ABSTRACT

WILLIAMS JR., DENNIS KEITH. Exploring Fundamental Aspects of Proteomic Measurements: Increasing Mass Measurement Accuracy, Streamlining Absolute Quantification, and Increasing Electrospray Response. (Under the direction of David Charles Muddiman).

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) offers unparalleled performance in terms of resolving power and mass measurement accuracy. In order to realize the highest achievable mass measurement accuracy for a given FT-ICR MS system, frequency shifts due to space-charge effects must be accounted for via external or internal calibration methods. Herein, a dual electrospray ionization source was coupled to a hybrid quadrupole Fourier transform ion cyclotron resonance mass spectrometer and utilized to incorporate internal calibrant to yield high mass measurement accuracy results. Frequency shifts from space-charge effects were also counteracted by accounting for total ion population and relative ion population on a MALDI-FT-ICR MS and a hybrid LTQ-FT-ICR MS equipped with automatic gain control. This was achieved through the use of multiple linear regression and artificial neural networks. These experiments resulted in mean mass measurement accuracies in the parts-per-billion range.

C-reactive protein (CRP) is an important clinical marker for inflammation, atherosclerosis, and has also been observed to be upregulated in patients with epithelial ovarian cancer (EOC). This dissertation contains results from a protein cleavage isotope dilution mass spectrometry method developed for the absolute quantification of CRP from human plasma. A total of 110 human plasma samples were analyzed including 54 samples from patients with EOC. The results were

compared to a CLIA certified ELISA test which showed high correlation but different absolute values, which suggested different reference ranges for different analytical techniques. In addition, a correlation was observed between the stage of diagnosis of cancer and CRP concentrations.

Chemical tags have long been utilized with mass spectrometry for a multitude of purposes, extending the effectiveness of the measurements dramatically. Electrospray ionization has been shown to preferentially ionize more hydrophobic species. Four new iodoacetamide derivatives, which react with the amino acid cysteine, were reacted with three peptides to determine their ability to increase electrospray response. This resulted in increases up to 2000-fold compared to alkylation of peptides with iodoacetamide. The combination of these results will aid future studies to identify and quantify proteins and peptides which contain cysteine and can be expanded to all peptides using amine-specific chemistry.

Exploring Fundamental Aspects of Proteomic Measurements: Increasing Mass Measurement Accuracy, Streamlining Absolute Quantification, and Increasing Electrospray Response

by
Dennis Keith Williams, Jr.

A dissertation submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

Chemistry

Raleigh, North Carolina

May 5, 2009

APPROVED BY:

_____	_____
David C. Muddiman Professor, Chemistry Committee Chair	Edmond F. Bowden Professor, Chemistry
_____	_____
Kenneth W. Hanck Professor, Chemistry	Lin He Assistant Professor, Chemistry

DEDICATION

This dissertation is dedicated to my parents, Dr. Dennis and Mrs. Sharon Williams, and my sister Ms. Carrie Williams. Without their influence, guidance, support, love, and friendship none of this would have been possible.

BIOGRAPHY

Dennis Keith Williams, Jr. was born on October 13, 1982 in Charlotte, NC to Dennis and Sharon Williams. He has one younger sister, Carrie who was born in 1988. He lived in Matthews, NC until 1991 when his family moved to the town of Weddington, NC. From a young age he was fascinated by cars (sports cars in particular) and the physical sciences. He played many sports as a youth but football became his favorite sport. He played offensive line for the football team at Sun Valley Middle and High Schools from 6th grade until graduation. He also held the bench press record at Sun Valley High School (420 lbs.). After graduating from high school he attended Hampden-Sydney College where he was a member of the football team and Master Alchemist (President) of his fraternity (ΑΧΣ). He graduated with a Bachelor of Science degree in Chemistry, under the direction of Dr. H.J. Sipe, Jr. During his studies at Hampden-Sydney College his interests in the physical sciences and specifically chemistry grew and he decided to attend graduate school. He has since been working to earn his Ph.D. in Analytical Chemistry under the direction of Dr. David C. Muddiman.

ACKNOWLEDGMENTS

I would like to acknowledge several people for their help along the way in my graduate school career. My advisor, Dr. David Muddiman, provided a lot guidance which helped me become a better scientist and person throughout my time in his research group. Seeing how hard he works to provide funding for his research group helped to inspire me to work hard throughout my graduate career. In addition, Dr. Adam Hawkridge has also provided much guidance along the way: talking about how to properly care for the superconducting magnets to discussions on nano-LC, it has all been greatly appreciated.

Also, I would like to thank Dr. Ed Bowden who aided greatly in my decision to come to NC State for graduate school. If it had not been for some of our discussions I may have not ended up at this institution!

My fellow graduate students in the Muddiman research group have also helped me along the path to graduation, especially Michael Bereman and Brent Dixon. Both of these guys have been great contemporaries to work with while on the job and friends both in and out of the laboratory.

Last but not least, I want to thank my family for their love and support throughout this process. My parents provided guidance and inspiration for me to perform at the top level and to always push myself for which I am very appreciative. Also, my sister has always been a good friend for me to lean on, which during my time in graduate school has been a wonderful thing to have.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF REACTION SCHEMES

LIST OF FIGURES

LIST OF PUBLICATIONS

The research in this dissertation has resulted in the following publications:

Peer reviewed manuscripts:

1. **Williams, D.K.**, et. al., Utilizing Artificial Neural Networks in MATLAB to Achieve Parts-per-Billion Mass Measurement Accuracy with a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer*. J. Am. Soc. Mass Spectrom*, 2009. Accepted.

2. **Williams, D.K.** and D.C. Muddiman, Absolute quantification of C-reactive protein in human plasma derived from patients with epithelial ovarian cancer utilizing protein cleavage isotope dilution mass spectrometry*. J Proteome Res*, 2009. **8**. 1085-90.

3. **Williams, D.K.**, et al., Synthesis, characterization, and application of iodoacetamide derivatives utilized for the ALiPHAT strategy*. J. Am. Chem. Soc.*, 2008. **130**. 2122-+.

4. **Williams, D.K.**, et al., Calibration laws based on multiple linear regression applied to matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry*. J. Mass Spectrom.*, 2008. **43**. 1659-1663.

5. **Williams, D.K.**, et al., Dual electrospray ion source for electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer*. Anal. Chem.*, 2007. **79**. 7916-7919.

6. **Williams, D.K.** and D.C. Muddiman, Parts-per-billion mass measurement accuracy achieved through the combination of multiple linear regression and automatic gain control in a Fourier transform ion cyclotron resonance mass spectrometer*. Anal. Chem.*, 2007. **79**. 5058-5063.

7. **Williams, D.K.**, A.M. Hawkridge, and D.C. Muddiman, Sub parts-per-million mass measurement accuracy of intact proteins and product ions achieved using a dual electrospray ionization quadrupole Fourier transform ion cyclotron resonance mass spectrometer*. J. Am. Soc. Mass Spectrom.*, 2007. **18**. 1-7.

Abstracts for presentation:

1. **Poster Presentation.** "Absolute Quantification of C-Reactive Protein in Plasma Utilizing Isotope Dilution Mass Spectrometry", American Society for Mass Spectrometry Meeting, Denver, CO, June, 2008 D. Keith Williams, Jr., Adam M. Hawkridge, and David C. Muddiman.

2. **Poster Presentation.** "Application of Iodoacetamide Derivatives Utilized to Increase Ion Abundance through the ALiPHAT Strategy", American Society for Mass Spectrometry Meeting, Denver, CO, June, 2008 David C. Muddiman, D. Keith Williams Jr., Corey W. Meadows, Ibrahim D. Bori, Adam M. Hawkridge, and Daniel L. Comins.

3. **Oral Presentation.** "Coupling Automatic Gain Control and Calibration Laws to Achieve Parts-Per-Billion Mass Measurement Accuracy Utilizing a FT-ICR Mass Spectrometer", Southeastern Regional Meeting of the American Chemical Society, Greenville, SC, October, 2007 D. Keith Williams, Jr., Adam M. Hawkridge, and David C. Muddiman.

4. **Poster Presentation.** "Investigation of Mass Measurement Accuracy of Intact Proteins and Product Ions using Hybrid Fourier Transform Ion Cyclotron Resonance Mass Spectrometry", American Society for Mass Spectrometry Meeting, Indianapolis, IN, June, 2007 D. Keith Williams, Jr., Adam M. Hawkridge, and David C. Muddiman.

5. **Oral Presentation.** "Parts Per Billion Mass Measurement Accuracy Achieved by Automatic Gain Control and Calibration Laws Utilizing FT-ICR Mass Spectrometry", Local Meeting of the American Chemical Society, Durham, NC April, 2007 D. Keith Williams, Jr. and David C. Muddiman.

6. **Oral Presentation.** "Achieving high mass measurement accuracy of intact proteins and product ions utilizing a dual electrospray ionization quadrupole Fourier transform ion cyclotron resonance mass spectrometer", Southeastern Regional Meeting of the American Chemical Society, Augusta, GA November, 2006 D. Keith Williams, Jr., Adam M. Hawkridge, and David C. Muddiman.

CHAPTER 1

Introduction

1.1 Brief Introduction to Proteomics

The word proteome was coined in the 1990's by Marc Wilkins and was defined as "the protein complement of the genome."[1] The genome is defined as "a complete single set of the genetic material of a cell or of an organism; the complete set of genes in a gamete."[2] Genomics is the study of the genome, as such proteomics is the study of the proteome; more specifically the currently accepted definition of proteomics is "Proteomics is the identities, quantities, structures, and biochemical and cellular functions of all proteins in an organism, organ, or organelle, and how they vary in space, time, and physiological state."[3] While much effort at the time was being spent on decoding the human genome, Wilkins felt that it was of utmost importance to identify the proteins expressed by the genome because the genome is static while the proteome is dynamic. In other words, to understand biological changes in an organism like disease progression it was necessary to begin to understand the proteome. The genome and proteome are related by the central dogma introduce by Crick in 1970.[4] Crick's idea of central dogma proposed that DNA, RNA, and proteins were all related by general and special transfers to allow for a biological specimen to function. The DNA, composed of four nucleotides (adenine, guanine, cytosine, thymine), is replicated and transcribed into messenger RNA (mRNA) by the enzyme RNA polymerase. The mRNA is composed of the

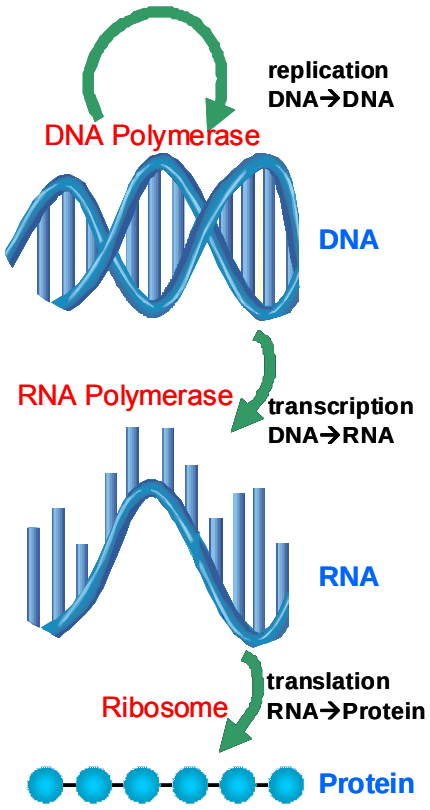


Figure 1.1 A diagram of the central dogma which shows how DNA, RNA, and proteins are related.

same building blocks as DNA, except uracil is substituted for thymine. Proteins are then created by the translation of mRNA by ribosomes, located in the cytosol and on the rough endoplasmic reticulum of the cell. A general overview of this process is illustrated in **Figure 1.1**. Translation occurs by the "reading" of triplets of nucleotides on mRNA termed codons. Each codon specifies a specific amino acid, and as the amino acids are linked together, forms a protein. Also important to note are the start (AUG) and stop (UAG, UGA, UAA) codons that instruct the translation process to begin or end.

1.1.1 Amino Acid and Protein Properties

There are twenty standard amino acid structures which serve as the building blocks of proteins. Proteins are the cellular vehicle from which genetic information expresses itself in living things.[5] Amino acids have a general structure which includes an amino and carboxyl group which are both attached to a central carbon termed the α-carbon. The general amino acid structure is illustrated in **Figure 1.2**. The chemical

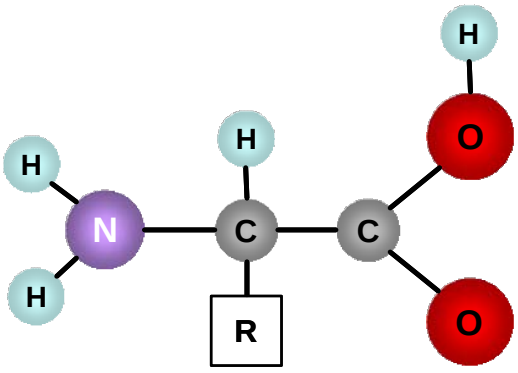


Figure 1.2 A general structure of an amino acid is illustrated. Different R-groups give each amino acid its own unique chemical properties.

constituent which gives each amino acid its own unique characteristics is the R-group which is also attached to the α-carbon, and can contrast in size from a hydrogen atom (glycine) to a complex heterocyclic group in tryptophan. These side chains can cause the amino acid to act as a weak acid, weak base, hydrophobe (nonpolar side chain), or hydrophile (polar side chain). **Figure 1.3** shows the twenty standard amino acids with the nine *essential amino acids*[6] for humans highlighted.

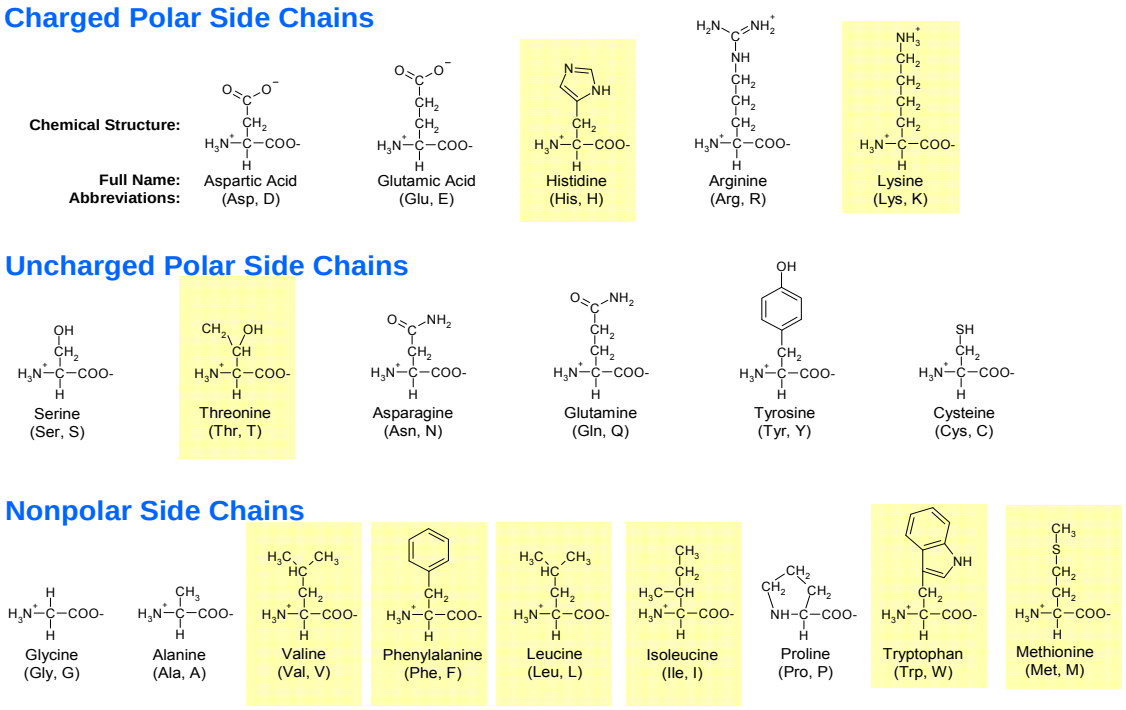


Figure 1.3 The structures, names, and abbreviations for the twenty amino acids are illustrated and divided into three subgroups. Nine essential amino acids for humans highlighted.

Since amino acids have both an amino and carboxylic acid functional group, they can function as both acids and bases at the same time; this is only true in either the solid phase or the solution phase, but not in the gas phase.[7] The pH at which a particular amino acid is neutral (equal number of ammonium groups with a positive

charge and carboxylic groups with negative charge) is called the isoelectric point. Zwitterion is a term that describes the ability to act in this way. In addition to the isoelectric point another chemical descriptor of amino acids is the grand average of hydropathy (GRAVY) score.[8] This describes how hydrophobic or hydrophilic a particular amino acid is. For a given peptide or protein the GRAVY scores for the individual amino acids can be average to determine whether it is hydrophobic (positive) or hydrophilic (negative). A summary of these values and other physical characteristics of the 20 natural amino acids can be found in **Table 1.1**.

Table 1.1 A summary of the chemical properties of the twenty amino acids is shown.

Amino Acid	Monoisotopic Mass[1]	pKa (C-terminus)[2]	pKa (N-terminus)[2]	pKa (Side chain)[2]	pI[2]	Hydropathy (Kyte-Doolittle)[3]	Average % Occurrence[2]
Aspartic Acid (D)	115.0270	1.99	9.90	3.90	2.94	-3.5	5.3
Glutamic Acid (E)	129.0426	2.10	9.47	4.07	3.08	-3.5	6.3
Histidine (H)	137.0589	1.80	9.33	6.04	7.68	-3.2	2.3
Arginine (R)	156.1011	1.82	8.99	12.48	10.74	-4.5	5.1
Lysine (K)	128.0950	2.16	9.06	10.54	9.80	-3.9	5.9
Serine (S)	87.0320	2.19	9.21		5.70	-0.8	6.8
Threonine (T)	101.0477	2.09	9.10		5.60	-0.7	5.9
Asparagine (N)	114.0429	2.14	8.72		5.43	-3.5	4.3
Glutamine (Q)	128.0586	2.17	9.13		5.65	-3.5	4.3
Tyrosine (Y)	163.0633	2.20	9.21	10.46	6.33	-1.3	3.2
Cysteine (C)	103.0092	1.92	10.70	8.37	5.14	2.5	1.9
Glycine (G)	57.0215	2.35	9.78		6.06	-0.4	7.2
Alanine (A)	71.0371	2.35	9.87		6.11	1.8	7.8
Valine (V)	99.0068	2.29	9.74		6.02	4.2	6.6
Phenylalanine (F)	147.0684	2.33	9.74		6.04	2.8	3.9
Leucine (L)	113.0841	2.32	9.76		6.04	3.8	9.1
Isoleucine (I)	113.0841	2.13	9.28		5.71	4.5	5.3
Proline (P)	97.0053	1.95	10.64		6.30	-1.6	5.2
Tryptophan (W)	186.0793	2.20	9.31		5.76	-0.9	1.4
Mehtionine (M)	131.0405	2.46	9.41		5.93	1.9	2.2

[1] National Institute of Standards and Technology (www.nist.gov)
[2] Voet D., Voet J. G. *Biochemistry*, 2nd ed., 1995, pp 58-59.
[3] Kyte J., Doolittle R.F., *Journal of Molecular Biology*, 1982, 157, 105-132.

1.1.2 Mass Spectrometry and Proteomics

Mass spectrometry's ability for molecular specific identification of analytes as well as its high sensitivity makes it an ideal tool to be utilized in proteomics.[9, 10] There are two main methodologies that are used to interrogate the properties of proteins. These methods are referred to as *top-down* and *bottom-up proteomics*. The top-down approach depends on the determination of an accurate intact mass and fragmentation species of the intact protein to yield identification.[11] The bottom-up approach first digests the protein with an enzyme, such as trypsin, before introduction into the mass spectrometer. The resulting peptide masses are measured and then subjected to MS/MS for sequence determination, from which an identity is established.[12] These types of studies are generally utilized to search a complex mixture to identify particular proteins and/or peptides that could be of possible interest (i.e. *biomarkers*).

Once a top-down or bottom-up investigation has identified a protein or peptide of interest it is then possible to use one of many techniques to quantify this target. If the goal is only for the *relative quantification*, which further aids in the elucidation of candidate *biomarkers*, there are several techniques that can be implemented: SILAC[13], ICAT[14], cleavable ICAT[15], $^{16}O/^{18}O$ labeling[16], iTRAQ[17], and IDBEST[18] are some of the more commonly utilized methods.[19] To confirm the status of a *biomarker absolute quantification* must be performed. *Absolute quantification* is performed utilizing a method known as *protein cleavage isotope dilution mass spectrometry* (*PC-IDMS*). This method was developed by Barr and was later given a

different acronym by Gygi and co-workers.[20, 21] This technique involves exploiting a stable isotope labeled (SIL) version of the peptide target as the internal standard. This serves two purposes: limiting any bias that might be present in the ionization process and ensuring identification of the correct peak in the LC-MS chromatogram since the natural peptide and the SIL peptide should co-elute. The quantification is completed by comparing the areas of both natural and SIL peptide in the LC-MS chromatogram. Several studies have been completed in this area of research within recent years.[22-26]

1.2 Electrospray Ionization

Electrospray ionization (*ESI*) was first coupled with mass spectrometry by Professor John B. Fenn in 1984.[27] In 2002 this research enabled Professor Fenn to receive the Nobel Prize. Initial research on electrospray reaches as far back as the late 1700's with the work of Bose[28], and undoubtedly to the early 1900's with the work of Zeleny.[29, 30] However, much early work with using electrospray to create ions can be attributed to Malcom Dole.[31, 32] The intention of this research was to produce gas phase macro-ions, of which experimental evidence was provided for ionization of zein and lysozyme.[33, 34] Many of the *ESI* advancements achieved up to this point and those to be developed in the future can thank the initial experiments mentioned above.

ESI is known as a "soft" ionization source, which means that it does not fragment the analyte during the ionization process to the extent of other ionization

methods. *ESI* is achieved by the application of a high voltage to a conductive solution flowing through a capillary. The end of the capillary is usually pulled to a tip with a smaller diameter than the rest of the capillary; it is here where the Taylor cone forms and the analyte begins its journey from liquid to gas phase ions.

There are two methods that scientists believe might be plausible for this journey to reach gas phase ions. One such mechanism is referred to as the charged residue model. This was developed by Dole and co-workers and presumes that solvent evaporation and an increase of surface charge density occurs until the Rayleigh limit is reached.[32] The Rayleigh limit is shown in **Equation 1.1**. At this point coulombic explosions take place which creates offspring droplets through the fission of the parent droplets. The offspring droplets typically contain 15% of the charge and 2% of the mass of the parent droplet.[35, 36] This continues until an "ultimate droplet" is created with one analyte within. The analyte then enters the gas phase as an ion when the remaining solvent is evaporated.[32]

Equation 1.1 $q = \sqrt{64\pi^2\varepsilon_0\gamma r^3}$

The second mechanism was proposed by Iribarne and Thompson and is known as the ion evaporation model. Parent droplets, in this model, undergo several coulombic explosions inherently creating smaller and smaller offspring droplets. This causes an increase of the electric field at the droplet's surface until ions are evaporated out of the droplet into the gas phase.[37, 38] The ion evaporation model is presently the more accepted theory for the electrospray process, except for

very large (> 100 kDa) molecules; however, it is important to emphasize the creation of gas phase ions through the electrospray process is still not fully understood.

1.3 Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry

Fourier transform-ion cyclotron resonance mass spectrometry was first reported by Comisarow and Marshall in 1974.[39, 40] The *resolving power* and *mass measurement accuracy* provided by this technology is paramount in the identification of unknowns (e.g. *biomarkers*). The performance this technology provides has only become almost equaled in recent years by the Orbitrap mass analyzer, which still relies on Fourier transform.[41]

Following ionization, charged analyte enters the orifice of the mass spectrometer. In *electrospray ionization* only a very small amount, 0.01% of the ionized molecules, actually reach the detector.[42] The ions that enter the instrument are collected in either a quadrupole or a linear quadrupole trap depending on whether a hybrid QFT-ICR or hybrid LTQ-FT-ICR mass spectrometer is being utilized.[43, 44] After this external accumulation of ions, they are transferred to the ICR cell, which is located inside of a superconducting magnet, through ion optics. Once in the ICR cell, ions assume a natural cyclotron motion because of the *Lorentz force* which is a result of charged particles residing in a magnetic field.[45, 46] Ions with positive charge will rotate in a clockwise manner while negatively charged ions will rotate counterclockwise, which follows the right hand rule.

This natural cyclotron motion by itself is not enough for detection by the ICR cell. The ions must be excited by one of a variety of waveforms. This excitation is performed for two main reasons: 1) increases the ions cyclotron radius (this radius is referred to as the *post-excite radius*) and 2) allows for all ions to reach phase coherence. Both of these events are critical for achieving the optimal induced image current on the detection plates. Broadband excitation occurs by differentially applying an rf voltage containing a band of frequencies matching the cyclotron frequencies of the *m/z* range desired to be detected. **Figure 1.4** shows the two excitation plates as well as the two detection plates, in a cylindrical ICR cell.[46] The image current observed by the detection plates results in the instrument recording a time domain which contains the frequencies of all the ions contained within the ICR cell. A fast Fourier transform is applied to convert the time domain to a frequency domain. This frequency domain is then converted to a mass spectrum by one of many calibration equations. The most elementary relationship between *cyclotron frequency* and mass-to-charge ratio is shown in **Equation 1.2**. This equation relates *cyclotron frequency* (ω) to the charge (q), magnetic field strength (β_0), and mass of the ion being measured (m).

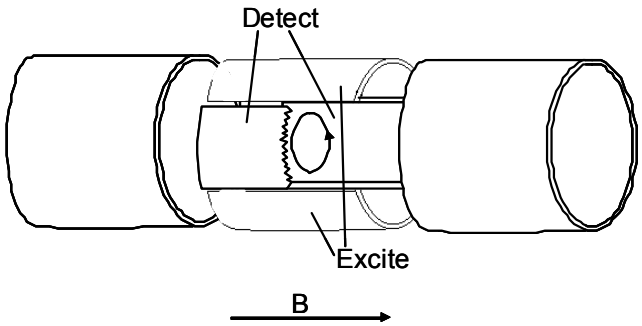


Figure 1.4 An ICR cell is illustrated with the detection and excitation plates identified.

Equation 1.2: $\omega = \frac{qB_0}{m}$

As illustrated in **Equation 1.2**, *cyclotron frequency* is linearly related to magnetic field strength. However, this is not the only characteristic of FT-ICR MS measurements that a stronger magnet can influence. *Resolving power*, data acquisition speed, and upper mass limit for *peak coalescence* also increased linearly with increased magnetic field strength.[47] Furthermore, some parameters increase quadratically with increase in magnetic field strength: upper mass limit due to trapping potential, maximum ion trapping duration, and maximum number of trapped ions.[45, 47]

1.3.1 Excitation

Ion cyclotron motion induced by a magnetic field alone is not useful in making mass spectrometric measurements. A main contributor to this fact is that the thermal cyclotron radius is too small to be detected by the detection plates of the ICR cell.

Equation 1.3 $$r_{thermal} = \frac{1}{q\beta_0}\sqrt{2kmT}$$

The thermal cyclotron radius is shown in **Equation 1.3**, and is related to the mass (m, kg), temperature (T, K), charge (q, C), magnetic strength (β_0, T), and Boltzmann's constant (k, m^2 kg s^{-2} K^{-1}). In order for measurements to be recorded by the ICR cell the ions must be excited. Excitement of ions in FT-ICR MS can have different results: (i) it accelerates ions coherently to a larger, and thus detectable radius; (ii) to increase the kinetic energy above a threshold at which ion dissociation can take place if collided with an inert gas; and (iii) to accelerate ions to a cyclotron radius larger than that of the ICR cell such that ions are consequently removed from

the instrument. For our purposes, the result of ions reaching a coherent and larger and detectable radius is most important for further discussion.

Excitement occurs during the application of a spatially uniform electric field oscillating at or close to that of the *cyclotron frequency* of ions of particular *m/z* values. This rf voltage ($V_{p\text{-}p}$, V) is applied to the two excitation plates of the ICR cell (**Figure 1.4)** for a certain time period (T_{excite}, s). The *post-excite radius* of the ions also depends of the magnetic field (β_0, T) and distance between the excitation plates (d, m), as shown in **Equation 1.4**. One important fact to note about **Equation 1.4** is that the *post-excite radius* has no dependence on *m/z*; as a consequence ions of all *m/z* can be excited to an identical *post-excite radius* with an rf signal that has constant magnitude across the frequency range to be observed.

Equation 1.4 $$r_{post-excite} = \frac{V_{p-p} T_{Excite}}{2d\beta_0}$$

1.3.2 Detection

After excitation the ions present in the ICR cell move as a coherent packet with much closer proximity to the detection plates. As ion packets pass by the detection plates they induce a current, termed an image current. The instantaneous charge induced on one of the detection plates at a time (t) is represented by Q(t) in **Equation 1.5**. The induced current is also dependent on the number of ions (N), charge (q), *post-excite radius* ($r_{post\text{-}excite}$), *cyclotron frequency* (ω), and distance between the detection plates (d).[48] The charge, *cyclotron frequency*, and distance between the plates are constant for a particular ICR cell geometry and molecule.

Equation 1.5 $$Q(t) = \frac{-Nqr_{post-excite}\cos(\omega t)}{d}$$

The number of ions can be changed by the concentration of the sample being introduced into the mass spectrometer and external collection time, via a linear ion trap or accumulation hexapole in hybrid *FT-ICR* instruments.[44, 49] Problems with *peak coalescence* and *cyclotron frequency* shifts can occur if too many ions are introduced into the cell.[50-52] The induced charge can also be varied by adjusting the *post-excite radius* of the ions in the ICR cell, which instrument permitting, is the most efficient method for increasing the induced charge on the detection plates.[53]

The image current that is produced is represented by the derivative of **Equation 1.5** with respect to time. This image current is detected as a sinusoidal wave as a function of time which is Fast Fourier transformed to retrieve the individual cyclotron frequencies. The frequencies are then converted to *m/z* through the utilization of **Equation 1.2**. One of the benefits of *FT-ICR* is that broadband excitation and detection can be utilized to excite and detect a wide range of *m/z* values simultaneously, which is called multichannel advantage.[54] This advantage is a consequence of using a Fourier transform and describes the ability to yield a spectrum with N number of points in 1/N the amount of time to scan the spectrum one channel at a time.[45, 54]

One problem that has been noted in detection of ions with FT-ICR MS is signal dampening. Signal dampening occurs when ion clouds fall out of phase (i.e. lose coherence) with one another, which causes a loss in detectable current. This effect can be quantified by τ, which is defined by the time it takes for the signal to reach 1/e of its original magnitude.[55] This phenomenon can also result from excess

pressure in the ICR cell and ion-ion and ion-neutral collisions; however, this method of dampening has been addressed in recent years by improvement in vacuum systems for application with FT-ICR MS. However, loss of coherence of the ion cloud can be affected by the density of the ion cloud.[55, 56]

It has been previously reported that an increased ion cloud density results in larger rotational frequencies. This offsets the shearing forces from magnetic and electric field inhomogeneities that adversely affect the ion cloud.[57] The rotational frequency of the cloud is a consequence of Coulombic interactions between ions of the identical *m/z* (i.e. same ion cloud), which results in the ion cloud revolving about its symmetry axis.[57]

1.3.3 Space-Charge Effects and Mass Measurement Accuracy

FT-ICR MS measures the *cyclotron frequency* of ions in an ICR cell to which trapping voltages are applied. The ICR cell is situated inside of a superconducting magnet in a region where there is high magnetic field homogeneity. These all create forces which act upon the ions and can perturb the observed frequency.[58, 59] Given that FT-ICR MS has high *resolving power*, the *m/z* values of ions can be determined very precisely because current electronics can record frequencies with many (>9) significant figures readily.[58] However, these high precision measurements are not sufficient by themselves; high *mass measurement accuracy* must be achieved in order to determine chemical composition of ions unambiguously. This means that all FT-ICR MS instrumentation must be calibrated to supply the end user with results of high accuracy.[58]

The idealized relationship between *cyclotron frequency* and *m/z* exemplified in **Equation 1.2** is only partially correct for ions in an ICR cell. This equation does not take into account other forces that are ongoing inside an ICR cell filled with ions. **Equation 1.6** introduced in 1983 by McIver and coworkers accounts for both magnetron motion of ions as well as *space-charge effects*.[58]

The first term in **Equation 1.6** is the unperturbed *cyclotron frequency*, as previously seen in **Equation1.2**. The second term in **Equation 1.6** is a correction for magnetron motion of ions. The magnetron motion is a result of the trapping voltages applied. The applied trapping voltage acts on the ions to produce an outward force that opposes the *Lorentz force* from the applied magnetic field, which creates magnetron motion.[45, 58, 59] The opposition to the *Lorentz force* is the reason magnetron motion decreases *cyclotron frequency*. For a given ICR cell geometry (α, cell geometry constant) (a, diameter of cell), applied trapping voltage (V), and magnetic field (β_0) the magnetron component is constant and can therefore be corrected by external calibration. For this correction trapping voltages must be the same between calibration spectrum and experimental spectra for this correction to occur.

Equation 1.6 $$\omega_{observed} = \frac{q\beta_0}{m} - \frac{2\alpha V}{a^2\beta_0} - \frac{q\rho G_i}{\varepsilon_0\beta_0}$$

The third term in **Equation 1.6** is a correction for *space-charge* induced frequency shifts. *Space-charge effects* occur from coulombic repulsions between ions in the ICR cell; these repulsions influence the detected frequency and as such need to be corrected for in order for the instrument to produce high accuracy data.

Space-charge effects also oppose the *Lorentz force* in the ICR cell; this causes a decrease in *cyclotron frequency* because by opposing the *Lorentz force* the effective magnetic field is decreased. Charge (q) and ion cloud geometry (G_i) have influence on the magnitude of the *space-charge effect*, but for a given set of experiments these two variables should remain approximately constant. However, the ion density (ρ) of the ICR cell can fluctuate; as such the space charge in the ICR cell is largely a function of the ion population.[45, 58, 59] The larger the ion population the larger the *space-charge effect* will be and as a consequence *mass measurement accuracy* will suffer. *Space-charge effect* can be compensated by external calibration to yield high *mass measurement accuracy* data if the ion population in both the calibration spectrum and experimental spectra are the same, which is the purpose of the recently developed *automatic gain control*.[44] Both internal and external calibration methods to correct for *space-charge effects* will be discussed in detail in later chapters of this dissertation.

1.4 Synopsis of Completed Research

The research described in this dissertation covers a wide variety of topics dealing with proteomic measurements. Methods for increasing *mass measurement accuracy* (*MMA*) by accounting for *space-charge effects* within *FT-ICR* instruments without *automatic gain control* as well as an instrument equipped with *automatic gain control* were investigated.[60-62] A dual electrospray source, first used for the improvement of *mass measurement accuracy*[60, 63], was modified to better perform

electron-transfer dissociation on a highly modified hybrid LTQ-Orbitrap mass spectrometer.[64] After the ability to increase *MMA* across a variety of *FT-ICR* platforms was completed, a method for the quantification of C-Reactive Protein (CRP), an *acute-phase protein biomarker*, was developed.[26] This method was free of sample clean up steps that many previous *absolute quantification* methods for human plasma have utilized.[65] In addition, newly developed cysteine specific chemical tags were tested on a variety of peptides to observe their ability to increase their instrumental response.[66] This was based on the tags' ability to increase the *hydrophobicity* of the analyte. This could aid future studies of low abundant cysteine containing peptides and proteins as well as be extended to other chemistries (e.g., amine) and biomolecules.

The ability of FT-ICR MS to evaluate analyte molecular masses with high *mass measurement accuracy* (*MMA*) is crucial for its utilization in proteomic experiments. The application of higher *MMA* techniques has the potential to increase the speed (throughput) of proteomic characterization and protein identification.[67-69] The more complex a sample is the higher instrumental demands are needed such as, *mass measurement accuracy*, speed of acquisition, resolution, and data-dependent acquisition.[69] **Chapter 2** describes an internal calibration method for both intact ions and *collision induced dissociation* product ions. This method employs a dual *electrospray ionization* source previously developed in the Muddiman laboratory.[63, 70] The use of the dual electrospray source enabled the introduction of calibrant ions simultaneously with analyte ions without adverse

influences such as ion suppression. The internal calibration yielded measurements in the part-per-billion range.

The same dual electrospray source was then modified and utilized as an ion source with the ability of applying voltages with opposite polarity to two different analyte solutions. **Chapter 3** details the research completed an a highly modified LTQ-Orbitrap mass spectrometer for which the dual electrospray enabled simultaneous introduction of a reagent necessary (9-anthracenecarboxylic acid) to implement *electron-transfer dissociation* (ETD) and peptides/proteins of interest to be interrogated with this technique of *tandem mass spectrometry*. This source was demonstrated to afford stable electrospray from both electrospray emitters, while enabling the voltage to remain constantly applied to both emitters. In previous sources for ETD the applied voltage was switched on and off because of the close proximity in space of the positive and negative emitter.[71, 72]

Improvement of *MMA* utilizing external calibration and *multiple linear regression* was the focus of **Chapter 4** and **Chapter 5. Chapter 4** applied external calibration laws accounting for *total ion population* and *relative ion population* to a FT-ICR MS system with a *MALDI* (matrix assisted laser desorption ionization) ionization system. The typical frequency range for MALDI FT-ICR data is larger than that of typical *ESI* data because *MALDI* inherently produces singly-charged species, thus reducing the *cyclotron frequency* of any given analyte that is observed as a multiply-charged ion when electrosprayed. As a result the *MMA* typically achieved with MALDI FT-ICR MS is not as high as that of ESI FT-ICR. The application of

these calibration laws enabled improved *MMA* for measurements performed with this system by almost 2 orders of magnitude when compared to the standard instrument calibration!

Chapter 5 details the application of calibration laws to data from a hybrid LTQ-FT-ICR MS. This instrument is equipped with *automatic gain control* (*AGC*), which meters the number of ions allowed into the instrument. This technology aims to place the same number of ions present in the ICR cell for both the calibration spectra and experimental spectra, with the goal being equal *space-charge effects*.[73] The calibration laws applied to this data account for the *total ion population* and *relative ion population* in the ICR cell by using *multiple linear regression* (*MLR*). Muddiman and Oberg had previously shown on a non *AGC* instrument that these calibration laws were able to improve *MMA*.[74] The combination of *AGC* and the external calibration laws were able to provide for *MMA* in the ppb range.

Chapter 6 uses the same instrumentation as discussed in **Chapter 5**, but approaches the calibration subject using *artificial neural networks* (*ANN*s). A benefit of using *ANN*s to fit data is that no physical model is assumed, as there is with *multiple linear regression*. The results achieved in this method were compared to results from *MLR*. The data used as inputs for the *ANN* were *cyclotron frequency*, *total ion population*, and *relative ion population*. This procedure produced a fit that was better than the fit provided by *MLR*; however, the average *MMA* was not as high. This result is interesting and will be the subject of future experiments.

With several experiments dedicated to improving the *mass measurement accuracy* of *FT-ICR* instruments, the decision was made to develop a method to be utilized for quantification of a protein of interest, C-Reactive Protein (CRP). **Chapter 7** discusses this set of experiments. Previous work has been completed utilizing *protein cleavage isotope dilution mass spectrometry* (*PC-IDMS*) for CRP; however this work involved many sample clean up steps to reduce the complexity of the matrix containing the analyte.[65] The experiments described within **Chapter 7** were the results of enzymatic digestion of human plasma followed by direct interrogation by nano-flow LC coupled to a triple quadrupole mass spectrometer utilized in single reaction monitoring mode. This study examined more than 100 different plasma samples, with each disease epithelial ovarian cancer plasma sample having its own age, menopause, and draw-match benign gynecologic tumor control. The results obtained by our mass spectrometric method were compared to the standard *ELISA* test for CRP and demonstrated high correlation. The CRP values showed an increase with an increase in EOC stage. In addition, it was also illustrated that the *ELISA* method and *PC-IDMS* method while having high correlation provided different absolute values, which indicates each having a different references range.

The peptides investigated in **Chapter 7** were free of the amino acid cysteine. **Chapter 8** probes the increase in *electrospray response* for peptides which contain cysteine. This is performed by application of several new iodoacetamide derivatives which are more hydrophobic than iodoacetamide itself and take advantage of the electrospray bias toward more hydrophobic analyte species.[75] An initial report by

Frahm et. al. describes the *electrospray response* enhancement of several peptides and one *hydrophobic tag*.[76] Besides increases in *electrospray response* an increase in column trapping efficiency is also observed in this method termed *ALiPHAT* (augmented limits of detection for peptides with hydrophobic alkyl tags).[76] The *hydrophobic tags* investigated in these experiments could have future application to improved limits of detection in targeted *absolute quantification* studies, such as **Chapter 7**, for peptides which contain cysteine.

1.5 References

1. Wilkins, M.R., et al., Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev*, 1996. **13**. 19-50.

2. Stenesh, J. and J. Stenesh, *Dictionary of biochemistry and molecular biology*. 2nd ed. 1989, New York: Wiley. vii, 525 p.

3. Kenyon, G.L., et al., Defining the mandate of proteomics in the post-genomics era: Workshop report. *Mol. Cell. Proteomics*, 2002. **1**. 763-780.

4. Crick, F., Central Dogma of Molecular Biology. *Nature*, 1970. **227**. 561-&.

5. Creighton, T.E., *Proteins : structures and molecular properties*. 2nd ed. 1993, New York: W.H. Freeman. xiii, 507 p.

6. Young, V.R., Adult Amino-Acid-Requirements - the Case for a Major Revision in Current Recommendations. *J. Nutr.*, 1994. **124**. S1517-S1523.

7. Remko, M. and B.M. Rode, Effect of metal ions (Li+, Na+, K+, Mg2+, Ca2+, Ni2+, Cu2+, and Zn2+) and water coordination on the structure of glycine and zwitterionic glycine. *Journal of Physical Chemistry A*, 2006. **110**. 1960-1967.

8. Kyte, J. and R.F. Doolittle, A Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology*, 1982. **157**. 105-132.

9. Aebersold, R. and M. Mann, Mass spectrometry-based proteomics. *Nature*, 2003. **422**. 198-207.

10. Mann, M., R.C. Hendrickson, and A. Pandey, Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.*, 2001. **70**. 437-473.

11. Mortz, E., et al., Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. *Natl. Acad. Sci. U. S. A.*, 1996. **93**. 8264-8267.

12. Henzel, W.J., et al., Identifying Proteins from 2-Dimensional Gels by Molecular Mass Searching of Peptide-Fragments in Protein-Sequence Databases. *Natl. Acad. Sci. U. S. A.*, 1993. **90**. 5011-5015.

13. Ong, S.E., et al., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics*, 2002. **1**. 376-386.

14. Gygi, S.P., et al., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.*, 1999. **17**. 994-999.

15. Li, J.X., H. Steen, and S.P. Gygi, Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents - The yeast salinity stress response. *Mol. Cell. Proteomics*, 2003. **2**. 1198-1204.

16. Mirgorodskaya, O.A., et al., Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using O-18-labeled internal standards. *Rapid Commun. Mass Spectrom.*, 2000. **14**. 1226-1232.

17. Ross, P.L., et al., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics*, 2004. **3**. 1154-1169.

18. Hall, M.P. and L.V. Schneider, Isotope-differentiated binding energy shift tags (IDBEST (TM)) for improved targeted biomarker discovery and validation. *Expert Rev. Proteomics*, 2004. **1**. 421-431.

19. Guerrera, I.C. and O. Kleiner, Application of mass spectrometry in proteomics. *Biosci. Rep.*, 2005. **25**. 71-93.

20. Barr, J.R., et al., Isotope dilution mass spectrometric quantification of specific proteins: Model application with apolipoprotein A-I. *Clin. Chem.*, 1996. **42**. 1676-1682.

21. Gerber, S.A., et al., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Natl. Acad. Sci. U. S. A.*, 2003. **100**. 6940-6945.

22. Barnidge, D.R., et al., Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal. Chem.*, 2003. **75**. 445-451.

23. Barnidge, D.R., et al., Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.*, 2004. **3**. 644-652.

24. Anderson, L. and C.L. Hunter, Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics*, 2006. **5**. 573-588.

25. Hawkridge, A.M., et al., Quantitative mass spectral evidence for the absence of circulating brain natriuretic peptide (BNP-32) in severe human heart failure. *Natl. Acad. Sci. U. S. A.*, 2005. **102**. 17442-17447.

26. Williams, D.K. and D.C. Muddiman, Absolute quantification of C-reactive protein in human plasma derived from patients with epithelial ovarian cancer utilizing protein cleavage isotope dilution mass spectrometry. *J Proteome Res*, 2009. **8**. 1085-90.

27. Yamashita, M. and J.B. Fenn, Electrospray Ion-Source - Another Variation on the Free-Jet Theme. *J. Phys. Chem.*, 1984. **88**. 4451-4459.

28. Bose, G.M., *Recherches sur le cause et sur la véritable théorie de l'électricité, Wittenberg*. 1745.

29. Zeleny, J., Instability of Electrified Liquid Surfaces. *Physcial Review*, 1917. **10**. 1-6.

30. Smith, R.D., et al., Principles and Practice of Electrospray Ionization - Mass-Spectrometry for Large Polypeptides and Proteins. *Mass Spectrom. Rev.*, 1991. **10**. 359-451.

31. Dole, M., L.L. Mack, and R.L. Hines, Molecular Beams of Macroions. *J. Chem. Phys.*, 1968. **49**. 2240-&.

32. Mack, L.L., et al., Molecular Beams of Macroions .2. *J. Chem. Phys.*, 1970. **52**. 4977-&.

33. Clegg, G.A. and M. Dole, Molecular Beams of Macroions .3. Zein and Polyvinypyrrolidone. *Biopolymers*, 1971. **10**. 821-&.

34. Gieniec, J., et al., Electrospray Mass-Spectroscopy of Macromolecules - Application of an Ion-Drift Spectrometer. *Biomedical Mass Spectrometry*, 1984. **11**. 259-268.

35. Taflin, D.C., T.L. Ward, and E.J. Davis, Electrified Droplet Fission and the Rayleigh Limit. *Langmuir*, 1989. **5**. 376-384.

36. Gomez, A. and K.Q. Tang, Charge and Fission of Droplets in Electrostatic Sprays. *Physics of Fluids*, 1994. **6**. 404-414.

37. Iribarne, J.V. and B.A. Thomson, Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.*, 1976. **64**. 2287-2294.

38. Thomson, B.A. and J.V. Iribarne, Field-Induced Ion Evaporation from Liquid Surfaces at Atmospheric-Pressure. *J. Chem. Phys.*, 1979. **71**. 4451-4463.

39. Comisarow, M.B. and A.G. Marshall, Fourier-Transform Ion-Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.*, 1974. **25**. 282-283.

40. Comisarow, M.B. and A.G. Marshall, Frequency-Sweep Fourier-Transform Ion-Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.*, 1974. **26**. 489-490.

41. Hardman, M. and A.A. Makarov, Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal. Chem.*, 2003. **75**. 1699-1705.

42. Smith, R.D., et al., New Developments in Biochemical Mass-Spectrometry - Electrospray Ionization. *Anal. Chem.*, 1990. **62**. 882-899.

43. Mciver, R.T., R.L. Hunter, and W.D. Bowers, Coupling a Quadrupole Mass-Spectrometer and a Fourier-Transform Mass-Spectrometer. *Int. J. Mass Spectrom. Ion Processes*, 1985. **64**. 67-77.

44. Syka, J.E.P., et al., Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.*, 2004. **3**. 621-626.

45. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.*, 1998. **17**. 1-35.

46. Hakansson, K., et al., High resolution tandem mass spectrometry for structural biochemistry. *Curr. Org. Chem.*, 2003. **7**. 1503-1525.

47. Marshall, A.G. and S.H. Guan, Advantages of high magnetic field for Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.*, 1996. **10**. 1819-1823.

48. Comisarow, M.B., Theory of Fourier-Transform Ion-Cyclotron Resonance Mass-Spectroscopy .2. Signal Modeling for Ion-Cyclotron Resonance. *J. Chem. Phys.*, 1978. **69**. 4097-4104.

49. Senko, M.W., et al., External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 1997. **8**. 970-976.

50. Gorshkov, M.V. and E.N. Nikolaev, Optimal Cyclotron Radius for High-Resolution Ft-Icr Spectrometry. *Int. J. Mass Spectrom. Ion Processes*, 1993. **125**. 1-8.

51. Uechi, G.T. and R.C. Dunbar, Space-Charge Effects on Relative Peak Heights in Fourier Transform-Ion Cyclotron-Resonance Spectra. *J. Am. Soc. Mass Spectrom.*, 1992. **3**. 734-741.

52. Taylor, P.K. and I.J. Amster, Space charge effects on mass accuracy for multiply charged ions in ESI-FTICR*. Int. J. Mass Spectrom.*, 2003. **222**. 351-361.

53. Frahm, J.L., C.M.C. Velez, and D.C. Muddiman, Understanding the influence of post-excite radius and axial confinement on quantitative proteomic measurements using Fourier transform ion cyclotron resonance mass spectrometry*. Rapid Commun. Mass Spectrom.*, 2007. **21**. 1196-1204.

54. Marshall, A.G. and F.R. Verdun, *Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook*. 1990, Amsterdam: Elsevier. 460.

55. Dekoning, L.J., et al., Segmented Fourier-Transform and Its Application to Fourier-Transform Ion-Cyclotron Resonance (Ft-Icr) Mass-Spectrometry - Ion Abundances and Mass Measurements*. Int. J. Mass Spectrom. Ion Processes*, 1989. **95**. 71-92.

56. Gordon, E.F. and D.C. Muddiman, Impact of ion cloud densities on the measurement of relative ion abundances in Fourier transform ion cyclotron resonance mass spectrometry: experimental observations of coulombically induced cyclotron radius perturbations and ion cloud dephasing rates*. J. Mass Spectrom.*, 2001. **36**. 195-203.

57. Mitchell, D.W. and R.D. Smith, Cyclotron Motion of 2 Coulombically Interacting Ion Clouds with Implications to Fourier-Transform Ion-Cyclotron Resonance Mass-Spectrometry*. Physical Review E*, 1995. **52**. 4366-4386.

58. Francl, T.J., et al., Experimental-Determination of the Effects of Space-Charge on Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 189-199.

59. Jeffries, J.B., S.E. Barlow, and G.H. Dunn, Theory of Space-Charge Shift of Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 169-187.

60. Williams, D.K., A.M. Hawkridge, and D.C. Muddiman, Sub parts-per-million mass measurement accuracy of intact proteins and product ions achieved using a dual electrospray ionization quadrupole Fourier transform ion cyclotron resonance mass spectrometer*. J. Am. Soc. Mass Spectrom.*, 2007. **18**. 1-7.

61. Williams, D.K. and D.C. Muddiman, Parts-per-billion mass measurement accuracy achieved through the combination of multiple linear regression and automatic gain control in a Fourier transform ion cyclotron resonance mass spectrometer*. Anal. Chem.*, 2007. **79**. 5058-5063.

62. Williams, D.K., et al., Calibration laws based on multiple linear regression applied to matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry. *J. Mass Spectrom.*, 2008. **43**. 1659-1663.

63. Hannis, J.C. and D.C. Muddiman, A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2000. **11**. 876-883.

64. Williams, D.K., et al., Dual electrospray ion source for electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.*, 2007. **79**. 7916-7919.

65. Kuhn, E., et al., Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and C-13-labeled peptide standards. *Proteomics*, 2004. **4**. 1175-1186.

66. Williams, D.K., et al., Synthesis, characterization, and application of iodoacetamide derivatives utilized for the ALiPHAT strategy. *J. Am. Chem. Soc.*, 2008. **130**. 2122-+.

67. Jensen, P.K., et al., Probing proteomes using capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 1999. **71**. 2076-2084.

68. Pasa-Tolic, L., et al., High throughput proteome-wide precision measurements of protein expression using mass spectrometry. *J. Am. Chem. Soc.*, 1999. **121**. 7949-7950.

69. Clauser, K.R., P. Baker, and A.L. Burlingame, Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS MS and database searching. *Anal. Chem.*, 1999. **71**. 2871-2882.

70. Nepomuceno, A.I., et al., Dual electrospray ionization source for confident generation of accurate mass tags using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 2003. **75**. 3411-3418.

71. McAlister, G.C., et al., Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.*, 2007. **79**. 3525-3534.

72. Xia, Y., X.R. Liang, and S.A. McLuckey, Pulsed dual electrospray ionization for ion/ion reactions*. J. Am. Soc. Mass Spectrom.*, 2005. **16**. 1750-1756.

73. Peterman, S.M., C.P. Dufresne, and S. Horning, The use of a hybrid linear trap/FT-ICR mass spectrometer for on-line high resolution/high mass accuracy bottom-up sequencing*. J Biomol Tech*, 2005. **16**. 112-24.

74. Muddiman, D.C. and A.L. Oberg, Statistical evaluation of internal and external mass calibration laws utilized in Fourier transform ion cyclotron resonance mass spectrometry*. Anal. Chem.*, 2005. **77**. 2406-2414.

75. Fenn, J.B., Ion Formation from Charged Droplets - Roles of Geometry, Energy, and Time*. J. Am. Soc. Mass Spectrom.*, 1993. **4**. 524-535.

76. Frahm, J.L., et al., Achieving augmented limits of detection for peptides with hydrophobic alkyl tags*. Anal. Chem.*, 2007. **79**. 3989-3995.

CHAPTER 2

Application of a Dual Electrospray Source for Internal Calibration of Top-Down Proteomic Data

2.1 Introduction

Hybrid quadrupole Fourier transform ion cyclotron resonance mass spectrometers (Q-FT-ICR MS)[1-5] are uniquely suited for making *top-down*[6-10] *proteomics* measurements. Their combination of high *resolving power* [11] and *mass measurement accuracy* [12] enables the confident determination of both the intact protein mass and the masses of its corresponding dissociation product ions. The combination of these experimental data sets provides confident and unambiguous identification of proteins. In order to realize the highest achievable *MMA* for a given FT-ICR MS system, frequency shifts due to *space-charge effects* must be accounted for via external or internal calibration methods.[13, 14]

External calibration has been used to account for frequency shifts and thereby improve the *MMA* of intact peptides and proteins.[15] For example, Amster and co-workers developed a calibration curve to account for the difference between the ion populations of the calibration spectrum and spectra that were to be externally calibrated resulting in <10 parts-per-million *MMA*.[16, 17] Hunt and co-workers have demonstrated the effectiveness of combining external calibration with *automatic gain control* (*AGC*) where the number of ions in the ICR cell were precisely controlled to fall within the external calibration range.[5] This approach allowed them to achieve <2

ppm *MMA*. Muddiman and co-workers were able to achieve < 1 ppm using this approach.[18] Based partly on the work of Amster[16, 17] and Smith[19], Oberg and Muddiman reported a novel external calibration law which afforded < 5 ppm *MMA*.[20]

Internal calibration has also been effective at providing high *MMA* for intact species. However, it can be difficult to implement in the context of modern proteomics experiments where introducing the calibrant ions concurrently with the analyte species can be problematic (e.g., ion suppression). One approach to internal calibration was reported by Smith and co-workers who developed an algorithm called deconvolution of Coulombic affected linearity (DeCAL).[21] Unfortunately, this approach required multiple charge states for the same species, which are not always present. Muddiman and co-workers have developed and utilized a dual electrospray emitter (dualESI) to achieve high *MMA.*[22-24] This approach is amenable to both direct infusion[25-27] and liquid chromatography experiments[28-30] without introducing measurement biases associated with ion suppression. Furthermore, this method does not compromise dynamic range by limiting the number of ions in the ICR cell (e.g., *AGC*). Despite the prevalence of multiple external and internal methods for calibrating *FT-ICR* data, *AGC*[5, 18] or internal calibration have thus far been the methods of choice for achieving high *MMA* for intact species.[21-24]

The interpretation and confident assignment of intact proteins and their product ions is a major challenge in *top-down proteomics*. Robust methods for providing high *MMA* of intact and dissociated proteins are sorely needed for *top-*

down proteomics to mature into a viable alternative to *bottom-up proteomics*.[31] A sampling of the current *top-down proteomics* literature clearly demonstrates the need to address this critical issue in *top-down proteomics*. To date, reported *MMA* of product ions range from -41 ppm to 10 ppm, all of which employed external calibration methods.[32, 33] In some cases the *MMA* is not reported for the product ions in a *tandem-MS* experiment.[34-36] Very little research thus far has focused on the development of strategies employing internal calibration as a means for achieving high *MMA* of *top-down proteomics* data.

DeCAL was recently applied to *tandem-MS* data which resulted in product ion *MMA*s of 1.59 ppm; however, it still required the presence of at least two charge-states for one species.[37] Kelleher and co-workers used "unmodified z ions" for internal calibration; however, the *MMA* reported had a range from 0.1 ppm to 44 ppm.[38] Furthermore, this method requires the identity of the protein in order to generate the exact mass of unmodified z ions which are used as the internal calibrants.[38] Several years ago, the Muddiman group developed and utilized a dual electrospray source to internally calibrate SORI-CID tandem mass spectrometry data of short oligonucleotides and achieved < 3.2 ppm *MMA* on a 7 Tesla FT-ICR MS.[24]

The results discussed here demonstrate internal calibration of intact proteins and their product ions formed by *collision induced dissociation* in a dual electrospray ionization[22-24] hybrid quadrupole FT-ICR MS using a modified pulse sequence. Internal standards of poly propylene glycol (PPG) and a mixture of four *iodopeptides*

that fall in *m/z* range which is unoccupied by natural (i.e., not chemically modified) peptide/proteins, were chosen for these MS/MS studies. This *m/z* range is also known as the forbidden zone and is resultant from the mass excess of the chemical constituents.[26] These results also present the capability of a dual ESI source coupled with a QFT-ICR MS to achieve sub-ppm *MMA*, which has significant impact for top-down proteomic applications.

2.2 Experimental

2.2.1 Materials

Polypropylene-Glycol (PPG) with an average molecular weight of 1000 Da, melittin, ubiquitin, ammonium acetate (>99.%), and formic acid were purchased from Sigma-Aldrich (St. Louis, MO). HPLC grade acetonitrile and high purity water were purchased from Burdick Jackson (Muskegon, MI). 2-propanol (HPLC Grade) was purchased from Fischer Scientific. The *iodopeptides* used were synthesized in the Mayo Proteomics Research Center (Rochester, MN). These four *iodopeptides* had the sequences (I_2Y)GK (**I_1**), (I_2Y)GK(I_2Y)G (**I_2**), (I_2Y)SR(I_2Y)GSYGSSI (**I_3**), and (I_2Y)SR(I_2Y)GSYGSSIGSY (**I_4**) with neutral *monoisotopic masses* of 617.9836, 1089.8617, 1742.1433, and 2049.2601 Da respectively. All materials were used as received.

2.2.2 Dual Electrospray Ionization Hybrid Quadrupole Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (dualESI-QFT-ICR MS)

A modified version of a dual ESI source developed previously in this laboratory[23] was coupled to a hybrid Ionspec (Lake Forest, CA) QFT-ICR MS equipped with an actively shielded 7T superconducting magnet (Cyromagnetics,

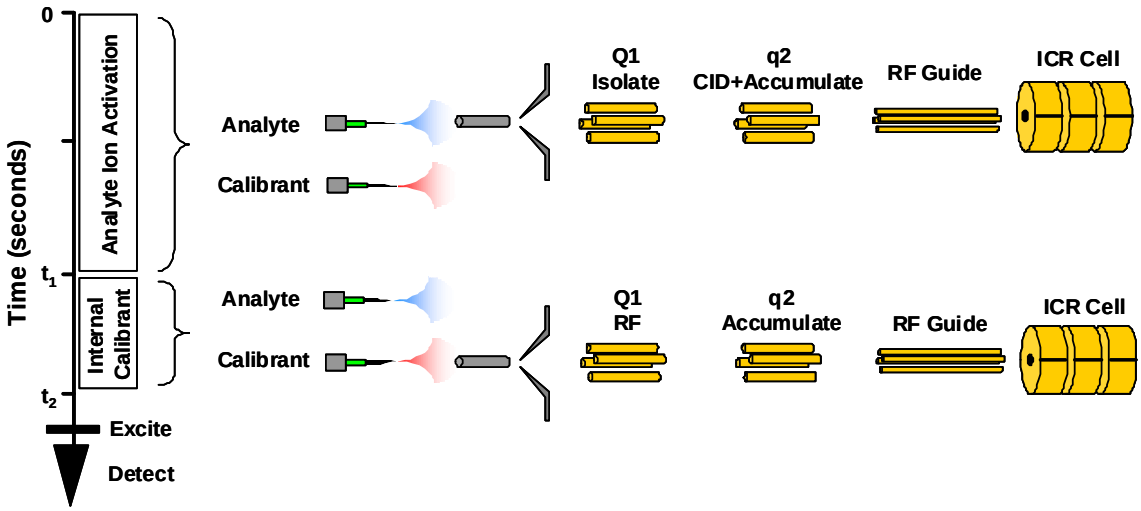


Figure 2.1 Schematic of QFT-ICR MS/MS experiment utilizing dual electrospray source to incorporate internal calibrants for increasing mass measurement accuracy of product ions.

Oak Ridge, TN) and a Z-spray *ESI* source (Waters Corp, Milford, MA). All spectra were acquired with 1024 k points and an ADC rate of 1 MHz. Samples were introduced by direct infusion using a 100 µl gas-tight syringe (Hamilton, Las Vegas, NV) and a Harvard Apparatus model PHD 2000 syringe pump (Holliston, MA) at a flow rate of 0.5 µl/min – 1.00 µL/min. The *ESI* emitter tips used were 360 µm o.d., 50 µm i.d. and tapered to 30 ± 1.0 µm i.d. (New Objective, Woburn, MA) and held at a constant potential of 2200V for all experiments. *Electrospray* solutions were composed of 1:1 (vol:vol) acetonitrile/water, with 0.1% formic acid, with the

exception of the ammonium-adducted polypropylene glycol solution which was composed of 70:30 2-propanol:water with 0.5 mM ammonium acetate (NH_4OAc). The generic experimental pulse sequence used for all experiments is shown in **Figure 2.1**.

2.3 External vs. Internal Calibration of Protein MS/MS Data

2.3.1 Mass Measurement Accuracy of MS/MS of Melittin

The primary sequence of melittin with illustration of the doubly charged y-ions that were observed is shown at the top of **Figure 2.2**. To evaluate the stability of the dualESI-QFT-ICR MS experimental pulse sequence, 50 sequential internally calibrated *collision induced dissociation* (*CID*) mass spectra were collected for two different internal calibrant mixtures. Both ammonium-adducted polypropylene glycol (NH_4-PPG-1000) and a mixture of four *iodopeptides* were used as internal calibrants. **Figure 2.2A** shows a representative "NH_4-PPG-1000" internally calibrated CID mass spectrum of melittin labeled as "Spectrum #2". The "NH_4-PPG-1000" ions were labeled as 12_A, 13_A, 14_A….20_A representing the number of propylene-glycol repeat units. The corresponding total ion abundance plot of the 50 sequential measurements is shown in **Figure 2.2B** with the arrows indicating every fifth spectrum. The arrow labeled "Spectrum #2" indicates the data plotted in **Figure 2.2A**.

Figure 2.2C shows a representative "*iodopeptide*" internally calibrated CID mass spectrum of melittin labeled also as "Spectrum #2. The "*iodopeptide*" ions

were labeled as I_1, I_2. I_3, and I_4 (see Section 2.2.1 for sequences). The corresponding total ion abundance plot of the 50 sequential measurements is shown in **Figure 2.2D** with the location of "Spectrum #2". In both spectra presented in this figure the y_{13}-ion is dominant, which is attributed to the presence of a proline (position 14 from the N-terminus) in the melittin sequence.[39] The data presented in **Figure 2.2B** and **2.2D** demonstrates the stability of the pulse sequence albeit the

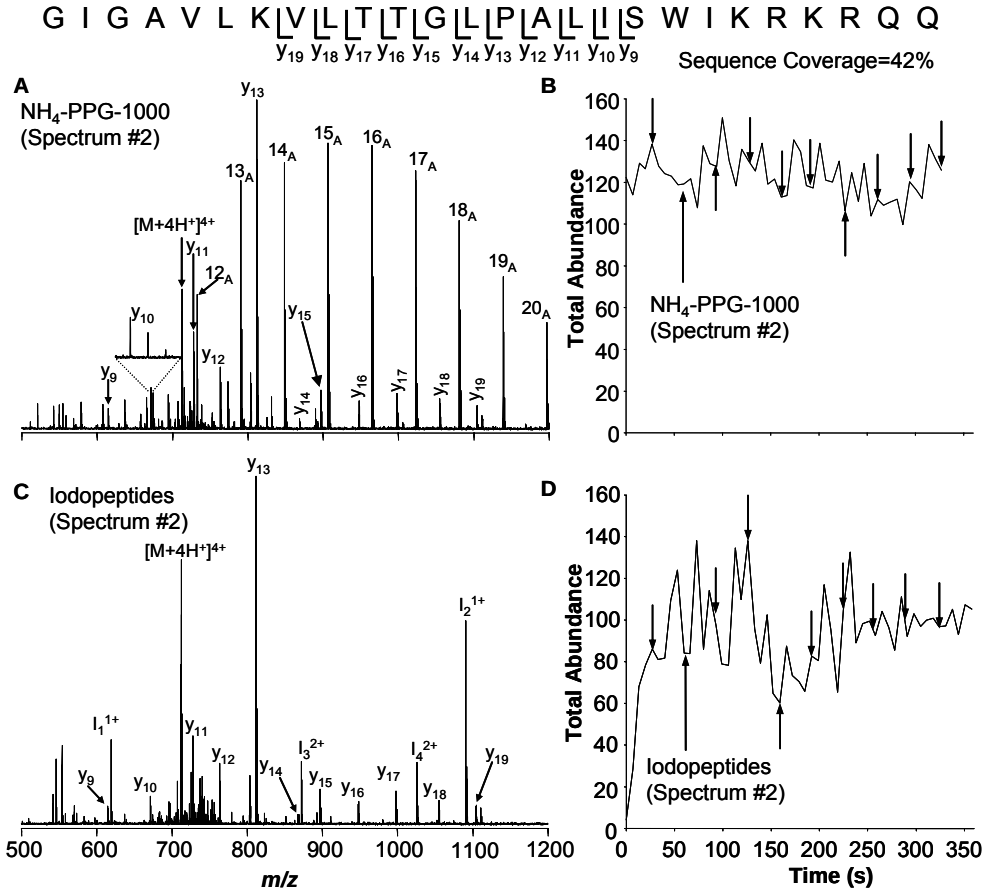


Figure 2.2 Primary sequence of melittin with observed y-ions. A) Internally calibrated dual ESI QFT-ICR tandem mass spectra of the 4+ charge-state of melittin using NH4-PPG-1000, B**)** total ion abundance plot of 50 sequential internally calibrated dual ESI QFT-ICR tandem mass spectra of the 4+ charge-state of melittin using NH4-PPG-1000, C) internally calibrated dual ESI QFT-ICR tandem mass spectra of the 4+ charge-state of melittin using iodopeptides as the internal standards, and D) total ion abundance plot of 50 sequential internally calibrated dual ESI QFT-ICR tandem mass spectra of the 4+ charge-state of melittin using four iodopeptides as the internal calibrants.

iodopeptide is more variable which was attribute to the *ESI* emitter.

The 4+ charge-state of the monoisotopic peak of melittin was observed to have a *MMA* ranging from -4.1 ppm to -1.85 ppm and -0.17 ppm to -0.31 ppm, respectively, for externally and internally calibrated data using NH_4-PPG-1000. Similarly, when the *iodopeptides* were used as calibrants, the *MMA* ranged from -1.71 ppm to -4.23 ppm and -0.17 ppm to -0.31 ppm, respectively, for externally and internally calibrated data.

Figure 2.3 presents the *MMA* data obtained from the tandem mass spectra of melittin shown in **Figure 2.2**. **Figure 2.3A** shows the relationship between *MMA* and the mass-to-charge ratio for external calibrated *tandem-MS* data using NH_4-PPG-1000 (left)) and *iodopeptides* (right) as calibrants. The *MMA* systematically improves with decreasing mass-to-charge ratio as expected given the inverse relationship between *cyclotron frequency* and *m/z*.[40]

Similarly, **Figure 2.3B** shows the internally calibrated data using NH_4-PPG-1000 (left) and *iodopeptides* (right) as internal calibrants. Importantly, the *MMA* is now centered about 0 ppm and there is no significant correlation between *MMA* and *m/z* indicating the systematic error has been removed. The data in **Figure 2.3A** and **Figure 2.3B** suggests that NH_4-PPG-1000 is a better internal calibrant than the *iodopeptides*. This observation was due to the fact that there are ten calibrant points for the NH_4-PPG-1000 as opposed to only four for the *iodopeptides*; however, this result could partly be related to the calibrants' frequency range relative to the product ions.[20]

Figure 2.3C shows the cumulative percentage plots as a function of *MMA* obtained for the externally (red) and internally calibrated (blue) data using NH_4-PPG-1000 (left) and four *iodopeptides* (right) as the calibrants. The dashed line in **Figure 2.3C** should theoretically intersect the curves shown at 50% if the *MMA* was exactly centered at 0 ppm. External calibration using either NH_4-PPG-1000 or the *iodopeptides* intersects this dashed line when 100% of the ions have already been accounted for, which clearly demonstrated the presence of systematic error. Internal calibration using either NH_4-PPG-1000 or the *iodopeptides* are approximately centered about 0 ppm *MMA* supporting our hypothesis that internally calibrated data would accurately remove the systematic bias. Furthermore, for internally calibrated data using the NH_4-PPG-1000 or the *iodopeptides*, the *MMA* ranged from -0.62 ppm to 0.64 ppm and -1.08 ppm to 1.09 ppm, respectively.

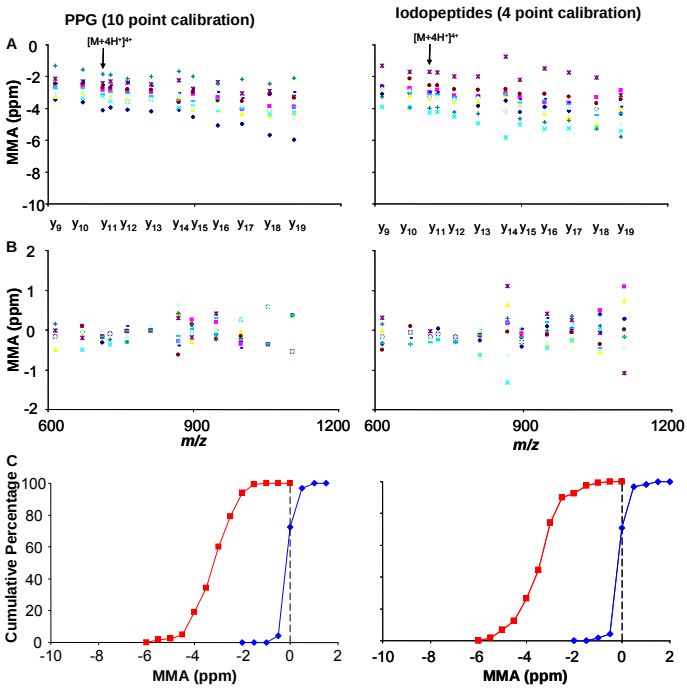


Figure 2.3 A) Mass measurement accuracy as a function of m/z for externally and **B)** internally calibrated data. The different symbols reflect the ten different spectra analyzed. **C)** Cumulative percentage plot of ions as a m/z for the melittin data (red is externally calibrated data and blue is internally calibrated data).

2.3.2 Mass Measurement Accuracy of Ubiquitin

Figure 2.4 shows the primary sequence for ubiquitin (top) along with the seven different y-ions observed (~9% sequence coverage). In these studies involving ubquitin, external and internal calibration were carried out only using the iodopeptide mixture to minimize interference in terms of complexity as well as the ability to access forbidden zones.[26] The single-acquisition QFT-ICR tandem mass spectrum is shown with the y-ions and internal calibrants labeled. The two insets show the expansion of two product ions (y_{58} and y_{40}). The dominant y_{58} and y_{40} product ions are attributed to the presence of two prolines (at positions 19 and 37 from the N-terminus) which may provide valuable bioinformatic constraints for top-down proteomic data.[39]

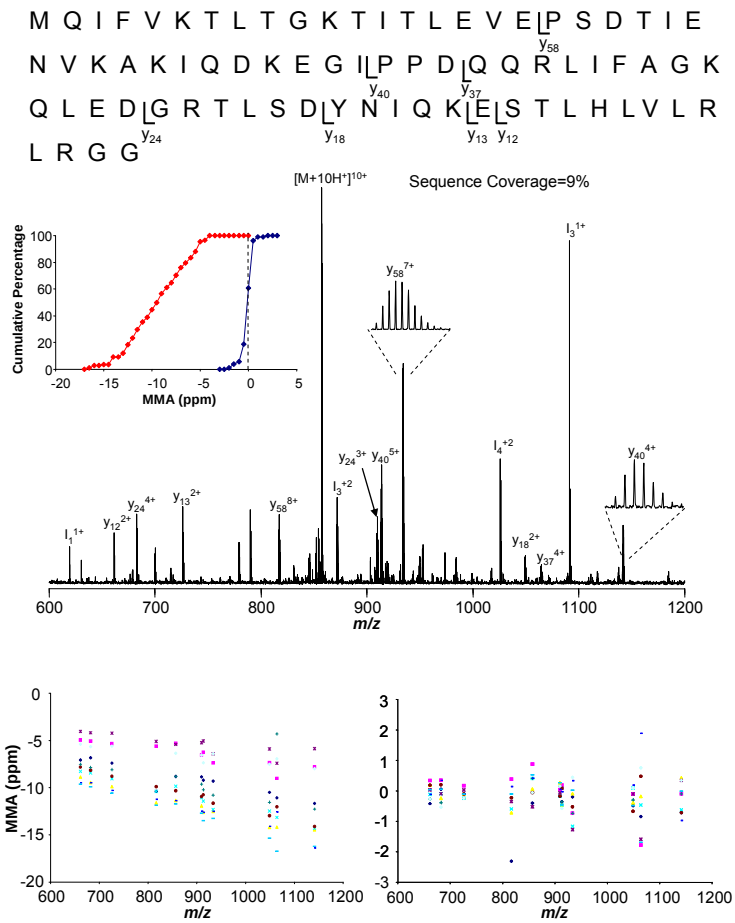


Figure 2.4 Ubquitin primary sequence and denotation of observed y-ions. Dual ESI QFT-ICR tandem mass spectra of the 10+ charge-state of ubiqutin using the four iodopeptides as internal standards. Plot of MMA as a function of *m/z* for externally (bottom left) and internally calibrated data (bottom right). The cumulative percentage as a function of MMA is shown (middle left) for the externally calibrated data (red) and internally calibrated data (blue).

The *MMA* achieved for the externally and internally calibrated data ranged from -16 to -4 ppm and from -2 to 1 ppm, respectively, with the exception of one peak that had a very low abundance. All *MMA* measurements were calculated from the monoisotopic peak for each of the isotope clusters being investigated. The *MMA* as a function of *m/z* for the externally (bottom left) and internally (bottom right) calibrated data is shown once again demonstrating that the systematic error has been removed using internally calibration. The cumulative percentage plot as a function of *MMA* is also shown for the external (red) and internal (blue) again demonstrating that high *MMA* can routinely be achieved.

2.4 Conclusions

The utilization of a dual electrospray source coupled to a QFT-ICR MS instrument enabled the generation and mass-selective dissociation of intact proteins while subsequently incorporating internal mass calibrants ions allowed for sub ppm *MMA* for melittin and low ppm *MMA* for ubiquitin. These data demonstrate that the dual electrospray source combined with tandem QFT-ICR mass spectrometry data can reliably yield extremely high *MMA* when compared to other methods of calibration. If coupled with stronger magnetic fields[40], this method of internally calibrating top-down proteomic data has the potential to improve *MMA* even further.

2.5 References

1. Mciver, R.T., R.L. Hunter, and W.D. Bowers, Coupling a Quadrupole Mass-Spectrometer and a Fourier-Transform Mass-Spectrometer. *Int. J. Mass Spectrom. Ion Processes*, 1985. **64**. 67-77.

2. Wang, Y., et al., Mass-selective ion accumulation and fragmentation in a linear octopole ion trap external to a Fourier transform ion cyclotron resonance mass spectrometer. *Int. J. Mass Spectrom.*, 2000. **198**. 113-120.

3. Harkewicz, R., et al., ESI-FTICR mass spectrometry employing data-dependent external ion selection and accumulation. *J. Am. Soc. Mass Spectrom.*, 2002. **13**. 144-154.

4. Patrie, S.M., et al., Construction of a hybrid quadrupole/Fourier Transform Ion Cyclotron Resonance Mass Spectrometer for versatile MS/MS above 10 kDa. *J. Am. Soc. Mass Spectrom.*, 2004. **15**. 1099-1108.

5. Syka, J.E.P., et al., Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.*, 2004. **3**. 621-626.

6. Loo, J.A., C.G. Edmonds, and R.D. Smith, Primary Sequence Information from Intact Proteins by Electrospray Ionization Tandem Mass-Spectrometry. *Science*, 1990. **248**. 201-204.

7. Kelleher, N.L., et al., Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J. Am. Chem. Soc.*, 1999. **121**. 806-812.

8. Kelleher, N.L., Top-down proteomics. *Anal. Chem.*, 2004. **76**. 196A-203A.

9. Reid, G.E. and S.A. McLuckey, 'Top down' protein characterization via tandem mass spectrometry. *J. Mass Spectrom.*, 2002. **37**. 663-675.

10. Bogdanov, B. and R.D. Smith, Proteomics by FTICR mass spectrometry: Top down and bottom up. *Mass Spectrom. Rev.*, 2005. **24**. 168-200.

11. Zubarev, R.A., et al., Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.*, 2000. **72**. 563-573.

12. Gerber, S.A., et al., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Natl. Acad. Sci. U. S. A.*, 2003. **100**. 6940-6945.

13. Jeffries, J.B., S.E. Barlow, and G.H. Dunn, Theory of Space-Charge Shift of Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 169-187.

14. Francl, T.J., et al., Experimental-Determination of the Effects of Space-Charge on Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 189-199.

15. Zhang, L.K., et al., Accurate mass measurements by Fourier transform mass spectrometry*. Mass Spectrom. Rev.*, 2005. **24**. 286-309.

16. Easterling, M.L., T.H. Mize, and I.J. Amster, Routine part-per-million mass accuracy for high-mass ions: Space-charge effects in MALDI FT-ICR*. Anal. Chem.*, 1999. **71**. 624-632.

17. Taylor, P.K. and I.J. Amster, Space charge effects on mass accuracy for multiply charged ions in ESI-FTICR*. Int. J. Mass Spectrom.*, 2003. **222**. 351-361.

18. Johnson, K.L., et al., Accurate mass precursor ion data and tandem mass spectrometry identify a class I human leukocyte antigen A*0201-presented peptide originating from vaccinia virus*. J. Am. Soc. Mass Spectrom.*, 2005. **16**. 1812-1817.

19. Masselon, C., et al., Mass measurement errors caused by 'local" frequency perturbations in FTICR mass spectrometry*. J. Am. Soc. Mass Spectrom.*, 2002. **13**. 99-106.

20. Muddiman, D.C. and A.L. Oberg, Statistical evaluation of internal and external mass calibration laws utilized in Fourier transform ion cyclotron resonance mass spectrometry*. Anal. Chem.*, 2005. **77**. 2406-2414.

21. Bruce, J.E., et al., Obtaining more accurate Fourier transform ion cyclotron resonance mass measurements without internal standards using multiply charged ions*. J. Am. Soc. Mass Spectrom.*, 2000. **11**. 416-421.

22. Hannis, J.C. and D.C. Muddiman, A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in fourier transform ion cyclotron resonance mass spectrometry*. J. Am. Soc. Mass Spectrom.*, 2000. **11**. 876-883.

23. Nepomuceno, A.I., et al., Dual electrospray ionization source for confident generation of accurate mass tags using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry*. Anal. Chem.*, 2003. **75**. 3411-3418.

24. Flora, J.W., J.C. Hannis, and D.C. Muddiman, High-mass accuracy of product ions produced by SORI-CID using a dual electrospray ionization source coupled with FTICR mass spectrometry. *Anal. Chem.*, 2001. **73**. 1247-1251.

25. Null, A.P., J.C. Hannis, and D.C. Muddiman, Genotyping of simple and compound short tandem repeat loci using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 2001. **73**. 4514-4521.

26. Frahm, J.L., et al., Accessible proteomics space and its implications for peak capacity for zero-, one- and two-dimensional separations coupled with FT-ICR and TOF mass spectrometry. *J. Mass Spectrom.*, 2006. **41**. 281-288.

27. Chalmers, M.J., et al., Protein kinase A phosphorylation characterized by tandem Fourier transform ion cyclotron resonance mass spectrometry. *Proteomics*, 2004. **4**. 970-981.

28. Nepomuceno, A.I., et al., Detection of genetic variants of transthyretin by liquid chromatography-dual electrospray ionization Fourier-transform ion-cyclotron-resonance mass spectrometry. *Clin. Chem.*, 2004. **50**. 1535-1543.

29. Johnson, K.L., et al., Analysis of the low molecular weight fraction of serum by LC-dual ESI-FT-ICR mass spectrometry: Precision of retention time, mass, and ion abundance. *Anal. Chem.*, 2004. **76**. 5097-5103.

30. Lee, S.W., et al., Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Natl. Acad. Sci. U. S. A.*, 2002. **99**. 5942-5947.

31. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.*, 2001. **73**. 5683-90.

32. Patrie, S.M., et al., Top down mass spectrometry of < 60-kDa proteins from Methanosarcina acetivorans using quadrupole FTMS with automated octopole collisionally activated dissociation. *Mol. Cell. Proteomics*, 2006. **5**. 14-25.

33. Guan, S.H., A.G. Marshall, and M.C. Wahl, Ms/Ms with High Detection Efficiency and Mass Resolving Power for Product Ions in Fourier-Transform Ion-Cyclotron Resonance Mass-Spectrometry. *Anal. Chem.*, 1994. **66**. 1363-1367.

34. Patrie, S.M., et al., Strategies for automating top-down protein analysis with Q-FTICR MS. *Int. J. Mass Spectrom.*, 2004. **234**. 175-184.

35. Roth, M.J., et al., Precise and parallel characterization of coding polymorphisms, alternative splicing, and modifications in human proteins by mass spectrometry. *Mol. Cell. Proteomics*, 2005. **4**. 1002-1008.

36. Fernandez, F.M., et al., Protein identification via surface-induced dissociation in an FT-ICR mass spectrometer and a patchwork sequencing approach. *J. Am. Soc. Mass Spectrom.*, 2006. **17**. 700-709.

37. Kaiser, N.K., G.A. Anderson, and J.E. Bruce, Improved mass accuracy for tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2005. **16**. 463-470.

38. Thomas, C.E., N.L. Kelleher, and C.A. Mizzen, Mass spectrometric characterization of human histone H3: A bird's eye view. *J. Proteome Res.*, 2006. **5**. 240-247.

39. Schwartz, B.L. and M.M. Bursey, Some Proline Substituent Effects in the Tandem Mass-Spectrum of Protonated Pentaalanine. *Biol. Mass Spectrom.*, 1992. **21**. 92-96.

40. Marshall, A.G. and S.H. Guan, Advantages of high magnetic field for Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.*, 1996. **10**. 1819-1823.

CHAPTER 3

Application of a Dual Electrospray Source for hybrid LTQ-Orbitrap Electron Capture Dissociation

3.1 Introduction

The ability of *electron-transfer dissociation* (*ETD*) to produce similar fragmentation patterns to *electron-capture dissociation* (*ECD*) in the context of the ubiquitous RF quadrupole ion trap mass spectrometer has garnered the technique considerable attention.[1-4] Among other attributes, *ETD* can allow for the identification of post-translational modification sites and is able to dissociate large peptides and entire proteins on a chromatographic time-scale.[5-9] The ability to identify post-translation modification sites with *ETD* is an artifact of its fragmentation pathway being a nonergodic process (i.e. not involving intramolecular vibrational energy redistribution).[2] This technique produces c- and z-type ions as opposed to the typical b- and y-type ions produced by *collision activated dissociation*. This nomenclature was first introduced by Roepstorff and Fohlman, and then later adjusted by Johnson et. al (**Figure 3.1**).[10, 11] Several ion

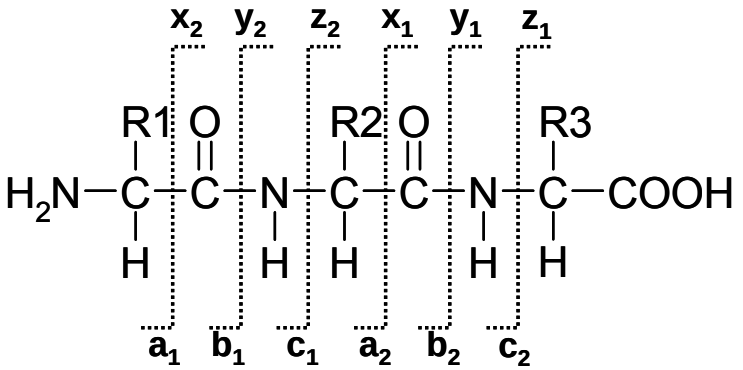


Figure 3.1 Peptide fragmentation nomenclature. x-, y-, and z-ions are formed if charge remains on C-terminal fragment. a-, b-, and c- ions are formed if charge remains on the N-terminal fragments.

trap MS systems offer a good platform for conducting ion/ion chemistry[12-16] (i.e. *ETD*); however, today numerous hybrid systems incorporate ion trapping devices as either mass analyzers or intermediate storage devices. In these hybrids, Orbitrap (LTQ-Orbitrap), time-of-flight (Q-TOF), and Fourier transform ion cyclotron resonance (LTQ-FT-ICR) mass analyzers are frequently employed for their ability to bring higher mass resolving power and *mass measurement accuracies*.

Originally, *ETD* was accomplished with reagent anions that were produced by a negative chemical ionization source (NCI), which had been fitted to the unoccupied end of a linear ion trap (LTQ).[2] However, because this modification took advantage of the unoccupied end of the stand-alone LTQ, it is not easily incorporated on hybrid instruments such as the LTQ-Orbitrap or LTQ-FT-ICR. To address the problem of implementing *ETD* and other ion/ion reactions on hybrid mass spectrometers, McLuckey and co-workers have recently developed different strategies for introducing both cations and anions through a single atmospheric pressure (AP) inlet. Their initial strategy utilized a nanoelectrospray source for generation of peptide cations and an atmospheric pressure chemical ionization (APCI) for anion production.[17, 18] The anions generated via APCI were nitrobenzene and azobenzene, with azobenzene having a slightly higher *ETD* efficiency.

McLuckey et al. also developed a dual AP *ETD* source that generates the anion reagent population via *ESI*, around the same time.[19] This scheme produces the same *ETD* anion as used in the initial report of *ETD*, but in this case a precursor anion 9-anthracenecarboxylic acid is electrosprayed and then subjected to *collision*

activated dissociation (*CAD*). These experiments were performed utilizing multiple independent AP sources in which cation and anion populations were constantly generated and gated by downstream ion optics.

Coon and co-workers have adapted that approach for use on a single inlet LTQ-Orbitrap hybrid mass spectrometer.[4] Briefly, two *ESI* emitters, close in proximity, were pulsed on and off to generate distinct populations of cations and anions for ion/ion reactions in the LTQ. An addition of 750 ms to the total scan time was necessary to pulse the high-voltage power supplies and inject the precursor anion population. Following the *ETD* reaction in the LTQ, product ions were *m/z* analyzed in either the LTQ or the Orbitrap, the latter produces high-resolution and high *m/z* measurement accuracy (~ 3 ppm) product ion spectra at the expense of time and sensitivity. McLuckey and co-workers have also developed a pulsed *ESI* source in which voltages were pulsed as quickly as 10 ms; however, this was utilized for ion/ion reactions other than *ETD* including charge reduction, charge inversion, and complex formation.[20]

The experiments described in this chapter detail the modification of a dual ESI source, previously developed by Muddiman and coworkers, [21-24] for the production of both the analyte cation population and the anion reagent population. The voltage on each of the two emitters was constant throughout a given experiment, translating to improved *ESI* stability (for both anions and cations) and much more rapid switching times. The fast physical switching time allows for more *ETD* tandem mass spectra to be collected during a given time period.

3.2 Experimental

3.2.1 Materials

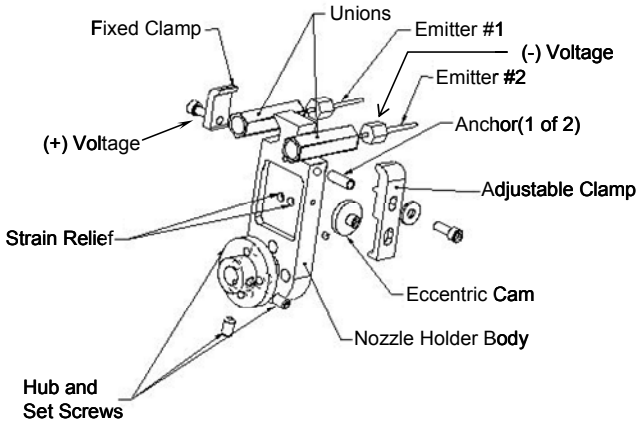
The peptide KAAAKAAAK was synthesized at the University of Wisconsin Biotechnology Center (Madison, WI). All other peptides and chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Peptide and protein solutions were prepared by dilution of stock solutions to ~3 pmol/μL in 60:39.9:0.1 acetonitrile/water/acetic acid by volume. The reagent anion 9-anthracenecarboxylic acid was dissolved in a solution of 50:50 acetonitrile/water to a concentration of 150 μmol/μL. The solution was centrifuged for ~5 minutes (~9000 rcf) to prevent particulates from clogging the *ESI* emitter tip during the process of direct infusion.

3.2.2 Dual Electrospray Ionization Hybrid Linear Ion Trap-Orbitrap

All data shown were collected on a highly modified LTQ-Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). A description of the internal instrument modifications necessary for *ETD* (i.e. ion optics and power supplies) have been previously published, as well as the voltage changes within the instrument throughout the experiment.[4] These modifications will not be covered in this document because they were performed prior to these experiments. A previously developed dual ESI source,[21-24] to which changes were made such that one emitter could produce cations while the other could simultaneously produce anions (**Figure 3.2**), was installed on the front-end of the instrument. The necessary changes to the source included replacement of the original electrical connector that connected the two emitters to a single voltage source. This piece was replaced with

two separate anchors for the fixed clamp and adjustable clamp that enabled the two emitters to be held at constant, but separate, voltages throughout the experiment. The *ion trap control language* (*ITCL*) was modified to trigger a TTL pulse, when desired, to induce emitter switching. The capability of this source to be modified to accommodate ion-ion chemistry was previously noted;[23] however, it was not adapted for such use until directly before the described experiments were performed.

A) Schematic of Dual ESI source Emitter Holder



B) Picture of Dual ESI source coupled to LTQ-Orbitrap

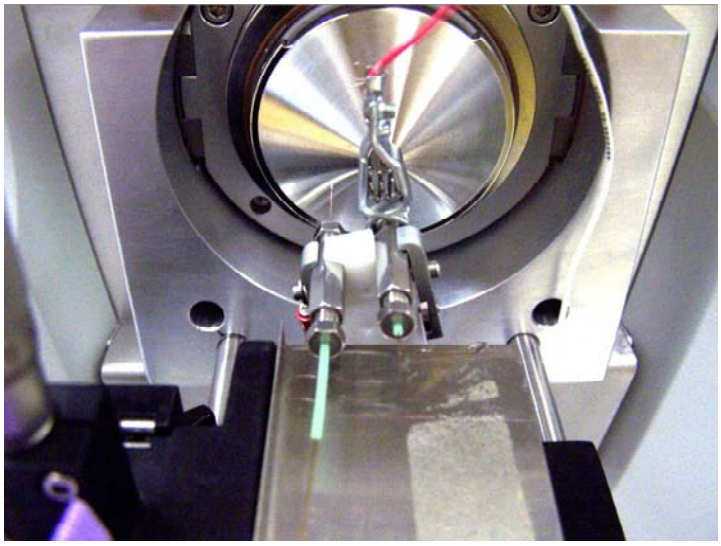


Figure 3.2 A) Expanded view of the emitter holder of the dual electrospray assembly. B) Picture of Dual ESI source coupled to LTQ-Orbitrap.

Samples were introduced by direct infusion using a 250 µL gas-tight syringe (Hamilton, Las Vegas, NV) at a flow rate of 1.5 µL/minute – 2.00 µL/min. The *ESI* emitter tips utilized were 360 µm o.d., 50 µm i.d. and tapered to 30 ± 1.0 µm i.d.

(New Objective, Woburn, MA). The cation emitter tip was held at a potential of +2.3-2.5 kV while the anion emitter tip was held at a constant potential of -2.7 kV for all experiments.

Since both *ESI* emitter tips were kept at a constant voltage, throughout a given experiment, there was no need to allow the electrospray to stabilize. The maximum total time for switching between the emitters and anion injection and for these experiments was ~85 ms – significantly faster than that of other dual ESI sources.[4]

3.3 Results and Discussion

Stability of signal is one important factor to consider when implementing a new ion source. **Figure 3.3** shows the stability of the *ETD* reagent at *m/z* value 177 (the *CAD* product of the anion 9-anthracenecarboxylic acid) over a fifteen minute period. During this collection time the source was operating in the switching mode, just as it would be utilized in a typical *ETD* experiment. The inset of **Figure 3.3** displays a single mass spectrum

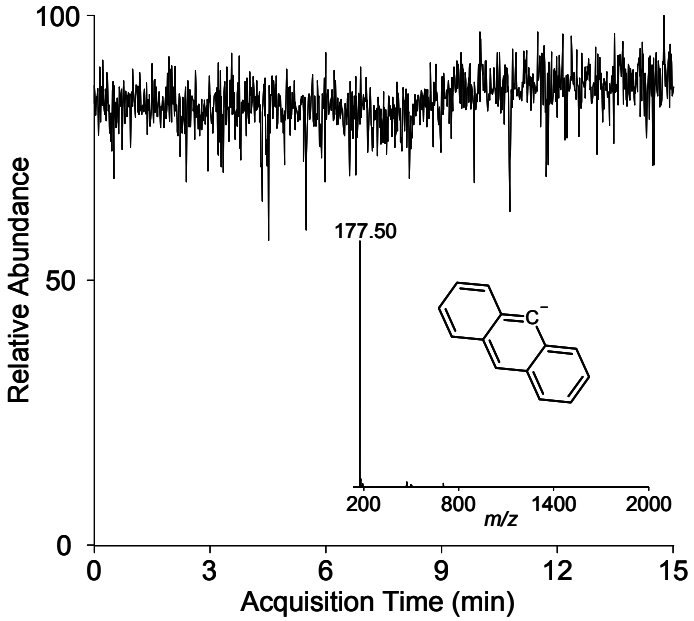


Figure 3.3 Total ion chromotogram for 9-anthracenecarboxylic acid undergoing CAD to produce the ETD anion during a pulsing experiment. The inset displays a single scan mass spectrum of the ETD reagent anion and its structure.

acquired by the LTQ and the structure of the *ETD* reagent. For this work anion stability is critical as automatic gain control was not employed to meter the number of anions allowed into the instrument.

Figure 3.4a shows a single scan LTQ-Orbitrap tandem mass spectrum produced following the reaction of a triply-protonated synthetic peptide (KAAAKAAAK) and the collisionally activated 9-anthracenecarboxylic acid anion. This reaction persisted for 100 ms in the LTQ. The reaction was then quenched and the product ions transferred to the Orbitrap mass analyzer for *m/z* analysis. Detected c- and z-type product ions are labeled on the mass spectrum and shown on the sequence. An inset showing the Orbitrap's ability to achieve isotopic resolution is also shown in **Figure 3.4a**.

Figure 3.4b illustrates the sequence and labeled *ETD* tandem mass spectrum of the benchmark *ECD* peptide, substance P. This reaction proceeded for 150 ms in the LTQ before product ions were transferred to the Orbitrap mass analyzer. Unlike the spectrum of KAAAKAAAK, where nearly every backbone bond cleavage was observed in both the c and z directions, the spectrum of substance P contained mostly c-type product ions. This result was similar to that achieved by Coon and co-workers utilizing their dual electrospray source.[4]

Coon and co-workers demonstrated the ability of their pulsed dual ESI to perform *ETD* and its compatibility with automated data acquisition was also shown. In that work, however, it was realized that a limiting factor was the "relatively long times required for anion injection" – about 750 ms, due to pulsing on and off of the

high voltages.[4] The source presented herein does not suffer from the same time limitations because the voltages are held constant throughout an experiment, and the time needed to physically switch the emitters is quite fast (~ 40 ms). The ten-fold decrease in switching and injection time reported for this source equates to a 1.9 to 2.4-fold increase in tandem MS per unit time (respectively, LTQ and Orbitrap analysis), which will aid greatly for conducting higher order experiments such as LC-MS/MS.

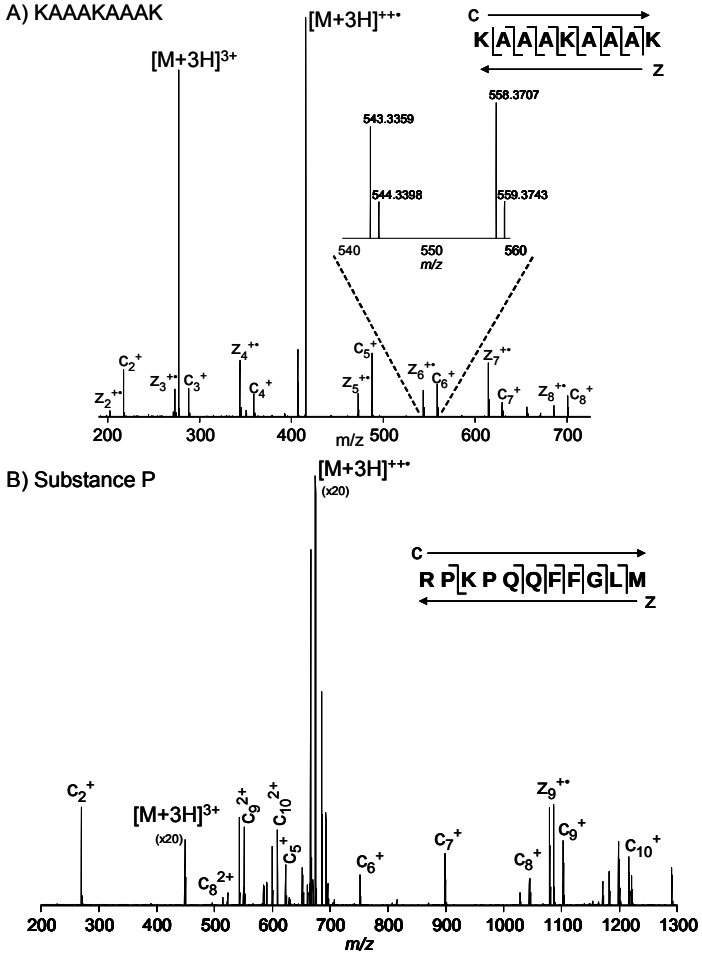


Figure 3.4 A) Dual ESI LTQ-Orbitrap ETD mass spectrum of the 3+ charge-state of the synthetic peptide KAAAKAAK, with c- and z-type ions labeled. The inset presents an expanded view of the c_6 and z_6 ions to highlight the isotopic resolution achieved by use of the orbitrap mass analyzer. B) Sequence of substance P and denotation of the observed c- and z-type product ions. Dual ESI LTQ-Orbitrap ETD mass spectrum of the 3+ charge-state of substance P.

3.4 Conclusions

The ability to couple a rapidly spatially-pulsed *ESI* source to a high resolving power instrument, such-as the hybrid LTQ-Orbitrap, and to perform *ETD* experiments is demonstrated. A maximum of 85 ms to inject the anion population and switch between the emitters was observed, which allows for

the collection of a larger amount of data during a given time period. This increased amount of data will lead to more identifications when coupled with online liquid chromatography. Some previously developed pulsed ESI sources can switch as quickly as 10 ms[20], but this does not consider the time it takes to develop a stable electrospray. A second major benefit of the constant voltage is electrospray stability, which results in more consistent anion densities and equates to more reproducible *ETD* experiments. This source presents a practical solution to the problems of previously developed dual AP ETD sources, allowing the focus of experiments to shift towards finding a more efficient ETD precursor anion that can be generated via *ESI*. Finally, we note operation of this device is far simpler and much more convenient to implement as the distance between *ESI* needles need not be manually optimized.

3.5 References

1. Coon, J.J., et al., Anion dependence in the partitioning between proton and electron transfer in ion/ion reactions. *Int. J. Mass Spectrom.*, 2004. **236**. 33-42.

2. Syka, J.E.P., et al., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Natl. Acad. Sci. U. S. A.*, 2004. **101**. 9528-9533.

3. Good, D.M. and J.J. Coon, Advancing proteomics with ion/ion chemistry. *Biotechniques*, 2006. **40**. 783-789.

4. McAlister, G.C., et al., Implementation of electron-transfer dissociation on a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.*, 2007. **79**. 3525-3534.

5. Gunawardena, H.P., J.F. Emory, and S.A. McLuckey, Phosphopeptide anion characterization via sequential charge inversion and electron- transfer dissociation. *Anal. Chem.*, 2006. **78**. 3788-3793.

6. Hogan, J.M., et al., Complementary structural information from a tryptic N-linked glycopeptide via electron transfer ion/ion reactions and collision-induced dissociation. *J. Proteome Res.*, 2005. **4**. 628-632.

7. Pitteri, S.J., et al., Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: Reactions of doubly and triply protonated peptides with SO2 center dot-. *Anal. Chem.*, 2005. **77**. 1831-1839.

8. Coon, J.J., et al., Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Natl. Acad. Sci. U. S. A.*, 2005. **102**. 9463-9468.

9. Swaney, D.L., et al., Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal. Chem.*, 2007. **79**. 477-485.

10. Roepstorff, P. and J. Fohlman, Proposal for a Common Nomenclature for Sequence Ions in Mass-Spectra of Peptides. *Biomedical Mass Spectrometry*, 1984. **11**. 601-601.

11. Johnson, R.S., et al., Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass-Spectrometer - Differentiation of Leucine and Isoleucine. *Anal. Chem.*, 1987. **59**. 2621-2625.

12. Herron, W.J., D.E. Goeringer, and S.A. McLuckey, Product ion charge state determination via ion/ion proton transfer reactions. *Anal. Chem.*, 1996. **68**. 257-262.

13. Stephenson, J.L. and S.A. McLuckey, Ion-ion proton transfer reactions of multiply-charged oligonucleotide cations. *Int. J. Mass Spectrom.*, 1997. **165**. 419-431.

14. Stephenson, J.L. and S.A. McLuckey, Adaptation of the Paul Trap for study of the reaction of multiply charged cations with singly charged anions. *Int. J. Mass Spectrom. Ion Processes*, 1997. **162**. 89-106.

15. McLuckey, S.A. and J.L. Stephenson, Ion ion chemistry of high-mass multiply charged ions. *Mass Spectrom. Rev.*, 1998. **17**. 369-407.

16. Syka, J.E.P., et al., Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.*, 2004. **3**. 621-626.

17. Liang, X.R., Y. Xia, and S.A. McLuckey, Alternately pulsed nanoelectrospray ionization/atmospheric pressure chemical ionization for ion/ion reactions in an electrodynamic ion trap. *Anal. Chem.*, 2006. **78**. 3208-3212.

18. Xia, Y., et al., Implementation of ion/ion reactions in a quadrupole/time-of-flight tandem mass spectrometer. *Anal. Chem.*, 2006. **78**. 4146-4154.

19. Huang, T.Y., et al., Electron-transfer reagent anion formation via electrospray ionization and collision-induced dissociation. *Anal. Chem.*, 2006. **78**. 7387-7391.

20. Xia, Y., X.R. Liang, and S.A. McLuckey, Pulsed dual electrospray ionization for ion/ion reactions. *J. Am. Soc. Mass Spectrom.*, 2005. **16**. 1750-1756.

21. Hannis, J.C. and D.C. Muddiman, A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2000. **11**. 876-883.

22. Flora, J.W., J.C. Hannis, and D.C. Muddiman, High-mass accuracy of product ions produced by SORI-CID using a dual electrospray ionization source coupled with FTICR mass spectrometry. *Anal. Chem.*, 2001. **73**. 1247-1251.

23. Nepomuceno, A.I., et al., Dual electrospray ionization source for confident generation of accurate mass tags using liquid chromatography Fourier

transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 2003. **75**. 3411-3418.

24. Williams, K., A.M. Hawkridge, and D.C. Muddiman, Sub parts-per-million mass measurement accuracy of intact proteins and product ions achieved using a dual electrospray ionization quadrupole Fourier transform ion cyclotron resonance mass spectrometer. *J. Am. Soc. Mass Spectrom.*, 2007. **18**. 1-7.

CHAPTER 4

Calibration Laws Based on Multiple Linear Regression Applied to Matrix-Assisted Laser Desorption / Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

4.1 Introduction

Since the advent of matrix-assisted laser desorption/ionization (*MALDI*), a soft ionization technique, mass spectrometry has become useful in the investigation of tissue and biological samples. [1-4] This is largely due to the protection of fragile bio-molecules by a matrix which serves the dual purpose of absorbing energy from a laser beam (thus avoiding direct impact with the sensitive biological material) while facilitating ionization. In addition to the benefits of soft-ionization, the pulsed nature of both *MALDI* and a *FT-ICR* mass analyzer is highly beneficial, as all desorbed material may be subjected to analysis by the mass spectrometer. Furthermore, MALDI-FT-ICR MS provides high mass resolving power and has the potential to provide high *mass measurement accuracy* (*MMA*); the latter can only be achieved if global and local *space-charge effects* are accounted for quantitatively.[5, 6]

External calibration has been used to account for frequency shifts and to thereby improve the *MMA* of measurements made by FT-ICR MS.[7] In 1990, Cody received a patent for accounting for the differences in *total ion population* (A_T) between calibration spectra and sample spectra.[8] The difference was accounted for by generating a calibration curve and comparing the trapping voltage to the (A_T) value for a calibrant compound. This was then utilized to correct the sample

spectra, based on relative ion current and trapping voltages. To account for abundance of a particular frequency observed, M.J. Smith derived a calibration law which has a calibration factor for *relative ion abundance* (A_S).[9] These two seminal works[8,9] have been oftentimes overlooked in the literature including our own previous publications related to this line of research for which we regret. Amster and co-workers developed a calibration curve to account for the difference between ion populations of the calibration spectrum and subsequent spectra resulting in <10 parts-per-million *MMA*.[10, 11] Based partly on the work of Amster[10, 11] and R.D. Smith[12], Oberg and Muddiman developed a novel external multiple linear calibration law for non-AGC ESI-FTICR measurements which allowed them to achieve <5 ppm *MMA*.[5] Amster and co-workers recently developed a novel stepwise-external calibration, in which calibration spectra are acquired at low trapping voltages that provide low ppm *MMA*.[13]

Previous studies have shown the global space-charge effect related to A_T in the ICR cell to be a linear function of the observed frequency, which decreases with increasing A_T.[10, 14] Similar studies have also shown that accounting for local space charge effects related to A_S can also improve *MMA*.[12] These perturbations in frequency due to global space charge effects (and local space charge effects, to a lesser extent) must be formally accounted for in the external calibration strategies in order to achieve high *MMA*.[5, 6]

The calibration laws applied herein have been previously applied to Fourier transform ion cyclotron resonance mass spectrometers utilizing an electrospray

ionization source, which because of multiple charging has a much smaller frequency range than MALDI-FT-ICR MS.[5, 6] The two external calibration laws were applied to MALDI-FT-ICR MS to account formally for only *total ion population* and both *total* and *relative ion population* to achieve high *mass measurement accuracy*.

4.2 Experimental

4.2.1 Materials

All materials were purchased from Sigma Aldrich (St. Louis, Missouri), with the exception of the HPLC grade acetonitrile (ACN) and HPLC grade water, which were obtained from Burdick Jackson (Muskegon, MI).

4.2.2 Polymer Sample Preparation

Polypropylene glycol oligomer with average molecular weight of 1000 Da was used as the calibrant in this experiment. A PPG-1000 stock solution was prepared by combining 10 µL of a solution containing 49.5 µL of PPG-1000 oligomer in 10 mL of ACN with 10 mL of a solution containing 7 mL ACN and 3 mL H_2O. The resulting PPG-1000 stock solution concentration was 5 ng/mL. A 100 mg/mL 2,5-dihydroxybenzoic acid (2,5-DHB) matrix stock solution was then prepared in 1:1 ACN : 50 mM NaCl / H2O. The PPG-1000 stock solution was then combined 1:1 with the 2,5-DHB matrix stock solution before spotting on the MALDI target.

4.2.3 Peptide Mixture Preparation

A peptide mixture (**Table 4.1**) containing bradykinin fragment 1-7, angiotensin II, P14R, ACTH fragment 18-39 and insulin oxidized B-chain was also analyzed, to

afford a wide range of *m/z* values. All individual peptide stock solutions were prepared by combining 10 nmoles of peptide with 1 mL of 0.1% TFA in H_2O, with the exception of bradykinin fragment 1-7 and insulin oxidized B-chain, which were prepared by combining 10 nmoles of each peptide with 1 mL of 1:1 ACN : 0.1% TFA

Table 4.1 Peptide calibration mixture information

Peptide	**MW**	**Solvent in Stock Peptide Solution**
Bradykinin Fragment 1-7	757.3997	ACN : 0.1%TFA = 1:1
Angiotensin II	1046.5423	0.1% TFA
P14R	1533.8582	0.1% TFA
ACTH Fragment 18-39	2465.1989	0.1% TFA
Insulin Oxidized B-Chain	3494.6513	ACN : 0.1%TFA = 1:1

in H2O. The peptide mixture was then created by combining 10 µL of each peptide solution. A 150 mg/mL 2,5-dihydroxybenzoic acid (2,5-DHB) matrix stock solution was then prepared in 1:1 ACN: 0.1% TFA in H2O. The peptide mixture (50 µL) was then combined with 50 µL of 150 mg/mL 2,5-DHB matrix solution, resulting in a 1:1 mixture of peptides : DHB matrix stock solution.

4.2.4 MALDI Target Preparation

Both PPG-1000 and peptide analyte/matrix mixtures were spotted by applying 0.8 µL onto a target obtained from Applied Biosystems (Foster City, CA). A passive drying method was used for the peptide/matrix mixture, whereas an active drying

method, using cool air from a heat gun to decrease sample drying time and prevent spot-spreading, was used for the PPG-1000 polymer/matrix mixture.

4.2.5 MALDI-FT-ICR MS Analysis

A Varian (Lake Forest, CA) FT-ICR MS comprised of a 9.4 Tesla horizontal bore superconducting magnet with a 128 mm bore (Cryomagnetics, Inc., Oakridge, TN), coupled with a ProMALDI ion source, was used to acquire all spectra in positive-ion mode. The power of the Nd:YAG frequency tripled-laser (355 nm) was varied from 17 to 20% to afford a range of *total ion populations* (A_T) to be analyzed. Several calibration spectra were acquired and processed using Varian Omega 2XP software and all data were acquired using a standard broadband pulsed sequence. Subsequent test spectra of the polymer and peptide mixture exhibiting A_T within the abundance range were acquired in order to determine *mass measurement accuracy* (*MMA*).

Law 4.1: $m/z = \dfrac{\beta_1}{f_{obs} - (\beta_0 + \beta_2 A_T)}$ Law 4.2: $m/z = \dfrac{\beta_1}{f_{obs} - (\beta_0 + \beta_2 A_T + \beta_3 A_S)}$

For external calibration Law 4.1, examined herein, β_0 accounted for electric field and β_1 for scaled magnetic field strength. Law 4.1 also formally took into account the *total ion population* in the ICR cell (A_T), which has the coefficient of β_2. Law 4.2 utilized an additional coefficient, β_3, which formally accounts for the *relative ion abundance* of a given species (A_S). A total of 80 *m/z* measurements were utilized for the quantification of the *MMA* results of each law and A_T range.

4.3 Results and Discussion

Figure 4.1 illustrates the eight sodium-adducted PPG-1000 oligomers utilized for calibration which range from *m/z* 911-1318. An example of the isotopic

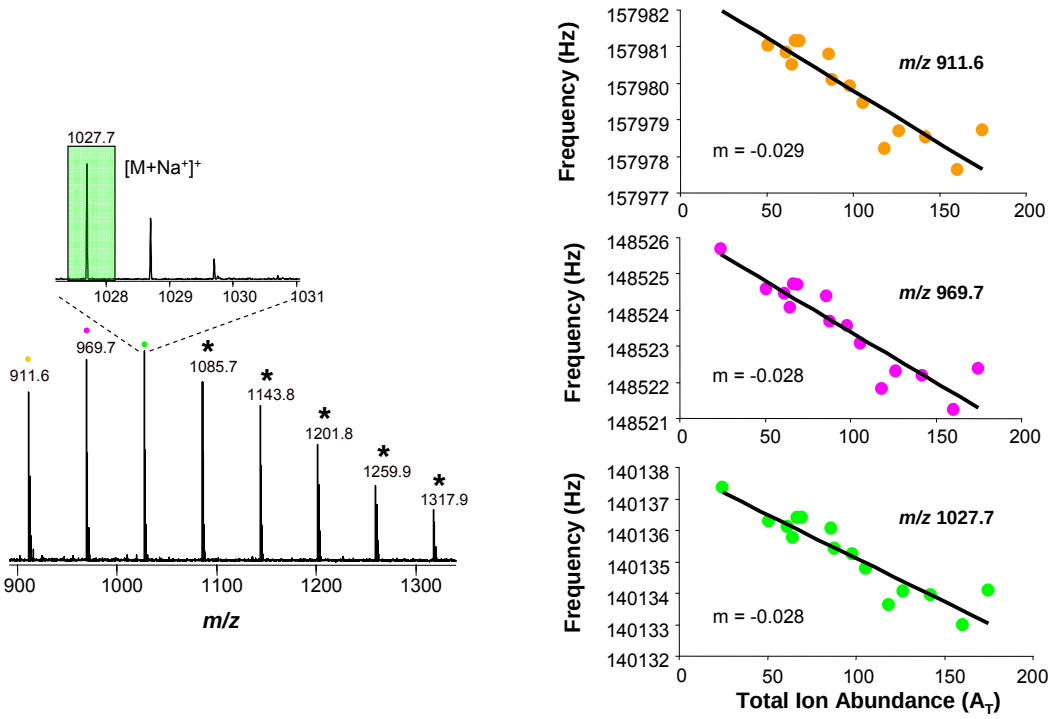


Figure 4.1 The left demonstrates the eight Na^+-PPG-1000 oligomers used (15-22), as well as the isotopic resolution achieved. The effects of total ion population on observed frequency (global space-charge effects) are demonstrated on the right. This linear relationship shows that, for a given *m/z* peak, observed frequency decreases as total ion abundance increases. The peaks corresponding to the graphs are for *m/z* = 911.6 (top), 969.7, and 1027.7 (bottom).

resolution achieved is demonstrated for the peak at *m/z* 1027.7. The frequency shifts induced by global *space-charge effects* are also shown. As A_T increased the observed frequency in the ICR cell decreased. The slope of the best fit line for each of the three *m/z* values is very similar because the shift in the observed frequencies should approximately be the same for a given increase in A_T.

The singly charged sodium-adducted PPG-1000 oligomers and peptide mixture (protonated with TFA) described a broad range of *m/z* values and *relative ion abundances* within each mass spectrum, as shown in **Figure 4.2**. Initially, the

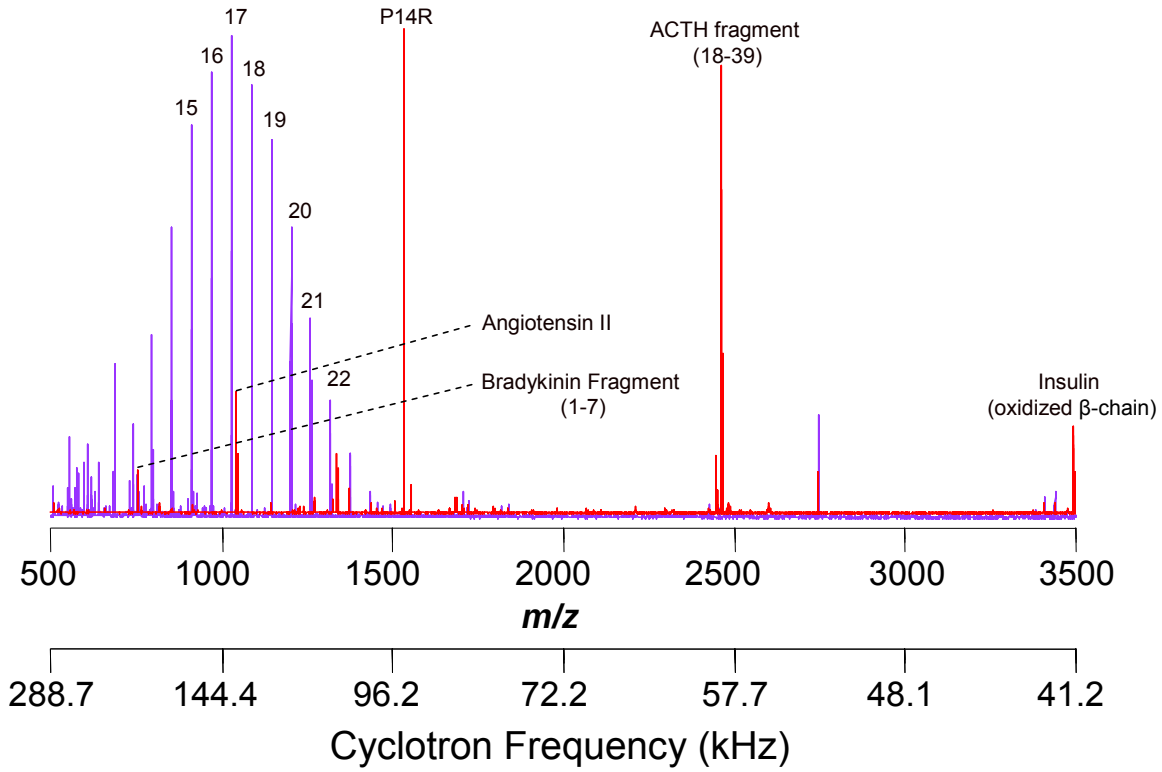


Figure 4.2 Sodium-adducted PPG-1000 oligomers n=15-22 and a peptide mixture are labeled. The peaks present described a broad range of *m/z* and A_S values.

instrument was calibrated utilizing the manufacturer's standard calibration procedure. This allowed for the definition of *mass measurement accuracy* (*MMA*) of all subsequent spectra utilizing the manufacturer's external calibration method. This method does take into account A_T, but the exact implementation of this information in the manufacturer's method is not known.

The goal of these experiments was to determine whether high *MMA* could be achieved by employing the calibration laws which utilize *multiple linear regression*

(*MLR*) to determine their coefficients, which formally account for A_T and A_S. Moreover, using a calibration set over a wide range of A_T values, it was possible to determine the constants (β_2 and β_3) needed for the *MLR* external calibration law. The results were then applied to a "test" set of polymer spectra, as well as a "test" set of peptide mixture spectra, in an effort to compare commonly used linear external calibration laws with our *MLR* calibration laws (**Figure 4.3**).

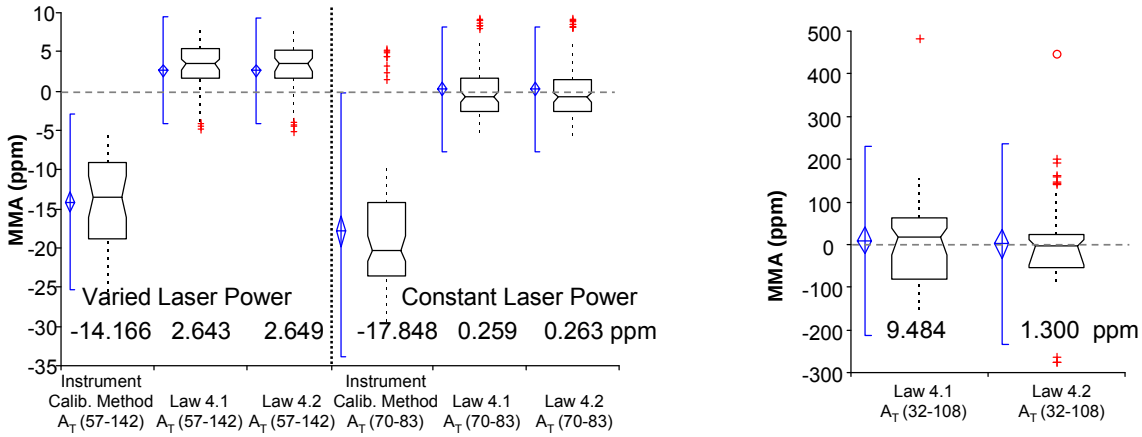


Figure 4.3 MMA Summary of Na^+-PPG-1000 and the peptide mixture are shown left and right, respectively. Box and whisker plots of MMA for the instrument calibration method (for Na^+-PPG-1000 only), Law 1, and Law 2 are illustrated, with mean MMA denoted. The blue diamond shows the mean and the 95% confidence interval of the mean, while the blue notched lines show the 95% parametric percentile range (2.5-97.5%). The notched box shows the median, lower and upper quartiles. The dotted-line connects the nearest observations within 1.5 IQRs (inter-quartile ranges) of the lower and upper quartiles. Red crosses (+) and circles (o) show observations more than 1.5 IQRs (near outliers) and 3.0 IQRs (far outliers) from the quartiles, respectively.

The application of the calibration Laws 4.1 and 4.2 (**Figure 4.3**) resulted in high *MMA* for the sodium-adducted PPG spectra. However, the highest *MMA* illustrated in **Figure 4.3** resulted from the constant laser power (17%) used to afford similar A_T values. Application of both calibration Law 4.1, accounting for A_T, and Law 4.2, accounting for both A_T and A_S, increased the *MMA* as compared to the default instrument calibration.

The instrument's calibration procedure produced data with mean *MMA* values of -14.166 ppm for the experiments where the *total ion population* (A_T) covered a large range. This large range (57-142) of A_T values was provided for by varying the laser power applied to the sample. The mean *MMA* achieved for the sodium-adducted PPG utilizing Law 4.1 was 2.643 ppm while Law 4.2 afforded a mean *MMA* value of 2.649 ppm. The application of the calibration laws shown in **Figure 4.3**, allowed for improvement of *MMA* values over the instrument's calibration method for this large A_T range. In addition, these calibration laws allowed for a decrease in the range of observed *MMA* values for a large A_T range.

The instrument's calibration method produced data with a mean *MMA* value of -17.848 ppm for experiments with a much smaller A_T range (70-83) than the previous set of data (vide supra) provided by the application of constant laser power. The mean *MMA* achieved for the sodium-adducted PPG was 0.259 ppm and 0.263 ppm for this data set utilizing Law 4.1 and Law 4.2, respectively. The constant laser power provided for the best overall *MMA* values for both Law 4.1 and Law 4.2; this data set also has zero included in its 95% confidence interval, which is indicative of a lack of systematic error. In addition, the data in **Figure 4.3** also display that the applied calibration laws reduce the range of *MMA* values observed when compared to the instrument's calibration procedure for the sodium-adducted PPG. These data demonstrate that, in order to achieve the best *MMA* routinely on MALDI-FT-ICR MS, it would be beneficial to optimize experimental conditions to produce similar A_T

values between different samples and utilize the calibration coefficients presented in **Figure 4.3**.

The peptide mixture in combination with Law 4.1 and Law 4.2 provided for mean *MMA* values of 9.484 and 1.300 ppm, respectively. The instrument's calibration method produced data with a mean *MMA* value of -84.807 ppm for the peptide mixture. **Figure 4.3** illustrates the large range observed in *MMA* values for the peptide data from Law 4.1 and Law 4.2. This is attributed to large frequency domain, 190152 – 41186 Hz for the peptide mixture versus 157981 – 109277 Hz for the sodium-adducted PPG-1000 that was covered by the peptide mixture calibration points, shown in **Figure 4.2**. Even with this large range of *MMA* values, the 95% confidence interval of the mean contained zero. This demonstrates a lack of systematic error for this data calibrated with Law 4.1 and Law 4.2. However, this data set reveals why MALDI-FT-ICR MS data reported to have high *MMA* values are typically over a small frequency scale, similar to what was utilized for the sodium-adducted PPG-1000 experiments described above.

4.4 Conclusions

MALDI-FT-ICR MS encompasses a much larger frequency range than that of a typical ESI-FT-ICR MS study. Through the application of novel calibration laws which account for both global and local space-charge effects, high *MMA* results were achieved, regardless of sample type (polymer or peptide mixture). The mean *MMA* values achieved in experiments with a smaller A_T range were in the parts-per-

billion range. This is an improvement of almost 2 orders of magnitude over the values achieved by the instrument's external calibration method. If the calibration laws presented herein are utilized and the instrument is tuned to produce a small A_T range this would surely aid in future experiments to identify unknowns by MALDI-FT-ICR MS.

4.5 References

1. Tanaka, K., et al., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.*, 1988. **2**. 151-153.

2. Hillenkamp, F., et al., Matrix-Assisted Laser Desorption Ionization Mass-Spectrometry of Biopolymers. *Anal. Chem.*, 1991. **63**. A1193-a1202.

3. An, H.J., et al., Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *J. Proteome Res.*, 2006. **5**. 1626-1635.

4. Caprioli, R.M., et al., Molecular imaging of biological samples by MALDI MS. *Abstr. Pap. Am. Chem. Soc.*, 1997. **214**. 113-ANYL.

5. Muddiman, D.C. and A.L. Oberg, Statistical evaluation of internal and external mass calibration laws utilized in Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 2005. **77**. 2406-2414.

6. Williams, D.K. and D.C. Muddiman, Parts-per-billion mass measurement accuracy achieved through the combination of multiple linear regression and automatic gain control in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem.*, 2007. **79**. 5058-5063.

7. Zhang, L.K., et al., Accurate mass measurements by Fourier transform mass spectrometry. *Mass Spectrom. Rev.*, 2005. **24**. 286-309.

8. Cody, R.B., *Method for External Calibration of Ion Cyclotron Resonance Mass Spectrometers*. 1990, Extrel FTMS, Inc.: United States of America. p. 15.

9. Smith, M.J., Intensity Dependence in FT-ICRMS Calibration. *ICR Ion Trap Newsletter*, 198940-43.

10. Easterling, M.L., T.H. Mize, and I.J. Amster, Routine part-per-million mass accuracy for high-mass ions: Space-charge effects in MALDI FT-ICR. *Anal. Chem.*, 1999. **71**. 624-632.

11. Taylor, P.K. and I.J. Amster, Space charge effects on mass accuracy for multiply charged ions in ESI-FTICR. *Int. J. Mass Spectrom.*, 2003. **222**. 351-361.

12. Masselon, C., et al., Mass measurement errors caused by 'local" frequency perturbations in FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2002. **13**. 99-106.

13. Wong, R.L. and I.J. Amster, Sub part-per-million mass accuracy by using stepwise-external calibration in Fourier transform ion cyclotron resonance mass spectrometry*. J. Am. Soc. Mass Spectrom.*, 2006. **17**. 1681-1691.

14. Francl, T.J., et al., Experimental-Determination of the Effects of Space-Charge on Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 189-199.

CHAPTER 5

Parts-Per-Billion Mass Measurement Accuracy Achieved through the Utilization of Multiple Linear Regression and Automatic Gain Control in a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer

5.1 Introduction

Fourier transform ion cyclotron resonance (*FT-ICR*), first developed by Comisarow and Marshall, provides the highest mass *resolving power* of all mass analyzers, which has the potential to translate into the highest *mass measurement accuracy* [1] over a wide *m/z* range.[2-6] High *MMA* is important for numerous applications of *FT-ICR* including the analysis of complex mixtures such as proteins[7-9], metabolites[10-12], and petroleum products. [13, 14] In order to realize the highest achievable *MMA* for a given FT-ICR MS system, frequency shifts due to *space-charge effects* must be accounted for using external or internal calibration strategies.[15, 16] While internal calibration[17-22] provides the best correction for *space-charge effects* it requires specialized hardware or software.

External calibration has been used to account for frequency shifts and to thereby improve the *MMA* of measurements made by *FT-ICR*.[23] For example, Amster and co-workers developed a calibration curve to account for the difference between ion populations of the calibration spectrum and subsequent spectra resulting in <10 parts-per-million *MMA*.[24, 25] Based partly on the work of Amster[24, 25] and Smith[26], Oberg and Muddiman reported a novel external calibration law which afforded <5 ppm *MMA*.[27] Amster and co-workers recently developed stepwise-

external calibration, in which calibration spectra are acquired at low trapping voltages that provide low ppm mass accuracy. This is then followed by collection of spectra at more routine trapping voltages where the major peaks, which appeared in the spectrum collected at low trapping voltages, were used as "internal calibrants" to calibrate the rest of the spectrum.[28] Smith and co-workers have demonstrated an external calibration method for liquid chromatography-mass spectrometry (LC-MS) that they base on mass accuracy histograms that are derived from sets of tentative assignments for species believed to be contained in the samples.[29]

The external calibration methods described above attempt to quantitatively account for differences in ion population between spectra. Recently, Hunt and co-workers have demonstrated the effectiveness of combining external calibration with *automatic gain control* (*AGC*) where the number of ions in the ICR cell are precisely controlled to fall within the external calibration range.[30] This approach allowed them to achieve *MMA* of <2 ppm. This external calibration strategy has been implemented on commercially-available FT-ICR MS instruments equipped with *AGC*.[30, 31] Muddiman and co-workers were able to achieve high *MMA* using this approach.[32] The Smith and Gygi research groups commonly report *MMA* values of ~5 ppm when utilizing *AGC* for external calibration procedures.[33-35]

While *AGC* does control the *total ion population* (A_T) present in the ICR cell, there is still variation in A_T between spectra. Gygi and co-workers have used the calibration coefficients given by the instrument to back calculate *cyclotron frequency*, and use that frequency to fit their data to a *multiple linear regression* calibration

equation based on that previously reported by Muddiman and Oberg[27], which accounts for A_T.[36]

Utilizing previously reported calibration laws in conjunction with *AGC*, we report improved *MMA* for a linear ion trap-Fourier transform ion cyclotron resonance mass spectrometer.[27, 37] These calibration equations formally account for *total ion population* due to the inability of *AGC* to precisely meter ions entering the ICR cell. In addition, one calibration equation also accounts for *relative ion abundance* of particular species. Finally, we demonstrate that the calibration strategies presented in this paper enable higher *AGC* limits (increased dynamic range), while still maintaining high *MMA*.

5.2 Experimental

5.2.1 Materials

Poly(propylene glycol) with an average molecular weight of 1000 Da, ammonium acetate (>99%), and formic acid were purchased from Sigma (St. Louis, MO). HPLC grade acetonitrile and high purity water were purchased from Burdick and Jackson (Muskegon, MI). 2-propanol (HPLC Grade) was purchased from Fischer Scientific (Fair Lawn, NJ). All materials were used as received.

A modified version of an electrospray ionization (*ESI*) source developed previously in this laboratory[38] was coupled to a hybrid Thermo-Electron (San Jose, CA) LTQ-FT Ultra MS equipped with an Oxford Instruments actively-shielded 7T superconducting magnet (Concord, MA). All spectra were acquired with a resolving

power at 400 *m/z* set to $100{,}000_{FWHM}$, along with *AGC* settings ranging from 5.0×10^5 to 3.0×10^7. Samples were introduced by direct infusion using a 100 µl gas-tight syringe (Hamilton, Las Vegas, NV) and the syringe pump on the LTQ-FT Ultra at a flow rate of 0.5 µl/min. The *ESI* emitter tips used were 360 µm o.d., 50 µm i.d. and tapered to 30 ± 1.0 µm i.d. (New Objective, Woburn, MA) and held at a constant potential of 2200 V for all experiments. Electrospray solutions were composed of 70:30 2-propanol:water with 0.5 mM ammonium acetate (NH_4OAc).

5.2.2 Experimental Design

Figure 5.1 shows the experimental design for the first series of experiments. Calibration of the instrument was carried out using a user-defined list of eleven monoisotoptic *m/z* values for ammonium-adducted PPG-1000 oligomers, which ranged from *m/z* = 732 to *m/z* = 1312. This calibration was conducted via the manufacturer's protocol which generates coefficients for five different *AGC* target values of which the largest 3 were used: 5.0×10^5, 1.0×10^6, and 3.0×10^6.

To enable the implementation of different calibration methods, the frequencies were obtained utilizing the diagnostic mode in the instrument software. The *AGC* level was first set to 5.0 x 10^5 and five calibration spectra were recorded, along with their respective cyclotron frequencies, to generate calibration coefficients (*vide infra*). This was followed by the collection of five test spectra including their cyclotron frequencies, to determine the achievable *mass measurement accuracy* (*MMA*). In an effort to reduce systematic error with respect to *AGC* levels, this procedure was then repeated for *AGC* target levels of 3.0 x 10^6 and 1.0 x 10^6 in that

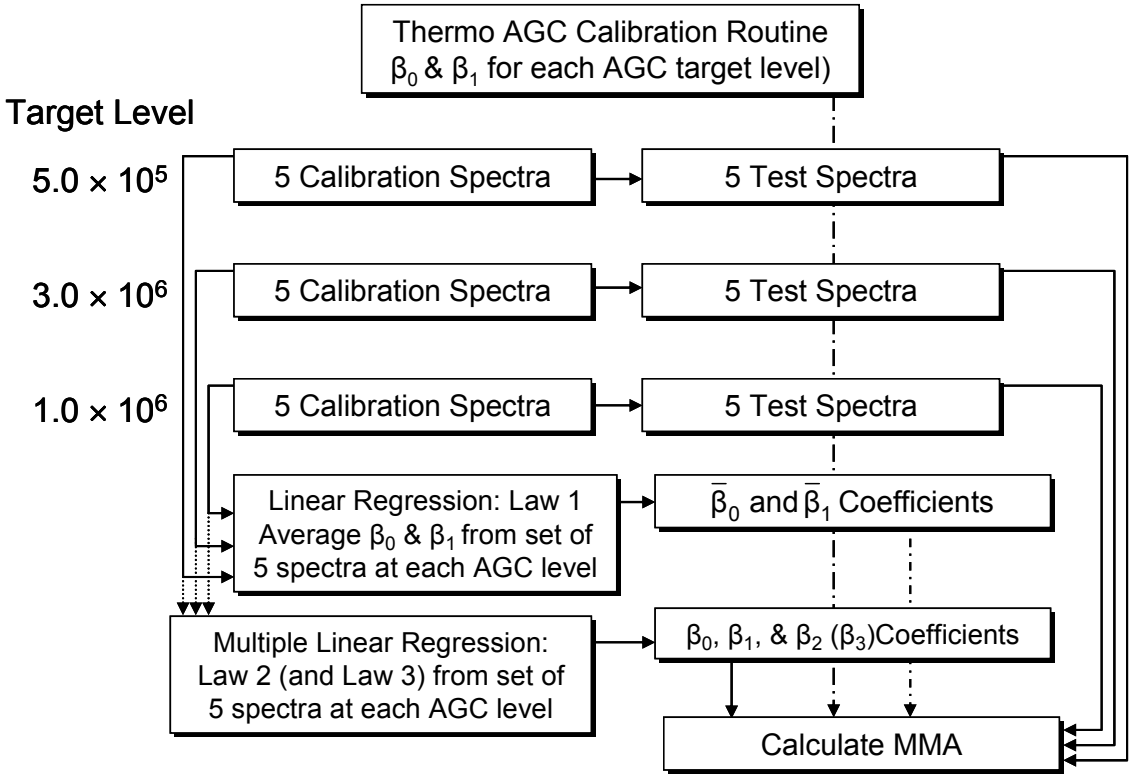


Figure 5.1 The experimental design is shown, which starts with the calibration of the instrument. This is followed by the collection of 5 calibration and 5 test spectra at each of the specified AGC target levels. Information from the calibration spectra was used in linear regression and multiple linear regression to determine calibration coefficients for Law 1, Law 2 and Law 3. Mass measurement accuracy was calculated using these coefficients and information from the test spectra. MMA was also calculated from the *m/z* reported by the instrument manufacturer's software.

order. The *mass measurement accuracy* for the five test spectra at each *AGC* level were first determined using the manufacturer's software.

The data from each of the five calibration spectra at each *AGC* target level were fit, by linear regression, to **Equation 5.1**, developed by Gross and co-workers[37], where the response is the product of the observed frequency (f_{obs}) and theoretical *m/z* values for each oligomer and the predictor is $1/f_{obs}$. The y-intercept yields β_0 (the scaled magnetic field strength) and the slope yields β_1, the contribution for electric fields including trapping voltages (constant during the measurement) and

the ions themselves (variable). All five calibration spectra were used to generate average values of β_0 and β_1. These average calibration coefficients were then used to determine the *MMA* for the five test spectra at each *AGC* level using **Equation 5.2**. The use of these two equations throughout this chapter will be denoted as **Law 5.1**.

Equation 5.1: $m/z_{theoretical} \times f_{obs} = \bar{\beta}_0 + \frac{\bar{\beta}_1}{f_{obs}}$

Equation 5.2: $m/z = \frac{\bar{\beta}_0}{f_{obs}} + \frac{\bar{\beta}_1}{f_{obs}^2}$

The data from the calibration spectra at each *AGC* target level were also fit using a *multiple linear regression* calibration law previously reported, which is shown in **Equation 5.3**.[27] β_0 accounts for electric field and β_1 is the scaled magnetic field strength. This calibration law also formally takes into account the *total ion population* in the ICR cell (A_T), which has the coefficient of β_2. These calibration coefficients were then used to determine the *MMA* for the five test spectra at each *AGC* level using **Equation 5.4**. The use of these two equations will be denoted as **Law 5.2**.

Equation 5.3: $f_{obs} = \beta_0 + \beta_1(z/m) + \beta_2 A_T$

Equation 5.4: $m/z = \frac{\beta_1}{f_{obs} - (\beta_0 + \beta_2 A_T)}$

The data from the calibration spectra at each *AGC* target level, where the instrument performs its calibration routine, were fit using a *multiple linear regression*

calibration law previously reported, which is shown in **Equation 5.5**.[27] The coefficients for this calibration law are the same as listed for **Law 5.2**, with the addition of β_3 which formally accounts for the *relative ion abundance* of a given species (A_S). These calibration coefficients were then used to determine the *MMA* for the five test spectra at each *AGC* level using **Equation 5.6**. The use of these two equations will be denoted as **Law 5.3**.

Equation 5.5: $f_{obs} = \beta_0 + \beta_1(z/m) + \beta_2 A_T + \beta_3 A_S$

Equation 5.6: $m/z = \dfrac{\beta_1}{f_{obs} - (\beta_0 + \beta_2 A_T + \beta_3 A_S)}$

The second set of experiments followed the same scheme except that they were performed at *AGC* levels that were above the range at which the instrument performs mass calibration: 5.0 x 10^6, 1.0 x 10^7, and 3.0 x 10^7; the latter is the highest *AGC* target level allowed by the software.

All statistical analysis was performed using Analyse-it® for Microsoft Excel®. It is important to note that the numerical values reported in this chapter should be reproducible (but not identical) in other laboratories.

5.3 Results and Discussion

The relative proportion of each monoisotopic *m/z* for each oligomer to A_T as a function of *m/z* for the three different *AGC* target values equal to the three highest settings at which the instrument's calibration routine runs are shown in **Figure 5.2**. The overall shape of the oligomer distribution did not change with respect to

changes in *AGC* target value. The coefficients from the instrument's calibration are shown for comparison to the calculated coefficients from **Law 5.1**. As expected, the β_0 coefficient remains essentially constant since magnetic field does not significantly

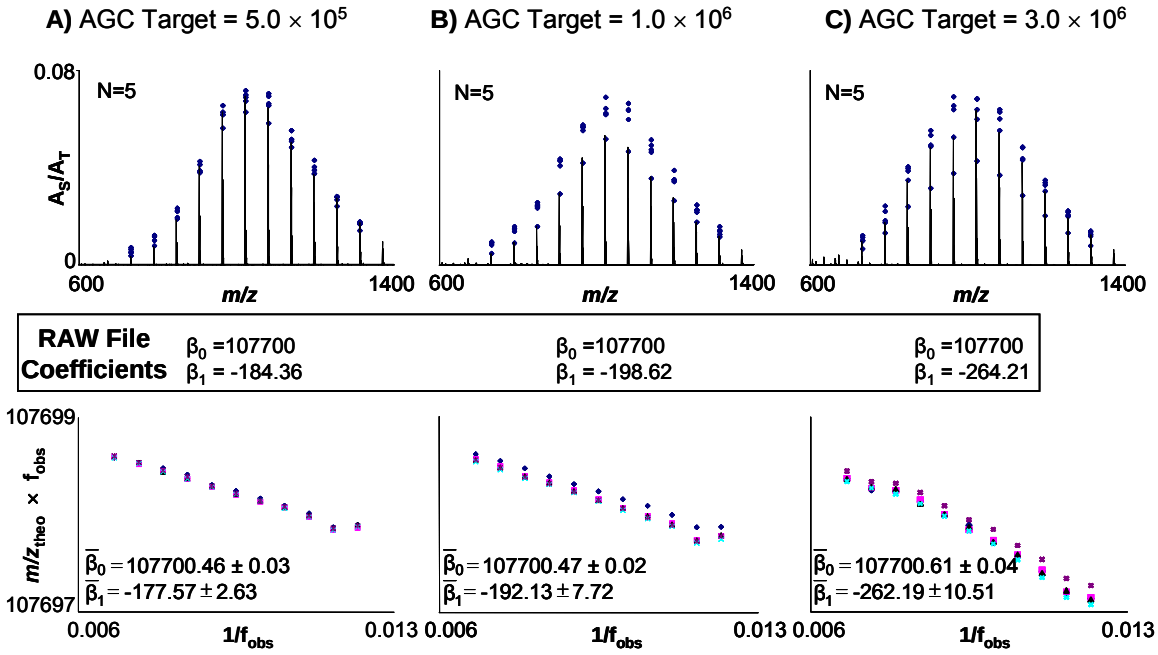


Figure 5.2 A) At the top, the range of A_S/A_T for calibrant peaks in each calibration spectrum are shown; overlaid is a representative spectrum from AGC target level of 5.0×10^5. Below, the calibration coefficients from the RAW file header are given. The bottom is a plot which displays β_0 as the y-intercept and β_1 as the slope, which were acquired from applying Law 1 to the data. Mean values of the coefficients with their 95% confidence interval are displayed. The same information is also shown for AGC target levels of B) 1.0×10^6 and C) 3.0×10^6.

decrease throughout the experiment. As the *AGC* increases, the magnitude of β_1 also increases, which accounts for electric field effects in the ICR cell. The variability of β_1 also increases which demonstrates that *AGC* has difficulty precisely reaching its target value as the *AGC* limit is increased. Also, the data becomes less linear as *AGC* target value increases which we attribute to the increase in *space-charge effects* at higher ion population.[15, 16]

The calculation of the total ion abundance for each spectrum, A_T, was carried out using two different methods. Method 1 summed the abundances of all peaks listed in the *m/z* list (SUM) and method 2 used the "total ion current" (TIC) from the RAW file header. The resultant from both the SUM and TIC were then multiplied by the “ion injection time” listed in the RAW file header to calculate the value of A_T. The reason for this multiplication is because both the TIC and individual abundances in the m/z list that make up the SUM method are normalized by the “ion injection time”. The top of **Figure 5.3** shows the individual A_T values for both methods 1 and 2 for the calibrant and test spectra at each *AGC* target. While the calculated A_T's were different, both methods provided statistically similar results when using the *multiple linear regression* strategy (data not shown). Method 2 is the preferred method, owing to its simplicity, and thus, this method was used herein.

Figure 5.3 shows a box and whisker plot of the resulting *MMA* data for the three highest *AGC* target values for which the instrument performs its calibration routine. The decreasing *MMA* using the manufacture’s calibration, as demonstrated by the decrease in *MMA* with increasing *AGC* target level, shows systematic error as the *AGC* target value increases. *MMA* from the instrument manufacturer’s *m/z* values was calculated to have a mean of -0.543, -0.513, and -1.117 ppm for the *AGC* target values of 5.0×10^5, 1.0×10^6, and 3.0×10^6, respectively. It is also important to note that the 95% confidence interval of the mean for these *MMA*’s does not include zero for any of the specified *AGC* target levels. The negative

values are a result of the ICR cell having fewer ions in the cell, at the time of analysis, than the calibration spectrum.

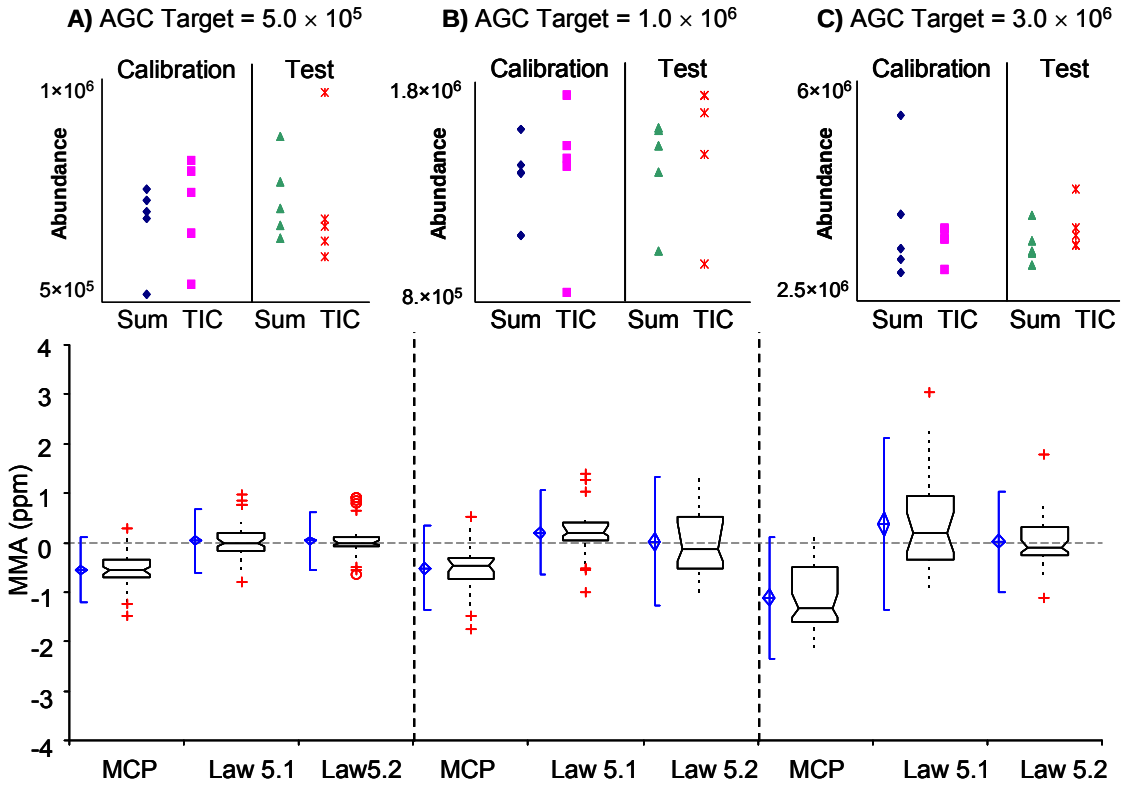


Figure 5.3 A) For an AGC target value of 5.0×10^5 the result from calculating A_T by methods 1 and 2 for both calibration and test spectra is shown at the top. Box and whisker plots of MMA for the manufacturer's calibration procedure (MCP), Law 1, and Law 2 are shown on the bottom. The blue diamond shows the mean and the 95% confidence interval of the mean, while the blue notched lines show the 95% parametric percentile range (2.5-97.5%). The notched box shows the median, lower and upper quartiles. The dotted-line connects the nearest observations within 1.5 IQRs (inter-quartile ranges) of the lower and upper quartiles. Red crosses (+) and circles (o) show observations more than 1.5 IQRs (near outliers) and 3.0 IQRs (far outliers) from the quartiles, respectively. The same information is also shown for AGC target levels of B) 1.0×10^6 and C) 3.0×10^6.

Using **Law 5.1** to calculate the average calibration coefficients (β_0 and β_1) for each of the five calibration spectra corrected for this systematic error, which is evident from the mean and median of the *MMA* data residing very close to zero for all *AGC* target levels, as shown in **Figure 5.3**. Even though the equation for fitting

and calculating *m/z* values for **Law 5.1** and the instrument manufacturer's calibration are the same, our data exhibited higher *MMA* because the calibration coefficients were calculated as the average for 5 different spectra at each *AGC* level, as shown by **Figure 5.1**. Using the average calibration coefficients determined from **Law 5.1,** we obtained *MMA* mean values of 0.036, 0.202, and 0.376 ppm at the *AGC* target levels 5.0×10^5, 1.0×10^6, and 3.0×10^6. The 95% confidence interval of the mean for this data includes zero for 5.0×10^5 (-0.053 to 0.126 ppm); however, not for 1.0×10^6 or 3.0×10^6 data sets.

Law 5.2 produced *MMA* mean values of 34, 23, and 27 parts-per-billion (ppb) for the same range of *AGC* target values as reported above. The 95% confidence interval of the mean includes zero for all *AGC* target levels in this range: -48 ppb to 116 ppb, -156 ppb to 203 ppb, and -114 ppb to 167 ppb for 5.0×10^5, 1.0×10^6, and 3.0×10^6, respectively. **Law 5.2** enhances *MMA* because it formally accounts for changes in A_T. **Law 5.1** is almost able to perform as well as **Law 5.2**, even while not formally accounting for A_T because **Law 5.1** is an average of 5 spectra instead of a single spectrum, as is the case in the instrument manufacturer's calibration procedure. However, since all of the results for **Law 5.2** had 95% confidence intervals that included zero, we can confidently state that **Law 5.2** performed better than alternative calibration laws because it accounts for A_T formally.

Gorshkov et. al. has previously shown that the frequency shift due to *space-charge effects* is dependant on the relative abundance of an ion in the ICR cell.[39] This *relative ion abundance* of a given species (A_S) has been previously utilized in

Law 5.3 on a FT-ICR MS without *AGC* by Muddiman and Oberg.[27] Moreover, Muddiman and Oberg demonstrated that both A_S and A_T were statistically significant in the calculation of *MMA* for data acquired without *AGC*.[27]

Law 5.3 formally accounts for both A_T and A_S, which narrows the range over which the *MMA* values are spread, especially for high *AGC* target values, when compared to **Law 5.2**, as shown in **Figure 5.4**. For this set of experiments, the

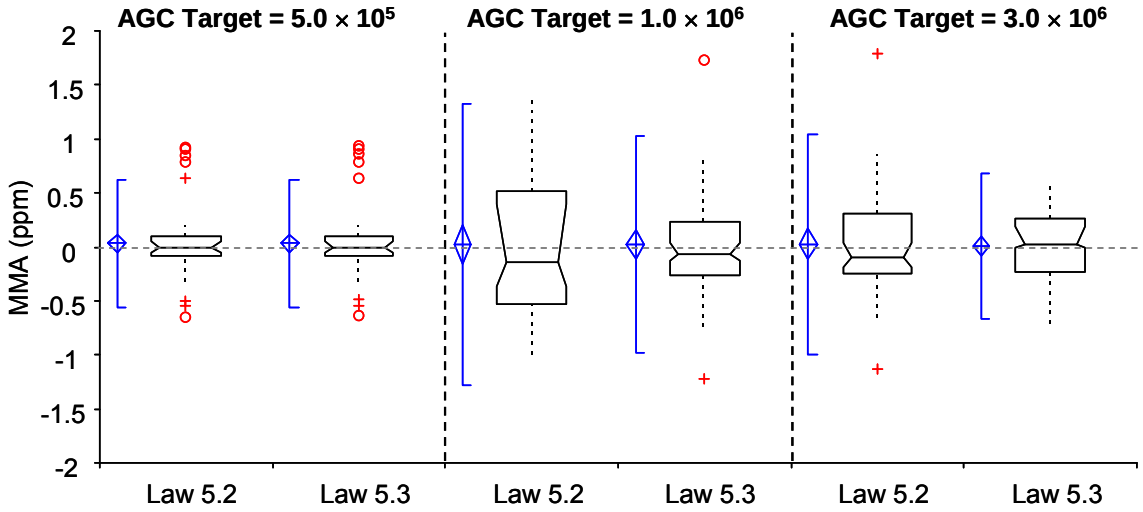


Figure 5.4 Box and whisker plots of MMA for Law 2 and Law 3 at each of the specified AGC target levels. The notation for these box and whisker plots is identical to that used in Figure 5.3.

calibration spectra and test spectra had similar A_S values, since both were ammonium-adducted PPG-1000 oligomers. The mean *MMA* values produced by the application of **Law 5.3** are 34, 24, and 9 ppb for *AGC* target levels of 5.0×10^5, 1.0×10^6, and 3.0×10^6, respectively. The 95% confidence interval of the mean includes zero for all *AGC* target levels in this range: -48 ppb to 116 ppb, -115 ppb to 163 ppb, and -84 ppb to 102 ppb for the same *AGC* target values as reported above.

Figure 5.5 displays the relative proportion of the monoisotopic *m/z* value for each oligomer to A_T as a function of *m/z* for the range of *AGC* target values outside of the calibration range of the instrument. The overall distribution does not change

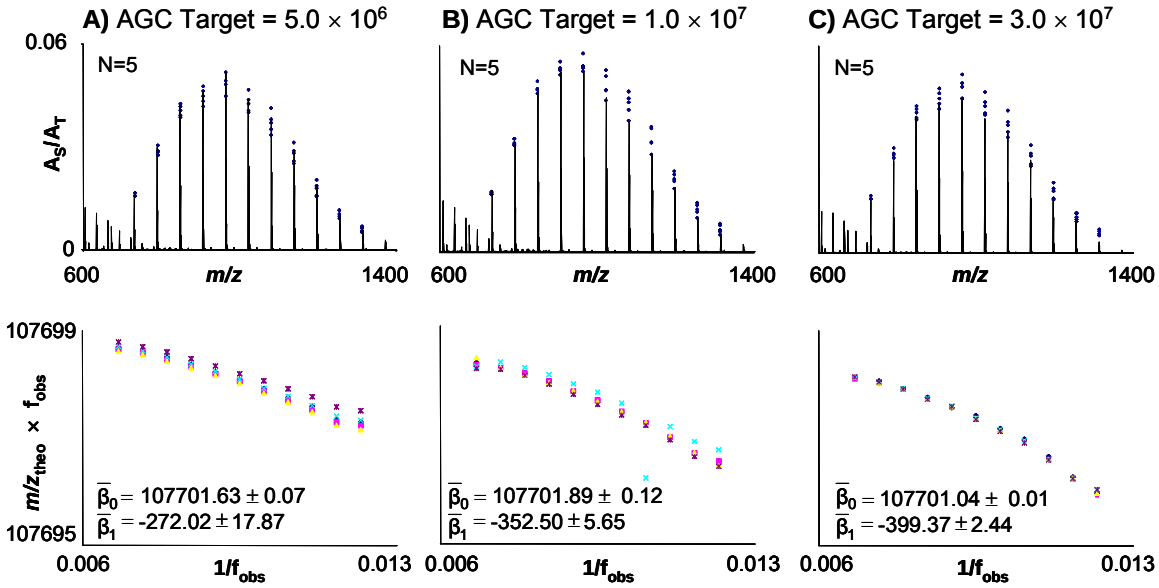


Figure 5.5 A) The top shows the range of A_S/A_T for calibrant peaks in each calibration spectrum; overlaid is a representative spectrum from the AGC target level of 5.0 × 10^6. Below, the calibration coefficients from the RAW file header are given. The bottom is a plot which displays β_0 as the y-intercept and β_1 as the slope, which were acquired from applying Law 1 to the data. Mean values of these coefficients with their 95% confidence interval are displayed. The same information is also shown for AGC target levels of B) 1.0 × 10^7 and C) 3.0 × 10^7.

over the range of *AGC* target values. As previously observed, β_0 remains similar throughout these experiments and the magnitude of β_1 does change systematically as A_T is increased.

There are no comparisons to be made to the instrument manufacturer's calibration coefficients for this data set because at *AGC* target levels above 3.0 × 10^6 the instrument manufacturer's coefficients remain the same because 3.0 × 10^6 is the highest *AGC* level at which the calibration routine operates. As the *AGC* target value increases for this data set, the departure from linearity of the data becomes

greater than that of the data in **Figure 5.2**; however, the variation of the β_1 coefficient surprisingly decreased, which is under further investigation in our laboratory.

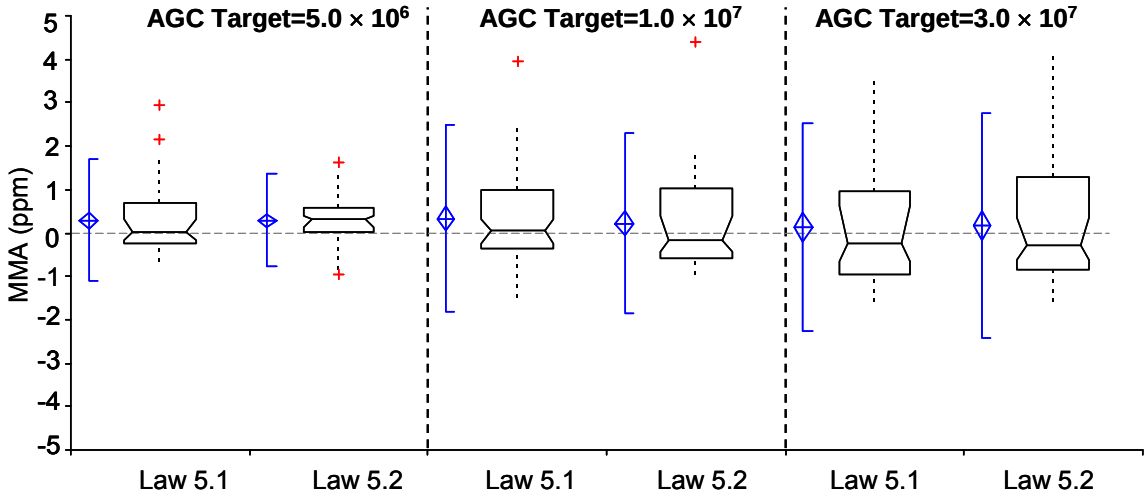


Figure 5.6 Box and whisker plots of MMA for Law 1 and Law 2 at each of the specified AGC target levels. The notation for these box and whisker plots is identical to that used in Figure 5.3.

MMA data from experiments with an *AGC* target value exceeding those at which the instrument manufacturer's software calibrates the instrument are presented in **Figure 5.6**. The resulting *MMA* from applying the average calibration coefficients obtained using **Law 5.1** for the five calibration spectra to the data at the higher *AGC* target levels is still able to maintain means close to the sub-ppm level. **Law 5.1** resulted in mean *MMA* values of 287, 336, and 137 ppb for the *AGC* target values of 5.0×10^6, 1.0×10^7, and 3.0×10^7, respectively. For the *AGC* target level of 3.0×10^7, the 95% confidence interval of the mean for **Law 5.1** was -185 ppb to 523 ppb; however, for the other two *AGC* target levels in this set of experiments the 95% confidence interval did not include zero.

Law 5.2 resulted in mean *MMA* values of 293, 219, and 169 ppb for the same range of *AGC* target values reported above. For *AGC* target levels 1.0×10^7 and 3.0

× 10^7 the 95% confidence interval of the mean included zero (-68 ppb to 506 ppb and -185 ppb to 523 ppb, respectively). Through the use of these calibration procedures we are able to extend the dynamic range of the LTQ-FT, while still maintaining high levels of *mass measurement accuracy*.

5.4 Conclusions

Data were acquired utilizing a LTQ-FT in diagnostic mode, which allowed for the recording of the *cyclotron frequency* of the ions in the ICR cell. Frequency data allowed for the calibration of spectra of ammonium-adducted PPG-1000 oligomers utilizing two different calibration laws: **Law 5.1** was employed with calibration coefficients determined for each of 5 spectra at a specific *AGC* target level and then averaged; and **Law 5.2**, a *multiple linear regression* law which formally accounts for the A_T at a specific *AGC* target level. These laws enabled parts-per-billion *MMA* over a wide range of *AGC* target levels even in excess of the *AGC* target level at which the instrument calibration routine permits, thus extending the dynamic range of the instrument. Between the two main laws investigated, **Law 5.2** allowed for the best *MMA* because it can correct for the inability of *AGC* to precisely aliquot a specified number of ions to the ICR cell, especially at higher *AGC* target levels. **Law 5.3**, applied only to the three lower *AGC* target levels examined in these experiments, was able to attain very high *MMA*, and fundamentally will likely prove to be the best since it formally accounts for both A_T and A_S. Coupled with stronger magnetic fields[40], this method of calibrating *FT-ICR* data has the potential to improve

MMA even further. These calibration laws should also translate to improving *MMA* on the LTQ-Orbitrap.[41] This work could be extended to intact proteins and protein fragments in MS^n data, which should benefit from the extended dynamic range which these calibration laws and procedures provide in future experiments.

5.5 References

1. Gerber, S.A., et al., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS*. Natl. Acad. Sci. U. S. A.*, 2003. **100**. 6940-6945.

2. Comisarow, M.B. and A.G. Marshall, Fourier-Transform Ion-Cyclotron Resonance Spectroscopy*. Chem. Phys. Lett.*, 1974. **25**. 282-283.

3. Comisarow, M.B. and A.G. Marshall, Frequency-Sweep Fourier-Transform Ion-Cyclotron Resonance Spectroscopy*. Chem. Phys. Lett.*, 1974. **26**. 489-490.

4. Gorshkov, M.V., S.H. Guan, and A.G. Marshall, Masses of Stable Neon Isotopes Determined at Parts-Per-Billion Precision by Fourier-Transform Ion-Cyclotron Resonance Mass-Spectrometry*. Int. J. Mass Spectrom. Ion Processes*, 1993. **128**. 47-60.

5. Rodgers, R.P., et al., Resolution, elemental composition, and simultaneous monitoring by Fourier transform ion cyclotron resonance mass spectrometry of organosulfur species before and after diesel fuel processing*. Anal. Chem.*, 1998. **70**. 4743-4750.

6. He, F., C.L. Hendrickson, and A.G. Marshall, Baseline mass resolution of peptide isobars: A record for molecular mass resolution*. Anal. Chem.*, 2001. **73**. 647-650.

7. Emmett, M.R., et al., Application of micro-electrospray liquid chromatography techniques to FT-ICR MS to enable high-sensitivity biological analysis*. J. Am. Soc. Mass Spectrom.*, 1998. **9**. 333-340.

8. Smith, R.D., et al., An accurate mass tag strategy for quantitative and high-throughput proteome measurements*. Proteomics*, 2002. **2**. 513-523.

9. Kujawinski, E.B., P.G. Hatcher, and M.A. Freitas, High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: Improvements and comparisons*. Anal. Chem.*, 2002. **74**. 413-419.

10. Brown, S.C., G. Kruppa, and J.L. Dasseux, Metabolomics applications of FT-ICR mass spectrometry*. Mass Spectrom. Rev.*, 2005. **24**. 223-231.

11. Breitling, R., A.R. Pitt, and M.P. Barrett, Precision mapping of the metabolome*. Trends Biotechnol.*, 2006. **24**. 543-548.

12. Sanders, M., et al., Utility of the hybrid LTQ-FTMS for drug metabolism applications. *Curr. Drug Metab.*, 2006. **7**. 547-555.

13. Qian, K.N., et al., Resolution and identification of elemental compositions for more than 3000 crude acids in heavy petroleum by negative-ion microelectrospray high-field Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuels*, 2001. **15**. 1505-1511.

14. Marshall, A.G. and R.P. Rodgers, Petroleomics: The next grand challenge for chemical analysis. *Acc. Chem. Res.*, 2004. **37**. 53-59.

15. Jeffries, J.B., S.E. Barlow, and G.H. Dunn, Theory of Space-Charge Shift of Ion-Cyclotron Resonance Frequencies. *Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 169-187.

16. Francl, T.J., et al., Experimental-Determination of the Effects of Space-Charge on Ion-Cyclotron Resonance Frequencies. *Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 189-199.

17. Kaiser, N.K., G.A. Anderson, and J.E. Bruce, Improved mass accuracy for tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2005. **16**. 463-470.

18. Hannis, J.C. and D.C. Muddiman, A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2000. **11**. 876-883.

19. Flora, J.W., J.C. Hannis, and D.C. Muddiman, High-mass accuracy of product ions produced by SORI-CID using a dual electrospray ionization source coupled with FTICR mass spectrometry. *Anal. Chem.*, 2001. **73**. 1247-1251.

20. Nepomuceno, A.I., et al., Detection of genetic variants of transthyretin by liquid chromatography-dual electrospray ionization Fourier-transform ion-cyclotron-resonance mass spectrometry. *Clin. Chem.*, 2004. **50**. 1535-1543.

21. Bruce, J.E., et al., Obtaining more accurate Fourier transform ion cyclotron resonance mass measurements without internal standards using multiply charged ions. *J. Am. Soc. Mass Spectrom.*, 2000. **11**. 416-421.

22. Williams Jr., D.K., A.M. Hawkridge, and D.C. Muddiman, Sub Parts-Per-Million Mass Measurement Accuracy of Intact Proteins and Product Ions Achieved Using a Dual Electrospray Ionization Quadrupole Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *J. Am. Soc. Mass Spectrom.*, 2007. **18**. 1-7.

23. Zhang, L.K., et al., Accurate mass measurements by Fourier transform mass spectrometry. *Mass Spectrom. Rev.*, 2005. **24**. 286-309.

24. Easterling, M.L., T.H. Mize, and I.J. Amster, Routine part-per-million mass accuracy for high-mass ions: Space-charge effects in MALDI FT-ICR. *Anal. Chem.*, 1999. **71**. 624-632.

25. Taylor, P.K. and I.J. Amster, Space charge effects on mass accuracy for multiply charged ions in ESI-FTICR. *Int. J. Mass Spectrom.*, 2003. **222**. 351-361.

26. Masselon, C., et al., Mass measurement errors caused by 'local" frequency perturbations in FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2002. **13**. 99-106.

27. Muddiman, D.C. and A.L. Oberg, Statistical evaluation of internal and external mass calibration laws utilized in Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 2005. **77**. 2406-2414.

28. Wong, R.L. and I.J. Amster, Sub part-per-million mass accuracy by using stepwise-external calibration in Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2006. **17**. 1681-1691.

29. Tolmachev, A.V., et al., Mass measurement accuracy in analyses of highly complex mixtures based upon multidimensional recalibration. *Anal. Chem.*, 2006. **78**. 8374-8385.

30. Syka, J.E.P., et al., Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.*, 2004. **3**. 621-626.

31. Peterman, S.M. and J.J. Mulholland, A novel approach for identification and characterization of glycoproteins using a hybrid linear ion trap/FT-ICR mass spectrometer. *J. Am. Soc. Mass Spectrom.*, 2006. **17**. 168-179.

32. Johnson, K.L., et al., Accurate mass precursor ion data and tandem mass spectrometry identify a class I human leukocyte antigen A*0201-presented peptide originating from vaccinia virus. *J. Am. Soc. Mass Spectrom.*, 2005. **16**. 1812-1817.

33. Dieguez-Acuna, F.J., et al., Characterization of mouse spleen cells by subtractive proteomics. *Mol. Cell. Proteomics*, 2005. **4**. 1459-1470.

34. Belov, M.E., et al., Automated gain control and internal calibration with external ion accumulation capillary liquid chromatography-electrospray

ionization-fourier transform ion cyclotron resonance*. Anal. Chem.*, 2003. **75**. 4195-4205.

35. Page, J.S., et al., Automatic gain control in mass spectrometry using a jet disrupter electrode in an electrodynamic ion funnel*. J. Am. Soc. Mass Spectrom.*, 2005. **16**. 244-253.

36. Haas, W., et al., Optimization and use of peptide mass measurement accuracy in shotgun proteomics*. Mol. Cell. Proteomics*, 2006. **5**. 1326-1337.

37. Ledford, E.B., D.L. Rempel, and M.L. Gross, Space-Charge Effects in Fourier-Transform Mass-Spectrometry - Mass Calibration*. Anal. Chem.*, 1984. **56**. 2744-2748.

38. Nepomuceno, A.I., et al., Dual electrospray ionization source for confident generation of accurate mass tags using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry*. Anal. Chem.*, 2003. **75**. 3411-3418.

39. Gorshkov, M.V., A.G. Marshall, and E.N. Nikolaev, Analysis and Elimination of Systematic-Errors Originating from Coulomb Mutual Interaction and Image Charge in Fourier-Transform Ion-Cyclotron Resonance Precise Mass Difference Measurements*. J. Am. Soc. Mass Spectrom.*, 1993. **4**. 855-868.

40. Marshall, A.G. and S.H. Guan, Advantages of high magnetic field for Fourier transform ion cyclotron resonance mass spectrometry*. Rapid Commun. Mass Spectrom.*, 1996. **10**. 1819-1823.

41. Hu, Q.Z., et al., The Orbitrap: a new mass spectrometer*. J. Mass Spectrom.*, 2005. **40**. 430-443.

CHAPTER 6

Utilizing Artificial Neural Networks in MATLAB to Achieve Parts-per-Billion Mass Measurement Accuracy with a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer

6.1 Introduction

Fourier transform ion cyclotron resonance (*FT-ICR*) mass spectrometry has the ability to detect ions with unparalleled *mass measurement accuracy*, even when working with complex analytes such as proteins, metabolites, and other molecules found in biological, biochemical, and physical science fields.[1-5] The fundamental relationship between *cyclotron frequency* and *m/z* has been under investigation since the introduction of FT-ICR MS in 1974 by Comisarow and Marshall.[1] In order to realize the highest *mass measurement accuracy* (*MMA*) achievable for a given FT-ICR MS system, frequency shifts due to *space-charge effects* must be accounted for using external or internal calibration strategies.[6, 7] While internal calibration[8-13] provides the best correction for *space-charge effects* it often requires specialized hardware or software.

In order to account for frequency shifts, external calibration has been utilized to improve the *MMA* of measurements made by *FT-ICR*.[14] Amster and co-workers developed a calibration curve to account for the differences in ion populations between the calibration spectrum and subsequent spectra acquired for data analysis resulting in <10 parts-per-million *MMA*.[15, 16] Based partly on both the work of Amster[15, 16] and Smith[17], Oberg and Muddiman reported a novel external calibration

law which provided data with <5 ppm *MMA*.[18] Amster and co-workers have also developed stepwise-external calibration, in which calibration spectra are acquired at low trapping voltages that provide low ppm mass accuracy. This is then followed by collection of spectra at more custom trapping voltages where the major peaks, which appeared in the spectrum collected at low trapping voltages, were used as "internal calibrants" to calibrate the rest of the spectrum.[19]

Hunt and co-workers have demonstrated the usefulness of combining external calibration with *automatic gain control* (*AGC*) where the number of ions in the ICR cell are controlled to fall within the external calibration range.[20] This approach allowed them to routinely achieve *MMA* of <2 ppm. This *automatic gain control* methodology has been implemented on commercially-available FT-ICR MS instruments.[20, 21] Muddiman and co-workers were able to achieve high *MMA* using this approach.[22] Both the Smith and Gygi research groups commonly report *MMA* values of ~5 ppm when utilizing *AGC* for external calibration procedures.[23-25] Recently, Muddiman and co-workers utilized a combination of *AGC* with calibration laws that used *multiple linear regression* and were able to achieve an external calibration with *mass measurement accuracy* in the parts-per-billion (ppb) range.[26]

Artificial neural networks (*ANN*s) utilize back propagation techniques to establish the weights and biases needed to fit a target output using measured parameters as input data. Unlike *multiple linear regression*, one does not need to specify a mathematical functional relationship between the input and target data – the *ANN* is trained to produce a set of outputs which minimize the difference

between the target data and the *ANN* output. The concept of an *ANN* has been applied in many different fields of chemistry including various problems in spectroscopy such as the calibration of previously existing spectra as input functions in NMR spectra,[27, 28] mass spectra,[29-32] infrared spectra,[33] and other forms.[34] They have also appeared in the study of chemical sensors applications[35, 36] and protein folding[37]. The concept of having a program that can generate solutions to an unknown via the input of a known is extremely useful.[18] Herein, data from a hybrid LTQ-FT-ICR mass spectrometer have been applied to an *artificial neural network* system in an attempt to generate *m/z* values with high *mass measurement accuracy* by utilizing parameters to account for *space-charge effects* within the ICR cell. These results are compared with that of previously published data utilizing *multiple linear regression* to provide a fit for this data.[26]

6.2 Experimental

6.2.1 Materials

Poly(propylene glycol) with an average molecular weight of 1000 Da, ammonium acetate (>99%), and formic acid were purchased from Sigma (St. Louis, MO). HPLC grade acetonitrile and high-purity water were purchased from Burdick and Jackson (Muskegon, MI). 2-Propanol (HPLC grade) was purchased from Fischer Scientific (Fair Lawn, NJ). All materials were used as received.

6.2.2 Instrument Parameters and Calibration Procedure

A modified version of an electrospray ionization (*ESI*) source developed previously in this laboratory was coupled to a hybrid ThermoFisher Scientific (San Jose, CA) LTQ-FT Ultra MS equipped with an Oxford Instruments actively shielded 7 T superconducting magnet (Concord, MA). All spectra were acquired with a resolving power at 400 *m/z* set to $100{,}000_{fwhm}$, along with *AGC* settings ranging from 5.0×10^5 to 3.0×10^6. Samples were introduced by direct infusion using a 100 μL gastight syringe (Hamilton, Las Vegas, NV) and the syringe pump on the LTQ-FT Ultra at a flow rate of 0.5 μL/min. The *ESI* emitter tips used were 360 μm o.d., 50 μm i.d. and tapered to 30 $\pm$ 1.0 μm i.d. (New Objective, Woburn, MA) and held at a constant potential of 2200 V for all experiments. Electrospray solutions were comprised of 70:30 2-propanol/water with 0.5mM ammonium acetate (NH_4OAc).

Calibration of the instrument was completed utilizing a user-defined list of eleven monoisotoptic *m/z* values for ammonium-adducted PPG-1000 oligomers, which ranged from *m/z* = 732 to *m/z* = 1312. This calibration was conducted using the manufacturer's protocol which generates coefficients for five different *AGC* target values of which the largest 3 were used: 5.0×10^5, 1.0×10^6, and 3.0×10^6.

The frequencies were obtained utilizing diagnostic mode in the instrument software, which allowed for calibration utilizing *artificial neural networks*. The *AGC* level was first set to 5.0×10^5 and five calibration spectra were recorded, along with their respective cyclotron frequencies. These frequencies provided 55 points (11 data points, 5 spectra) which were used to train the *neural network*. This was

followed by the collection of five validation spectra including their cyclotron frequencies, to determine the achievable *MMA* through the use of *neural networks*. In an effort to reduce systematic error with respect to *AGC* levels, this procedure was then repeated for *AGC* target levels of 3.0×10^6 and 1.0×10^6 in that order. Values for *total ion population* (A_T) and *relative ion population* (A_S) were calculated as previously described in Chapter 5.

5.2.3 Artificial Neural Networks in MATLAB

The basic computational unit of a *neural network* is the neuron. The network utilized in these studies consisted of six neurons arranged in three layers. The initial layer contained one neuron with a logsig output function. The middle layer contained four neurons each with logsig output function. The outer layer contained a single neuron with a linear output. Each neuron computes the value of N via the equation: $N = W_i*P_i+B_i$ and applies N to the input of its output function. W_i is a weight applied to the data by the neuron and B_i is a bias. The input data (matrix P) contains one row for each mass spectrum and one column for each experimentally measured independent variable upon which the calibration is based. Our experiments typically utilized 55 rows and either 1, 2 or 3 columns (i.e. frequency, A_T and A_S). A simplified diagram of the *neural networks* used in this work is shown in **Figure 6.1**. The maximum number of weights/biases available to our *ANN* was 9W/6B, 10W/6B or 11W/6B depending on the number of independent variables used.

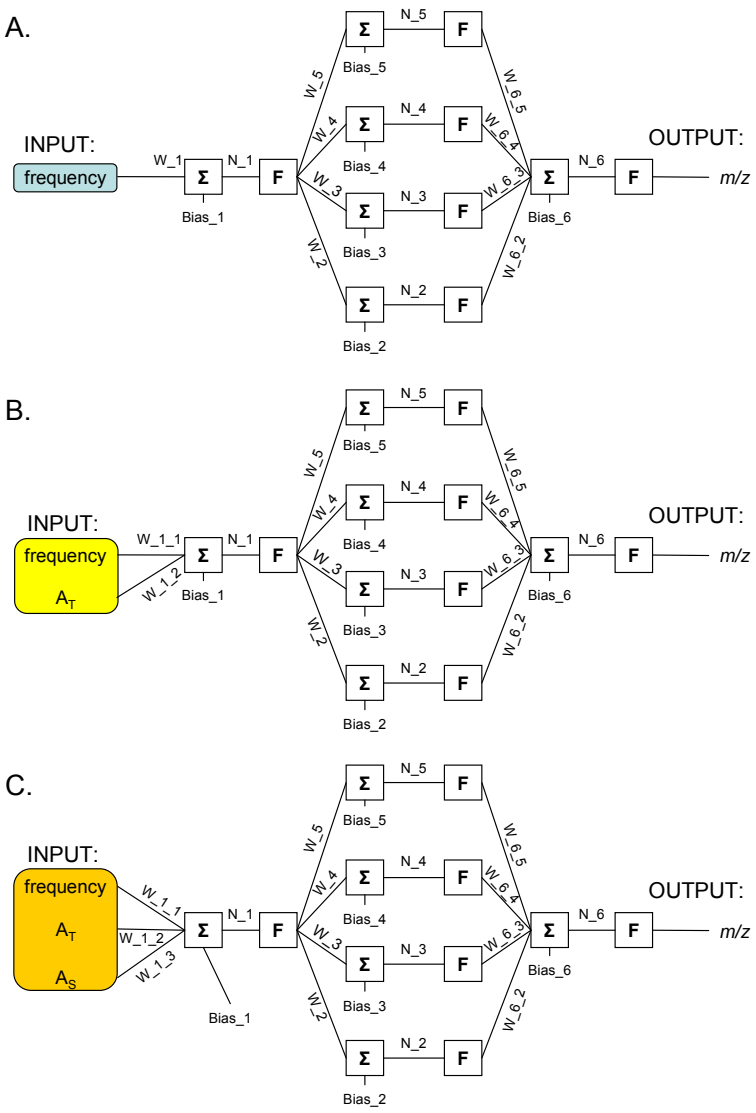


Figure 6.1 A) Neural network map utilizing frequency data with a single neuron, B) Neural Network map utilizing frequency along with total ion population (A_T), C) Neural Network map utilizing frequency, total ion population (A_T), and relative ion abundance of a given species (A_S). The first F is the logsig function and last F is the purelin function, as stated in text.

For training of the *neural network* the input layer used the calibration data set, which was obtained through a set of known compounds (PPG-1000 oligomers) that were analyzed by FT-ICR MS. This calibration data allowed the *neural network* to calculate an output which was compared to the target (*m/z* data) and through several iterations determined the weights and biases which provided outputs closest to those provided in the calibration data set. Once the desired output is obtained, the same weights and biases are applied to the validation data. This then results in *m/z* values upon which to base the *mass measurement accuracy* for the validation data.

Regardless of the number of inputs or outputs, the number of neurons can be changed which may change the quality of the fit provided by the *artificial neural network*. Usually, there are an optimal number of neurons that would allow for the best fit within a given system, and provide for acceptable computational time[38]; in this case it was determined to be 6 neurons. This allowed for a high quality, smooth fit of the line through the known points as seen in **Figure 6.2A.** Too few neurons (1 or 2) and the *ANN* did not converge well or took an extensive amount of time trying to find what would be the best solution. The counter example is taking too many neurons (≥8) which resulted in the *neural network* over fitting the line which leads to inaccurate prediction of the *m/z* values corresponding to particular frequency points, which is shown in **Figure 6.2B**. For the experiments

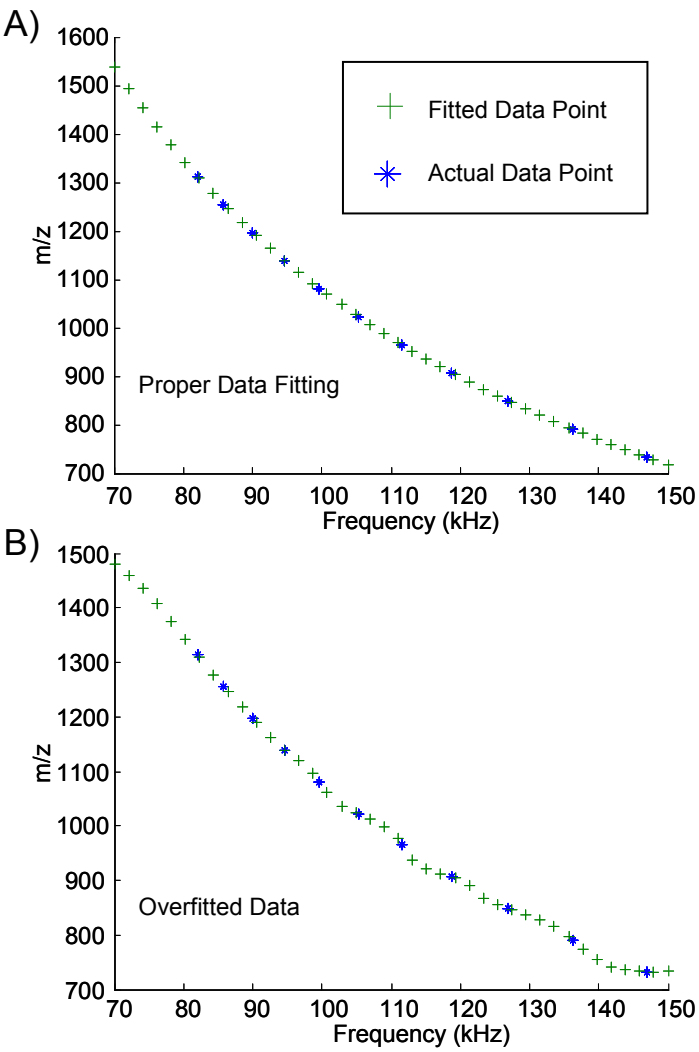


Figure 6.2 A) Graph consisting of an appropriate number of neurons to provide a proper fit, note that the line is a smooth curved line. B) Graph consisting of too many neurons which resulted in an over fitting of the data.

described in this chapter, all runs and data available here in were completed utilizing 6 neurons. Statistically the residuals were interpreted as errors and as such would be expected to follow a normal distribution. Several steps were taken to examine the residuals and evaluate whether they were normally distributed.

The data for the *Neural Networks* program was collected and applied using MATLAB 7.1.4 R14 SP3 for Windows XP or MATLAB 7.2.0.294 (R2006A) for Redhat Enterprise Linux 5. Input data was from *AGC* target values of 5.0×10^5, 1.0×10^6, and 3.0×10^6. The *trainbr neural network* process from the Matlab Neural Network Toolbox was utilized. Trainbr is a network training function that updates the weight and bias values according to the Levenberg-Marquardt optimization algorithm.[38] It minimizes a combination of squared errors and then determines the correct combination so as to produce a network that generalizes well. This process is called Bayesian regularization. The output also contained an estimate of the effective number of parameters needed for the fit.

The input and target data were normalized to fall between +1 and -1 using the Matlab routines *mapminmax* prior to initiation of the *neural network* training, but upon output these values were converted back to their proper numerical values. In addition, the training performance target was set to zero and was terminated when validation performance decreased more than two times since the last decrease in performance. Initial *neural network* weights and biases were assigned using random numbers generated by Matlab. Typically thirty training runs were made per data set. This allowed a wide range of initial weights and biases to be explored. Our criteria

for the network which provided the best fit was one which produced the lowest mean square of residuals, termed MSE, while producing residuals which passed at least 4 out of 5 normalcy tests.

Data was analyzed through variations of a script which utilized Matlab to train a *neural network* from one data set, and exploit the trained *neural network* to produce *m/z* values from another data set, referred to as the "test" data set. These programs were based on equations previously published by Muddiman[18, 26] which utilized the *relative ion abundance* of a given species (A_S) and the *total ion population* (A_T) of a FT-ICR MS data set and their effect on *mass measurement accuracy*. The first program utilized *cyclotron frequency* alone and did not incorporate *relative ion abundance* of a given species nor the *total ion population*. The second program incorporated the *cyclotron frequency* and *total ion population* of a given species (A_T), but not the *relative ion abundance* and the third program incorporated the *cyclotron frequency*, *total ion population*, and the *relative ion abundance* of a given species.

Key data that was determined via the Matlab Neural Networking program was saved in a data matrix which included values for the calibration data set including *MMA*, median, standard deviation and other values that would allow for interpretation of the *neural network* training. The Matlab program also exported values for the test data set, which are the values of particular interest since these are the points which the trained *neural network* actually predicts.

6.3 Results and Discussion

Prior to applying our *ANN* to *FT-ICR* data, the effectiveness of the *ANN* was evaluated by applying the network to the Filip dataset available from NIST.[39] The initial 82 points were divided into 54 points for training and 28 for validation. The results are summarized in **Table 6.1**. The residuals produced by both *MLR* and *ANN* were normally distributed and in excellent agreement with the NIST certified

Table 6.1 Summary of ANN and MLR Filip dataset Evaluation

Table 6.1 Comparison of MLR and ANN fits to NIST data

	MSE	Residuals	Number Parameter	Function
MLR calibration	1.14E-05	normal	11	10th order polynomial
MLR validation	1.07E-05	normal		10th order polynomial
ANN training	1.00E-05	normal	9.7	none needed
ANN validation	2.31E-05	normal		none needed

values. The MSE values produced by *ANN* were about twice as large as those obtained by *MLR*. The effective number of parameters utilized by the *ANN* for the Filip dataset was 9.7 compared to the 11 used by *MLR*. The effective number of parameters utilized by the *ANN* with regard to the FT-ICR MS data were 9.7 (frequency alone), 10.6 (frequency and A_T), and 12.3 (frequency, A_T, and A_S). The fact that the effective number of parameters increases by about 1 parameter each time another independent variable is added is consistent with the importance of A_T and A_S in the calibration of *FT-ICR* experiments. This finding is in agreement with previously published data that stated the statistical importance of A_T and A_S.[18]

Several tests were applied to assess whether the residuals were normally distributed. These were the Lilliefors test for residuals,[40] the Shapiro-Wilk parametric hypothesis test of composite normality,[41] the D'Agostino-Pearson's L2 test for assessing normality of data using skewness of kurtosis,[42] the Jarque-Bera test, [43] and the Anderson-Darling test for assessing normality of sample data.[44] The results from these tests, histograms and normal probability plots demonstrated that a fit using *artificial neural networks* provided a better fit of data than the *MLR* method previously published. This is because in all but one case the data provided by the *neural network* passed more of these normalcy tests than the previously produced *MLR* results. Also of importance, is the mean square of error values (MSE) calculated for the results of each run of the *ANN*; as this value approaches zero it indicates a better fit.

Figure 6.3 demonstrates one comparison between the previous scheme utilizing *MLR*[26] and the results achieved by the *ANN* method to perform a fit from the calibration data set. This calibration data was also utilized to generate the coefficients for the *MLR* equations and to train the *ANN*s. **Figure 6.3** contains data acquired at an *AGC* level of 5.0×10^5 and utilized only frequency. **Figure 6.3A** illustrates the residuals generated by the fitting of the calibration data by *MLR*. For this example the points on the normal probability plot deviate from the line that represents where theoretical residuals representing a normal distribution would be located. In addition, the box plot identifies several outlier points (+) that are produced by utilizing the *MLR* methodology on this particular data. In contrast,

Figure 6.3B shows that utilizing *ANN* to fit the same data produces residuals with a much more normal distribution. Many more data points of the normal probability plot fall along the line that represents a normal distribution of residuals, and there are no outliers present in the boxplot. In addition, the residuals from *ANN* fits pass at least 4 of the 5 normality tests, which is an improvement over the residuals produced by *MLR* (failed all 5).

Included in **Figure 6.3** are the average *mass measurement accuracy* (*MMA* Avg) and the mean square of error (MSE). If systematic error is eliminated *MMA* average should approach zero and the MSE should also approach zero. The *MLR* method was able to achieve 24 ppb for *MMA* average versus -0.05 ppb for the fit provided by the *ANN*, for this data set. In addition, the MSE value for the *ANN* fit is also smaller than that provided by the *MLR* fit of this data. **Figure 6.3** demonstrates only one example of the many data sets that were analyzed in this study.

In addition, data from the validation data set was utilized to test how well the *ANN* performed versus the *MLR* calibration method. The data from the validation data set also illustrated that *artificial neural network* provided a better fit for the data. However, the *MMA* average provided by the *ANN* did not achieve the same levels as that provided by *MLR*, with one exception. The *MMA* average for every *ANN* fit did remain sub-ppm for the validation data set, ranging from 602 ppb to 40ppb. The results of all data analyzed are summarized in **Table 6.2**.

All *MMA* values achieved through the use of *artificial neural network* calibration and the application of A_T and A_S parameters to account for global and local *space-charge effects* improved those provided strictly from the instrument.[26]

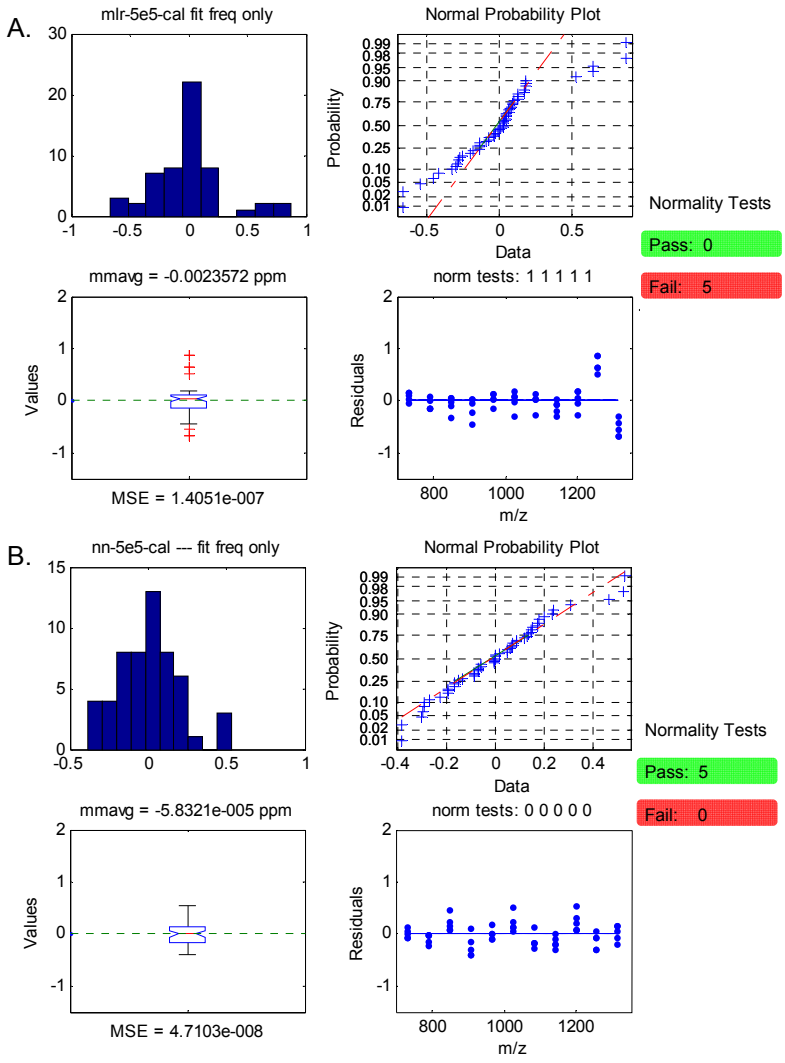


Figure 6.3 A) For an AGC target value of 5.0 × 10^5, fit using multiple linear regressions, a histogram of the residuals, normal probability plot, MMA box plot (outliers marked by +), and residual scatter plot are shown. These plots illustrate how well the MLR model utilizing frequency alone fit the data. The results of normalcy tests are also illustrated (0 indicates a pass and a value of 1 indicates fail). B) The same model using frequency alone is fit using artificial neural networks. This fit provided a more normal distribution of residuals, as demonstrated by the increased number of normality tests passed, and the decrease in outliers present in the box plot.

Table 6.2 Summary of all LTQ-FT-ICR MS Calibration for ANN and MLR. Normality test n=# normality tests failed. Fitted parameters are Frequency (F), total ion population (A_T), and relative ion abundance (A_S).

Comparison of Calibration Results

mmavg nn	mmavg mlr	MSE nn	MSE mlr	norm nn	norm mlr	Sample	Fitted Params
-5.832E-05	-2.357E-03	4.710E-08	1.405E-07	0	5	5e5	F
2.881E-04	-2.611E-03	3.268E-08	1.257E-07	0	5	5e5	F, A_T
-1.706E-05	3.879E-04	1.900E-08	1.244E-07	0	5	5e5	F, A_T, A_S
1.102E-02	-8.009E-04	3.466E-07	3.936E-07	1	3	1e6	F
-1.808E-02	-1.079E-03	1.935E-07	1.618E-07	0	5	1e6	F, A_T
-6.621E-03	-1.509E-03	1.133E-07	1.617E-07	0	5	1e6	F, A_T, A_S
1.231E-03	4.130E-02	4.774E-07	6.001E-07	1	0	3e6	F
-4.222E-04	4.166E-02	4.540E-08	3.312E-07	0	5	3e6	F, A_T
1.048E-06	3.296E-03	2.710E-08	1.226E-07	0	0	3e6	F, A_T, A_S

Comparison of Validation Results

mmavg nn	mmavg mlr	MSE nn	MSE mlr	norm nn	norm mlr	Sample	Fitted Params
-4.355E-02	3.407E-02	5.650E-08	1.614E-07	1	5	5e5	F
-1.356E-01	3.110E-02	5.288E-08	1.412E-07	0	5	5e5	F, A_T
-1.465E-01	3.380E-02	4.321E-08	1.408E-07	0	5	5e5	F, A_T, A_S
3.960E-02	2.016E-01	2.311E-07	3.127E-07	0	3	1e6	F
1.350E-01	2.335E-02	4.661E-07	3.577E-07	0	5	1e6	F, A_T
1.453E-01	2.263E-02	3.775E-07	3.564E-07	1	5	1e6	F, A_T, A_S
6.016E-01	4.174E-01	1.171E-06	9.818E-07	0	3	3e6	F
9.893E-02	3.378E-02	8.295E-08	3.617E-07	0	3	3e6	F, A_T
5.721E-02	5.355E-04	6.785E-08	1.282E-07	0	0	3e6	F, A_T, A_S

Notes --
1. MSE of neural network calibration & validation better than MSE mlr for all but 4 cases
2. mmavg of neural network calibration smaller than mmavg of mlr with 3 exceptions; underlined
3. mmavg of mlr validation smaller than mmavg of nn with 1 exception; double underlined

6.4 Conclusions

The methods reported herein continue to reinforce the importance of accounting for global and local *space-charge effects* in FT-ICR MS to achieve the best possible *MMA* values. It was demonstrated that mean *MMA* values in the ppb range can be achieved across a range of *AGC* target levels utilizing *artificial neural networks* to calibrate data from a hybrid LTQ-FT-ICR MS. Taking into account the *automatic gain control* settings, it is important to note that the calibration provides the most improvement when used at the highest possible population of ions, as expected. Also of importance is the fact that calibration with *ANN* did provide a better overall fit for these data (as shown by lower MSE values for the *ANN* fit data) , even though the average *MMA* was not as good as that provided for by the *MLR* methodology.[26] When these methods described herein are executed properly, we are able to obtain average *MMA*s between 600 and 40 ppb. Future experiments could explore the application of this calibration method to a variety of other analytes and other FT-ICR MS systems.

6.5 References

1. Comisarow, M.B. and A.G. Marshall, Fourier-Transform Ion-Cyclotron Resonance Spectroscopy*. Chem. Phys. Lett.*, 1974. **25**. 282-283.

2. Comisarow, M.B. and A.G. Marshall, Frequency-Sweep Fourier-Transform Ion-Cyclotron Resonance Spectroscopy*. Chem. Phys. Lett.*, 1974. **26**. 489-490.

3. Gorshkov, M.V., S.H. Guan, and A.G. Marshall, Masses of Stable Neon Isotopes Determined at Parts-Per-Billion Precision by Fourier-Transform Ion-Cyclotron Resonance Mass-Spectrometry*. Int. J. Mass Spectrom. Ion Processes*, 1993. **128**. 47-60.

4. Rodgers, R.P., et al., Resolution, elemental composition, and simultaneous monitoring by Fourier transform ion cyclotron resonance mass spectrometry of organosulfur species before and after diesel fuel processing*. Anal. Chem.*, 1998. **70**. 4743-4750.

5. He, F., C.L. Hendrickson, and A.G. Marshall, Baseline mass resolution of peptide isobars: A record for molecular mass resolution*. Anal. Chem.*, 2001. **73**. 647-650.

6. Jeffries, J.B., S.E. Barlow, and G.H. Dunn, Theory of Space-Charge Shift of Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 169-187.

7. Francl, T.J., et al., Experimental-Determination of the Effects of Space-Charge on Ion-Cyclotron Resonance Frequencies*. Int. J. Mass Spectrom. Ion Processes*, 1983. **54**. 189-199.

8. Kaiser, N.K., G.A. Anderson, and J.E. Bruce, Improved mass accuracy for tandem mass spectrometry*. J. Am. Soc. Mass Spectrom.*, 2005. **16**. 463-470.

9. Hannis, J.C. and D.C. Muddiman, A dual electrospray ionization source combined with hexapole accumulation to achieve high mass accuracy of biopolymers in fourier transform ion cyclotron resonance mass spectrometry*. J. Am. Soc. Mass Spectrom.*, 2000. **11**. 876-883.

10. Flora, J.W., J.C. Hannis, and D.C. Muddiman, High-mass accuracy of product ions produced by SORI-CID using a dual electrospray ionization source coupled with FTICR mass spectrometry*. Anal. Chem.*, 2001. **73**. 1247-1251.

11. Nepomuceno, A.I., et al., Detection of genetic variants of transthyretin by liquid chromatography-dual electrospray ionization Fourier-transform ion-cyclotron-resonance mass spectrometry. *Clin. Chem.*, 2004. **50**. 1535-1543.

12. Bruce, J.E., et al., Obtaining more accurate Fourier transform ion cyclotron resonance mass measurements without internal standards using multiply charged ions. *J. Am. Soc. Mass Spectrom.*, 2000. **11**. 416-421.

13. Williams, K., A.M. Hawkridge, and D.C. Muddiman, Sub parts-per-million mass measurement accuracy of intact proteins and product ions achieved using a dual electrospray ionization quadrupole Fourier transform ion cyclotron resonance mass spectrometer. *J. Am. Soc. Mass Spectrom.*, 2007. **18**. 1-7.

14. Zhang, L.K., et al., Accurate mass measurements by Fourier transform mass spectrometry. *Mass Spectrom. Rev.*, 2005. **24**. 286-309.

15. Easterling, M.L., T.H. Mize, and I.J. Amster, Routine part-per-million mass accuracy for high-mass ions: Space-charge effects in MALDI FT-ICR. *Anal. Chem.*, 1999. **71**. 624-632.

16. Taylor, P.K. and I.J. Amster, Space charge effects on mass accuracy for multiply charged ions in ESI-FTICR. *Int. J. Mass Spectrom.*, 2003. **222**. 351-361.

17. Masselon, C., et al., Mass measurement errors caused by 'local" frequency perturbations in FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2002. **13**. 99-106.

18. Muddiman, D.C. and A.L. Oberg, Statistical evaluation of internal and external mass calibration laws utilized in Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.*, 2005. **77**. 2406-2414.

19. Wong, R.L. and I.J. Amster, Sub part-per-million mass accuracy by using stepwise-external calibration in Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2006. **17**. 1681-1691.

20. Syka, J.E.P., et al., Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J. Proteome Res.*, 2004. **3**. 621-626.

21. Peterman, S.M. and J.J. Mulholland, A novel approach for identification and characterization of glycoproteins using a hybrid linear ion trap/FT-ICR mass spectrometer. *J. Am. Soc. Mass Spectrom.*, 2006. **17**. 168-179.

22. Johnson, K.L., et al., Accurate mass precursor ion data and tandem mass spectrometry identify a class I human leukocyte antigen A*0201-presented peptide originating from vaccinia virus. *J. Am. Soc. Mass Spectrom.*, 2005. **16**. 1812-1817.

23. Dieguez-Acuna, F.J., et al., Characterization of mouse spleen cells by subtractive proteomics. *Mol. Cell. Proteomics*, 2005. **4**. 1459-1470.

24. Belov, M.E., et al., Automated gain control and internal calibration with external ion accumulation capillary liquid chromatography-electrospray ionization-fourier transform ion cyclotron resonance. *Anal. Chem.*, 2003. **75**. 4195-4205.

25. Page, J.S., et al., Automatic gain control in mass spectrometry using a jet disrupter electrode in an electrodynamic ion funnel. *J. Am. Soc. Mass Spectrom.*, 2005. **16**. 244-253.

26. Williams, D.K. and D.C. Muddiman, Parts-per-billion mass measurement accuracy achieved through the combination of multiple linear regression and automatic gain control in a Fourier transform ion cyclotron resonance mass spectrometer. *Anal. Chem.*, 2007. **79**. 5058-5063.

27. Radomski, J.P., H. Vanhalbeek, and B. Meyer, Neural-Network-Based Recognition of Oligosaccharide H-1-Nmr Specter. *Nature Structural Biology*, 1994. **1**. 217-218.

28. Hare, B.J. and J.H. Prestegard, Application of Neural Networks to Automated Assignment of Nmr-Spectra of Proteins. *Journal of Biomolecular Nmr*, 1994. **4**. 35-46.

29. Gasteiger, J., et al., Neural Nets for Mass and Vibrational-Spectra. *Journal of Molecular Structure*, 1993. **292**. 141-159.

30. Werther, W., et al., Classification of Mass-Spectra - a Comparison of Yes/No Classification Methods for the Recognition of Simple Structural-Properties. *Chemometrics and Intelligent Laboratory Systems*, 1994. **22**. 63-76.

31. Goodacre, R., et al., Rapid and Quantitative-Analysis of Recombinant Protein Expression Using Pyrolysis Mass-Spectrometry and Artificial Neural Networks - Application to Mammalian Cytochrome B(5) in Escherichia-Coli. *Journal of Biotechnology*, 1994. **34**. 185-193.

32. Goodacre, R., M.J. Neal, and D.B. Kell, Rapid and Quantitative-Analysis of the Pyrolysis Mass-Spectra of Complex Binary and Tertiary Mixtures Using

Multivariate Calibration and Artificial Neural Networks. *Anal. Chem.*, 1994. **66**. 1070-1085.

33. Alam, M.K., S.L. Stanton, and G.A. Hebner, Near-Infrared Spectroscopy and Neural Networks for Resin Identification. *Spectroscopy*, 1994. **9**. 30-&.

34. Svozil, D., V. Kvasnicka, and J. Pospichal, Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, 1997. **39**. 43-62.

35. Kavuri, S.N. and V. Venkatasubramanian, Using Fuzzy Clustering with Ellipsoidal Units in Neural Networks for Robust Fault Classification. *Computers & Chemical Engineering*, 1993. **17**. 765-784.

36. Hoskins, J.C. and D.M. Himmelblau, Artificial Neural Network Models of Knowledge Representation in Chemical-Engineering. *Computers & Chemical Engineering*, 1988. **12**. 881-890.

37. Qian, N. and T.J. Sejnowski, Predicting the Secondary Structure of Globular-Proteins Using Neural Network Models. *Journal of Molecular Biology*, 1988. **202**. 865-884.

38. Hagan, M.T. and M.B. Menhaj, Training Feedforward Networks with the Marquardt Algorithm. *Ieee Transactions on Neural Networks*, 1994. **5**. 989-993.

39. Filippelli, A., Statistical Reference Datasets, Online DG. *NIST*, http://www.itl.nist.gov/div898/strd/lls/data/Filip.shtml.

40. Lilliefo.Hw, On Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association*, 1967. **62**. 399-&.

41. Shapiro, S.S. and M.B. Wilk, An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 1965. **52**. 591-&.

42. D'Agostino, R.B. and E.S. Pearson, Testing for Departures From Normality. 1. Fuller Empirical Results for the Distribution of b, and √b, . *Biometrika*, 1973. **60**. 613-622.

43. Jarque, C.M. and A.K. Bera, Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. *Economics Letters*, 1980. **6**. 255-259.

44. Anderson, T.W. and D.A. Darling, Asymptotic Theory of Certain Goodness of Fit Criteria Based on Stochastic Processes*. Annals of Mathematical Statistics*, 1952. **23**. 193-212.

CHAPTER 7

Absolute Quantification of C-Reactive Protein in Human Plasma Utilizing Protein Cleavage Isotope Dilution Mass Spectrometry

7.1 Introduction

C-reactive protein (CRP) is an important clinical marker for inflammation and atherosclerosis, a chronic inflammatory response in the walls of the arteries. In addition, CRP has been used as both a diagnostic and prognostic marker for cardiovascular disease, where it is used to identify acute cardiac trauma, or in conjunction with other markers to determine mortality.[1-6] CRP typically circulates in the plasma at concentrations of <10 µg/mL to >5 mg/mL as defined by *ELISA*.[1] Considering that human serum albumin (HSA) alone has a reference range of 3.5-5.5 g/dL,[7] this means that CRP, at its highest level is one order of magnitude less, and at its lowest level is present at a concentration 5000 times lower than HSA! There are presently several immunoassays utilized to detect CRP.[8] In addition to these immunoassays, CRP has recently been analyzed by mass spectrometry.[9, 10] These methods both employ immuno-affinity to capture CRP from human serum. Kuhn and co-workers used size exclusion chromatography with immuno-affinity chromatography to further fractionate serum and quantify CRP.[9] This study examined the CRP concentration of 12 pooled plasma samples. This provided the motivation to develop a method utilizing mass spectrometry for quantification that

would enable the simplification of the plasma preparation before analysis and the use of this developed method to analyze a much larger sample set.

The *absolute quantification* of a protein from its digestion products dates back to Barr et. al.[11] They used the combination of proteolysis and MS for *absolute quantification* of a European Community Bureau of Reference (BCR) certified apolipoprotein A-1 standard that was proteolyzed with trypsin and quantified with a stable isotope labeled internal standard peptide with LC- flow-FAB MS/MS.[11] Their results demonstrated that the use of protein cleavage coupled with isotope dilution mass spectrometry (*PC-IDMS*) was valid methodology for the standardization of measurements of particular proteins in a clinical environment. *PC-IDMS* is ideally completed when protein cleavage is complete, as this produces a 1:1 molar ratio between the initial intact protein and the peptide or peptides to be analyzed. The quantification using IDMS is based on the ratio of the response of the labeled internal standard peptide to that of the unlabeled peptide resulting from the digestion of the particular protein of interest. IDMS has been utilized on a range of different analytes for almost 40 years and still maintains its utility for the quantification of many different molecules.[12]

After this initial study[11] several applications utilizing *PC-IDMS* have been published for the quantification of peptides and proteins and are present in the current literature.[13-18] The ability to quantify proteins with post-translation modifications (PTM's) has been demonstrated by having PTM's incorporated onto

synthetic internal standard peptides.[19, 20] Other reports have completed *PC-IDMS* as a means to compare immunoassay analysis to that of mass spectrometry.[9, 10, 21, 22]

The driving force for the development of this simplified quantification method for a *biomarker*, is the fact that mass spectrometry is constantly discovering and investigating the possibility of new and exciting *biomarkers*, many of which do not have *ELISA* methods for their quantification. Thus, if we can demonstrate and provide verification that a relatively simple, molecularly specific method (i.e. utilizing mass spectrometry) for the quantification of protein *biomarkers* is robust, it can be applied in the future to newly discovered *biomarkers* that do not have existing methods for their quantification.

In this chapter, a method utilizing *PC-IDMS* is described for the *absolute quantification* of C-reactive protein in human plasma without the use of immuno-affinity chromatography or other separation techniques to simplify the plasma before introduction to the LC-MS/MS system for analysis. This method was developed utilizing nano-flow liquid chromatography and a triple quadrupole mass spectrometer operating in *selective reaction monitoring* (*SRM*) mode. The results achieved through the utilization of *PC-IDMS* were compared to results from a CLIA approved independent laboratory measurement using a wide range hsCRP immunoassay test. This method was developed such to continue the investigation of CRP levels in ovarian cancer patients, since the Muddiman group has previously reported that CRP is up-regulated in patients with ovarian cancer,[23] and to also probe possible affects smoking may have on CRP levels in plasma.

7.2 Experimental

7.2.1 Synthetic Peptides

Peptides were synthesized representing four tryptic peptides, from C-Reactive protein, whose sequences are as follows: APLTKPLK, ESDTSYVSLK, GYSIFSYATK, and QDNEILIFWSK. An additional set of peptides with identical sequences, but incorporation of ^{13}C and ^{15}N were synthesized for use as internal standards: APLTKP[$^{13}C_5$, $^{15}N_1$]LK, ESDTSYV[$^{13}C_5$, $^{15}N_1$]SLK, GYSIF[$^{13}C_9$, $^{15}N_1$]SYATK, and QDNEILF[$^{13}C_9$, $^{15}N_1$]WSK. All synthetic peptides were purchased from Mayo Clinic Proteomics Center (Rochester, MN). After dilution of these peptides their concentrations were confirmed utilizing UV-Vis spectroscopy and the *Scopes method*.[24]

7.2.2 Plasma Processing

Plasma digestion protocol followed a previously published methodology by Bondar, et. al.[25] Briefly, 125 µL aliquots of plasma were added to dry urea to create 6M urea in plasma to denature proteins; this was followed by addition of an internal standard peptide, reduction with dithiothreitol (DTT), and alkylation with iodoacetamide. Finally, proteins present in plasma were digested with TPCK treated trypsin. These samples were then diluted and used directly for interrogation with nano-HPLC *SRM*. A total of 110 human plasma samples were analyzed: 1 pooled sample purchased from Innovative Research, Inc. (Novi, Michigan), 1 pooled sample from the Red Cross, and 108 samples procured from the Mayo Clinic (Rochester, MN). The 108 plasma samples were collected at the Mayo Clinic under IRB number

026-06-1. These samples were then transferred to North Carolina State University and analyzed under IRB number 000-00-330. All reagents utilized in the digestion procedure were purchased from Sigma-Aldrich (St. Louis, MO).

The plasma was processed in eleven “experimental cycles” each containing eleven plasma samples; however, in each experimental cycle one of these samples was always a pooled sample purchased from Innovative Research, Inc., such that the analytical variation of the method could be examined. This allowed for the analysis of 110 different plasma samples over the eleven experimental cycles. It was necessary to break the processing of these samples up, such to minimize their time between chemical processing (i.e. reduction, alkylation, and digestion) and analysis.

7.2.3 LC-MS/MS

Solvents for liquid chromatography were purchased from Burdick and Jackson (Muskegon, MI). Reversed phase liquid chromatography was performed using a 75 μm i.d. PicoFrit capillary column (New Objective, Woburn, MA) with a 15 μm emitter tip packed in-house with 4 μm Jupiter Proteo C12 90Å stationary phase (Phenomenex, Torrance, CA). The packed volume had dimensions 75 μm i.d. × 100 mm and was operated at room temperature. Processed plasma injections of 10 μL were loaded using a PAL Autosampler (LEAP Technologies, Carrboro, NC) and over the course of 2 minutes trapped and washed on a custom built Jupiter Proteo C12 OPTI-PAK trap cartridge (Optimize Technologies; Oregon City, OR) with 100% Mobile Phase A (95/5 water/acetonitrile) at 10 uL/min. Then a 10 port switching

valve (VICI, Houston, TX) was triggered to move the sample in-line with the gradient. Elution was carried out by a Chorus 220 nano-flow pump (CS Analytics, Zwingen, Switzerland) at 500 nL/min with mobile phases containing 95/5 (v/v) (Mobile Phase A) and 5/95 (Mobile Phase B) water and acetonitrile, respectively. The ion pairing reagent used was 0.2% formic acid (Sigma Aldrich, St. Louis, MO) in both mobile phases. The LC gradient was held at initial conditions of 2% B for 6 minutes followed by a ramp to 95% B over 14 minutes and held for an additional three minutes before re-equilibrating at 2% B, for a total gradient time of 30 minutes.

All quantification was carried out on a triple quadrupole mass spectrometer (Thermo Scientific, San Jose, CA) that was operated in *SRM* mode to monitor transitions of the peptides of interest and their stable isotope labeled counterparts. Each transition was monitored for 50 milliseconds and a resolution of 0.7 was employed for both Q1 and Q3 quadrupoles. Calibration was carried out utilizing 10 fmoles of the stable isotope labeled peptide internal standard on column. The natural peptide had 0.5, 1, 2, 5, 10, 15, 20, and 50 fmoles on column. Each ratio of natural to stable isotope labeled peptide was analyzed in duplicate to ensure the accuracy and reproducibility of the measurements.

7.3 Results and Discussion

Four peptides were chosen to be synthesized and utilized as possible internal standards based on a tryptic digestion of purified CRP analyzed on a hybrid LTQ-FT-ICR. The synthesized peptides were then dissolved and examined for purity by

direct infusion. This initial investigation demonstrated the most suitable peptide to use as an internal standard would be ESDTSYVSLK, the naturally occurring form will herein be referred to as $tpCRP_{14-23}$ (tryptic peptide from CRP, amino acids 14-23). The other peptides, initially of interest as possible internal standards, presented challenges that prevented their further utilization in this study, such as solubility issues and calibration curves that did not perform as well as those generated from $tpCRP_{14-23}$. The concentrations of both the $tpCRP_{14-23}$ and the stable isotopically labeled form used as an internal standard were verified by the *Scopes method*[24] which allowed for quantification of the stock solutions. Dilutions of these stock solutions allowed for the generation of a calibration curve and were eventually added as the internal standard to plasma samples analyzed.

Quantitation of CRP using LC-MRM from human serum[9] has been previously demonstrated; however, the method described herein does not require the depletion of abundant proteins from plasma nor does it use size exclusion chromatography to further purify and simplify the sample. The reduction in sample processing minimizes the loss of the analyte of interest, and furthermore minimizes the variance in analytical results caused by the implementation of immuno-affinity and size exclusion separations. In addition, in the previously published method,[9] the internal standard was not added until after fractionation of the plasma sample, which makes potential losses of CRP hard to quantify. However, the method developed in these experiments incorporates the internal standard directly after the addition of plasma to urea, such that any losses in processing can be tracked.

The method presented herein utilized transitions of 564.78 → 347.23, 696.39 for the naturally occurring peptide ($tpCRP_{14-23}$) and 567.78 → 347.23, 702.41 for the internal standard (IS) peptide. Both sets of transitions minimized the interferences from other possible tryptic peptides that occur from tryptic digestion of human plasma. This was determined through the implementation of *SRM* Workflow software, provided by ThermoScientific, and the IPI human protein database, in addition to experimental results. The software searches the database for other tryptic fragments that could produce similar precursor and product *m/z* that would affect the *SRM* transitions monitored by the instrument. **Figure 7.1** demonstrates the ability of the instrument's *SRM* mode to monitor the transitions specified and

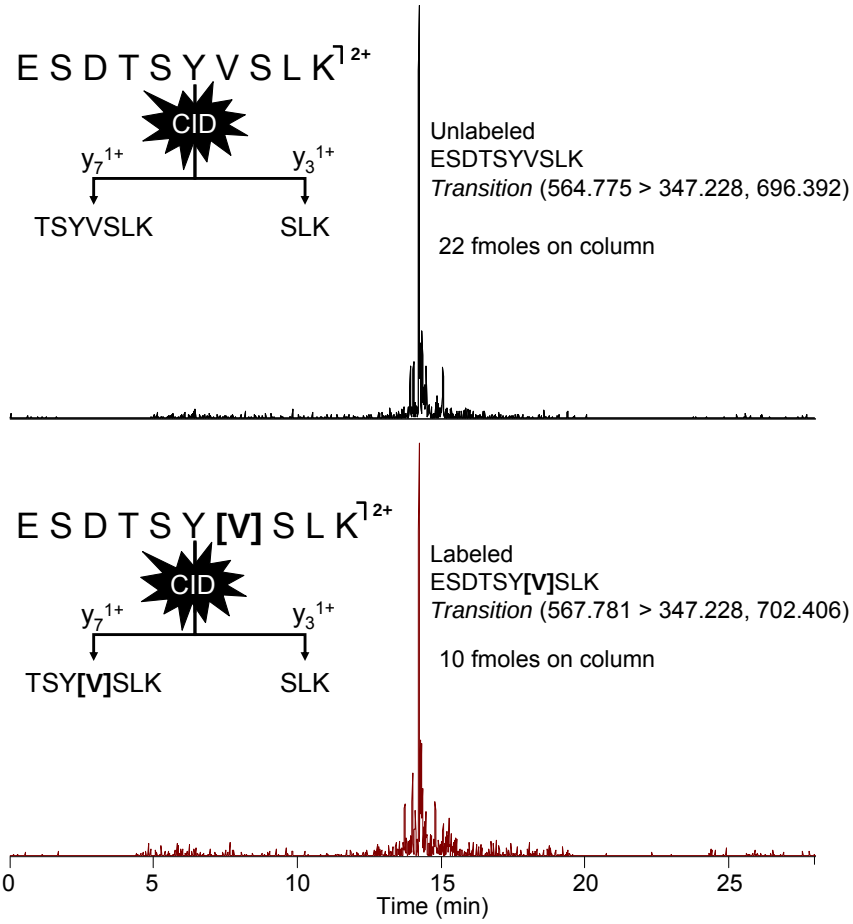


Figure 7.1 Extracted ion chromatograms from the LC-MS/MS analysis of a plasma digest are shown. The chromatograms illustrate the ability of SRM to observe only the analytes of interest, tpCRP and IS peptide, from an untreated human plasma tryptic digestion.

detect only the analytes of interest directly from a tryptic digestion of untreated human plasma.

The linear regression shown in **Figure 7.2** was derived from the LC-SRM data of 8 different concentrations of synthetic peptide with 10 fmoles of IS peptide on column. Each concentration was analyzed in duplicate, for a total of 16 points in the calibration. The x and y used for the linear regression represent the ratio of the concentration of $tpCRP_{14-23}$ to the IS peptide, and the ratio of their measured response from the mass spectrometer, respectively. Shown in **Figure 7.2** is the equation of the line describing the results from the linear regression, as well as the R^2-value. For an ideal system the line generated by linear regression should have both a slope of the 1.00 and intersect the origin. The slope generated from these experiments was 1.062 ± 0.023, with a y-intercept of 0.0497 ± 0.0408. The LOD for the $tpCRP_{14-23}$ was calculated to be 0.038 nmole/L, as determined by a previously utilized methodology, [14] in which y was set equal to the upper 95% confidence interval value of the y-intercept and a

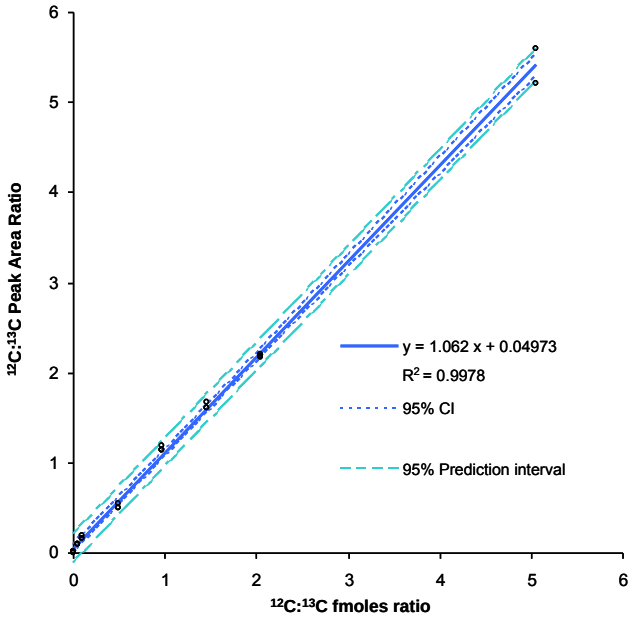


Figure 7.2 Calibration curve using the peptide ESDTSYVSLK generated by linear least squares regression analysis of SRM data. The graph shows the ratio of tpCRP peptide response/IS peptide response vs. ratio of their know concentrations.

concentration value for x was calculated using the equation generated by the previously completed least squares regression.

A total of 110 different plasma samples were analyzed in this study. These samples were distributed through eleven experimental cycles, in which each sample was analyzed in duplicate. An aliquot of the Innovative Research pooled plasma sample was processed for each experimental cycle to serve as a control through each cycle of sample processing. For cycles 1-10, ten single patient samples provided by the Mayo Clinic were analyzed (5 samples from patients with EOC and 5 age matched controls) in addition to the Innovative Research control. For cycle 11, eight single patient samples, a pooled sample from the Red Cross, and the Innovative Research control were evaluated. **Figure 7.3** shows the distribution of samples analyzed in each experimental cycle, excluding the age matched controls.

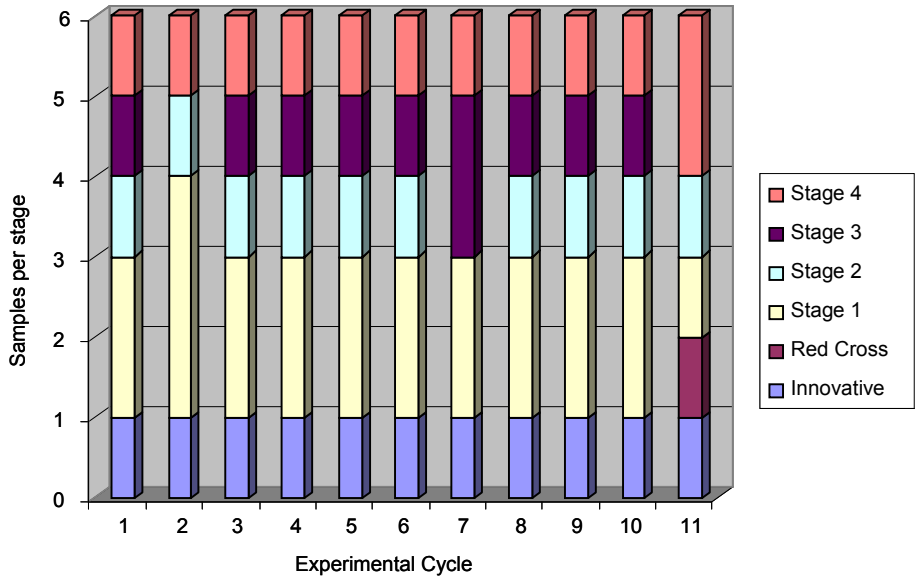


Figure 7.3 Samples analyzed in each experimental cycle, excluding age match controls for EOC samples: 11 samples for cycles 1-10 and 10 samples for cycle 11.

The retention time reproducibility for the first three cycles is shown in **Figure 7.4A**. The retention times of these analyses produced a mean of 13.7 minutes ± 6 seconds. Retention time reproducibility was similar through the remaining cycles. Column changes were required during analysis of the remaining cycles, which influenced the absolute retention time between cycles with different columns; however, the range of retention times remained similar on runs completed on the same column. This slight deviation in retention time between columns could be due to minor variations in the length of the self-packed columns utilized in this method, but were adequately accounted for because the IS peptide and $tpCRP_{14-23}$ coeluted. This reproducibility in conjunction with the use of the IS peptide and the specificity afforded by the utilization of *SRM* allows for the quantification of the IS peptide and the $tpCRP_{14-23}$ in the presence of other components from the plasma samples.

The reproducibility of this method is also demonstrated in **Figure 7.4B**, which shows the CRP concentration of the control plasma sample analyzed in each experimental cycle. This box plot contains 22 points, which result from the analysis of the control sample for each of the 11 experimental cycles, in duplicate. This was necessary to evaluate the analytical variability of the developed method. The values range from about 0.4 mg/dL to 1.5 mg/dL; however, upon further investigation the two lowest points resulted from problems with the analytical column clogging, and the highest point directly followed the highest CRP concentration detected. With these three points removed the range of the data narrows significantly, resulting in a range from ~0.6-1.2 mg/dL. It is also important to note that although the same

plasma sample was utilized, a new aliquot was processed for each experimental cycle (11 times); thus, there are several likely causes for this range of values (e.g. reduction, alkylation, digestion, etc.). However, correlation values between all CRP values attained by this *PC-IDMS* method and the orthogonal *ELISA* method were high (*vide infra*) which demonstrates that the method developed herein is robust.

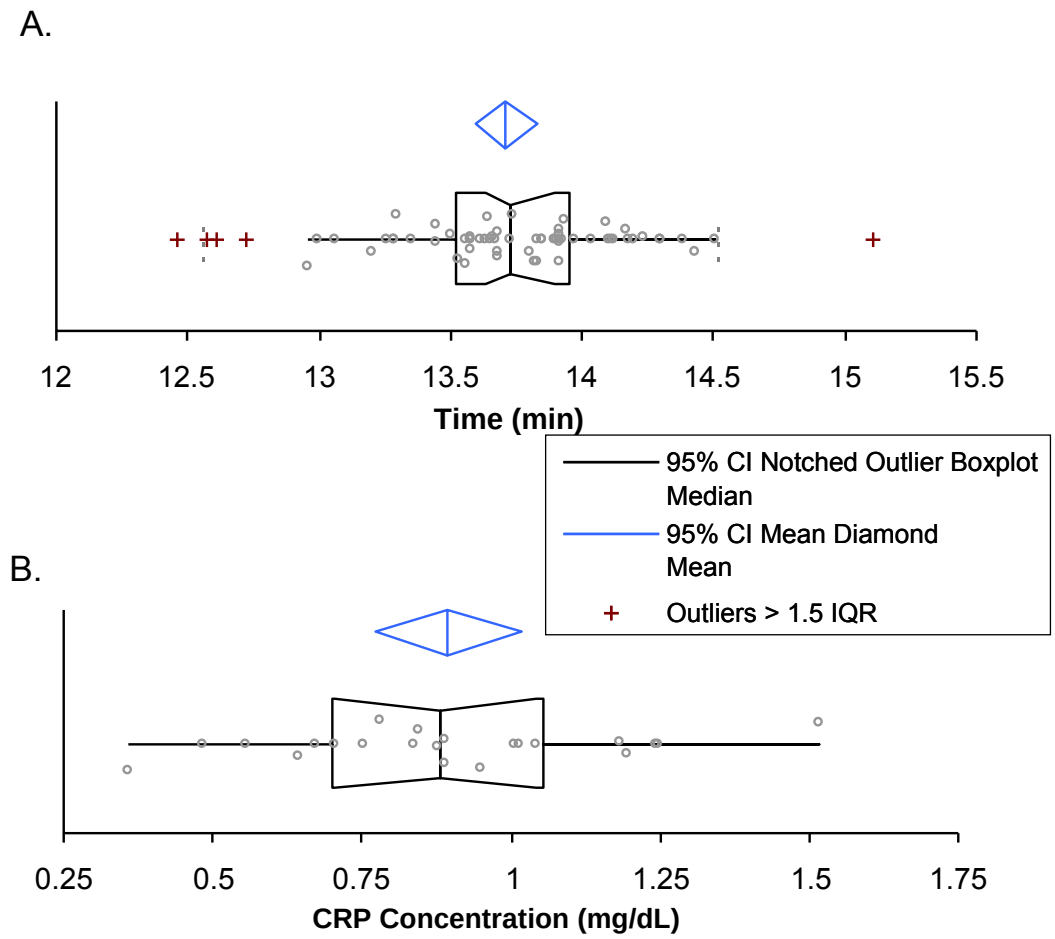


Figure 7.4 A) A box and whisker plot showing retention time reproducibility from Cycles 1-3. B) Concentration of control plasma sample throughout the 11 cycles is shown. The graph illustrated the mean (center line of diamond), 95% C.I. of the mean (blue diamond), as well data points classified as outliers (>1.5 times inner quartile range).

All samples analyzed by the *PC-IDMS* method were sent to an independent lab (CLIA certified) to have an *ELISA* analysis performed for validation. The analysis performed utilized a polyclonal antibody and was designated as a wide range/high sensitivity CPR test. This was able to detect low amounts of CRP (0.01 mg/dL), but had a higher upper limit than a typical hsCRP test (>0.30 mg/dL). The results from the *PC-IDMS* method developed herein are plotted against the *ELISA* results from the independent lab in **Figure 7.5**. When linear regression was performed on this scatter plot, an R^2 = 0.9708 was generated. This demonstrates a strong correlation between CRP concentrations determined utilizing the *PC-IDMS* method and the *ELISA* wide range hsCRP test. Discrepancies between the values of these two methods are immediately noticeable when **Figure 7.5** is examined, with the *PC-IDMS* values being larger. Previously published studies of assay development and methodology have shown results can differ by as much as two-orders of magnitude for different detections schemes and the same protein target.[26-29] This disparity in the values reported from different

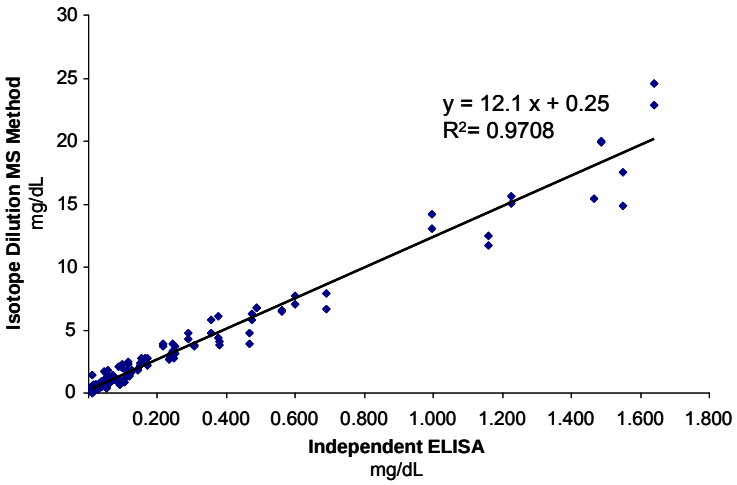


Figure 7.5 Plot of the concentration of CRP in plasma determined by PC-IDMS versus an ELISA carried out by an independent lab. A correlation coefficient of 0.9708 was achieved when performing least squares linear regression on the plot which demonstrates the successful quantification of CRP by the PC-IDMS method.

detection methods means that a reference range must be determined for each method, such that the values attained can provide appropriate information.

The reference range for the developed *PC-IDMS* method described in this chapter is different than that of *ELISA* method. The range detected utilizing the wide range hsCRP *ELISA* methodology was from <0.01 mg/dL to 1.64 mg/dL. The *PC-IDMS* method provided a range from 0.13 mg/dL to 22.8 mg/dL, of which the upper limit is much higher than the highest level in the calibration curve (10 mg/dL). On average the majority of the measurements were 1 order of magnitude greater utilizing the *PC-IDMS* method (within the calibration curve), which in effect would shift the reference range by an equal amount if this method were to be utilized in the clinical setting.

CRP has been previously found to be up-regulated in patients with epithelial ovarian cancer (EOC).[23, 30] Our individual patient samples had either a

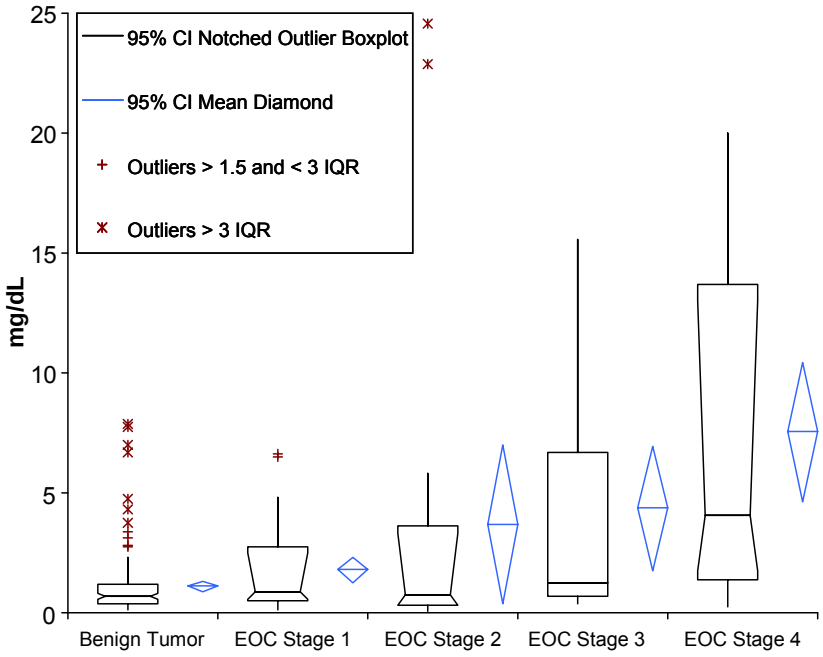


Figure 7.6 Concentration of all single patient plasma samples are shown. The graph illustrates the mean (center line of diamond), 95% C.I. of the mean (blue diamond), as well data points classified as outliers (>1.5 times inner quartile range and >3 times IQR) for patients with benign tumors followed by stage I, stage II, stage III, and stage IV EOC.

benign gynecological tumor, stage I, stage II, stage III, or stage IV EOC. Out of these groups, those with benign gynecological tumors were found to have the lowest levels of CRP. Also, as the EOC stage classification increased from I to IV so did the CRP levels found in these patients. **Figure 7.6** is a box plot containing the CRP levels found through the *PC-IDMS* developed herein. The mean levels ± 95% confidence interval found through this analysis were as follows: 1.10 ± 0.24 mg/dL for benign, 1.78 ± 0.53 mg/dL for stage I, 3.68 ± 3.30 mg/dL for stage II, 4.35 ± 2.60 mg/dL for stage III, and 7.54 ± 2.90 mg/dL for stage IV. While there was a wide range of CRP concentrations found, as evidence by the 95 % confidence interval, the overall mean does increase with the stage of the cancer present.

The *ROC* (*receiver operator characteristic*) *curve* in **Figure 7.7** illustrates that CRP is not a good indicator for early stage ovarian cancer. The p-value for significance was determined to be 0.003125. The lack of sufficient *specificity* and *sensitivity* lead to CRP's

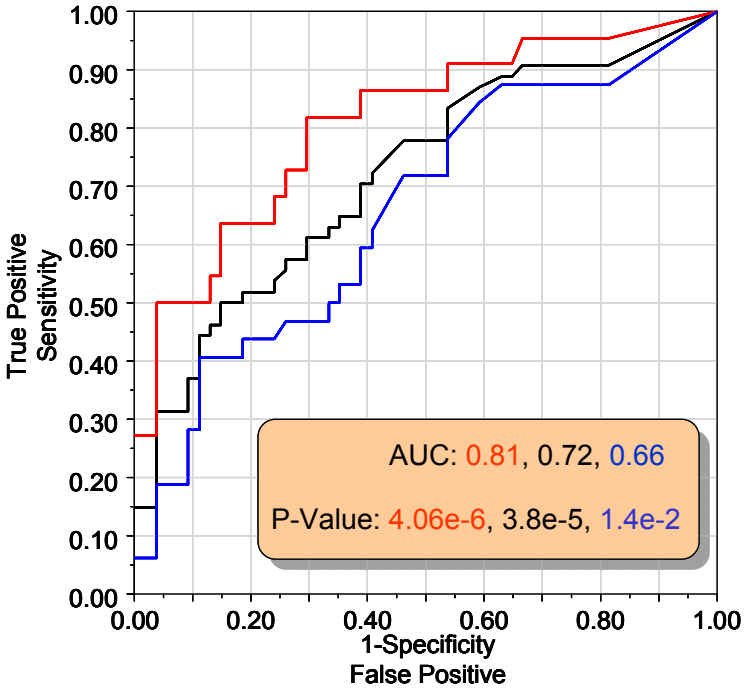


Figure 7.7 An ROC Curve containing data from CRP levels and EOC. Curves representing all samples (black), only early (blue), and only late (red) stage samples are shown. Area under the curve and the p-values are also displayed.

inability to be an acceptable marker for EOC. However, the findings from these experiments do suggest the possibility of utilizing CRP in conjunction with other markers for the early detection of ovarian cancer.

7.4 Conclusions

The method and results presented in this chapter continue to support the fact that *PC-IDMS* is a viable method for the *absolute quantification* of proteins. In addition, this chapter illustrates that it is possible to directly quantify a *biomarker* from plasma without having to use additional steps to simplify/purify the matrix containing the analyte. An International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) approved reference method for the measurement of glycohemoglobin in the blood of patients with diabetes has demonstrated that the combination of proteolysis and LC-MS can be used for clinically applicable tests.[31] This demonstrates that other methods utilizing MS, such as this, could eventually be used in the clinical environment as a way of standardizing commercially available immunoassays.

In this study, *absolute quantification* of C-reactive protein in human plasma was carried out utilizing *protein cleavage isotope dilution mass spectrometry* carried out by nanoLC-MS/MS. The samples were also analyzed by an independent laboratory using a typical clinical method. The absolute values between these two data sets were quite different; however, upon further inspection the values correlated well, as indicate by a $R^2 = 0.9708$. With this methodology validated, we were able to

quantify the CRP levels present in patients with EOC and found that as the severity of the cancer increased, the level of CRP present also increased. Thus, quantification of C-reactive protein in human plasma without the use of multiple clean up steps such as size exclusion chromatography and depletion of abundant proteins from plasma was completed successfully, verified by an orthogonal method, and utilized to find data of interest which will be evaluated in future experiments.

7.5 References

1. Anderson, L., Candidate-based proteomics in the search for biomarkers of cardiovascular disease*. J. Physiol. (London, U.K.)*, 2005. **563**. 23-60.

2. Rifai, N. and P.M. Ridker, High-sensitivity C-reactive protein: A novel and promising marker of coronary heart disease*. Clin. Chem.*, 2001. **47**. 403-411.

3. Tarkkinen, P., T. Palenius, and T. Lovgren, Ultrarapid, ultrasensitive one-step kinetic immunoassay for C-reactive protein (CRP) in whole blood samples: Measurement of the entire CRP concentration range with a single sample dilution*. Clin. Chem.*, 2002. **48**. 269-277.

4. Ridker, P.M., Clinical application of C-reactive protein for cardiovascular disease detection and prevention*. Circulation*, 2003. **107**. 363-369.

5. Verma, S. and E.T. Yeh, C-reactive protein and atherothrombosis--beyond a biomarker: an actual partaker of lesion formation*. Am J Physiol Regul Integr Comp Physiol*, 2003. **285**. R1253-6; discussion R1257-8.

6. Verma, S., P.E. Szmitko, and P.M. Ridker, C-reactive protein comes of age*. Nat. Clin. Practice Cardiovasc. Med.*, 2005. **2**. 29-36.

7. Braunwald, E., et al., *Harrison's Manual of Medicine, 15th ed.* 2002, New York: McGraw-Hill. 938-940.

8. Armani, A. and R.C. Becker, The biology, utilization, and attenuation of C-reactive protein in cardiovascular disease: Part I*. Am. Heart J.*, 2005. **149**. 971-976.

9. Kuhn, E., et al., Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and C-13-labeled peptide standards*. Proteomics*, 2004. **4**. 1175-1186.

10. Kiernan, U.A., et al., Quantitative multiplexed C-reactive protein mass spectrometric immunoassay*. J. Proteome Res.*, 2006. **5**. 1682-1687.

11. Barr, J.R., et al., Isotope dilution mass spectrometric quantification of specific proteins: Model application with apolipoprotein A-I*. Clin. Chem.*, 1996. **42**. 1676-1682.

12. de Leenheer, A.P. and L.M. Thienpont, Applications of Isotope-Dilution Mass-Spectrometry in Clinical-Chemistry, Pharmacokinetics, and Toxicology*. Mass Spectrom. Rev.*, 1992. **11**. 249-307.

13. Barnidge, D.R., et al., Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal. Chem.*, 2003. **75**. 445-451.

14. Barnidge, D.R., et al., Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.*, 2004. **3**. 644-652.

15. Anderson, L. and C.L. Hunter, Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics*, 2006. **5**. 573-588.

16. Pieper, R., et al., Multi-component immunoaffinity subtraction chromatography: An innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics*, 2003. **3**. 422-432.

17. Anderson, N.L., et al., Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA). *J. Proteome Res.*, 2004. **3**. 235-244.

18. Hawkridge, A.M., et al., Quantitative mass spectral evidence for the absence of circulating brain natriuretic peptide (BNP-32) in severe human heart failure. *Natl. Acad. Sci. U. S. A.*, 2005. **102**. 17442-17447.

19. Gerber, S.A., et al., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Natl. Acad. Sci. U. S. A.*, 2003. **100**. 6940-6945.

20. Kirkpatrick, D.S., S.A. Gerber, and S.P. Gygi, The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods*, 2005. **35**. 265-273.

21. Kippen, A.D., et al., Development of an isotope dilution assay for precise determination of insulin, C-peptide, and proinsulin levels in non-diabetic and type II diabetic individuals with comparison to immunoassay. *J. Biol. Chem.*, 1997. **272**. 12513-12522.

22. Fierens, C., et al., Standardization of C-peptide measurements in urine by method comparison with isotope-dilution mass spectrometry. *Clin. Chem.*, 2003. **49**. 992-994.

23. Ogata, Y., et al., Elevated levels of phosphorylated fibrinogen-alpha-isoforms and differential expression of other post-translationally modified proteins in the plasma of ovarian cancer patients. *J. Proteome Res.*, 2006. **5**. 3318-3325.

24. Scopes, R.K., Measurement of Protein by Spectrophotometry at 205-Nm. *Analytical Biochemistry*, 1974. **59**. 277-282.

25. Bondar, O.P., et al., LC-MS/MS quantification of Zn-alpha 2 glycoprotein: A potential serum biomarker for prostate cancer. *Clin. Chem.*, 2007. **53**. 673-678.

26. Hafner, G., et al., Comparison of diagnostic performance of cardiac troponin I on the IMMULITE system with other automated troponin I assays in minor myocardial damage. *Scand. J. Clin. & Lab. Invest.*, 2001. **61**. 227-235.

27. Maggiore, U., et al., Comparison of 3 automated assays for C-reactive protein in end-stage renal disease: Clinical and epidemiological implications. *J. Lab. & Clin. Med.*, 2005. **145**. 305-308.

28. Christenson, R.H., et al., Cardiac troponin I measurement with the ACCESS (R) immunoassay system: analytical and clinical performance characteristics. *Clin. Chem.*, 1998. **44**. 52-60.

29. Apple, F.S., et al., Multicenter clinical and analytical evaluation of the AxSYM troponin-I immunoassay to assist in the diagnosis of myocardial infarction. *Clin. Chem.*, 1999. **45**. 206-212.

30. Williams, T.I., et al., Epithelial ovarian cancer: Disease etiology, treatment, detection, and investigational gene, metabolite, and protein biomarkers. *J. Proteome Res.*, 2007. **6**. 2936-2962.

31. Jeppsson, J.O., et al., Approved IFCC reference method for the measurement of HbA(1c) in human blood. *Clinical Chemistry and Laboratory Medicine*, 2002. **40**. 78-89.

CHAPTER 8

Application of Iodoacetamide Derivatives for Increased Electrospray Response Utilizing the ALiPHAT Strategy

8.1 Introduction

Since the completion of the human genome project and with subsequent advances in databases, mass spectrometry has emerged as a leading technology in proteomics.[1-3] The use of iodoacetamide for alkylation of peptides and proteins has become common practice in the field of proteomics.[4-6] Dickens first discovered the reaction between iodoacetic acid and cysteine in 1933.[7] After a protein's disulfide bonds are reduced to their thiol forms, the cysteine residues are alkylated by reaction with iodoacetamide. Typically, this process serves to remove a protein's tertiary structure and prevent the reformation of disulfide bridges. There are multiple benefits to such a modification including increased efficiency of enzymatic digestions, enhanced chemical separations, and simplified interpretation of data for protein identification and characterization.

Chemical tags have long been utilized in the field of mass spectrometry for a multitude of purposes, extending the effectiveness of the measurements dramatically. Early in the history of chemical tagging for mass spectrometry, investigators demonstrated trimethylsilyl (TMS) derivatization improved the volatility of alcohols, thus allowing faster analysis of alcohols in hydrocarbon solutions without prior separation.[8] Later, the utility of TMS derivatization was applied to GC-MS

analysis[9], which is now ubiquitous. Chemical derivatization has also been utilized to allow the ionization of typically nonionizable compounds by electrospray ionization.[10-12] Sulfonic acid tags added to the N-terminus of tryptic peptides has been shown to result in a single PSD fragment ion series, which improves identification.[13]

Chemical derivatization of peptides and proteins has even been utilized to aid quantitative mass spectrometry. Cysteine residues have been modified via alkylation[14] including the addition of a chlorine-containing tag[15] to improve identification, while ICAT™[16], solid-phase ICAT[17], cleavable ICAT[18], and IDBEST™ (Target Discovery)[19] have allowed for the relative quantification of proteins in two different "states". Additionally, reagents that target primary amine groups (i.e., N-termini and lysine side chains) have been utilized for relative quantitation,[20, 21] including the iTRAQ reagent.[22] C-terminal tagging has also been explored; guanidination of C-terminal lysine for relative quantification is one such C-terminal tagging strategy.[23] Moreover, tagging of phosphorylation sites has also been developed for the relative quantification of phosphoproteins.[24] The numerous chemical tagging approaches for quantitative proteomics are presented in several recent reviews.[25-27] The basic concept of combining chemical derivatization and mass spectrometry has a long and productive history.

Electrospray ionization (*ESI*) is a powerful ionization method allowing for the characterization of biological macromolecules by mass spectrometry.[28, 29] In 2002, Professor John Fenn received the Nobel Prize in Chemistry for this invention given

its tremendous impact in chemistry, biology, drug discovery, medicine, and energy. However, *ESI* is a biased analytical technique dependent upon the various physical and chemical properties of the analyte. In 1993, Fenn demonstrated that the response factors for tetraalkyl ammonium halides increased as the hydrocarbon chain length increased from methyl to heptyl by nearly two orders of magnitude.[30]

The Muddiman group previously reported that a quantitative ESI-FTMS investigations of intact proteins[31] showed inherent biases which were attributed to the relative *hydrophobicity* of the analyte and internal standard. Specifically, the two dominant charge-states detected were the $[M+7H^+]^{7+}$ and $[M+8H^+]^{8+}$ for bovine and equine heart cytochrome C, respectively. Analysis of only the 7+ or the 8+ charge-state resulted in different sensitivities (slope of the calibration curve). However, when both charge-states were used in the linear least squares regression analysis, an ideal slope (m = 1.00) was achieved. These observations were explained based on the different hydrophobicities of equine and bovine heart cytochrome C and that a more hydrophobic species will have a lower average charge state.

Zhan and Fenn later analyzed two cyclic peptides, cyclosporin A and gramicidin S, and showed that the charge-state selectivity was related to the hydrophobic effect.[32] In these studies, they used reverse-phase liquid chromatography to demonstrate the relative hydrophobic character of these two peptides, cyclosporin A having a longer retention time than gramicidin S. The direct infusion *ESI* mass spectra of these two species were very different. Cyclosporin A was observed as both a singly-charged and doubly-charged species; each charge-

state was about 50% of the total ion abundance. However, for the less hydrophobic gramicidin S, over 95% of the total ion abundance was for the doubly-charged species.

In later work, the Enke group published a series of papers discussing the role that *hydrophobicity* has on *ESI response*.[33-35] In their first studies, they modeled experimental data obtained from a series of 3-mer's and convincingly demonstrated that an increase in non-polar character of small peptides resulted in a higher droplet surface activity, which translated into a higher *ESI response* factor.[33] They further studied the correlation of reverse-phase isocratic retention time with relative response for 6 different 3-mer peptides,[35] an approach Fenn[32] and the Muddiman[36] group utilized for larger biomolecules. Importantly, even with a relatively narrow retention window for the 3-mer peptides, good correlation was observed.[35]

Alkylation of peptides and proteins is undoubtedly utilized in many chemical tagging strategies because the reaction is simple and efficient.[5] Hydrophobic tagging of large biomolecules (>500 Da) was first described by Null et al.[36] One tagging strategy is the strategy (*a*ugmented *li*mits of detection for *p*eptides with *h*ydrophobic *a*lkyl *t*ags),[37] which is a method previously developed by Frahm et al. in which *electrospray response* of peptides was increased via a cysteine specific (iodoacetamide derivative) *hydrophobic tag*. In the previous study, peptides modified with one *hydrophobic tag*, 2-Iodo-*N*-octylacetamide, were demonstrated to have improved limits of detection relative to peptides alkylated with iodoacetamide.[37]

This chapter describes the application of four newly synthesized *hydrophobic tags* as well as the use of a previously developed tag for the application of the *ALiPHAT* strategy[37] on a new set of peptides.

8.2 Experimental

8.2.1 Materials

All peptides utilized in this study, iodoacetamide, and formic acid were purchased from Sigma-Aldrich (St. Louis, MO). Tris-HCL buffer (1M, pH 8.0) and ethanol were acquired from Fischer Scientific, Inc. (Pittsburgh, PA). Tris(2-carboxyethyl)phosphine was purchased from Pierce Biotechnology, Inc. (Rockford, IL). Burdick and Jackson HPLC-grade water, acetonitrile, and 2-proponol were purchased from VWR International (West Chester, PA). 2-Iodo-*N*-octylacetamide, 2-Iodo-*N*-dodecylacetamide, 2-Iodo-*N*-benzylacetamide, 2-Iodo-*N*-(phenethyl)acetamide, and 2-Iodo-*N*-(4-phenylbutyl)acetamide were synthesized and characterized in the Comins laboratory (North Carolina State University, Raleigh, NC). In addition, accurate mass was determined to ensure identity of synthesized *hydrophobic tags* (data not shown). All regents were utilized as received unless otherwise noted.

8.2.2 Reduction

A solution of 30 mM tris(2-carboxyethyl)phosphine (TCEP) was added to a 0.5 mg/mL standard peptide solution to achieve a final concentration of 9 mM; TCEP

and standard peptides were dissolved in 100 mM Tris buffer (pH 8.0). Samples were then incubated for 20 min at 37°C.

8.2.3 Alkylation with Iodoacetamide

A 20 mM stock solution of iodoacetamide was prepared in 100 mM Tris buffer. This solution was added to the peptide mixtures to a final concentration of 5 mM, vortexed, and allowed to incubate for 1 hour at 37°C in the dark. Each peptide was reacted with iodoacetamide in a different vial.

8.2.4 Alkylation with Hydrophobic Tags

A 20 mM stock solution of each hydrophobic alkylating reagent was prepared by first dissolving the appropriate amount in 200 µL of ethanol, to aid in solvation. Then, added 800 µL of 100 mM Tris was added. After adding the Tris buffer, some 2-iodo-*N*-octylacetamide and 2-Iodo-*N*-dodecylacetamide did precipitate out of solution; to account for this, hydrophobic alkylating reagents were added to reduced peptide solutions to a final concentration of 10 mM. This mixture was vortexed and allowed to incubate for 1 hour at 37°C in the dark. Each peptide and *hydrophobic tag* combination was reacted in a separate vial.

8.2.5 LC-ESI-FT-ICR MS

All peptides modified with *hydrophobic tags* were individually combined with peptides modified by iodoacetamide in a 1:1 volumetric ratio. Reversed-phase separation (*RP-HPLC*) was achieved on a 150 mm long × 0.5 mm inner diameter Jupiter Proteo column packed with 4 µm, 90 Å particles (Phenomenex, Torrance, CA). Mobile phase A was 98% water/ 1% isopropanol/ 1% acetonitrile/ 0.2% formic

acid and mobile phase B was 80% acetonitrile/ 10% water/ 10% isopropanol/ 0.2% formic acid. Pumps (Shimadzu LC-20AD, Shimadzu, Columbia, MD) were set to a 20 µL/min total flow rate. Gradient elution proceeded from 2% B at 2 minutes to a linear increase to 95% B at 60 minutes, maintained at 95% B for 10 minutes, and finally back to 2% B followed by a 10 minute re-equilibration period. Mass spectra of eluted species were acquired in MS mode with an LTQ-FT Ultra (Thermo Scientific, San Jose, CA) utilizing the IonMax Source. The AGC limit was set to 1×10^6 and mass resolving power (at *m/z* 400) was set to $100{,}000_{FWHM}$. External calibration was implemented according to the manufacturer's protocol.

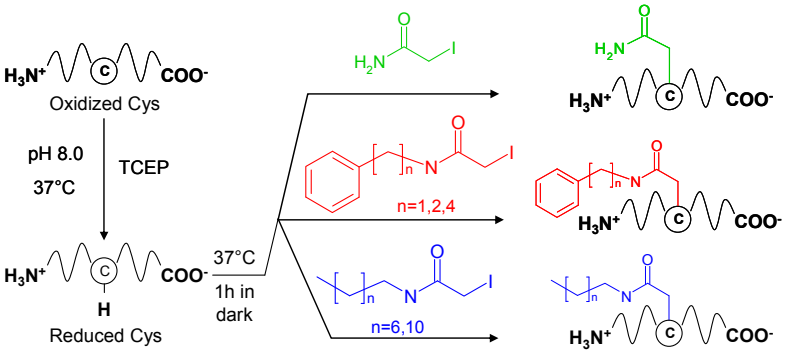
8.2.6 Determination of Relative Increase in Electrospray Response

Extracted ion chromatograms (XIC's) of the entire isotopic distribution of all charge states of each alkylated peptide were attained. Peak areas for each XIC were determined utilizing the ICIS peak detection method in Xcalibur software (Thermo Scientific) provided with the instrument. The relative increase reported is the ratio of the XIC area of a particular peptide alkylated with a *hydrophobic tag* to the same peptide alkylated with iodoacetamide, as shown in **Equation 8.1**.

$$\text{Equation 8.1: peak area ratio} = \frac{(\sum_{z=1}^{i} \text{peak area})_{\text{Hydrophobic Tag}}}{(\sum_{z=1}^{i} \text{peak area})_{\text{Iodoacetamide}}}$$

8.3 Results and Discussion

The application of these *hydrophobic tags* was completed on peptide E-76 (**Table 8.1**, peptide 1), whose amino acid sequence is shown in **Figure 8.1**, and two additional peptides utilizing liquid chromatography Fourier-transform ion cyclotron resonance mass spectrometry (LC-FT-ICR-MS). The E-76 peptide was chosen because of its biological significance as a potent inhibitor of coagulation factor VIIIa[38] and because its amino acid sequence contained cysteines. Iodoacetamide was the standard alkylating agent to which the other tags were compared. The reaction of iodoacetamide with cysteine groups results in a carboxyamidomethyl (*CAM*) modification. The peptides modified with the *hydrophobic tags* were combined with their *CAM* modified counterpart such that identical molar amounts of each were injected on-column. **Scheme 8.1** shows general reduction of protein and alkylation reactions of the *hydrophobic tags* used in these experiments.



Scheme 8.1 General reduction and alkylation of peptides

The ratio of peak areas from the extracted ion chromatograms (XICs) describes both the chromatographic and electrospray ionization consequences of addition of the *hydrophobic tag*. This ratio was utilized to quantify the results described herein, as described above.[7]

E-76 peptide alkylated with 2-Iodo-*N*-(4-phenylbutyl)acetamide (**Scheme 8.1**, n = 4) was observed to elute 6.05 minutes after the *CAM* modified peptide. This represents an increase in mobile phase B by 8.5%, which demonstrates the fact that this modified peptide is indeed more hydrophobic than the *CAM* modified peptide. The overlaid XICs of this and the *CAM* modified peptides are shown in **Figure 8.1**. The ratio of the area of these peaks demonstrates an improvement of 429 for the peptide modified with 2-Iodo-*N*-(4-phenylbutyl)acetamide versus the standard *CAM* modified peptide. This *hydrophobic tag* provided the greatest improvement over the *CAM* modification for the E-76 peptide.

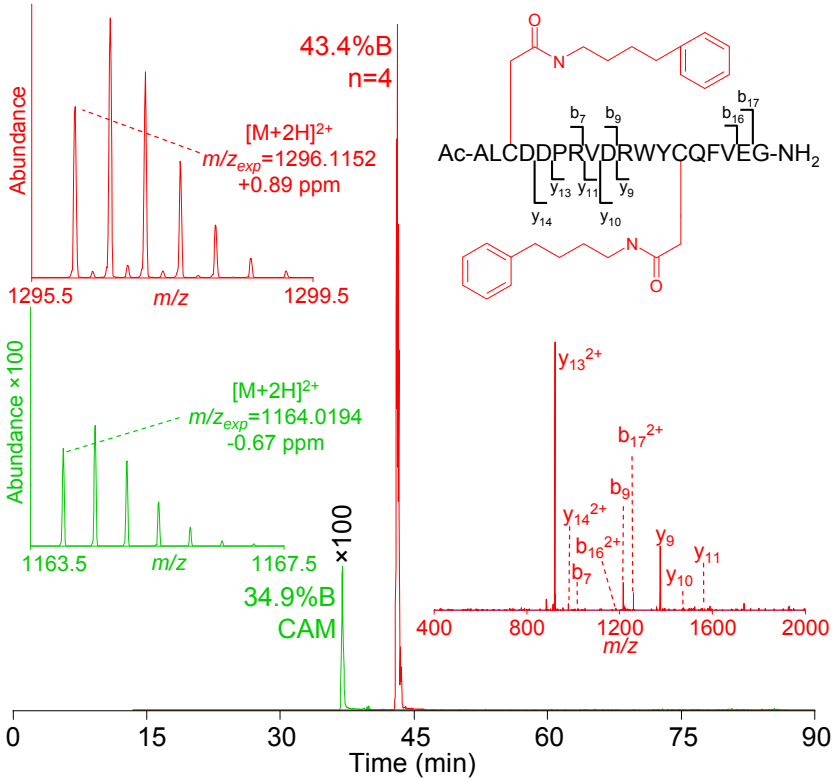


Figure 8.1 The overlaid XICs of the CAM (green) and n = 4 (red) modified peptide are shown as well as their mass spectra. Denoted is the charge state, theoretical m/z, and observed mass measurement accuracy. The sequence of the E-76 peptide as well as a labeled MS/MS spectrum demonstrating the n = 4 tag does not fragment under conditions of CID nor hinder MS/MS experiments are also shown

Also shown in **Figure 8.1** are mass spectra; one representing the *CAM* modified peptide and one representing the peptide modified with 2-Iodo-*N*-(4-phenylbutyl)acetamide which have theoretical *m/z*'s of 1164.0202 and 1296.1140 respectively, for the 2+ charge state. MS/MS data is shown for this modified peptide which demonstrates that the tag itself does not fragment. This is important to note because if the tag were to fragment under conditions of *collision induced dissociation* (*CID*), it would make the interpretation of MS/MS data much more difficult. Tandem MS data were analyzed for all peptides with each of the tags discussed herein. None of the tags were observed to fragment. The fragmentation pattern remained the same between peptides modified with iodoacetamide and the new *hydrophobic tags*, which demonstrates that the new tags do not adversely affect the *CID* mechanism for fragmentation.

Table 8.1 Investigated Peptide Sequences

Peptide #	Sequence
Peptide 1 (E-76)	Ac-ALCDDPRVDRWYCQFVEG-NH_2
Peptide 2	SCSLPQTSGLQKPES-NH_2
Peptide 3	CYFQNCPRG-NH_2

E-76 peptide alkylated with 2-Iodo-*N*-benzylacetamide (Scheme 1, n = 1) improved the chromatography and *ESI response*. **Figure 8.2** shows the improvement of 69. Modification of the peptide with this tag resulted in a shift in retention time by 2.85 minutes. This shift in retention time is a result of the peptide being more hydrophobic which causes it to elute in a concentration of mobile phase B increased by 3.4% when compared to the *CAM* modified peptide. The

improvement in chromatography is visible by a narrower peak which results from the more hydrophobic modified peptide being able to be captured more efficiently at the head of the column which favors sample concentration prior to elution, as compared to the *CAM* modified peptide. This phenomenon is also observed with the other tags examined in this study.

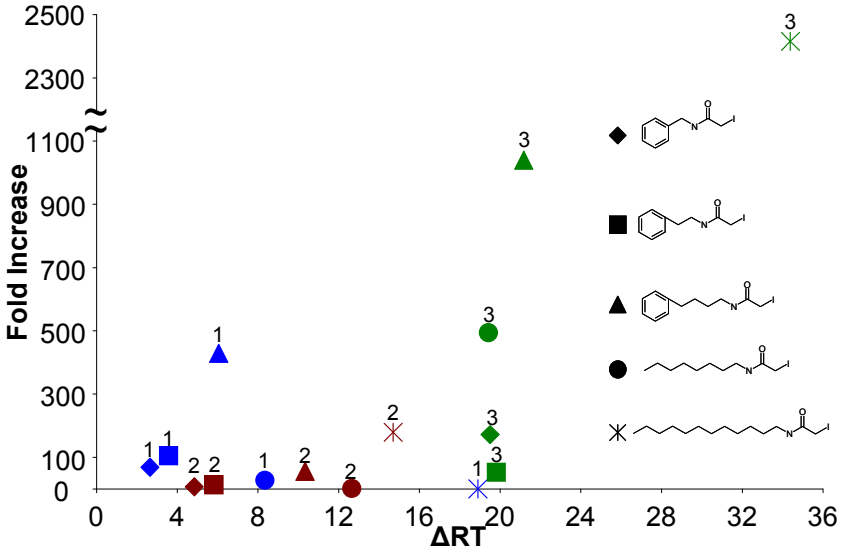


Figure 8.2 ESI response fold improvement vs. change in retention time in relation to the CAM modified peptide. Tags are represented by point shape and peptides represented by numbers corresponding to Table 1.

2-Iodo-*N*-(phenethyl)acetamide reacts with cysteines to create a modification observed in Scheme 1 with n = 2. This modified E-76 peptide resulted in a slightly more hydrophobic peptide than when modified with 2-Iodo-*N*-benzylacetamide. This increased *hydrophobicity* results in an increase in retention time of 3.57 minutes over the *CAM* modified peptide. This difference in retention time corresponded to the elution of this tagged peptide in mobile phase with an increase of 4.5% B over the *CAM* modified peptide. The ratio of the peak areas demonstrated an improvement of 105 for the modified peptide, shown in **Figure 8.2**.

As previously shown, alkylation with 2-Iodo-*N*-octylacetamide (Scheme 1, n=6) creates an octylcarboxyamidomethyl (*OCAM*) modification to cysteines in a

peptide or protein sequence.[37] The addition of the *OCAM* tag to the cysteines of E-76 improved both the chromatography and *electrospray response* of the peptide which is evident by the increase in peak area by 27 times versus the *CAM* modified peptide, shown in **Figure 8.2**. The difference in retention time was 8.33 minutes which corresponds to an increase of 13.4% B for the *OCAM* peptide to elute. The increase in *hydrophobicity* is evident by the increased retention time; however, the *OCAM* tag does not improve the *ESI response* as well as the tags with the phenyl group.

The final tag examined was 2-Iodo-*N*-dodecylacetamide (Scheme 1, n = 10), which generates a dodecylcarboxyamidomethyl (*DCAM*) modification to cysteines. The *DCAM* modified E-76 peptide lead to an increase in retention time of 16.90 minutes versus the *CAM* modified peptide. This difference in retention time represents an increase in mobile phase B by 25.9%. *DCAM* modification of the E-76 peptide results in the most hydrophobic of any of the tagged peptides observed in this study. The peak ratio of the XIC between the *DCAM* and *CAM* modified peptide yielded a result of 0.6.

Figure 8.2 demonstrates the relationship between the differences in retention time between the different peptides (Table 1) with the iodoacetamide derived tags and their *CAM* modified counterpart. The modifications with the phenyl terminal group yielded better improvements when compared to the tags with only alkyl chains for the E-76 peptide. 2-Iodo-*N*-dodecylacetamide performed the best for peptides 2

and 3 and provided for *electrospray response* improvements of 179 and 2441, respectively.

8.4 Conclusions

Five *hydrophobic tags* have been applied to three peptides whose *electrospray response* improvements are summarized in **Figure 8.2**. The data, presented in this chapter, clearly shows alkylation of the E-76 peptide, with the *hydrophobic tags* applied in this study, have the capability to improve *ESI response* >400 fold as well as provide for improved chromatographic behavior. Peptides 2 and 3 showed improvement in *electrospray response* for all *hydrophobic tags* in comparison to their *CAM* modified counterpart. Furthermore, peptide 3 was able to achieve an improvement of >2000 fold improvement over its *CAM* modified counterpart! These improvements in *electrospray response* come at essentially no experimental cost since the alkylation step is facile and carried out in nearly all *bottom-up proteomic* analyses. These benefits will be able to aid in the investigation of cysteine containing peptides and proteins that have low *electrospray response* or concentration. In addition, their future application to targeted proteomic studies may aid in lowering limits of detection for cysteine containing peptides of interest.

8.5 References

1. Kenyon, G.L., et al., Defining the mandate of proteomics in the post-genomics era: Workshop report. *Mol. Cell. Proteomics*, 2002. **1**. 763-780.

2. Aebersold, R. and M. Mann, Mass spectrometry-based proteomics. *Nature*, 2003. **422**. 198-207.

3. Mann, M., R.C. Hendrickson, and A. Pandey, Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.*, 2001. **70**. 437-473.

4. Moritz, R.L., et al., S-pyridylethylation of intact polyacrylamide gels and in situ digestion of electrophoretically separated proteins: A rapid mass spectrometric method for identifying cysteine-containing peptides. *Electrophoresis*, 1996. **17**. 907-917.

5. Herbert, B., et al., Reduction and alkylation of proteins in preparation of two-dimensional map analysis: Why, when, and how? *Electrophoresis*, 2001. **22**. 2046-2057.

6. Sechi, S. and B.T. Chait, Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.*, 1998. **70**. 5150-5158.

7. Dickens, F., Interaction of halogenacetates and sh compounds. The reaction of halogenacetic acids with glutathione and cysteine. The mechanism of iodoacetate poisoning of glyoxalase. *Biochem. J.*, 1933. **27**. 1141-1151.

8. Sharkey, A.G., R.A. Friedel, and S.H. Langer, Mass Spectra of Trimethylsilyl Derivatives. *Anal. Chem.*, 1957. **29**. 770-776.

9. Sweeley, C.C., et al., Gas-liquid chromatography of trimethylsilyl derivatives of sugars and related substances. *JACS*, 1963. **85**. 2497-&.

10. Quirke, J.M.E., C.L. Adams, and G.J. Vanberkel, Chemical derivatization for electrospray-ionization mass-spectrometry .1. Alkyl-halides, alcohols, phenols, thiols, and amines. *Anal. Chem.*, 1994. **66**. 1302-1315.

11. Van Berkel, G.J. and K.G. Asano, Chemical derivatization for electrospray-ionization mass-spectrometry .2. Aromatic and highly conjugated molecules. *Anal. Chem.*, 1994. **66**. 2096-2102.

12. Van Berkel, G.J., et al., Derivatization for electrospray ionization mass spectrometry. 3. Electrochemically ionizable derivatives. *Anal. Chem.*, 1998. **70**. 1544-1554.

13. Keough, T., R.S. Youngquist, and M.P. Lacey, A method for high-sensitivity peptide sequencing using postsource decay matrix-assisted laser desorption ionization mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.*, 1999. **96**. 7131-7136.

14. Sechi, S. and B.T. Chait, Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.*, 1998. **70**. 5150-5158.

15. Goodlett, D.R., et al., Protein identification with a single accurate mass of a cysteine-containing peptide and constrained database searching. *Anal. Chem.*, 2000. **72**. 1112-1118.

16. Gygi, S.P., et al., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.*, 1999. **17**. 994-999.

17. Zhou, H.L., et al., Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat. Biotechnol.*, 2002. **20**. 512-515.

18. Li, J.X., H. Steen, and S.P. Gygi, Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents - The yeast salinity stress response. *Mol. Cell. Proteomics*, 2003. **2**. 1198-1204.

19. Hall, M.P. and L.V. Schneider, Isotope-differentiated binding energy shift tags (IDBEST (TM)) for improved targeted biomarker discovery and validation. *Expert Rev. Proteomics*, 2004. **1**. 421-431.

20. Geng, M.H., J.Y. Ji, and F.E. Regnier, Signature-peptide approach to detecting proteins in complex mixtures. *J. Chromatogr., A*, 2000. **870**. 295-313.

21. Chakraborty, A. and F.E. Regnier, Global internal standard technology for comparative proteomics. *J. Chromatogr., A*, 2002. **949**. 173-184.

22. Ross, P.L., et al., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics*, 2004. **3**. 1154-1169.

23. Cagney, G. and A. Emili, De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.*, 2002. **20**. 163-170.

24. Goshe, M.B., et al., Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal. Chem.*, 2001. **73**. 2578-2586.

25. Julka, S. and F. Regnier, Quantification in proteomics through stable isotope coding: A review*. J. Proteome Res.*, 2004. **3**. 350-363.

26. Leitner, A. and W. Lindner, Current chemical tagging strategies for proteome analysis by mass spectrometry*. J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.*, 2004. **813**. 1-26.

27. Ong, S.E. and M. Mann, Mass spectrometry-based proteomics turns quantitative*. Nat. Chem. Biol.*, 2005. **1**. 252-262.

28. Yamashita, M. and J.B. Fenn, Electrospray ion-source - another variation on the free-jet theme*. J. Phys. Chem.*, 1984. **88**. 4451-4459.

29. Fenn, J.B., et al., Electrospray ionization for mass-spectrometry of large biomolecules*. Science*, 1989. **246**. 64-71.

30. Fenn, J.B., Ion formation from charged droplets - roles of geometry, energy, and time*. J. Am. Soc. Mass Spectrom.*, 1993. **4**. 524-535.

31. Gordon, E.F., et al., Hydropathic influences on the quantification of equine heart cytochrome c using relative ion abundance measurements by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry*. J. Mass Spectrom.*, 1999. **34**. 1055-1062.

32. Zhan, D.L. and J.B. Fenn, Gas phase hydration of electrospray ions from small peptides*. Int. J. Mass Spectrom.*, 2002. **219**. 1-10.

33. Cech, N.B. and C.G. Enke, Relating electrospray ionization response to nonpolar character of small peptides*. Anal. Chem.*, 2000. **72**. 2717-2723.

34. Cech, N.B. and C.G. Enke, Effect of affinity for droplet surfaces on the fraction of analyte molecules charged during electrospray droplet fission*. Anal. Chem.*, 2001. **73**. 4632-4639.

35. Cech, N.B., J.R. Krone, and C.G. Enke, Predicting electrospray response from chromatographic retention time*. Anal. Chem.*, 2001. **73**. 208-213.

36. Null, A.P., A.I. Nepomuceno, and D.C. Muddiman, Implications of hydrophobicity and free energy of solvation for characterization of nucleic acids by electrospray ionization mass spectrometry*. Anal. Chem.*, 2003. **75**. 1331-1339.

37. Frahm, J.L., et al., Achieving augmented limits of detection for peptides with hydrophobic alkyl tags*. Anal. Chem.*, 2007. **79**. 3989-3995.

38. Dennis, M.S., et al., Peptide exosite inhibitors of factor VIIa as anticoagulants. *Nature*, 2000. **404**. 465-470.

APPENDIX

APPENDIX A

Summary of protocols utilized in this dissertation are listed on the following pages.

Laboratory Protocol: PPG-1000 Electrospray Calibration Solution

PPG-1000 – Aldrich 202320-250G
NH_4OAC – Aldrich 431311-50G
Solvents: Burdick-Jackson HPLC Water (365-4), Fisher 2-Propanol (A416-4)

Solution 1: 10 mL of a 5 mg/mL Stock Solution of PPG-1000 in 2-propanol
Density PPG-1000: 1.01 g/mL

Pipet 49.5 µL of PPG-1000 and dilute to 10 mL with 2-propanol in blue top Falcon Tube

---

Solution 2: 10 mL of a 70/30 vol/vol of 2-propanol/water with 0.5 M NH_4OAc

Weigh out 0.3854 grams of NH_4OAc and place into 10 mL blue top Falcon Tube

Add 3 mL of water, cap and vortex for 1 minute

Add 7 mL of 2-propanol, cap and vortex for 1 minute

Solution 3: 10 mL of the PPG-1000 ESI Calibration Solution at 5 ng/µL

Take 10 µL of solution 1 and add it to solution 2. Cap and vortex for 3 minutes

Aliquot solution into 5 × 1 mL plastic centrifuge tubes; 1 mL for laboratory purposes and other 4 should be stored in refrigerator. Discard remaining solution.

Theoretical *m/z* values for NH_4^+ Adducted Oligomers

PPG Oligomer Number	Theoretical *m/z*
12_A	732.54677
12_{A+1}	733.55012
13_A	790.58863
13_{A+1}	791.59199
13_{A+2}	792.59534
14_A	848.63050
14_{A+1}	849.63385
14_{A+2}	850.63721
15_A	906.67236
16_A	964.71423
17_A	1022.75609
18_A	1080.79796
19_A	1138.83982
19_{A+1}	1139.84318
19_{A+2}	1140.84653
20_A	1196.88169

Laboratory Protocol: Tryptic Plasma Digestion for PC-IDMS Experiments

Acetic Acid – Fisher A38-212-2.5L
$CaCl_2$ – Sigma C1016-100g
DTT – Sigma D0632-5g
Formic Acid – Fluka 56302-50mL
Iodoacetamide – Sigma I1149-5g
Trypsin – Sigma T1426-250g
TRIS-HCL – Fisher BP1758-100mL
Urea – Fluka 51456-500g
Solvents: Burdick-Jackson HPLC Water (365-4) & Methanol (230-4)

1) Dry urea made by diluting 60 mg of urea in ~500 µL HPLC Methanol (~2M) in 1.5 mL eppendorf tube. Then use Speed-Vac to dry down completely.
2) 125 µL of plasma is added to dry urea to create 6M solution – vortex.
   NOTE - volume will increase to ~ 150 µL
3) Add 5 µL of internal standard peptide mix and 150 µL 50 mM TRIS pH 8.0 containing 10 mM $CaCl_2$ - vortex
4) Add 5 µL of 1M DTT (made fresh!) and heat at 45°C for 1hour – vortex*
5) Add 30 µL of 0.5M Iodoacetamide (made fresh!) and let react in the dark for 1 hour – vortex*
6) Dilute urea to 2M by adding 120 µL of 50 mM TRIS - vortex
7) Add 10 µL of 0.1 mg/ µL trypsin stock (1 mg trypsin per sample) – vortex**
8) Digest at 37°C for 12 hours
9) Add 5 µL of formic acid to stop the reaction – vortex: Final volume ~ 480 µL
10) Dilute to necessary concentration for desired application.

As Adapted from: Bondar et al., Clin. Chem. 2007 Apr;53(4):673-8

*DTT and Iodoacetamide are dissolved with 50mM Tris buffer
**Trypsin is dissolved in 50 mM Acetic Acid and heated 30 min @ 37°C

APPENDIX B

Glossary: Definition of terms and abbreviations italicized in dissertation.

A_s- *relative ion population*

A_T- *total ion population*

Absolute quantification- peptide or protein ion abundance is compared with a stable isotope labeled standard, which is added prior to mass spectrometric interrogation, at a known concentration; results usually reported as specific concentration.

Acute-phase protein- protein whose concentration varies by more than 25% in response to an inflammatory stimulus.

ALiPHAT- augmented limits of detection for peptides with hydrophobic alkyl tags.

Artificial neural network- ANN; computer model which contains an interconnection of neurons, which have weights and biases; utilized for data fitting; assumes no physical model.

Automatic Gain Control (AGC)- method by which a mass spectrometer can meter the number of ions reaching the detection cell.

Biomarker- a peptide or protein that is an indicator of disease.

Bottom-up proteomics- proteins are digested utilizing enzymes or other chemical processes. The resulting peptides are then identified from these peptides' masses and their MS/MS data.

Box and Whisker Plot- blue diamond shows the mean and the 95% confidence interval of the mean, while the blue notched lines show the 95% parametric percentile range (2.5-97.5%). The notched box shows the median, lower and upper quartiles. The dotted-line connects the nearest observations within 1.5 IQRs (inter-quartile ranges) of the lower and upper quartiles. Red crosses (+) and circles (o) show observations more than 1.5 IQRs (near outliers) and 3.0 IQRs (far outliers) from the quartiles, respectively.

CAM- modification resulting from the alkylation of cysteines with iodoacetamide.

Collision Activated Dissociation (CAD)- see collision induced dissociation.

Collision Induced Dissociation (CID)- low energy dissociation method which yields y- and b-type peptide fragment ions.

Cyclotron Frequency- number of times, in one second, that ions orbit inside of the ICR cell. This is inversely proportional to the particle's mass-to-charge ratio.

DCAM- modification resulting from the alkylation of cysteines with 2-Iodo-*N*-dodecylacetamide.

Electron-capture dissociation (ECD)- a method which transfers an electron to a cation to produce dissociation along the peptide backbone.

Electron-transfer dissociation (ETD)- a method which uses an anion to transfer an electron to a cation to produce dissociation along the peptide backbone.

Electrospray Response- ability of an ion to escape the droplet in electrospray ionization. Measured via ion abundance.

ELISA- enzyme-linked immunosorbant assay; utilized to detect a particular antigen or antibody in a sample.

ESI- electrospray ionization; a soft ionization method used in mass spectrometry.

Essential Amino Acid- amino acid not synthesized by an organism; as such must be included in diet.

Forbidden zone- *m/z* region that is not occupied by proteins or peptides.

Free energy of solvation- energy required to transfer the ion from solution to the gas phase.

FT-ICR- Fourier transform ion cyclotron resonance.

Hydrophobicity- determined by the amino acids that make up a protein or peptide. A GRAVY (grand average of hydropathy) score is given to each amino acid, which are summed up across the protein/peptide sequence and divided by total number of amino acids. The scale ranges -4.6 to 4.6 of which 4.6 is the most hydrophobic and -4.6 is the most hydrophilic.

Hydrophobic Tag- chemical tag utilized to increase the hydrophobicity of a biomolecule.

Iodopeptides- a set of peptide internal standards with 3,5-diiodotyrosine incorporated to shift the *mass defect* into *forbidden zones*.

Ion trap control language (ITCL)- computer control language utilized to change the operation of ThermoScientific linear ion trap.

Lorentz Force- force exerted upon a charged particle in an electromagnetic field resulting from both the electric field and magnetic field forces present.

Multiple linear regression (MLR)- a regression technique which determines the relationship between multiple independent variables and a dependant variable.

MALDI- matrix assisted laser desorption ionization.

Mass defect- difference between the *monoisotopic mass* and the mass of the nucleons.

Mass Measurement Accuracy (MMA)- accuracy of the mass measured compared to the theoretical mass. In FT-ICR it is usually reported in terms of parts-per-million, but can be in the parts-per-billion range.

Monoisotopic mass- comprised of the sum of the most abundant naturally occurring isotope of each of the elements in the chemical formula.

Neural Network- see *artificial neural network*.

OCAM- modification resulting from the alkylation of a cysteine with 2-Iodo-*N*-octylacetamide.

Peak coalescence- similar *m/z* peaks combine to become indistinguishable.

Post excite radius- orbital radius of the ions in the ICR cell after application of excitation voltage.

Protein cleavage isotope dilution mass spectrometry (PC-IDMS)- a method utilized to perform *absolute quantification* utilizing stable isotope labeled peptides as internal standards.

Receiver operator characteristic (ROC) curve- a graph that illustrates sensitivity and specificity of *biomarkers*.

Relative ion population- A_S; the ion population of a particular *m/z* value present in detector; also relative ion abundance.

Relative quantification- ion abundances of the analyte are compared to a standard with results usually reported as fold improvement.

Resolving Power (RP)- this is normally defined at the full width at half maximum (FWHM). High resolving power aids in achieving high MMA.

RP-HPLC- reverse phase high performance liquid chromatography

Scopes method- method utilized to quantify concentration of neat peptide samples; performed by measuring absorbance at 205 nm; concentration in mg/L equal to absorbance multiplied by pathlength divided by 31.

Selective Reaction Monitoring (SRM)- method by which a particular product ion is monitored from the fragmentation of a particular parent ion; both parent and product *m/z* are specified by user.

Sensitivity- true positive fraction of results for a particular test or *biomarker*.

Space-charge effects- caused by columbic repulsion between ion packets in the ICR cell, which shifts the frequency of these ions. The result is an adverse effect on *mass measurement accuracy*.

Specificity- true negative fraction of results for a particular test or *biomarker*.

Total ion population- A_T; the total number of ions (all *m/z*'s) present in detector.

Tandem Mass Spectrometry (MS/MS, MS^n)- process that fragments ions in the mass spectrometer. Can utilize a number of different dissociation techniques. (e.g., *collision induced dissociation*)

Top-Down Proteomics- method for investigating proteins in which an accurate intact mass is determined; followed by gas-phase dissociation. This information can then be used to search a database that can identify the protein(s) in question.