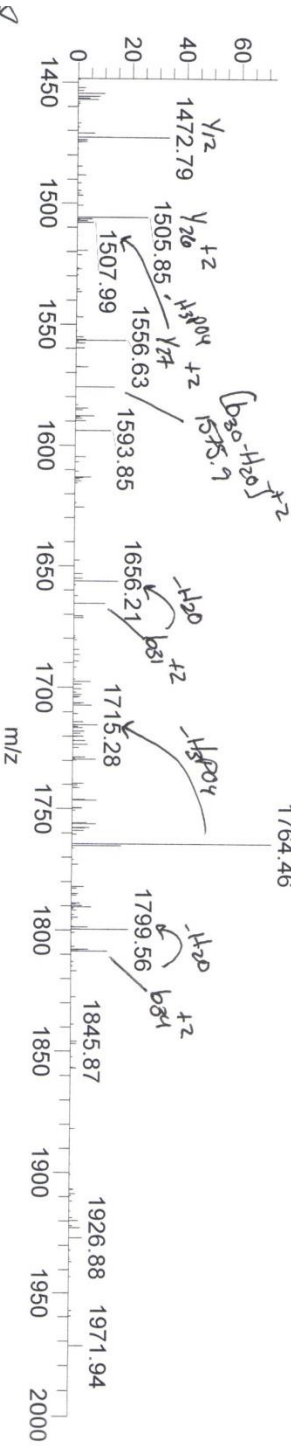
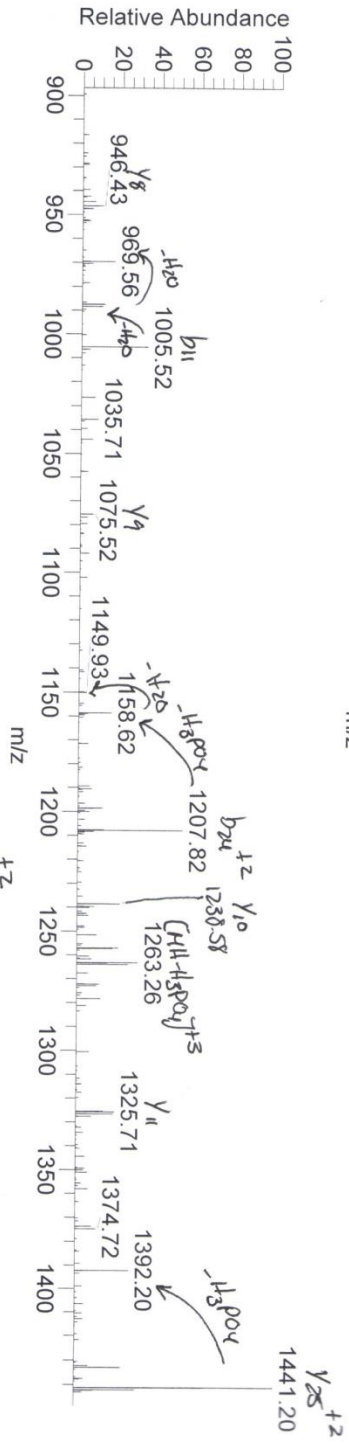
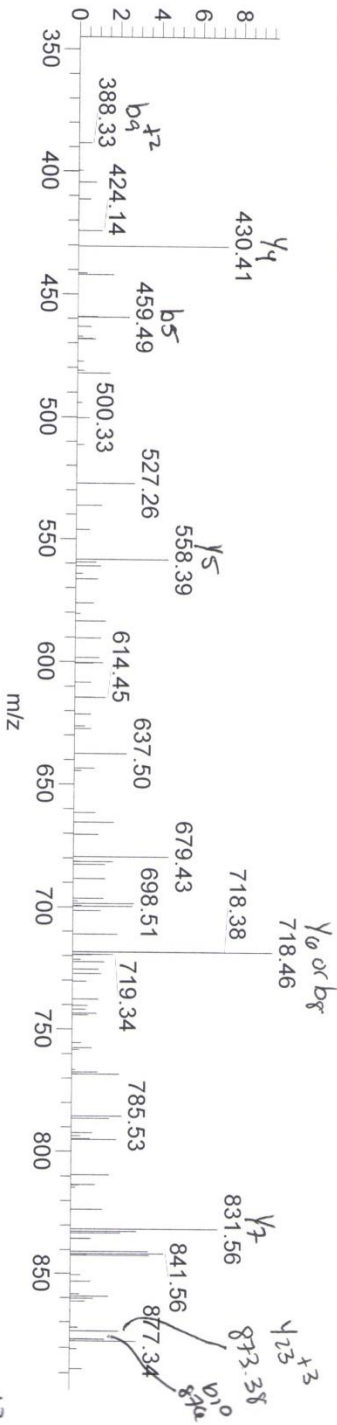
AS NVSS D667EPs ALVDDN6SEEDFSYEDLQASPK

sh_2796_KH_081312_ASEF_amm_bic_digest
 8/13/2012 1:07:47 PM
 5.000000
 sh_2796_KH_081312_ASEF_amm_bic_digest #18130-19255 RT: 43.1443.43 AV: 4 NL: 2.29E4
 F: TMS + c NSI d Full ms2 1296.50



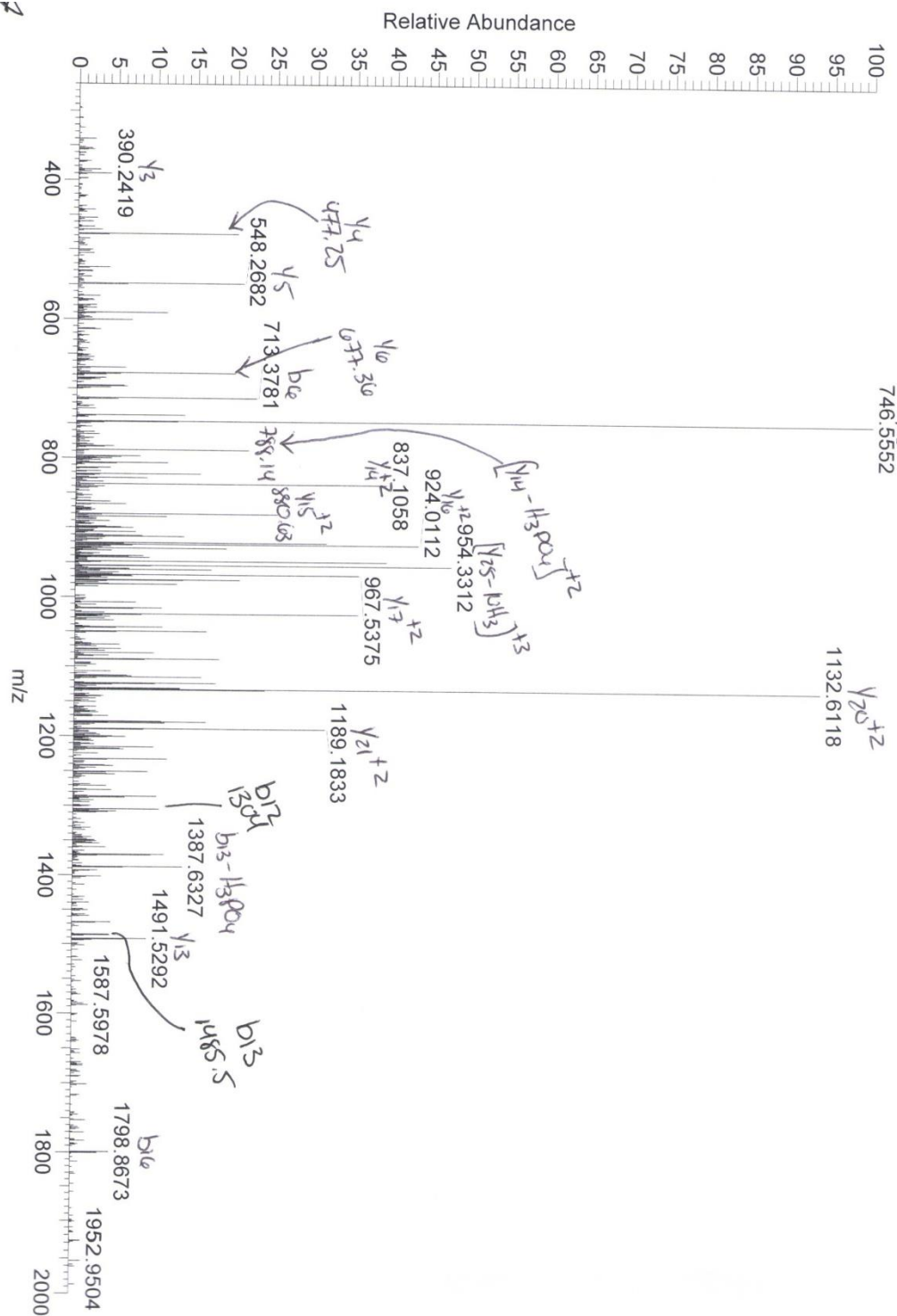
ASUVSSJ2667EP5ALVDDN6SSEEDFSYEDLQASPR

sh_2796_KH_081312_ASEF_ACN_digest
 8/13/2012 4:42:04 PM
 5.000000
 sh_2796_KH_081312_ASEF_ACN_digest #18276-18306 RT: 43.1943.21 AV: 2 NL: 1.28E4
 F: ITMS + c NSI d Full ms2 1295.86



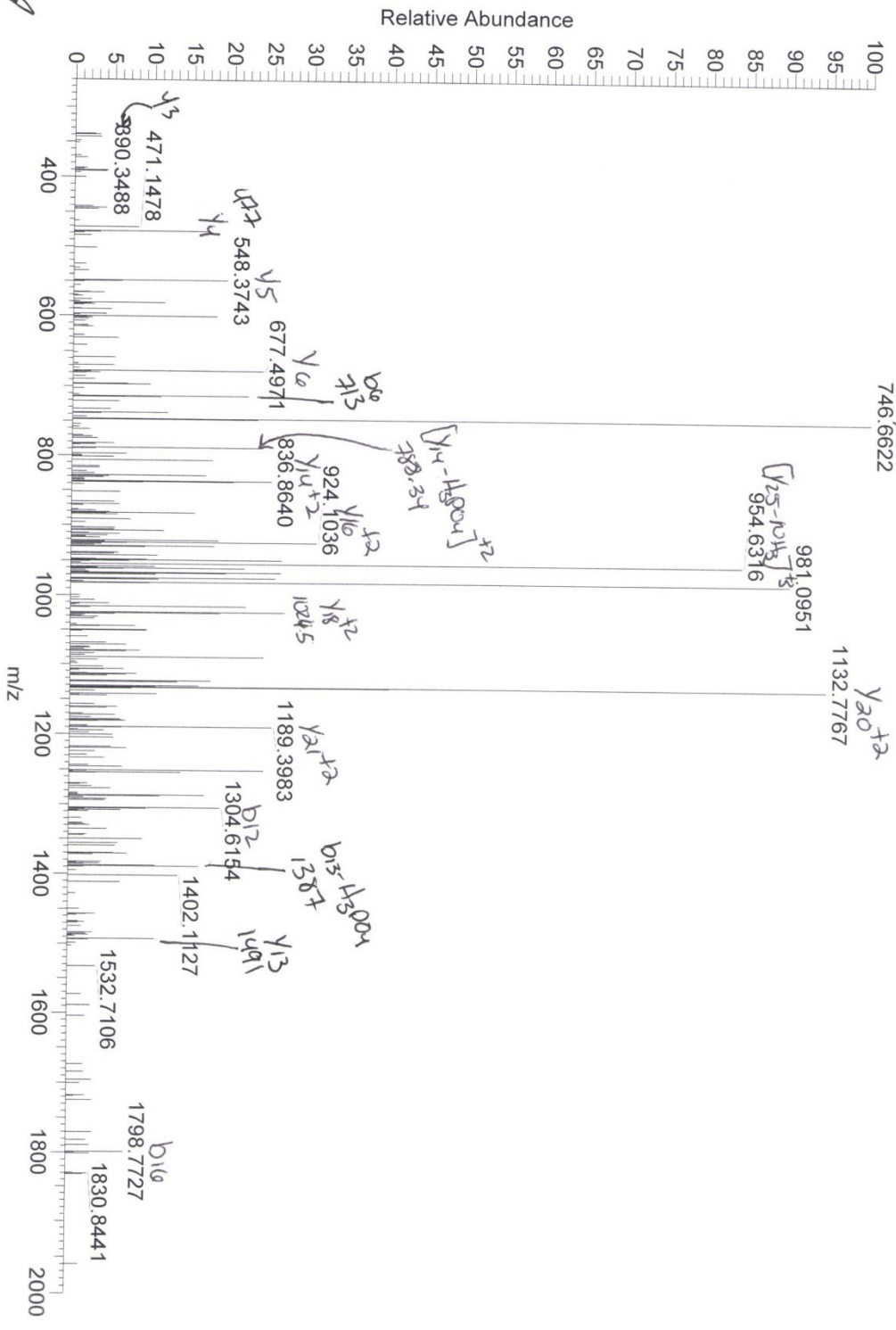
sh_2514_JF_012412_Asef2_tryp 1/25/2012 8:44:34 AM WQEELEUSSS&PSFEQDDEASQSR
 2A01 5.000000

sh_2514_JF_012412_Asef2_tryp #12441 RT: 41.51 AV: 1 NL: 7.75E3
 F: ITMS + c NSI d Full ms2 993.07@cid35.00 [260.00-2000.00]

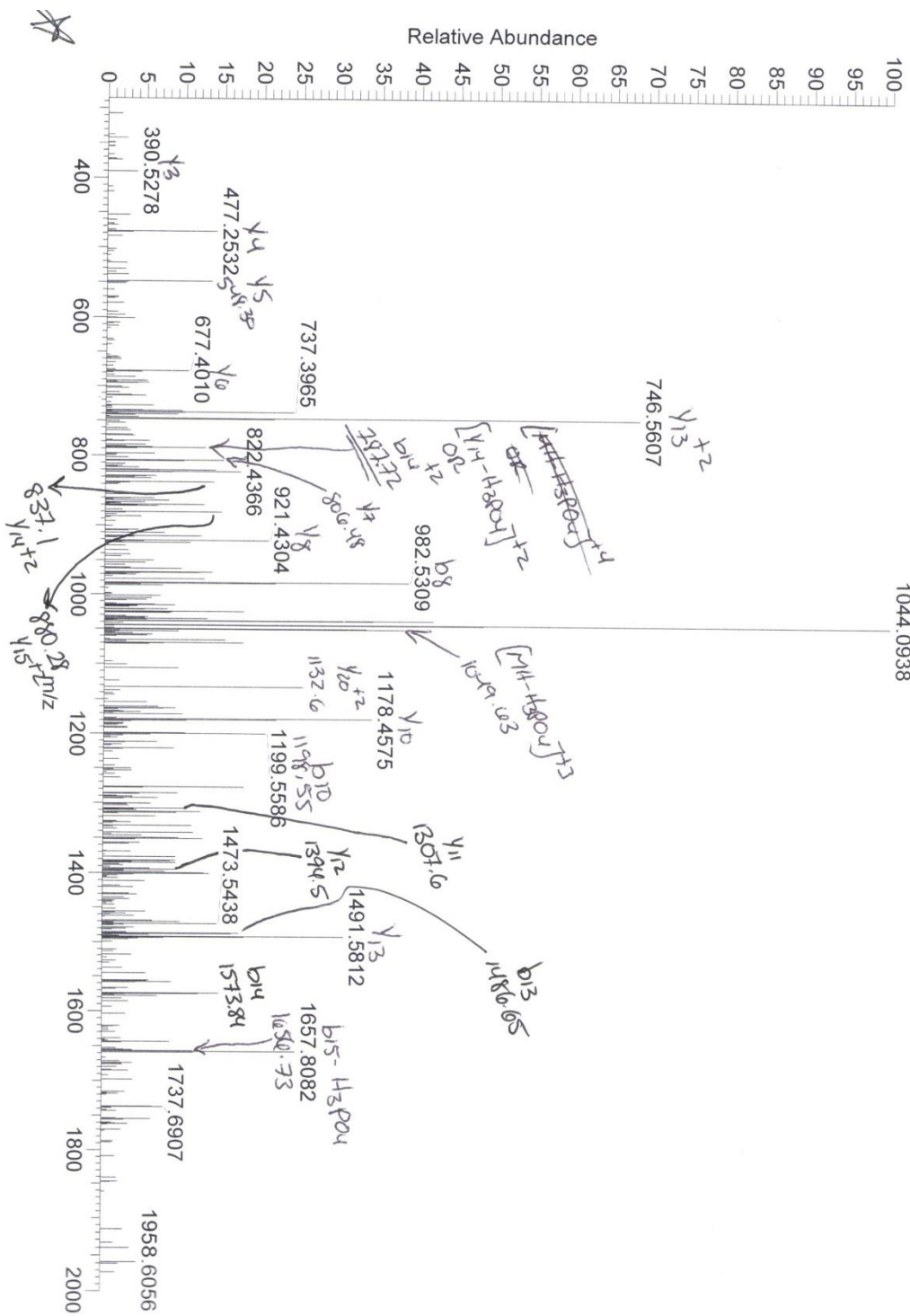


sh_2796_KH_081312_ASEF_amm_bic_digest 8/13/2012 1:07:47 PM WHEELSENSSSTPSEEDDEASQR
ZE05 5.000000

sh_2796_KH_081312_ASEF_amm_bic_digest#13962 RT: 32.60 AV: 1 NL: 4.61E4
F: ITMS + c NSI d Full ms2 993.01 [260.00-2000.00]



sh_2796_KH_081312_ASEF_ACN_digest 8/13/2012 4:42:04 PM LRVIGEELSENSSTPSEE QDEEASOS R
 ZE06 5.000000
 sh_2796_KH_081312_ASEF_ACN_digest #14975 RT: 35.06 AV: 1 NL: 1.87E4
 F: ITMS + c NSI d Full ms2 1082.80@cid35.00 [285.00-2000.00]

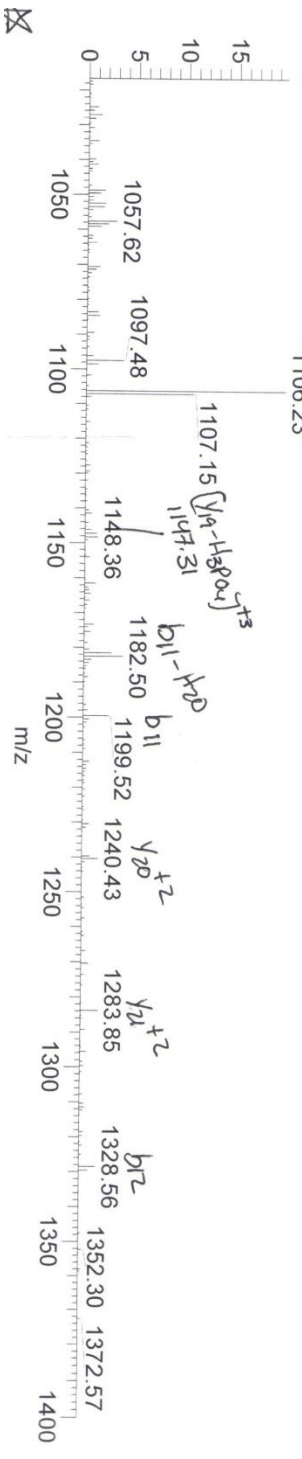
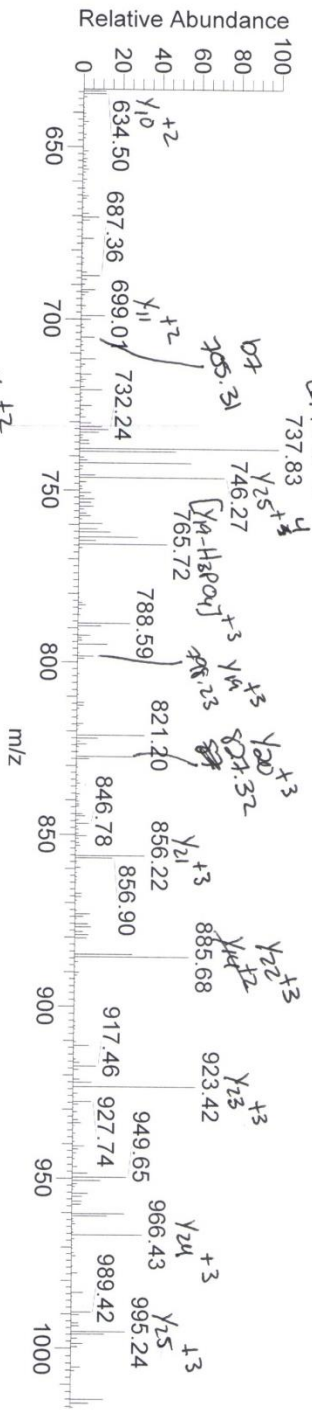
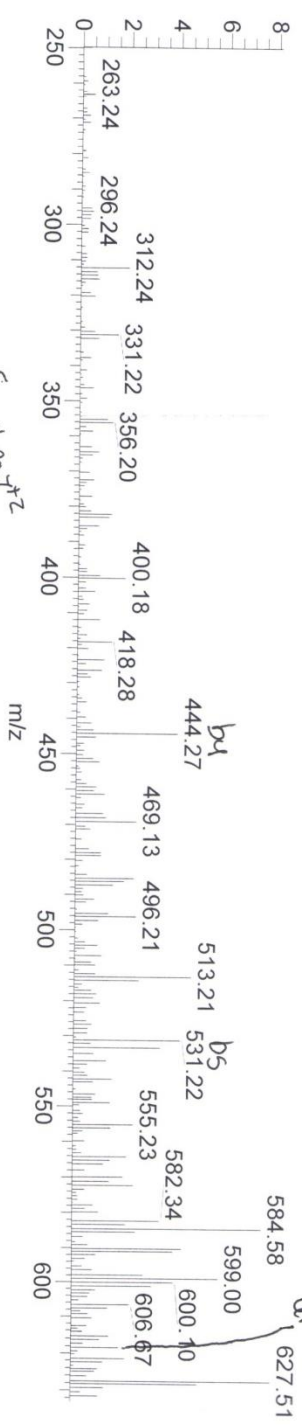


sh_2514_JF_012412_Asetf2_GluC
2A01

sh_2514_JF_012412_Asetf2_GluC #11528 RT: 37.16 AV: 1 NL: 2.69E4
F: ITMS + c NSI.d Full ms2 775.09

1/25/2012 10:56:12 AM
5.000000

SENSSE PSEEWDEENASRHRHCE

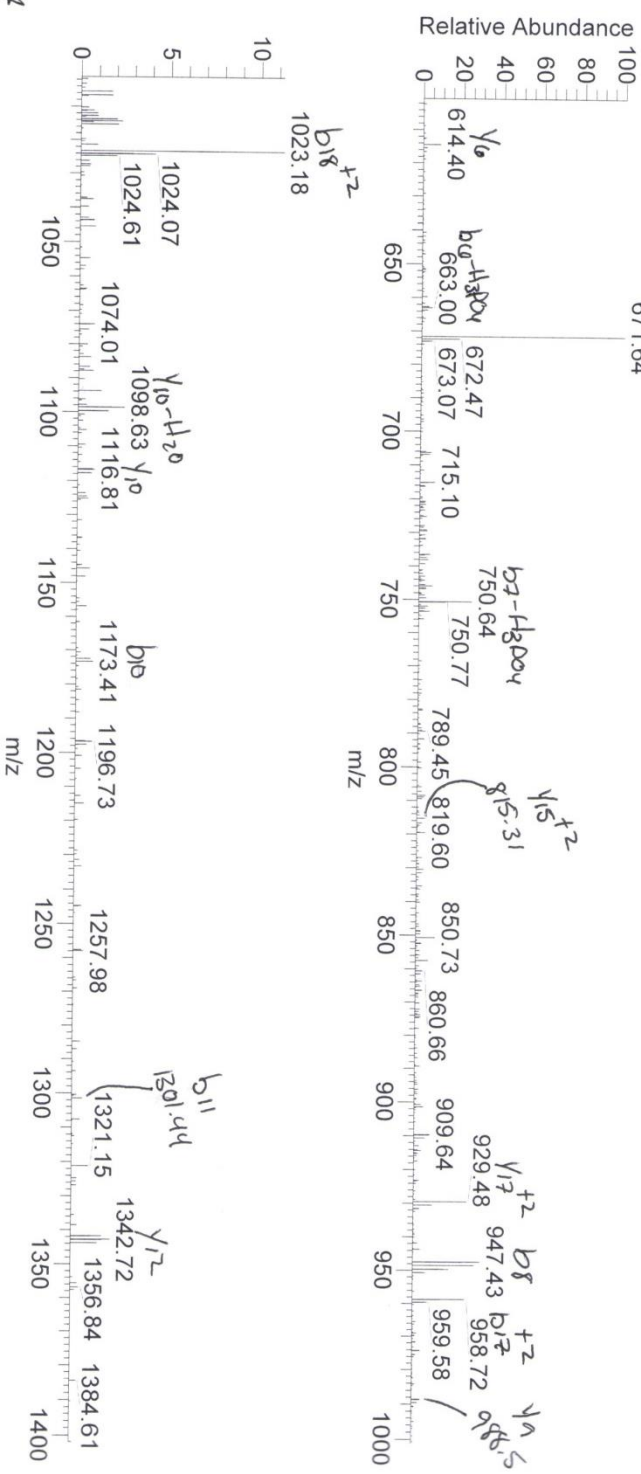
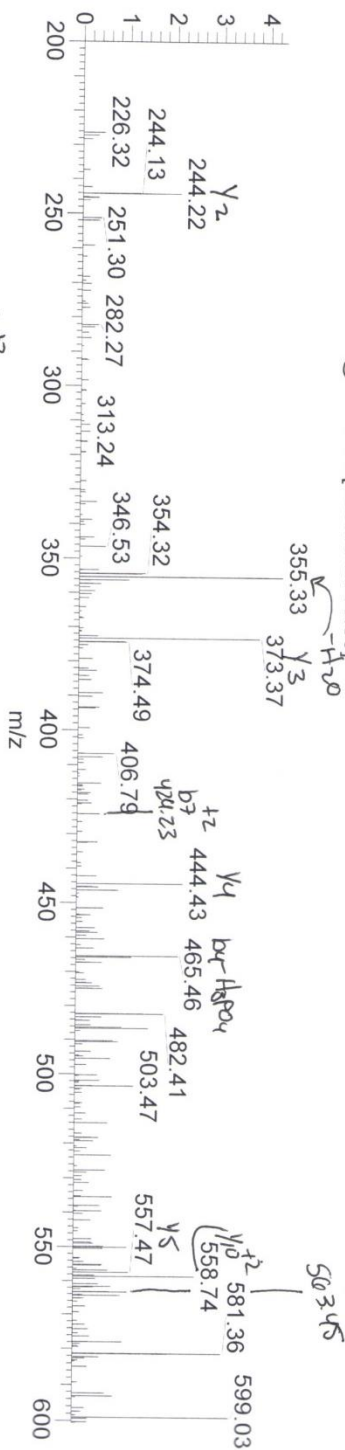


sh_2796_KH_081312_ASEF_ACN_digest
2E06

8/13/2012 4:42:04 PM
5.000000

H16MPTSVPAQAV6LAEPRK

sh_2796_KH_081312_ASEF_ACN_digest #19009-19081 RT: 45.08-45.13 AV: 2 NL: 4.51E4
F: ITMS + c NSI d Full ms2 763.76@cid35.00 [200.00-2000.00]



APPENDIX E

CURRICULUM VITAE

Kelly M. Hines, Ph.D.

Vanderbilt University Department of Chemistry
7330 Stevenson Center, Station B 35-1822
Nashville, TN 37235
Cell: (954)-540-3005
kelly.m.hines@vanderbilt.edu

EDUCATION

Vanderbilt University
Ph.D. in Chemistry

Nashville, TN
Expected: 05/2014

Dissertation: Biomolecular Signatures of Disease via Ion Mobility
and Mass Spectrometry Techniques

Mentor: John A. McLean

University of Florida
Bachelor of Science in Chemistry, *Cum Laude*

Gainesville, FL
05/2009

RESEARCH EXPERIENCE

Graduate Research Assistant

- Designed bottom-up proteomic experiments incorporating strongly denaturing preparations and multiple proteases to characterize the phosphorylation of Asef2. Achieved 95% sequence coverage and 100% coverage of serine, threonine and tyrosine residues with a liquid chromatography-tandem MS (LC-MS/MS) approach and multiple digestion and denaturation strategies. Identified six sites of phosphorylation. Mutation analysis of pS106 revealed impacts on cell adhesion turnover and cell migration.
- Investigated the molecular features distinguishing diabetic from non-diabetic wound fluids isolated from a model system of diabetic wound healing using an untargeted ion mobility-mass spectrometry (IM-MS) platform. Performed targeted

analyses for the most statistically significant molecular features, including additional ultra-performance liquid chromatography (UPLC) separations, MS/MS experiments and validation with standards.

- Investigated the metabolic profiles of human breast cancer tissues in an untargeted manner using polar extraction strategy and UPLC-IM-MS/MS. Performed multivariate statistical analyses to reveal differentially expressed features. Characterized molecular features with MS/MS to generate tentative identifications. Distinguished cancerous and noncancerous breast tissues based on biomolecular signatures characteristic of metabolism in cancer.

PUBLICATIONS

6. **Kelly M. Hines**, Billy R. Ballard, Dana M. Marshall and John A. McLean, "UPLC-IM-MS/MS of Tissue Extracts to Distinguish Cancerous and Non-cancerous Breast Diseases," In preparation for *Analytical Chemistry*, 2014.
5. J. Corey Evans, **Kelly M. Hines**, Jay G. Forsythe, Begum Erdogan, Mingjian Shi, Salisha Hill, Kristie L. Rose, John A. McLean and Donna J. Webb, "Phosphorylation of Serine 106 in Asef2 Regulates Cell Migration and Adhesion Turnover," Submitted to *Journal of Proteome Research*, February 2014.
4. Sarah M. Stow, Nichole M. Lareau, **Kelly M. Hines**, C. Ruth McNees, Cody R. Goodwin, Brian O. Bachmann, and John A. McLean, "Structural separations for natural product characterization by ion mobility-mass spectrometry: Fundamental theory to emerging applications," Invited chapter for inclusion in "Natural Products Analysis: Instrumentation, Methods and Applications," Vladimir Havlicek and Jaroslav Spizek, Eds. *John Wiley & Sons, Inc.* In press 2013.
3. **Kelly M. Hines**, Samir Ashfaq, Jeffrey M. Davidson, Susan R. Opalenik, John P. Wikswa and John A. McLean, "Biomolecular Signatures of Diabetic Wound Healing by Structural Mass Spectrometry," *Analytical Chemistry*, 2013, 85 (7), 3651-3659.

06/2013 Highlighted in an interview article in *The Analytical Scientist* "Wound Healing in Diabetes." Circulation of 52,000 electronic copies and 21,000 printed copies internationally.

2. **Kelly M. Hines**, Jeffrey R. Enders, and John A. McLean, "Multidimensional Separations by Ion Mobility-Mass Spectrometry," Invited chapter for inclusion in "Encyclopedia of Analytical Chemistry (online)," Robert Myers and David Muddiman, Eds. *John Wiley & Sons, Ltd.* December 2012.

1. Alexander P. Lamers, Mary E. Keithly, Kwangho Kim, Paul D. Cook, Donald F. Stec, **Kelly M. Hines**, Gary A. Sulikowski, and Richard N. Armstrong, "Synthesis of Bacillithiol and the Catalytic Selectivity of FosB-Type Fosfomycin Resistance Proteins," *Organic Letters*, 2012, 14 (20), 5207-5209.

PRESENTATIONS

6. **Kelly M. Hines**, Billy R. Ballard, Dana R. Marshall, and John A. McLean, *UPLC-IM-MS/MS Profiling of Metabolites in Human Breast Cancer Tissues*, Vanderbilt Institute of Chemical Biology Student 2013 Symposium, Nashville, TN (August 2013).
5. **Kelly M. Hines**, Billy R. Ballard, Dana R. Marshall, and John A. McLean, *Metabolomic Profiling of Human Breast Cancer Tissues by UPLC-IM-MS*, 61st American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics, Minneapolis, MN (June 2013).
4. **Kelly M. Hines**, Billy R. Ballard, Dana R. Marshall, and John M. McLean, *Metabolomic Profiling of human breast cancer tissues by UPLC-IM-MS*, Vanderbilt Institute of Chemical Biology 10th Anniversary Celebration, Nashville, TN (March 2013).
3. **Kelly M. Hines**, Samir Ashfaq, Jeffrey M. Davidson, Susan R. Opalenik, and John A. McLean, *Identification of diabetic wound healing biomolecular signatures by UPLC-IM-MS*, 60th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics, Vancouver, BC, Canada (May 2012).
2. **Kelly M. Hines**, Samir Ashfaq, Jeffrey M. Davidson, Lily Wang, Susan R. Opalenik, and John A. McLean, *Analysis of wound fluid by IM-MS for biomolecular signatures of diabetic wound healing*, 59th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics, Denver, CO (June 2011).
1. **Kelly M. Hines**, Samir Ashfaq, Jeffrey M. Davidson, Susan R. Opalenik, and John A. McLean, *Analysis of wound fluid by IM-MS for protein signatures of wound healing*, Vanderbilt Institute of Chemical Biology Student 2010 Symposium, Nashville, TN (August, 2010).

TEACHING EXPERIENCE

- **Teaching Assistant**, Organic Chemistry Laboratory, August 2009 – August 2010, August 2011 – December 2011.

- **Teaching Assistant**, Instrumental Analysis Laboratory, August 2010 – December 2010.
- **Teaching Assistant**, Forensic Chemistry Laboratory, January 2011 – May 2011, January 2012 – May 2012.

FUNDING PROPOSAL CURRENTLY UNDER REVIEW

Title: Geospatial and longitudinal mapping of antigenic drift in H1N1 HA by top-down MS
Agency: NIH/NIAID, Ruth L. Kirschstein NRSA Postdoctoral Fellowship (PA-11-113)
Period: 07/01/2014-06/31/2017
Notes: Pending IRG Review (Scheduled: March 20, 2014)

PROFESSIONAL MEMBERSHIPS

The American Chemical Society
The American Society for Mass Spectrometry

REFERENCES

John A. McLean

Dissertation Advisor
Stevenson Associate Professor of Chemistry
Vanderbilt University
7330 Stevenson Center
Station B 35-1822
Nashville, TN 37235
(615) 322-1195
john.a.mclean@vanderbilt.edu

Donna J. Webb

Research Collaborator and Committee Member
Associate Professor of Biological Sciences
Assistant Professor of Cancer Biology
Vanderbilt University
U6211 MRBIII
Station B 35-1634
Nashville, TN 37235
(615) 936-8274
donna.webb@vanderbilt.edu

John P. Wikswo

Research Collaborator

Gordon A. Cain University Professor

A.B. Learned Professor of Living State Physics

Director, Vanderbilt Institute for Integrative Biosystems Research and Education

Professor of Biomedical Engineering

Professor of Molecular Physiology and Biophysics

Professor of Physics

Vanderbilt University

6301 Stevenson Center

Station B 35-1807

Nashville, TN 37235

(615) 343-4124

john.wikswo@vanderbilt.edu

David E. Cliffl

Committee Member

Associate Professor of Chemistry

Vanderbilt University

7330 Stevenson Center

Station B 35-1822

Nashville, TN 37235

(615) 322-2861

d.cliffl@vanderbilt.edu