USING ION MOBILITY SPECTROMETRY TO RESOLVE ISOTOPOMERS, IDENTIFY ISOMERS BY ISOTOPOLOGIC SHIFTS, AND PREDICT MOBILITY CHANGES FOR POST-TRANSLATIONALLY MODIFIED PEPTIDES

A Thesis by

Julia L. Kaszycki

Master of Arts, Wichita State University, 2013

Bachelor of Science, Mount Union University, 2012

Submitted to the Department of Chemistry
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Master of Science

May 2017

USING ION MOBILITY SPECTROMETRY TO RESOLVE ISOTOPOMERS, IDENTIFY ISOMERS BY ISOTOPOLOGIC SHIFTS, AND PREDICT MOBILITY CHANGES FOR POST-TRANSLATIONALLY MODIFIED PEPTIDES

The following faculty members have examined the final copy of this thesis for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Science, with a major in Chemistry.

_____
Alexandre Shvartsburg, Committee Chair

_____
David Eichhorn, Committee Member

_____
Katie Mitchell-Koch, Committee Member

_____
William Groutas, Committee Member

_____
Elizabeth Behrman, Committee Member

ACKNOWLEDGMENTS

I would like to thank my family including my parents and all four siblings for their support throughout my extensive stay in school. I would like to thank my lab partners Matt Baird and Andrew Bowman for their humor and patience with my happy morning personality. I would also like to thank my advisor, Dr. Alexandre Shvartsburg, for allowing me to be a contributing member of the group as his lab started.

I would like to thank my committee members including Dr. David Eichhorn, Dr. William Groutas, Dr. Elizabeth Behrman, Dr. Katie Mitchell-Koch, and Dr. Alexandre Shvartsburg for agreeing to be on my committee. Other thanks include Dr. Kevin Langenwalter for all of the electrical help so early in the labs creation and lending random tools. Also, Susan McCoy and Mary Cambridge in the stockroom for their help when different varieties of equipment were needed.

Outside of WSU, Dr. Keqi Tang of PNNL and Dr. Gordon A. Anderson of GAACE both contributed to the installation FAIMS device and the ion funnel utilized in the experiments and the refining of how we utilized them, for which I am grateful.

A kind thanks is also extended to the funding used for this research including NIH K-INBRE (P20 GM103418), NIH COBRE (P30 GM 110761), NSF FIRST (ESP-0903806), and NSF CAREER (CHE-1552640).

ABSTRACT

Differential or field asymmetric waveform ion mobility spectrometry (FAIMS) operating at high electric fields fully resolves isotopic isomers for a peptide with labeled residues. The naturally present isotopes, alone and together with targeted labels, also cause spectral shifts that approximately add for multiple heavy atoms. Separation qualitatively depends on the gas composition and field strength. These findings may enable novel strategies in proteomic and metabolomic analyses using stable isotope labeling.

FAIMS delves more deeply into the naturally present isotopes and shows a method for characterization of isomers based upon their isotopologic shifts in a high electric field. It looks at the differentiating +1 Da shifts for identification with monohalogenated anilines. This produces an identifying fingerprint potentially for any ion based on isotopologic shifts.

The rising profile of ion mobility spectrometry (IMS) in proteomics has driven the efforts to predict peptide cross-sections. In the simplest approach, these are derived by adding the contributions of all amino acid residues and post-translational modifications (PTMs) defined by their intrinsic size parameters (ISPs). It shows that the ISPs for PTMs can be calculated from properties of constituent atoms, and introduce the “impact scores” that govern the shift of cross-sections from the central mass-dependent trend for unmodified peptides. The ISPs and scores tabulated for 100 more common PTMs enable predicting the domains for modified peptides in the IMS/MS space that would guide subproteome investigations.

PREFACE

This thesis is a compilation of several projects. Two have been published in the Journal of American Society for Mass Spectrometry, the first being *Ion Mobility Separation of Peptide Isotopomers* Volume 27, page 795, published May 2016. The second is *A priori Intrinsic PTM Size Parameters for Predicting the Ion Mobilities of Modified Peptides* Volume 28, page 294, published January 2017. The third is in preparation for publication in a journal, but has not yet been submitted as of the completion of this document. The publications have been adapted for this document with permissions from the publisher.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS

MS	Mass Spectrometry
GC	Gas Chromatography
LC	Liquid Chromatography
IMS	Ion Mobility Spectrometry
NMR	Nuclear Magnetic Resonance
ESI	Electrospray Ionization
DC	Direct Current
RF	Radio Frequency
CV	Compensation Voltage
FAIMS	Field Asymmetric Waveform Ion Mobility Spectromtry
N_2	Nitrogen Gas
CO_2	Carbon Dioxide Gas
He	Helium Gas
H_2	Hydrogen Gas
ETD	Electron Transfer Dissociation
MBA	Monobromoaniline
MCA	Monochloroaniline
ISP	Intrinsic Size Parameter
pal	Palmitoylation
cam	Carboxyamidomethylation
p	Phosphorylation
CID	Collision Induced Dissociation

LIST OF ABBREVIATIONS (continued)

R	Resolving Power
AA	Dialanine unlabeled
A*A	Dialanine heavy label on the first alanine (amine end)
AA*	Dialanine heavy labeled on the second alanine (carboxyl end)
AAA	Trialanine unlabeled
A*AA	Trialanine heavy labeled on the first alanine (amine end)
AA*A	Trialanine heavy labeled on the middle alanine
AAA*	Trialanine heavy labeled on the third alanine (carboxyl end)
Da	Dalton (unit mass)
STD	Standard Deviation

LIST OF NOMENCLATURE

Stoichiometry	Atomic makeup of a molecule
Nominal Mass	Atomic mass based upon integer values
Exact Mass	Atomic mass with as many significant figures as measurable
Isobaric	Same exact mass
Isomers	Molecules with identical stoichiometry varying by structure or isotopic makeup.
Isotopologues	Isomers with different isotopic makeup
Isotopomers	Isobaric isotopologues
Isotopic Envelope	Isotopic pattern from naturally occurring isotopes seen in Mass Spectrometry
Pseudo Isotopomers	Isomers with the same nominal mass but not the same exact mass.

LIST OF SYMBOLS

m/z	Mass to charge ratio
K	Absolute Mobility
E	Electric Field Strength
g	Gap Width
Ω	Cross Section
Ω_{IMP}	Cross Section Impact
mM	Millimolar
V	Volts
V/cm	Volts per centimeter
E_C	Compensation field
Ω/z	Cross section to charge ratio
q	Charge
m	Mass of ion
M	Gas molecule mass
T	Gas Temperature
k	Boltzmann constant
E_D	Dispersion field
S_{iso}	Isotopomeric shift
ΔE_{iso}	Isotopic shifts
w	Peak Width
µM	Micromolar ($1x10^{-6}$ moles/liter)
Å	Angstrom ($1x10^{-10}$ meters)

CHAPTER 1

INTRODUCTION

Mass spectrometry (MS) originated in the late 1800s with experiments ranging from electric discharges at low pressure (Thompson) to positive ray deflection by magnetic fields (Wein) [1]. The instruments tasked to measure objects as small as ions when the machining capabilities were extremely limited were understandably finicky. Francis William Aston, credited with the creation of the first mass spectrograph [1], has summed thus:

> "*The mass spectrometer behaves at times in the most capricious and unaccountable manner...When by good fortune all is well, the arrangement is capable of good performance. Thus, after a favorable setting of the apparatus, six elements were successfully analyzed in as many working days. On the other hand, after dismantling became imperative, and it had to be cleaned and rebuilt, exactly as before as far as one could tell, no results of any value could be obtained during weeks of work*."

Despite those challenges, MS soon enabled a crucial discovery of the existence of isotopes [2]. The finding that elements possess more than one naturally occurring mass and the corresponding nuclei can be resolved has enabled numerous new scientific directions and given MS the first major practical application - preparative isotope separation in the 1940s [3]. Today, isotopes have various uses. In particular, many quantification and pathway-tracking strategies in proteomics and metabolomics employ stable isotope labeling [4 - 10]. Natural isotopic patterns, including the quantitative abundances, are employed to confirm stoichiometric assignments by MS.

Mass spectrometry still faces substantial limitations for complex mixtures. Even tandem MS spectra are often congested and burdened by uninformative fragments. These problems can be alleviated by separations prior to MS, but each separation method also has constraints. Historically, MS was preceded by gas chromatography (GC), liquid chromatography (LC), or

electrophoresis. All chromatographic methods separate compounds based on the affinity for a stationary phase in the column. The presence and particularly location of neutrons on a molecule (except for hydrogen/deuterium substitution) do not substantially affect its binding to a stationary phase and hence separations. The same applies to electrophoretic approaches that separate molecules by mobility in liquids and charge-related phenomena. Identification and quantification of isotopically labeled molecules is important for the elucidation of reaction mechanisms. Determination of the site of stable isotope incorporation is a key step in the understanding of mechanisms and structures. One approach is Nuclear Magnetic Resonance (NMR) which is primarily used for structural characterization or molecular identification [11]. However, NMR requires large amounts of pure compounds and is thus not suitable for trace amounts or complex mixtures.

A recently emerged approach of ion mobility spectrometry (IMS) offers faster separation based on the mobility of ions in gases under the influence of an electric field [12]. All IMS methods can be classified as linear or nonlinear. Linear IMS can provide only minor separation for ions differing solely in the number of neutrons and none for those with only different neutron placement. In contrast, nonlinear IMS holds significant potential for diverse isotopic separations.

1.1 Instrumentation

Characterization of isotopically labeled sites involves instruments of different types. The principles and capabilities of each are outlined below.

1.1.1 Mass Spectrometer

A mass spectrometer separates and identifies charged species by the mass to charge ratio (*m/z*). The specific methods for separation and detection depend on the instrument type. Here, we

utilized a linear ion trap (LTQ XL, Thermo Scientific) with a split aperture/electrodynamic ion funnel interface and electrospray ionization (ESI) source.

In ESI, a sample solution is pushed through a narrow capillary ending in a sharp needle. The needle is biased at ~2.5 - 4 kV above the instrument inlet, which produces a Taylor Cone in front of the tip [1]. The droplets formed there shrink as the solvent evaporates, bringing the charges closer to each other. The increasing Coulomb repulsion separates the ions, creating microdroplets containing one ion each. These are further desolvated in the heated capillary or curtain plate/ orifice inlet with gas counter-flow and enter the MS instrument via an electrodynamic ion funnel.

The funnel comprises ring electrodes with gradually narrowing internal diameter and decreasing voltage along the axis. The ions are drawn in by the front dc potential below the MS inlet and gas flow due to the pressure dropping from ambient in the lab to vacuum inside the mass spectrometer. The dc voltage ladder with maximum on the front electrode and minimum on the back creates the potential gradient pulling ions through the funnel. In the previously standard skimmer-cone arrangement, ions leaving the MS inlet spread out, with most hitting the chamber walls. With the funnel, a radio frequency (rf) voltage applied to the alternating electrodes prevents that by pushing ions away from the walls. While nearly all ions thus exit the funnel into the first quadrupole, the neutrals entering the funnel are not so focused and are largely destroyed on the walls.

Per the name, the quadrupole consists of four identical equally spaced “poles”. While their theoretically ideal shape is hyperbolic, cylindrical and rectangular rods deliver reasonable performance at much lower cost. The identical rf voltages loaded on the opposite poles create the field guiding all ions through the quadrupole. Addition of dc voltages (positive to two opposing

poles and negative to the other two) turns a quadrupole into a filter that passes only ions with certain *m/z*. Thus mass-selected ions are injected into the quadrupole ion trap (Paul trap). This trap accumulates ions over a specified time, confining them with rf field, then ejects them to the detector in the order of increasing *m*/*z* upon appication of and raising DC voltage.

The detector is a dual electron multiplier, comprising a resistive conducting surface held at a voltage. Ions smashing into it eject electrons that start a cascade down the detector, with the signal continuously amplified and eventually turned into the current recorded by the instrument.

1.1.2 Ion Mobility Spectrometry

While MS measures the ratio of mass to charge (*m/z*), linear IMS measures that of ion-molecule collision cross section (*Ω*) to charge. The classic implementation of linear IMS is drift tube IMS. The *Ω/z* is found from the absolute mobility (*K*), defined as the ratio of ion drift velocity *v* to electric field *E* determined at moderate *E*, according to the Mason-Schamp equation [13]:

$$K = \frac{3q}{16N\Omega}\sqrt{\frac{2\pi mM}{(m+M)kT}} \tag{1.1}$$

where *m* and *q* are the ion mass and charge, *M* is the gas molecule mass, *T* and *N* are the gas temperature and number density, and *k* is the Boltzmann constant. Eq (1.1) does not feature the field strength, *E*, i.e., *K* is deemed independent of *E* [13]. The mobility depends on the ion mass per eq. (1.1), but only slightly except for the lightest ions with masses comparable to the gas molecule mass. Hence linear IMS can separate only non-isobaric isotopic ions, and only to a very limited extent. While ion mobilities depend on the gas temperature and pressure, those dependences for isotopic ions are (nearly) identical and do not allow improving isotopic separations. Indeed, the correlation between mass and **Ω** in linear IMS is especially limiting for most similar species such as isotopic ions. The measured cross-sections are informative of ion

structure and can be related to specific geometries by calculations of varied sophistication. This makes linear IMS useful for many analytical and structural applications, but not isotopic analyses.

The ion drift velocities become nonlinear with respect to E in strong fields [14]. This regime is enabled in the field asymmetric waveform IMS (FAIMS) where the increment of K between two E levels ("high" and "low") is extracted by an asymmetric voltage waveform (commonly with a bisinusoidal profile). The resolution rapidly improves with increasing peak field strength [14].

Since FAIMS is based on the field-dependent increment of ion mobility, it lacks the direct connection to Ω found in linear IMS. While rendering FAIMS ill-suited for structural elucidation, it weakens the connection to mass to make FAIMS more orthogonal to MS than linear IMS [14] and thus improve the overall separation. One can group ions by the trend of mobility with increasing E: increasing (type A), increasing and then decreasing (B), and decreasing (C).

High-resolution planar FAIMS analyzers feature two parallel mirror-like metal electrodes spaced by ceramic washers that control the gap width and allow varying it. To pass ions exiting the device to the MS instrument, those electrodes are biased relative to its inlet positively (for cations) or negatively (for anions). The periodic asymmetric voltage waveform and dc compensation voltage (CV) are loaded on the opposite electrodes, adding to the bias. The rf waveform causes ions to oscillate with net drift across the gap (toward either electrode) that would have caused all ions to be neutralized on the electrodes. For any particular species, application of specific CV can offset that drift and bring the species into equilibrium, permitting it to pass the gap and be detected in the ensuing MS stage. The FAIMS spectrum is collected by

scanning the CV, which reveals the full range of species present (Figure 1). Only one species is equilibrated, with others ideally filtered out.

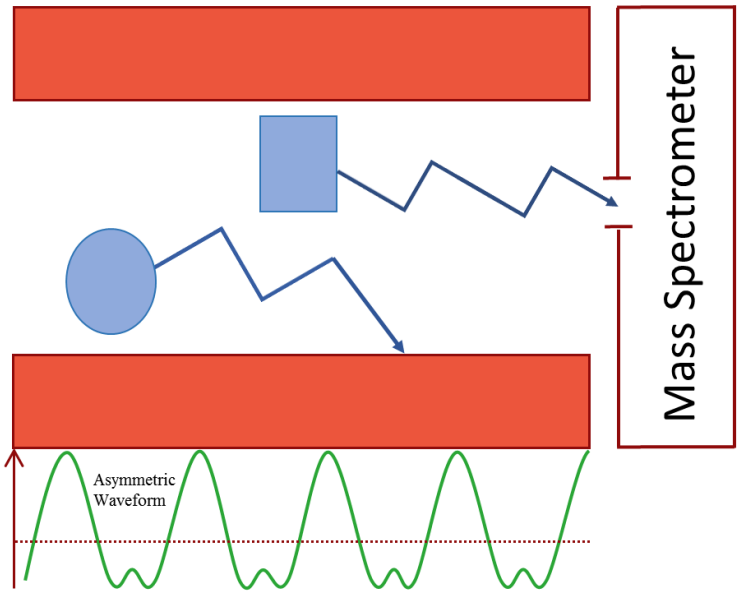

Figure 1: Visual representation of FAIMS (using a bisinusoidal waveform) at a particular CV.

Ions are carried through the gap by buffer gas flow, and the gas selection strongly affects the separation properties. For example, high molecular polarizability of CO_2 results in stronger attraction to ions with localized partial charges [15]. Addition of light gases (He or H_2) to heavy gases (N_2 or CO_2) generally improves resolution, in particular for peptides, proteins, and lipids [16 - 21]. This happens because lighter gases increase the ion mobility, while mixtures of light and heavy gases produce prominent non-Blanc effects (deviations from the Blanc's law stipulating that the mobility of any ion in a mixture of multiple gases is the weighted geometric average of those in individual components [22]). The resolution generally improves with increasing fraction of the lighter gas (up to ~80%). However, the threshold for avalanche electrical breakdown is much lower in H_2 and (especially) He than N_2 or CO_2, which limits the maximum content of the lighter gas. The breakdown depends on gas composition, gas pressure,

gap width, and field strength, creating a multi-dimensional "safe zone" for FAIMS operation. The device used here operates at ambient pressure, but the other three parameters are variable.

1.1.3 Nuclear Magnetic Resonance

A common technique for structural determination is NMR. Certain atomic nuclei, including ^{1}H and ^{13}C, have a spin number of ½ and thus interact with a magnetic field. This interaction depends on the bonding of atom involved and its position in the molecule (mainly the identity of neighboring nuclei). These factors combine to create unique spectra for different molecules, allowing *de novo* structural determination of unknown compounds and assignment of analytes to known ones. Most NMR analyses employ ^{1}H and ^{13}C, as other nuclei with half-spins are rare. The proton NMR spectra are quick to obtain (as nearly all hydrogens are ^{1}H), but usually most complicated (as organic compounds comprise numerous hydrogen atoms). The carbon NMR spectra are much simpler but, with only ~1% of C nuclei being ^{13}C, require longer acquisitions to accumulate enough signal. The 2-D NMR combining two spectral types offers the most complete data. Performing NMR requires samples on the mg scale [11], to compare with ng for IMS/MS.

1.2 Isomers and their Separation Using Ion Mobility

1.2.1 Terminology

Isomers are molecules with identical stoichiometry (and thus mass) but different structures. Most isomers vary in the atomic geometry, but isotopic isomers (isotopomers) feature labeled nuclei in different locations. There also are isotopologues that comprise different isotopes of same atom (in any position). Distinguishing the exact masses (values with the maximum measurable number of decimals) and nominal masses (obtained from integer masses for all atoms) is important in this context. Structural isomers obviously have equal exact and

nominal masses. Isotopologues generally have different nominal and exact masses, producing the isotopic envelopes seen in mass spectra. However, some with different isotopic substitutions ("pseudoisotopomers") are nominally isobaric with different exact masses. For example, the amino acid alanine ($NH_2CH_2CH_2CO_2H$) may feature ^{13}C instead of ^{12}C or ^{15}N instead of ^{14}N. With those atoms having slightly unequal mass defects, these pseudoisotopomers differ by 0.6 mDa [23]. The number of isotopologues and isotopomers for a compound depends on its stoichiometry, geometry, and number of stable isotopes for each constituent atom. While MS of sufficient resolving power can, in principle, disentangle all isotopologues, it cannot separate isotopomers.

1.2.2 Previous Isomer Separations

Present MS capabilities have been considered impossible not long time ago. While early mass spectrometers with unit resolution separated atomic isotopes, modern instruments (time-of-flight and Fourier-transform) with resolving power (R) of ~10^5 - 10^6 can separate isotopologues of macromolecules and nearly all pseudoisotopomers of smaller peptides [23]. These successes shifted the focus to isotopomers. Some can be distinguished by tandem MS methods such as collision-induced dissociation (CID) and electron transfer dissociation (ETD), but these approaches are hampered by uninformative fragments and spectral congestion. Complex mixtures require prior separation, but condensed-phase methods (chromatography and electrophoresis) have not been able to resolve isotopomers (although LC can resolve some non-isobaric isotopologues for labeled peptides [24-25]). While not a separation technique, NMR can distinguish some isotopomers - but only in simple mixtures [7, 26].

Linear IMS broadly separates structural isomers [27-28]. It should also be able to resolve non-isobaric isotopologues based on the reduced mass term of Mason-Schamp equation. That

was just demonstrated in experiment, but the capability rapidly diminishes for heavier ions: e.g., separating isotopologues that differ by 1 Da at $m > 100$ to even see the isotopic envelope requires $R > 10^3$ that is beyond realistic in the foreseeable future [7, 29].

Nonlinear IMS that may separate isotopologues by factors besides the mass term is not subject to same limitations. High orthogonality to MS has allowed FAIMS to broadly outperform linear IMS in isomer separations for lipids [16], peptides [17-19, 30], and proteins [20, 21]. This advantage extends to pseudoisotopomers and true isotopomers. Protonated alanines and glycines labeled by D, ^{13}C, or ^{15}N in one or two specific locations were partly separated by FAIMS in a manner not obviously correlated to the ion mass, but controlled by the label position [29]. In particular, the CV shifts were often smaller between non-isobaric isotopologues than isotopomers. Isotopologues also have slightly unequal bond lengths and angles, resulting in marginally differing cross sections. As one would expect from the largest isotope mass ratio, the greatest geometry difference normally results from the H/D substitution. However, that has not always caused the largest peak shifts in FAIMS, indicating that other phenomena are in play [29].

One key finding was that the unequal peak shifts due to two specific ^{13}C labels add to the shift upon same two ^{13}C labeled together. This additivity of shifts strongly implicates the move of ion center of mass within the geometry frame as the cause of observed effect. Indeed, labeling the C atom near amino acid nitrogen (i.e., the furthest C from ion center of mass) produced largest CV shifts [29]. In this work, the FAIMS separations of isotopomers and isotopologues will be greatly expanded to polypeptides and natural (not artificially labeled) species.

1.2.3 Linear Mobility and Intrinsic Size Parameters

The reliability of assignments for peptides in proteomic analyses using IMS/MS can be improved and false discovery rate decreased by using the constraint of measured mobilities. One approach is comparing those with the values previously measured with high confidence, but often no such data exist. Another way is to compare the measurements with calculated mobilities. The low-field mobilities of ions relevant to linear IMS can be calculated using *a priori* physical methods such as trajectory propagation [31], scattering on electron density isosurfaces [31, 32], and ultimately their hybrid [33]. These computationally intensive techniques are suitable for fundamental investigations of targeted species, but not global high-throughput proteomic analyses involving thousands of peptides. A cruder but tremendously faster practical alternative is the Intrinsic Size Parameter (ISP) approach, where the cross section is determined as a sum of contributions from constituents such as amino acid residues (with the contribution of each constituent proportional to unique ISP).

To match the central trend for peptides in IMS/MS space, the ISPs of amino acid residues are normalized to the mean of unity. (This ignores the fact that different amino acids have unequal masses and abundances across the proteome, which has little consequence as the abundance and ISP are not materially correlated). The ISPs for all residues were initially extracted by fitting the measured mobilities for diverse tryptic peptides [34]. However, the values calculated from the mean residue density inferred simply from its atomic composition were as accurate as those derived from extensive experimental statistics for 1+ ions [35]. This indicates generally compact close-packed peptide ion geometries with the ratio of cross section to mass controlled mainly by the density of all building blocks involved.

The initial ISP treatment was for unmodified peptides, but many proteins and thus proteolytic peptides feature post-translational modifications (PTMs). These PTMs often have great biomedical impact but the peptides bearing them are challenging to identify because of low abundance (largely due to sub-stoichiometric site occupation) and difficulty of MS/MS analysis (due to PTM lability). Hence additional information from IMS data would particularly help in this context. Some PTMs, most prominently phosphorylation, broadly shift peptides away from the central IMS/MS trend: the cross sections of monophosphorylated peptides are statistically smaller than those of isobaric unmodified peptides by ~3% [36, 37]. One can account for that by ascribing ISPs to PTMs. Then the ISP of phosphorylation must be below one - the regression to experimental data yields 0.64 ± 0.05 [38]. Other ISPs extracted from experiments were for carboxyamido-methylation (0.92 ± 0.04) and palmitoylation (1.26 ± 0.04) [39]. Here we explore whether the ISPs for PTMs can be derived similarly to the values for amino acid residues - from the densities of constituent atoms.

1.2.4 Research Projects

This thesis comprises three distinct projects. The first is FAIMS separation of isotopomeric ions labeled by stable isotopic atoms at specific sites, specifically for small peptides. The second investigates the isotopologic shifts in FAIMS spectra (specifically, for monohalogenated anilines) to identify structural isomers by fingerprint isotopic shift patterns in a manner reminiscent to NMR. The third project examines the capability to predict the effect of post-translational modifications on peptides on their absolute mobilities (measured in linear IMS) via simple first-principle calculations employing the concept of intrinsic size parameters.

CHAPTER 2

ARTIFICIAL LABELING

2.1 Experimental

The high resolution FAIMS device is paired to the Thermo LTQ XL ion trap modified with a slit aperture interface with an ion funnel courtesy of Heartland Mobility (Wichita KS, USA). The gap width (g) on the FAIMS is adjustable by varying the thickness of the ceramic washers at 1.25 mm or 1.88 mm. This allows for adjustment of the field strength. The high definition waveform generator (Heartland Mobility, Wichita, KS USA) provides an adjustable waveform for another option for varying field intensity. The fields used included: 4.00 kV waveform with g = 1.88 mm making a dispersion field (E_D) of 21.3 kV/cm "moderate field", 3.2 kV with g = 1.25 mm (E_D = 25.6 kV/cm) "high field", and 3.94 kV with g = 1.25 mm (E_D = 31.5 kV/cm) "maximum field."

The buffer gas was formulated by digital flow meters (MKS Instruments) controlled from a PC and delivered at a rate of 1.1-3.5 L/min to balance resolution and sensitivity. With electrical breakdown of helium being the limiting factor for different fields of He/N_2 and He/CO_2 mixtures: 70 % He was the maximum for moderate field, ~35% for high field, and no helium could be added to maximum field.

The regular dialanine (AA 160.1 Da) and analogs labeled by $\{^{12}C_3;^{15}N\}$ (mass increment = 4 Da) on the first residue (A*A) and second residue (AA*) along with regular trialanine (AAA, 231. Da) and analogs with an identical label on the first residue (A*AA), second residue (AA*A), and third residue (AAA*) were obtained from Thermo Fisher (Waltham, MA). The ~0.2 mM solutions of standard (unlabeled analog) with labeled dialanine (A*A/AA*) were in a

50:49:1 methanol/water/acetic acid solvent. These solutions were infused at 0.4 µL/min to the electrospray source (as protonated ions), emitter biased ~2.5 kV above the FAIMS inlet.

The mass spectra for the different compounds were dominated by the protonated analogs. Figure 2 has example spectra for both di- (a) and tri- (b) alanines as unlabeled along with a labeled analog. A notable peak in these spectra is the +1 Da isotopologue caused mostly by ^{13}C o

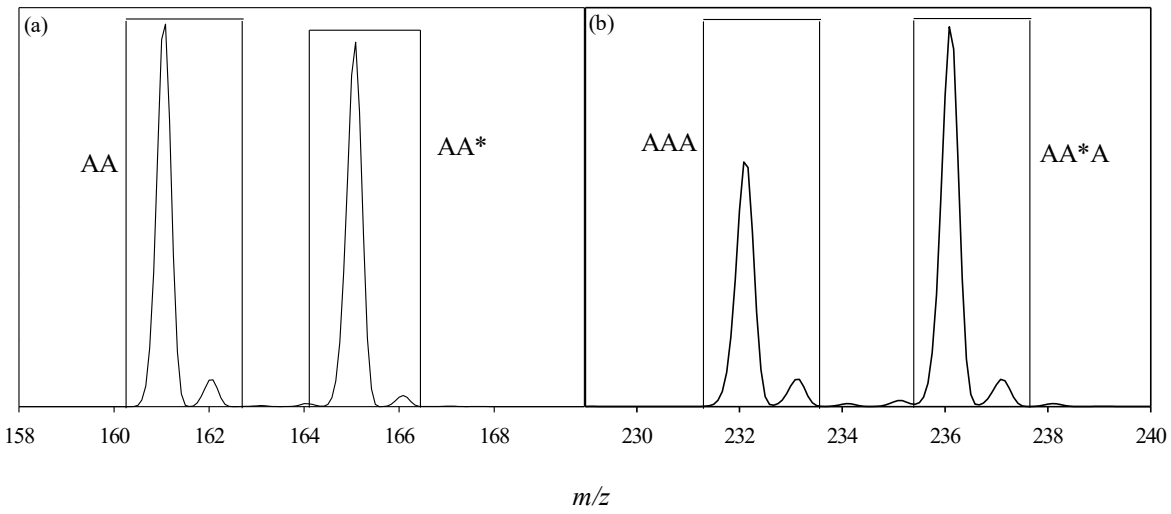

Figure 2: Example spectra of (a) dialanine (AA unlableled and AA* labeled) and (b) trialanine (AAA unlabeled and AA*A labeled).

The FAIMS spectra were acquired by scanning the CV at the speed of 0.3-1 V/min. Labeled analogs were measured separately with their corresponding unlabeled analog for an internal standard. This allowed for the measurement of CV shifts with the precision of ~0.01V (per replicate statistics).

For consistency across systems CV is expressed as compensation field E_C, set as positive when the voltage is below the bias. Isotopic shifts (E_{iso}) are measured as the difference between the labeled and unlabeled analog.

$$\Delta E_{iso} = E_C(labeled) - E_C(unlabeled) \quad (2.1)$$

The isotopomeric shift (S_{iso}) is defined as the difference in isotopic shifts.

$$S_{iso} = \Delta E_{iso}(AA^*/A^*AA) - \Delta E_{iso}(A^*A/A^*AA) \quad (2.2)$$

2.2 Dialanine Results

Dialanine initially expressed itself as a type A ion with E_C being negative and becoming more negative at higher E_D (Figure 3). The addition of helium changed the interaction of the ions. The low polarizability of helium compared with N_2 or CO_2 makes for a shallower interaction potential with any ion and thus harder scattering, so adding helium eventually converts [40,41] ions to type C with E_C>0 and increases at higher E_D. At moderate E_D, the E_C turns slightly more negative between 0-20% helium because of prominent non-Blanc Effects [29]. These non-Blanc effects scale with more dissimilar components thus are greater for He/CO_2 than He/.

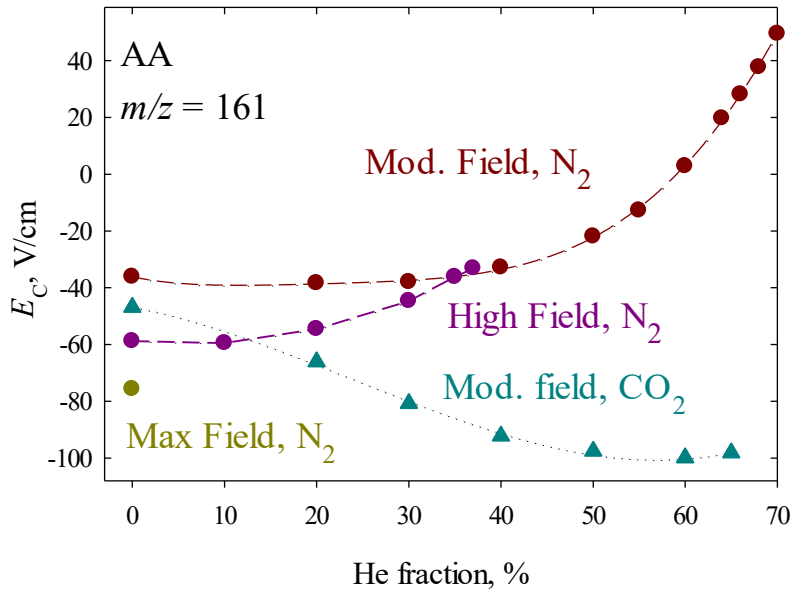

Figure 3: Tracking elution position of unlabeled dialanine through different gases (circles-N_2 triangles-CO_2) and different field intensities.

In He/N_2 mixtures at moderate field, the shift for A*A is positive (less negative E_C against the standard AA) and substantial (1.4-2.1 V/cm) at all helium fractions with a maximum

at 55% helium. The shift for AA* is small nudging from 0.5 V/cm in N_2 to 0 at 60% and -0.3 V/cm at 65 % helium. Accordingly the isotopomeric shift increased from 1.0 V/cm in N_2 to 2 V/cm at 65 % helium

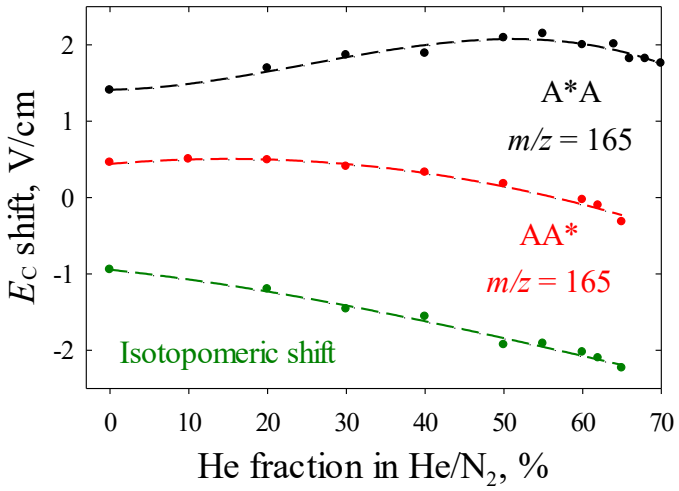

Figure 4: Moderate field isotopic and isotopomeric shifts for the labeled dialanines with He/N_2. Lines are the quadratic regressions through each curve.

Indeed at the same E_D the isotopic shifts for first (^{13}C1) isotopomers of glycine and alanine were positive with a slight maximum (~1 V/cm) near 50% helium whereas those for the second (^{13}C2 for G; ^{13}C3 for alanine) isotopomers were near zero, gradually decreasing from 0.3 V/cm in N_2 to -0.4 V/cm at 65% helium (all values for glycine) [29]. Thus S_{iso} values were similarly negative, with the magnitude increasing from 0.6 V/cm in N_2 to 1.3 V/cm at 60% helium. The ΔE_{iso} value for N-terminally labeled species is about doubled here, which may reflect a greater transposition of the ion's center of mass due to higher total mass of the label (4 vs. 1 Da) and perhaps its more distant move in an ion of larger dimensions compared to the shifts of singular alanine and glycine. This translates into a greater S_{iso} that improves the isotopomer resolution. As is common [18,29], the optimum (though still incomplete) separation is achieved

somewhat below the maximum helium fraction because the signal declines with rising helium content.

The He/N_2 and He/CO_2 mixtures elicited the opposite peak order for many glycine and alanine isotopologues (including isobaric ones) [29], but not isotopomers labeled by $^{13}C1$ and $^{13}C3$. Here, the isotopic shifts are almost switched: ΔE_{iso} for A*A is near-zero increasing from -0.2 V/cm in CO_2 to 0.7 V/cm at 65% He, and ΔE_{iso} for AA* is positive and substantial with a maximum of 1.7 V/cm at ~55% helium (Figure 5). The isotopomeric shift thus amounts to 0.7-1.5 V/cm, maximizing

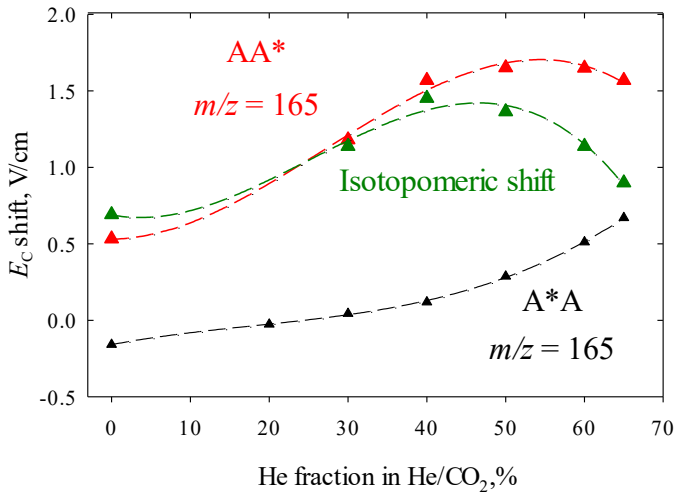

Figure 5: Moderate field isotopic and isotopomeric shifts for the labeled dialanines with He/CO_2. Lines are the quadratic regressions through each curve.

The peaks are wider in He/CO_2 than in He/N_2 at equal helium percentage (Figure 6). This is because the peak widths in planar FAIMS devices scale at $\sim K^{1/2}$, and all ions are less mobile in CO_2 than N_2 with smaller and lighter molecules. In Figure 6, the peak widths in He/N_2 and He/CO_2 converge at higher helium percent as helium becomes the dominant gas. The labeled dialanine's peak widths were averaged since the label should cause the same change in peak width regardless of position of the label.

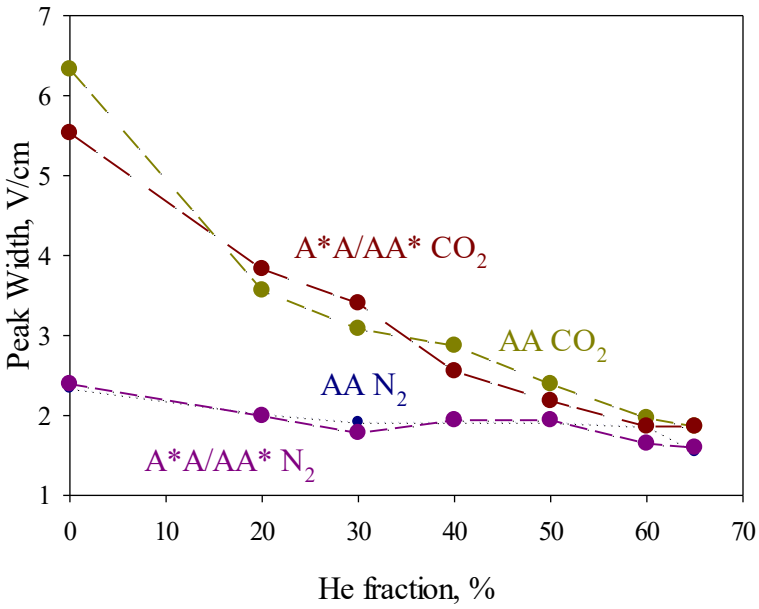

Figure 6: Peak widths of dialanine at half max in different gas mixtures. Labeled isotopomers averaged to one value for the individual gas mixtures.

A lower S_{iso} and broader peaks with He/CO_2 versus those with He/N_2 create worse separation, thus we did not pursue the analysis in He/CO_2 mixtures in detail. Though this data demonstrates the likely utility of varying the gas composition for isotopomer separations. Figure 7 simulates the best separation achieved at moderate field in He/N_2 at 60% helium with the flow rate at 2 L/min, indicating where A*A and AA* would be found in relation to one another based upon their locations relative to AA.

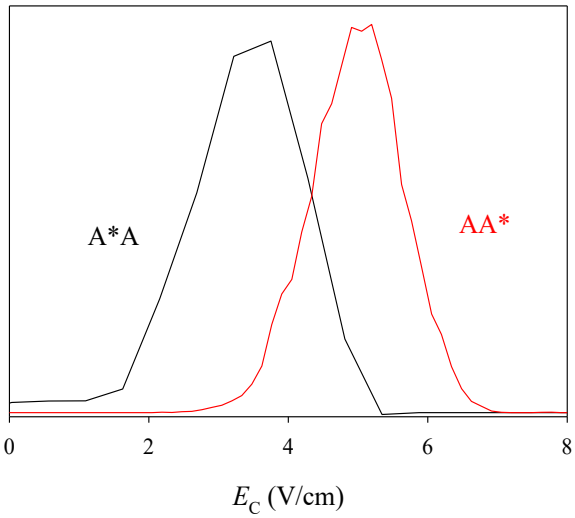

Figure 7: Moderate field simulated max separation at 60 % He/N_2.

Our standards were not isotopically depleted at unlabeled atoms, as noted in Figure 2, and thus include isotopologues conforming to the stoichiometry ($C_6H_{12}N_2O_3$) and natural isotopic abundances for unlabeled atoms of constituent elements: 0.011% (D), 1.07 % (^{13}C), 0.37 % (^{15}N), 0.05% (^{17}O), and 0.20% (^{18}O). Hence, nearly all species at 1 Da increment (96% for AA and 92% for A*A and AA*) are due to the doping by ^{13}C or ^{15}N that compose the labels in A*A and AA* (Table 1).

Table 1

Fractions of heavy isotopologues with +1 mass increment in unlabeled (top) and labeled (bottom) dialanine ions.

Element	Fraction of isotope with +1 mass increment	Atoms in H^+AA	Fraction of atoms with +1 mass increment	Contribution to ions with +1 increment
D	0.011%	13	0.14%	1.9%
C	1.07%	6	6.42%	86.1%
N	0.37%	2	0.74%	9.9%
O	0.04%	4	0.16%	2.1%
All		25	7.46%	100%

Table 1 (continued)

Element	Fraction of isotope with +1 mass increment	Atoms in H^+A^*A or H^+AA^*	Fraction of atoms with +1 mass increment	Contribution to ions with +1 increment
D	0.011%	13	0.14%	3.6%
C	1.07%	3	3.21%	82.7%
N	0.37%	1	0.37%	9.5%
O	0.04%	4	0.16%	4.1%
All		21	3.88%	100%

One of these isotopes replacing one ^{12}C or ^{14}N in alanine ($NH_2CH_2CH_2CO_2H$) creates eight possible +1 Da isotopomers for AA at the nominal 161 Da mass and for A*A or AA* at 165 Da. The intensities of MS peaks at *m/z* 162.1 (vs 161.1) and 166.1 (versus 165.1) in Figure 1 match the calculated 7% and 4% from Table 1. The absolute E_C shifts for the features at *m/z* 166.1 versus those from the same sample at *m/z* 165.2 (Figure 8a) are less than Figure 4, which is reasonable considering the 4 fold smaller mass increment. The shifts for A*A and AA* differ as well: the first decreases from 0.2 V/cm in N_2 to 0 at 65% helium, whereas the second increas the AA fea ting to 0.2 '

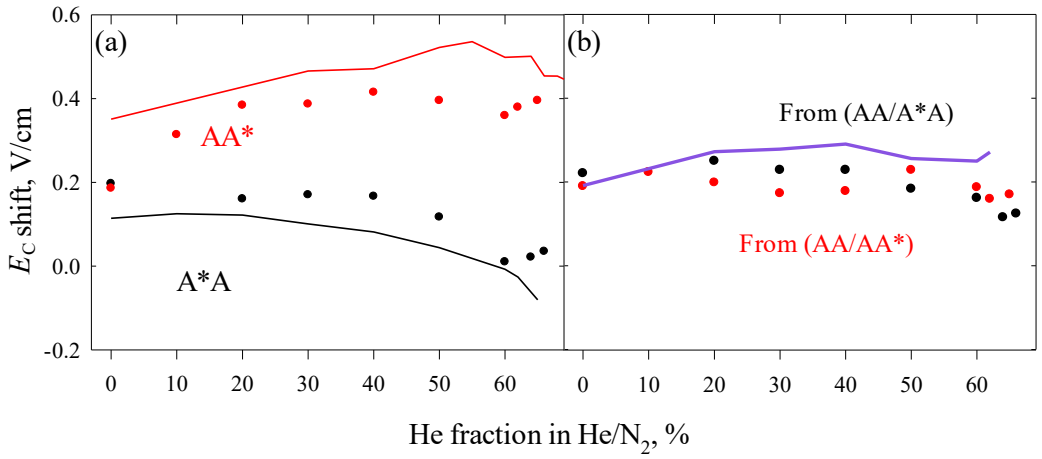

Figure 8: Isotopic shifts for A*A and AA* (a) and for AA (b) with measured (points) and modeled values (lines).

The features at *m/z* = 166.1 are not split or much broadened: at 50% helium, the mean peak width exceeds those at *m/z* = 161.1 or 165.1 by ~15% (Table 2). Hence all fairly abundant isotopomers have close E_C values. Then, if the isotopic shifts are additive, ΔE_{iso} from ^{13}C or ^{15}N in either residue would equal 1/4 of ΔE_{iso} from the $\{^{13}C_3;^{15}N\}$ block in the *same* residue. The curve thus obtained from experimental ΔE_{iso} for A*A and AA* at *m/z* = 165.1 clearly tracks the observations for, respectively, AA* and A*A at *m/z* 166.1 (Figure 8a). The small discrepancies are within the combined uncertainty of two ΔE_{iso} measurements and the error margin of our approximation of equal shifts from all ^{13}C and ^{15}N labels within a residue, not even considering the currently unquantifiable contribution of up to 8% of isotopes at *m/z* = 166.1 comprising ^{17}O or D (Table 1). The ensemble of AA isotopomers at *m/z* = 162.1 is a 1:1 mixture of the pools doped on the first and second residues. Then the mean ΔE_{iso} for AA should equal the average of E_C shifts for A*A and AA* in Figure 8a, in line with the data (Figure 8b). These findings uphold a wide additivity of isotopic shifts.

Table 2

Peak positions ($-E_C$), shifts, and widths (w), all in V/cm, for protonated AA and AA* species in 15 spectral replicates measured at the moderate field in 1:1 He/N_2 buffer with the flow rate of 2 L/min. The mean shifts and widths with error margins (at 95% confidence) are in the last column. Similar data were obtained at 60% He.

AA (*m/z* = 161.1)		A*A (*m/z* = 165.1)		A*A (*m/z* = 166.1)		Isotopic E_C shifts	
E_C	w	E_C	w	E_C	w	(165.1–161.1)	(166.1–165.1)
21.67	1.99	21.51	2.01	21.08	2.31	0.16	0.43
21.68	1.99	21.51	2.03	21.13	2.26	0.17	0.38
21.68	2.02	21.51	2.07	21.12	2.39	0.18	0.39
21.70	1.98	21.56	1.99	21.16	2.23	0.14	0.40
21.71	1.99	21.47	1.86	21.10	2.23	0.24	0.37
21.84	2.02	21.68	2.01	21.34	2.27	0.16	0.34
21.88	1.95	21.65	1.95	21.27	2.39	0.23	0.38
21.91	2.06	21.73	1.98	21.34	2.25	0.18	0.39
21.96	2.01	21.77	2.01	21.31	2.14	0.19	0.45
21.97	1.97	21.79	2.03	21.37	2.43	0.18	0.42
22.03	1.98	21.87	1.99	21.45	2.35	0.15	0.43
22.07	1.97	21.93	2.03	21.49	2.31	0.14	0.44

Table 2 (continued)

AA (m/z = 161.1)		A*A (m/z = 165.1)		A*A (m/z = 166.1)		Isotopic E_C shifts	
E_C	w	E_C	w	E_C	w	(165.1–161.1)	(166.1–165.1)
22.26	2.03	22.08	1.96	21.72	2.31	0.18	0.36
22.52	2.07	22.33	2.04	21.91	2.42	0.19	0.42
22.96	1.91	22.77	1.95	22.41	2.20	0.18	0.36
	2.00 ± 0.02		1.99 ± 0.03		2.30 ± 0.05	0.18 ± 0.015	0.40 ± 0.019

Table 2 is an example set of replicate statistics to ensure accurate data. 15 replicates at 2 L/min in a 50% He/N_2 were obtained, noting positions and peak widths of AA, AA*, and AA* +1 Da errors for widths and shifts are at 95% confidence level on the bottom row.

For labeled G and A raising E_D to 28.7 kV/cm in He/N_2 buffers retained the CV spectra overall while expanding isotopic shifts (despite lower helium content limited by electrical breakdown) [29]. The same happens with dialanine at E_D = 25.6 kV/cm "high field". ΔE_{iso} goes up to 3.3 V/cm for A*A and 0.8 V/cm for AA* (m/z = 165.1), with the absolute S_{iso} growing to ~2-3 V/cm (Figure 9).

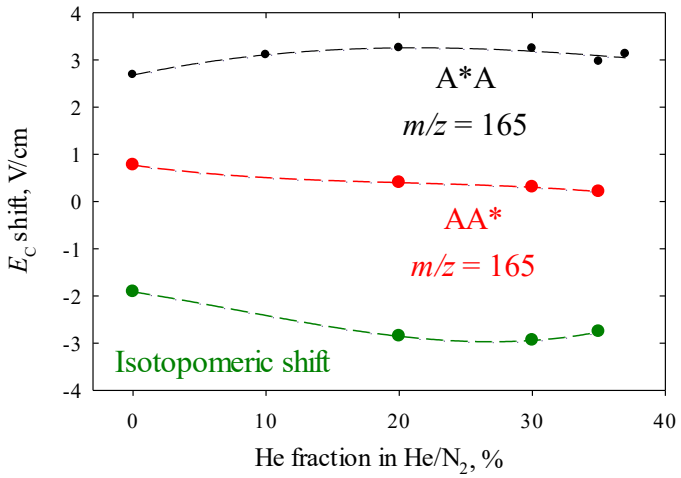

Figure 9: High field isotopologic shifts for labeled dialanine against unlabeled dialanine.

The extra shifts for A*A and AA* of *m/z* 166.1 similarly expand up to 0.7 V/cm (for AA*) and roughly match the values derived from measuring ΔE_{iso} at *m/z* 165.1 as described earlier (Figure 10), wi

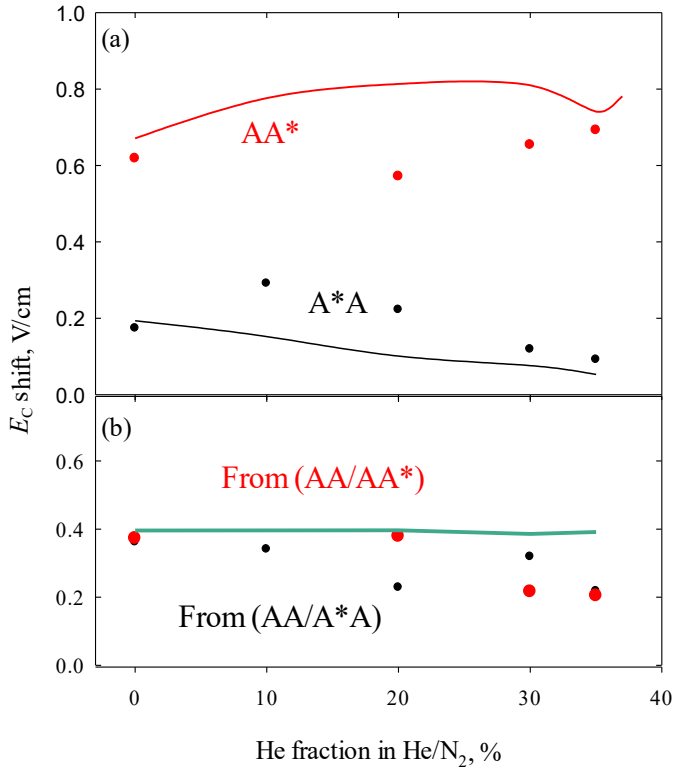

Figure 10: In high field, the +1 shifts of the labeled dialanines from experiment (dots) and modeled (lines) (a) and the +1 shifts of unlabeled dialanine from both mixtures (b).

At maximum field (E_D = 31.5 kV/cm) only one point was gathered since no helium could be added due to electrical breakdown. The isotopic shift for A*A reached 4.2 V/cm, where as that for AA* remains at 0.7 V/cm; thus absolute S_{iso} expands to 3.5 V/cm. The regime has provided the best (better than half maximum) resolution of A*A and AA*, enabling perfect filtering of either at its peak apex even at different ion ratios including 1:1 (Figure 11a) and 1:2 (Figure 11b).

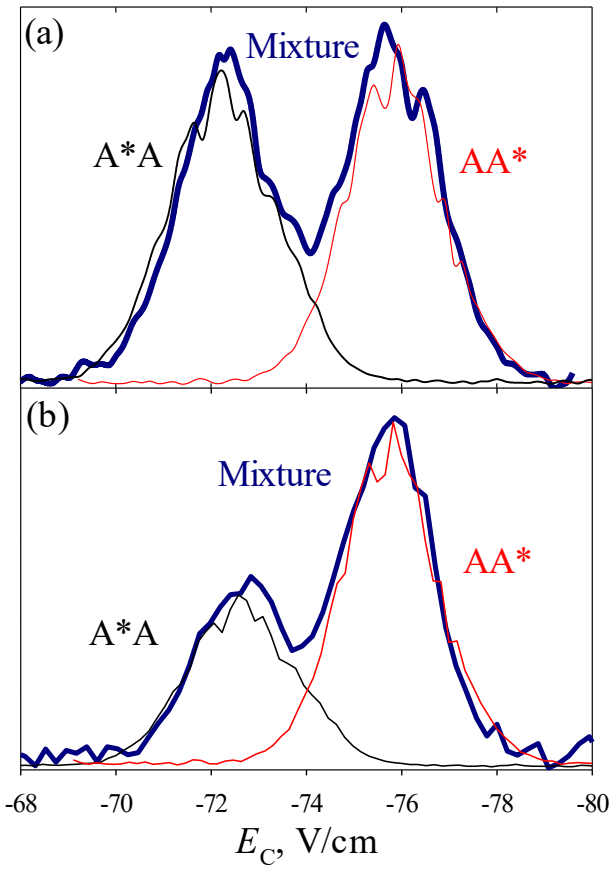

Figure 11: Maximum field separation of isotopomers AA and AA* at different ion intensities 1:1 (a) and 1:2 (b).

The E_C shifts at *m/z* 166.1 also increase to 0.4 V/cm (A*A) and 0.8 V/cm (AA*), whereas ΔE_{iso} for AA at *m/z* = 162.1 increase to 0.5 V/cm. The relative isotopic shift for AA* and A*A at *m/z* = 166.1 versus 165.1 is of opposite sign compared with S_{iso} at *m/z* 165.1. This shrinks the isotopomeric shifts at *m/z* = 166.1 at any E_D hence resolving A*A and AA* at this mass was not sought. This makes sense given that random doping by natural isotopes blunts the distinction between deliberately labeled species.

2.3 Trialanine Results

The unlabeled analog of trialanine presented itself only as a type A ion at all fields with its positive E_C increasing with greater helium in nitrogen. At moderate field in carbon dioxide

trialanine was a C type ion with the E_C becoming slightly more negative; then with greater helium percentages it slowly started to creep back to zero. Figure 12 offers a visual representation of t ttings.

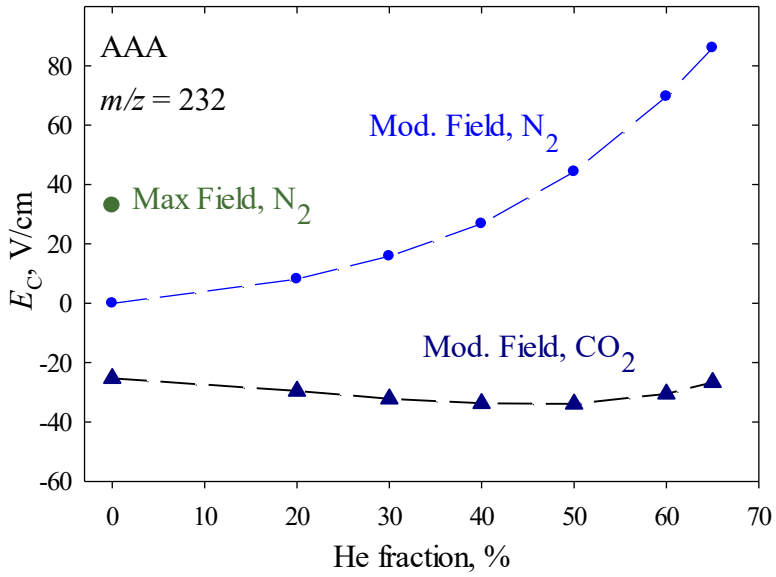

Figure 12: Unlabeled trialanine at various field strengths and gas compositions.

The separation of the labeled trialanine and the unlabeled trialanine in moderate field and nitrogen gas is shown in Figure 13. A*AA has the largest separation with as much as 0.78 V/cm between the labeled and unlabeled analogs. AA*A and AAA* have virtually overlapping positions in comparison to AAA at all percentages of helium, varying at most by 0.05 V/cm at 40% He. The isotopomeric shift calculated as the shift between AAA* and A*AA has a consistent difference between isomers after the initial addition of helium of around -0.5 V/cm.

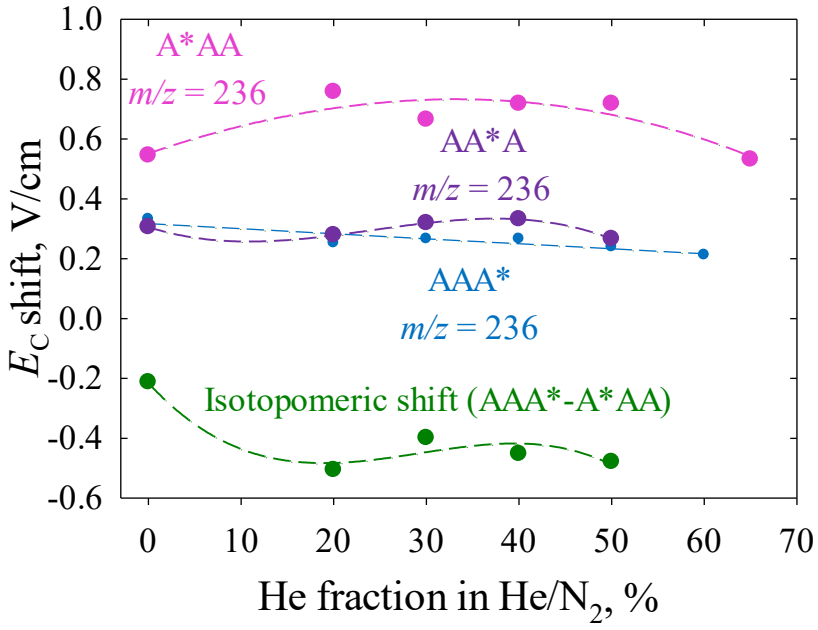

Figure 13: Moderate field isotopologic shifts and isotopomeric shift for labeled trialanine in nitrogen base gas.

After switching the gas to carbon dioxide, there were very similar results to dialanine, where there is little differentiation between any of the isotopologues at moderate field. Figure 14 shows the near uniform increasing of distance between the labeled trialanine and the unlabeled analog. This uniform increase means the isotopomeric shift between isomers maxes out near 0.2 V/cm which in an absolute sense is less than half of the maximum separation achieved in nitrogen, furthering the decision to not continue exploring carbon dioxide in this project.

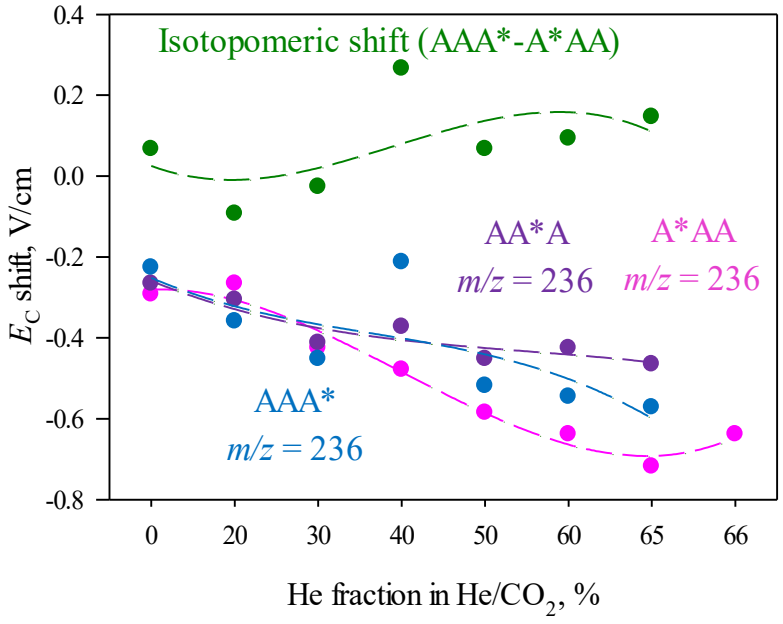

Figure 14: Moderate field isotopologic shifts and isotopomeric shift for labeled trialanine in carbon dioxide base gas.

Following the path in the dialanines the next step was to increase the field intensity. With dialanine maximum field had the best separation; this was tried for trialanine as well. This had similar results to the maximum field dialanine in that the isotopomeric shift reported increased to 1.9 V/cm. However it remained consistent with the moderate field results such that AA*A and AAA* still overlapped while A*AA separated by nearly 2 V/cm. Figure 15 is a simulated FAIMS spectrum for each species showing the overlap and the separation for all three isotopomeric species.

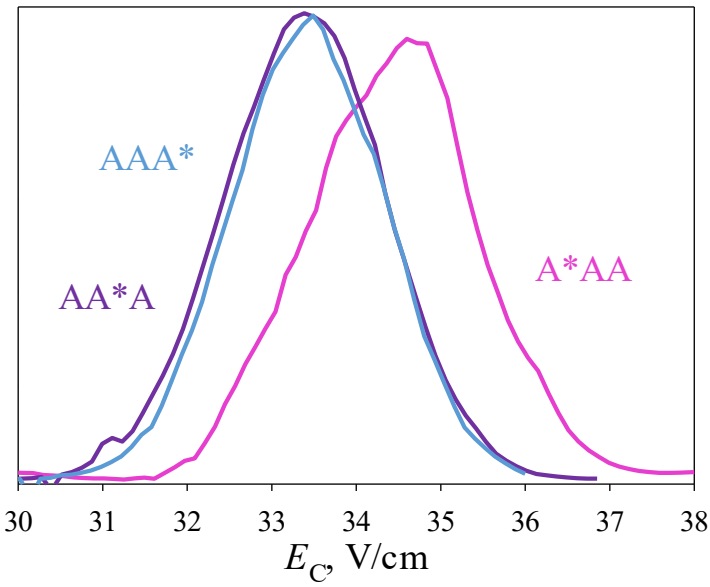

Figure 15: Simulated FAIMS spectra for all labeled isotopomeric species at maximum field in pure nitrogen buffer gas.

2.4 Conclusions

High definition FAIMS can resolve isotopomers for species over 100 Da, such as peptides with labeled amino acids. The optimum condition is near the breakdown threshold in N_2, which differs from the He/N_2 mixtures with 50-75% helium that worked well thus far [16-20]. This indicates testing the new regime in other analyses such as proteomic and metabolmic practices. This separation in He/CO_2 buffers is mostly opposite to that in He/N_2, showcasing the importance of gas properties and suggesting the exploration of other gases. The isotopomer separation is better for dipeptides than for labeled single amino acids, suggesting a substantial and perhaps greater effect for larger biomolecules more relevant to topical applications. The limited affect in trialanine suggests that more exploration is needed to understand why this effect is limited to the amine of the amino acid. We also observed the previously unknown shifts for species with natural and hybrid natural/labeled isotopic atoms that expand the analytical scope of

approach. These shifts apparently exhibit the pair wise additivity noted for labeled amino acids [29]. This shows the broad validity that should help rationalize the isotopic effects in FAIMS.

Trialanine did not completely resolve for its isotopomers, though this effect is not unexpected. Following the trend in dialanine it was expected that AAA* would move the least, partly because the heaviest end is weighted more and increasing the heaviest end does little to affect the ion in its travel through the gas and the electric field. When AA*A also did not alter much in its movement, it is presumed that adding the extra weight in the middle of the peptide did little to change its own reactions to the gas and electric field. Weighting the light end offered a far more drastic change. This leads to the idea that the center of mass shifts could be the indicating factor in how much difference there will be between isotopomers and isotopologues.

CHAPTER 3

NATURAL ISOTOPES

3.1 Experimental

The high resolution FAIMS device is paired to the Thermo LTQ XL ion trap modified with a slit aperture interface with an ion funnel courtesy of Heartland Mobility (Wichita, KS USA). The gap width set at 1.88 mm. The waveform had two levels of strength 3.75 kV and 4.00 kV, making two field strengths-minimum (E_D = 20.0 V/cm) and moderate (E_D = 21.3 V/cm), respectively. Gas combinations consisted of He/N_2 and He/CO_2 run at 3 L/min configured by digital flow meters (MKS Instruments) controlled by a PC. Electrical breakdown limited He percentage to 65% in N_2 and 63% in CO_2.

Samples were mono-halogenated anilines with chlorine (m/z = 127.6) and bromine (m/z = 172.1), MCA and MBA respectively, as the selected halogens. They were selected for their natural isotopic abundances. Chlorine has two isotopes ^{35}Cl (76%) and ^{37}Cl (24%). Bromine also has two isotopes ^{79}Br (51%) and ^{81}Br (49%). Positions on the ring included 2, 3, and 4 (5 matches position 3 and 6 matches position 2) (Figure 16). For each main halogen peak there is a +1 Da isotopologue, of which over 93% comes from a ^{13}C in various positions. Table 3 presents the percentages for each atom and its contributions to the +1 Da isotopologues.

Table 3

Fractions of heavy isotopologues with +1 mass increment in MCA (top) and MBA (bottom).

Element	Fraction of isotope with +1 mass increment	Fraction of atoms with +1 increment	Contribution to ions with +1 increment
D	0.01%	0.08%	1.12%
C	1.07%	6.42%	93.49%
N	0.37%	0.37%	5.39%
Cl	0%	0.00%	0.00%
All		6.87%	100.00%

Table 3 (continued)

Element	Fraction of isotope with +1 mass increment	Fraction of atoms with +1 increment	Contribution to ions with +1 increment
D	0.01%	0.08%	1.12%
C	1.07%	6.42%	93.49%
N	0.37%	0.37%	5.39%
Br	0%	0.00%	0.00%
All		6.87%	100.00%

There are no noticeable +2 Da contributions from atoms other than the halogens. Even two ^{13}C atoms on the same molecule have a 0.3 % occurrence making the contributions negligible overall. The standard for anchoring throughout the experiment was 2,5 di-choloraniline (m/z = 162.02). Samples were solvated at 100 μM in 99:1 methanol and formic acid.

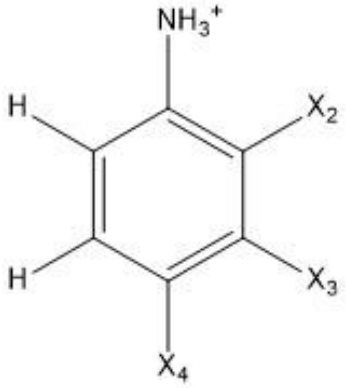

Figure 16: Illustration of a protonated halogenated aniline molecule with the halogens at specified positions.

The mass spectra for each compound obtained matched the isotopic pattern expected at the masses of interest. Figure 17 has spectra for 2 MCA (a) and 2 MBA (b) for visibility of each important isotopologue peak. For each FAIMS spectrum obtained, each peak was measured at 25, 50, and 75% intensity (and averaged) to ensure accurate peak widths and peak positions. Separate FAIMS peaks were extracted for each isotopologue and the shifts between isotopologues were obtained. Isotopologue shifts fall into three categories: halogen to halogen isotope (base to +2 Da), halogen to halogen +1 Da isotope (base to +1 Da or +2 Da to +3 Da),

and "carbon to carbon isotope" (+1 Da to +3 Da). Replicate analysis used the modified Tau test [42] to eliminate outliers, with all error bars indicated a standard error of mean for remaining values.

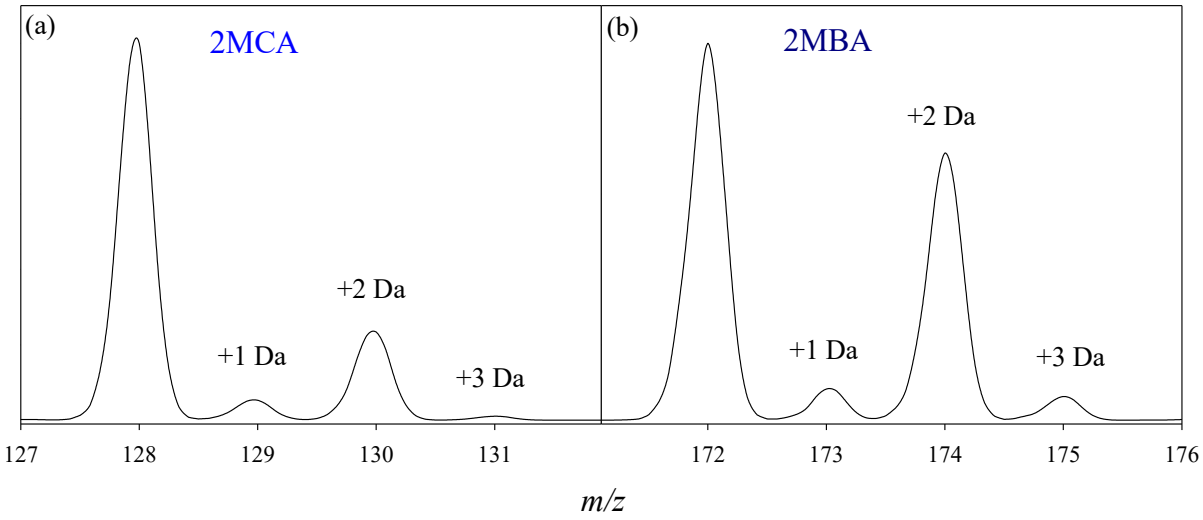

Figure 17: Mass spectra for 2MCA (a) and 2MBA (b) showing intensities for each relevant peak.

3.2 Monochloroaniline Results

In N_2 at minimum field, MCA's presented themselves as type A with increasing helium becoming more positive. All MCAs separated with as little as 20% helium with -10 V/cm separation between isomers 2MCA and 3MCA as well as between 3MCA and 4MCA at 30%. This separation continues to increase for 3MCA vs 4MCA, but shrinks for 2MCA vs 3MCA with increasing helium percentages until an overlap at 60% helium (Figure 18).

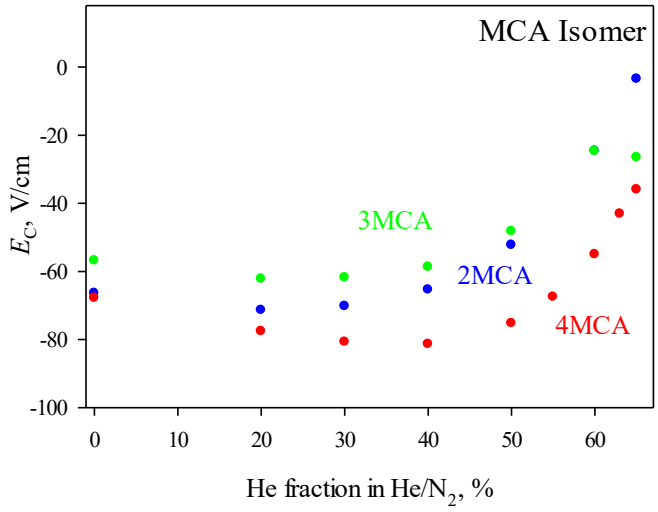

Figure 18: Monochloroaniline isomer FAIMS localization in various helium percentages in the base gas nitrogen at minimum field.

The internal shifts between various isotopologues presented small differences between isomers. While the base peak to the +1 Da shift (128-129 *m/z*) has some scattering present, the scaling of the shifts ranged only from 0 to 0.2 V/cm. The highest shift between any isotopologue occurred between the base peak and the +3 Da peak (128-131 *m/z*) at 0.95 V/cm at 65% helium. While some differences can be seen between different isomers the intensities of the shifts are not suitable for identification of individual isomers.

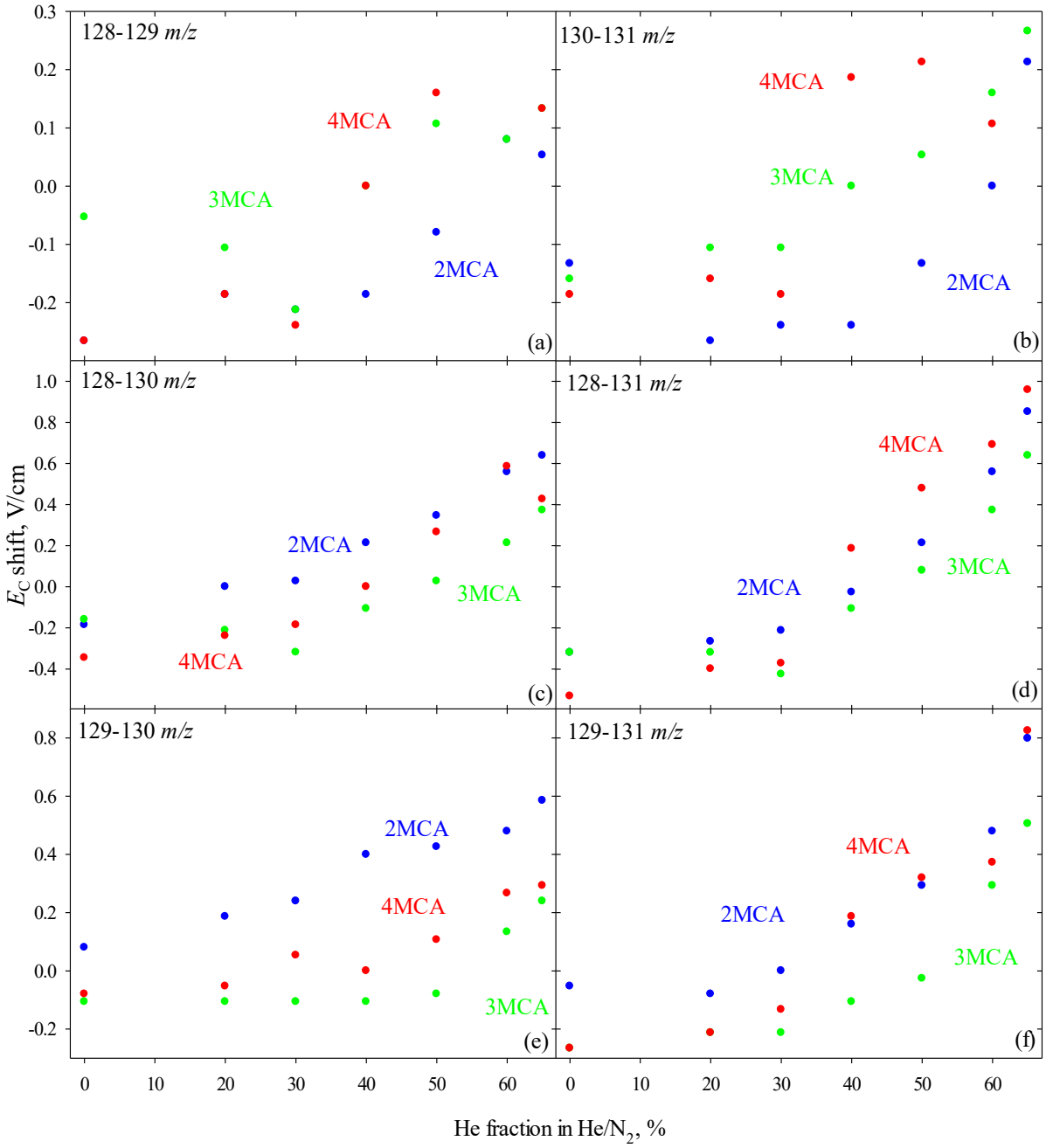

Figure 19: Various shifts between isotopologues of MCA in He/N_2 buffer gas at minimum field.

The CO_2 data presented itself in a different manner. At both minimum and moderate fields, the isomers of MCA presented as C type ions throughout all helium percentages. The isomers separated completely starting at 50% in minimum field and at 60% for moderate field (Figure 20). Separation of isomers was by as much as 12 V/cm at 50% He in minimum field

already larger than the separation in nitrogen at the same field intensity. This maximizes at 15 V/cm at the settings of 63% helium in carbon dioxide at moderate field. Isotopologue shifts were very different from N_2 at both field intensities.

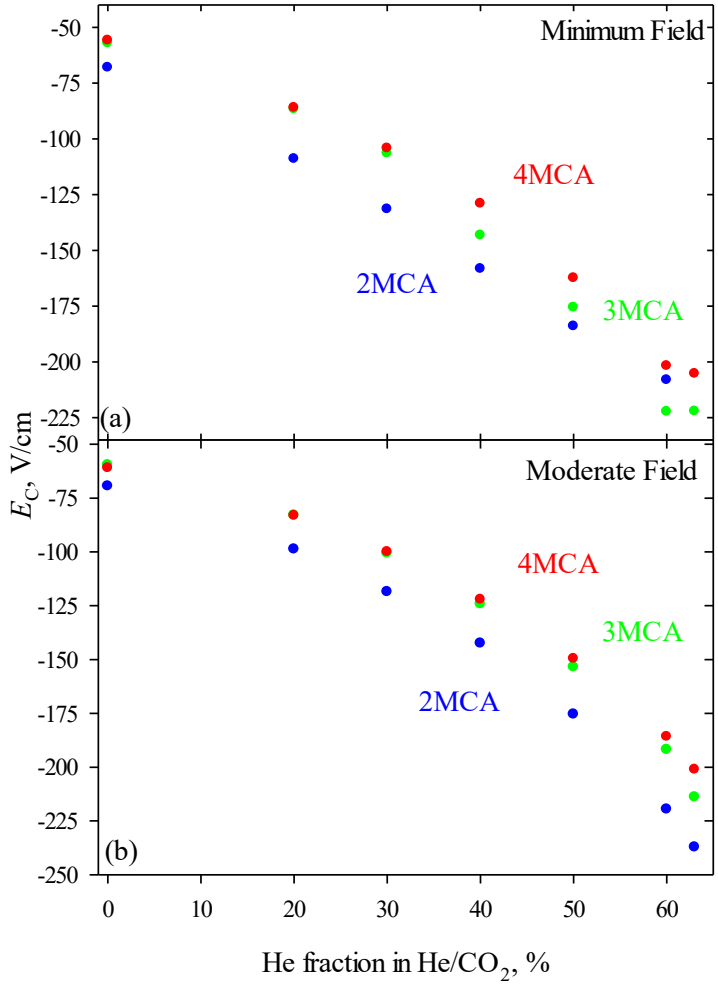

Figure 20: Minimum field (a) and moderate field (b) absolute position of the three monochloroaniline isomers.

Starting with minimum field, the shifts between the base peak and +1 Da (128-129 *m/z*) at 40% helium the internal shift separated by 0.5 V/cm for each isomer (Figure 21a). 2 MCA started at - 0.6 V/cm and increased to -1.4 V/cm at 50% before decreasing to -1.2 V/cm at 63% helium. 3 MCA started at -0.5 V/cm, increasing to -1.4 V/cm at 50% before decreasing to -1.2 V/cm matching 2 MCA. 4 MCA had the smallest relative shift starting at -0.1 V/cm increasing to

-0.8 V/cm at 60% but decreasing again to -0.6 V/cm at 63%. The chlorine isotopes shifts (128-130 *m/z*) all went from -0.5 V/cm increasing to a high point at 60% helium and keep decreasing until 63% with all three isomers following very similar patterns(Figure 21c). The carbon isotopologue shifts (129-131 *m/z*) start at -0.4 V/cm for all 3 isomers increasing until 63% between -1.6 and -1.8 V/cm for the final shifts. The only shifts that showed major differences between isomers were those between a chlorine and its natural isotopologue from carbon. The increase in isotopologic shift intensities by changing the main gas from nitrogen to carbon dioxide allows for the idea of identification of isomers based upon their isotopologic shift pattern. Nitrogen is not pursued at high field intensities due to the smaller isotopologic shifts.

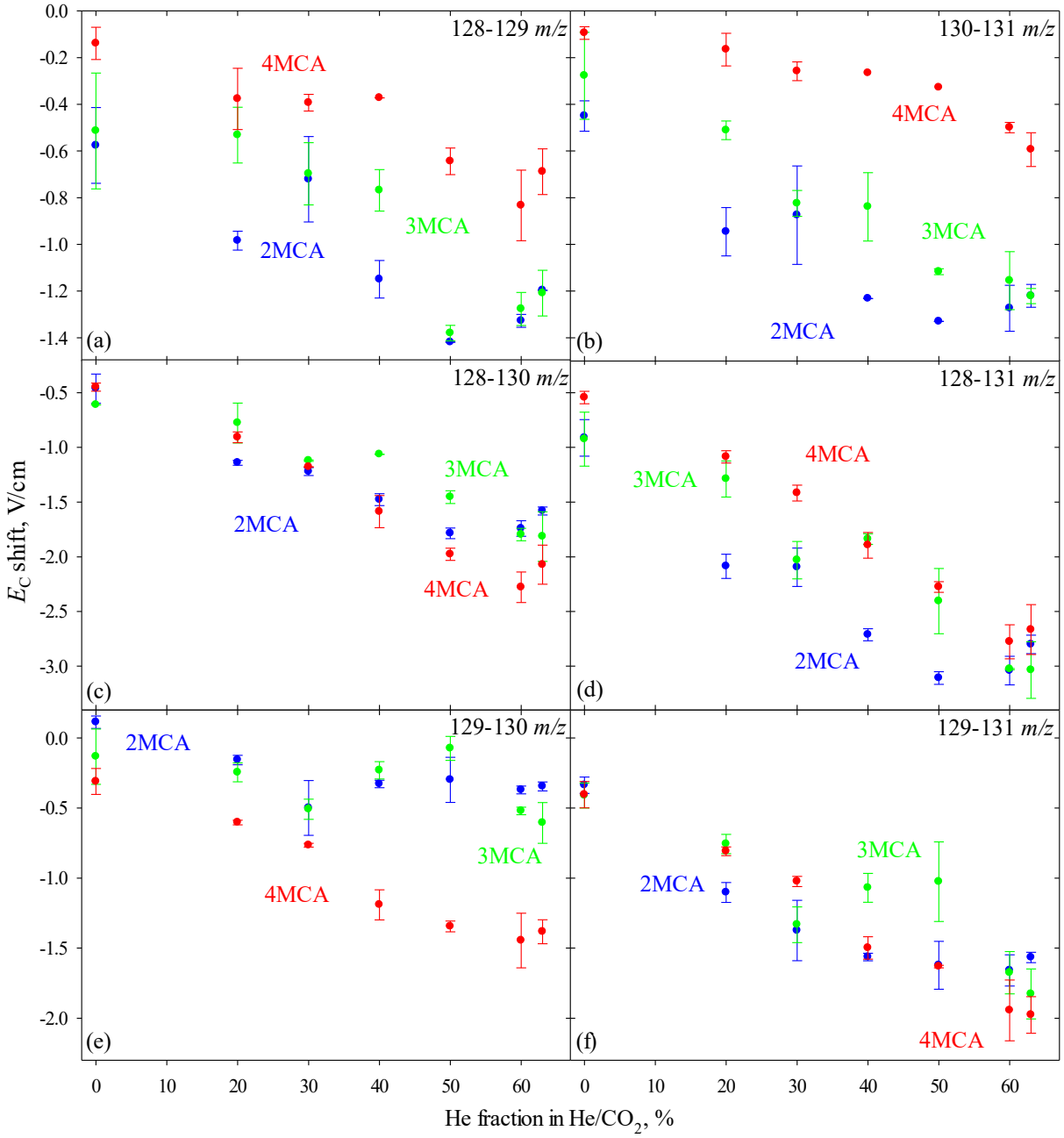

Figure 21: All shifts for MCA isomers[4MCA (red), 3MCA (green), and 2MCA (blue)] at minimum field in He/CO_2 gas buffer: +1 Da shifts from base peak ^{35}Cl (a) and ^{37}Cl (b), +2 Da shifts between Cl isotopes (c), +3 Da shift between ^{35}Cl and $^{37}Cl+^{13}C$ (d), +1 Da shifts between $^{35}Cl+^{13}C$ and ^{37}Cl (e), +2 Da shift between $^{35}Cl+^{13}C$ and $^{37}Cl+^{13}C$ (f).

The overall isotopologic shift trends continue with the transition from minimum to moderate field. Moderate field altered the shifts of the isomers, slightly spreading out the different shifts and further differentiating isomers. The shifts of isomers between the base peak

and the +1 Da isotopologue (128-129 *m/z*) start overlapped at around -0.4 V/cm at 0% helium but with added helium the shifts expand to as much as 1.0 V/cm apart at 50% helium. At 40, 50, and 63 percent helium in carbon dioxide all three isomers have distinct isotopologic shifts for all three isomers (Figure 22a). The shifts between the +2 Da and +3 Da (130-131 *m/z*) follow a similar pattern where, at 0% helium shifts of the different isomers overlap but increasing helium percentage separated the isomer shifts at 30% and they remain separated until 63% helium (Figure 22b).

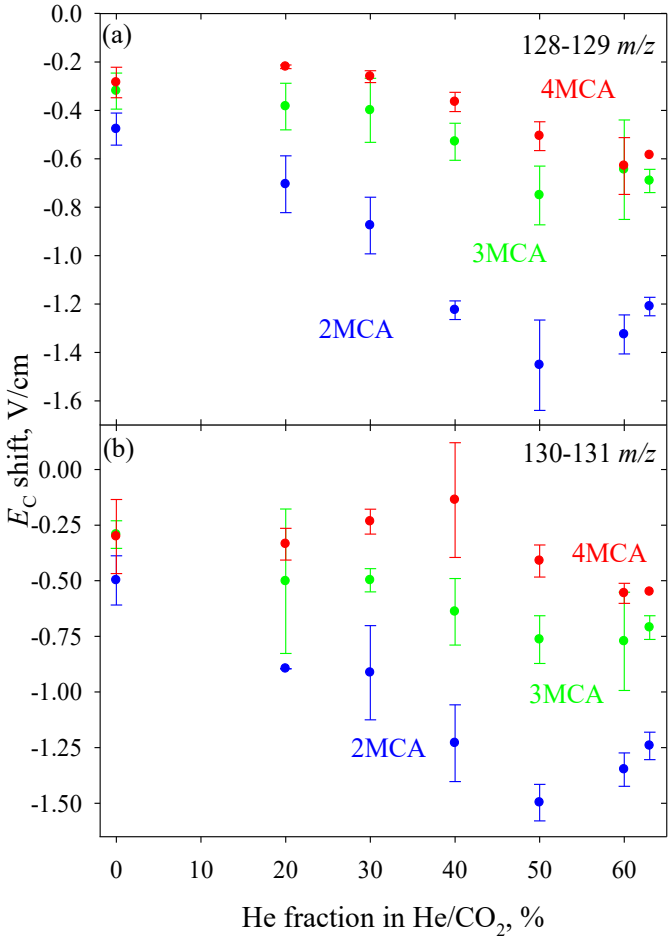

Figure 22: +1Da isotopologue shifts for MCA with ^{35}Cl (a) and ^{37}Cl (b) at moderate field in He/CO_2 gas buffer.

The halogen to halogen peaks exhibited similar behavior between minimum and moderate field. Any isotopologic shifts relating to an increase of +2 Da resulted in nearly identical patterns for the different isomers (Figure 23). For the ^{35}Cl-^{37}Cl isotopologic shift all isomers increased with the increase in helium following nearly identical paths. Starting at -0.5 V/cm the shifts expand as high as -1.9 V/cm, though differentiation between the isomers is not possible at most helium percentages. A very similar increase for the +1 isotopologue versions of the same shift (129-131 *m/z*) starting around -0.4 V/cm increasing to -1.8 V/cm with increasing helium (Figure 23).

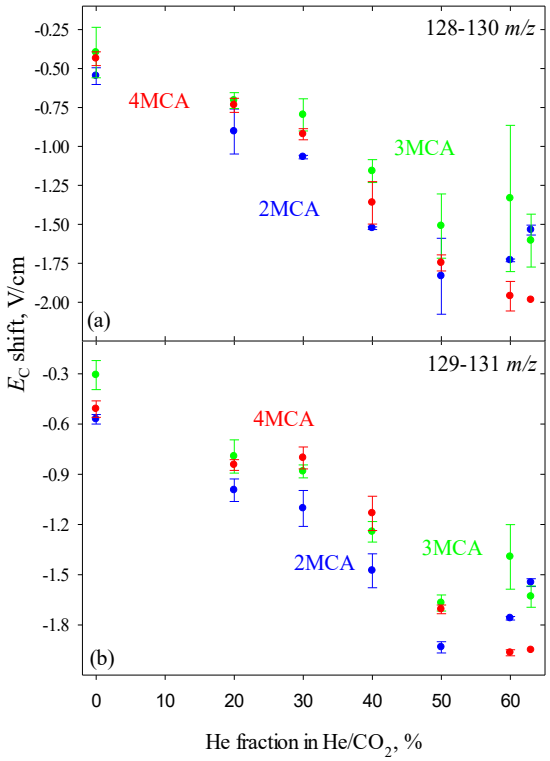

Figure 23: Halogen to halogen shifts for MCA at moderate field for base peaks (a) and +1 peaks (b).

This leaves two remaining shifts, a +1 shift from ^{35}Cl+^{13}C to ^{37}Cl and a +3 shift from ^{35}Cl to ^{37}Cl+^{13}C. For the first one there is a distinct pattern change from the previous +1 Da shifts. The pattern seems to have inverted itself such that now 2MCA<3MCA<4MCA in shift

magnitude (Figure 24). Like other +1 shifts there are distinct differences between the isomers starting at 40% helium and continuing through 63% helium. The +3 Da shift has 2 MCA easily separating from the other isomers at all amounts of helium, though 3MCA and 4MCA nearly overlap the entire time. The shift intensities increase, maxing out around 50% before shrinking slightly approaching 63%.

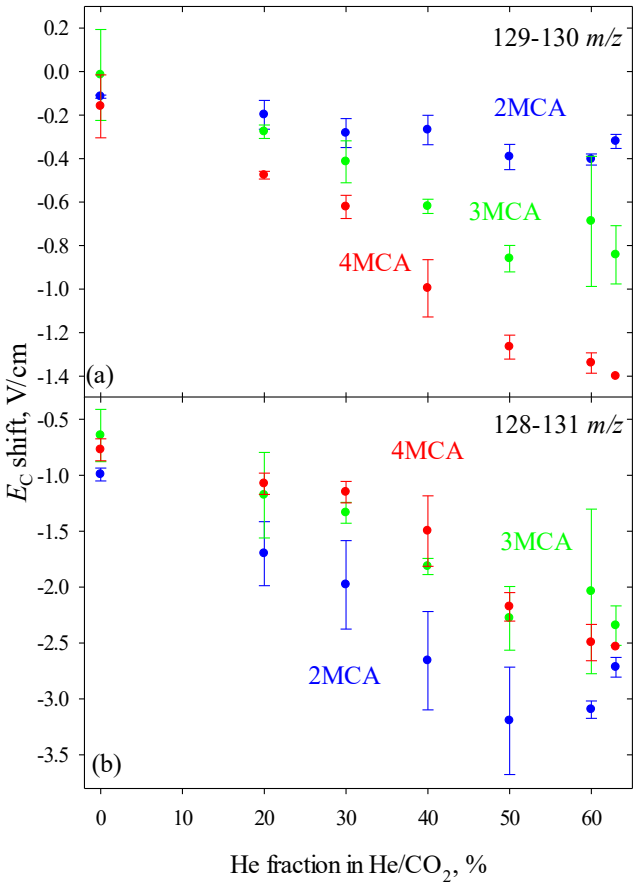

Figure 24: Moderate field in He/CO_2 gas buffer MCA isomer shifts +1 Da shift between $^{35}Cl+^{13}C$ and ^{37}Cl (a) and +3 Da shift between $^{35}Cl+^{13}C$ and ^{37}Cl (b).

The shift intensity noted earlier is expected due to shift additivity. Noted originally with singular alanine and glycine a +2 Da shift is the sum of its two +1 Da shifts [29]. This was also seen with dialanine where the +1 Da shift was about 1/4 of the +4 Da shift. Each of the multiple Da shifts can be matched to the addition of the individual shifts of its components. Within each

spectrum, these match perfectly. Table 4 has the additions of each shift after statistical workup, and they still fall within matching ranges to the measured shift. Thus when the 128-129 *m/z* shift had a distinct isomeric pattern and the 128-130 *m/z* had little differentiation between isomers it indicates that the 129-130 *m/z* shifts would have to be inverted to match peak additivity.

Table 4

At moderate fields in He/CO_2, the peak additivity for MCA +2 Da shifts (first) and the +3 Da shifts (second).

+2 Da Shifts		Shifts Added		Measured Shift			Shifts Added		Measured Shift	
Isomer	%He	128-129 129-130	STD	128-130	STD		129-130 130-131	STD	129-131	STD
2MCA	0	-0.5929	0.0666	-0.9942	0.0576		-0.6137	0.1109	-0.5718	0.0285
2MCA	20	-0.9043	0.1347	-1.7009	0.2861		-1.0942	0.0661	-0.9954	0.0672
2MCA	30	-1.1583	0.1348	-1.9799	0.3951		-1.1957	0.2220	-1.1042	0.1074
2MCA	40	-1.4934	0.0777	-2.6585	0.4398		-1.4982	0.1850	-1.4766	0.1012
2MCA	50	-1.8449	0.1956	-3.1959	0.4808		-1.8897	0.1005	-1.9337	0.0338
2MCA	60	-1.7294	0.0842	-3.0952	0.0777		-1.7526	0.0792	-1.7606	0.0103
2MCA	63	-1.5309	0.0499	-2.7172	0.0887		-1.5628	0.0699	-1.5474	0.0234
3MCA	0	-0.3358	0.2220	-0.6452	0.2333		-0.3081	0.2180	-0.3081	0.0865
3MCA	20	-0.6605	0.1009	-1.1781	0.3823		-0.7787	0.3261	-0.7939	0.0994
3MCA	30	-0.8144	0.1636	-1.3361	0.0935		-0.9124	0.1097	-0.8832	0.0390
3MCA	40	-1.1492	0.0835	-1.8161	0.0716		-1.2587	0.1531	-1.2441	0.0613
3MCA	50	-1.6113	0.1356	-2.2795	0.2847		-1.6246	0.1232	-1.6692	0.0480
3MCA	60	-1.3337	0.3630	-2.0390	0.7355		-1.4609	0.3721	-1.3938	0.1927
3MCA	63	-1.5337	0.1422	-2.3444	0.1765		-1.5527	0.1442	-1.6312	0.0638
4MCA	0	-0.4445	0.1580	-0.7725	0.0982		-0.4610	0.2207	-0.5109	0.0492
4MCA	20	-0.6960	0.0190	-1.0760	0.0956		-0.8118	0.0735	-0.8455	0.0320
4MCA	30	-0.8827	0.0586	-1.1514	0.0954		-0.8561	0.0772	-0.8028	0.0655
4MCA	40	-1.3617	0.1374	-1.4991	0.3155		-1.1339	0.2895	-1.1339	0.1027
4MCA	50	-1.7729	0.0810	-2.1764	0.1271		-1.6777	0.0907	-1.7072	0.0261
4MCA	60	-1.9694	0.1263	-2.4956	0.1622		-1.8967	0.0648	-1.9663	0.0178
4MCA	63	-1.9858	0.0000	-2.5355	0.0000		-1.9504	0.0000	-1.9504	0.0000

Table 4 (continued)

+3 Da Shifts		Shifts Added		Shifts Added		Shifts Added		Measured Shift	
Isomer	%He	128-129 129-131	STD	128-130 130-131	STD	128-129 129-130 130-131	STD	128-131	STD
2MCA	0	-1.0494	0.0721	-1.0471	0.1229	-1.0913	0.1291	-0.9942	0.0576
2MCA	20	-1.7009	0.1353	-1.7996	0.1444	-1.7996	0.1347	-1.7009	0.2861
2MCA	30	-1.9799	0.1589	-1.9814	0.2120	-2.0714	0.2510	-1.9799	0.3951
2MCA	40	-2.7017	0.1082	-2.7549	0.1724	-2.7234	0.1890	-2.6585	0.4398
2MCA	50	-3.3862	0.1898	-3.3301	0.2571	-3.3422	0.2121	-3.1959	0.4808
2MCA	60	-3.0860	0.0809	-3.0796	0.0755	-3.0780	0.1128	-3.0952	0.0777
2MCA	63	-2.7575	0.0446	-2.7787	0.0695	-2.7729	0.0795	-2.7172	0.0887
3MCA	0	-0.6283	0.1143	-0.6904	0.1735	-0.6283	0.2305	-0.6452	0.2333
3MCA	20	-1.1781	0.1382	-1.2083	0.3286	-1.1628	0.3400	-1.1781	0.3823
3MCA	30	-1.2831	0.1377	-1.2958	0.1155	-1.3123	0.1717	-1.3361	0.0935
3MCA	40	-1.7738	0.0982	-1.7977	0.1666	-1.7884	0.1713	-1.8161	0.0716
3MCA	50	-2.4205	0.1304	-2.2757	0.2324	-2.3759	0.1728	-2.2795	0.2847
3MCA	60	-2.0390	0.2815	-2.1061	0.5187	-2.1061	0.4250	-2.0390	0.7355
3MCA	63	-2.3227	0.0798	-2.3151	0.1780	-2.2442	0.1519	-2.3444	0.1765
4MCA	0	-0.7958	0.0799	-0.7383	0.1722	-0.7459	0.2295	-0.7725	0.0982
4MCA	20	-1.0654	0.0329	-1.0714	0.0845	-1.0317	0.0739	-1.0760	0.0956
4MCA	30	-1.0637	0.0699	-1.1563	0.0663	-1.1170	0.0810	-1.1514	0.0954
4MCA	40	-1.4991	0.1101	-1.4991	0.2918	-1.4991	0.2922	-1.4991	0.3155
4MCA	50	-2.2135	0.0650	-2.1588	0.0887	-2.1841	0.1085	-2.1764	0.1271
4MCA	60	-2.5957	0.1187	-2.5179	0.1053	-2.5261	0.1341	-2.4956	0.1622
4MCA	63	-2.5355	0.0000	-2.5355	0.0000	-2.5355	0.0000	-2.5355	0.0000

3.3 Monobromoaniline Results

Since results for MCA in nitrogen showed an inability to separate any isotopologue shifts amongst isomers to the same extent as in CO_2; nitrogen was not pursued when switching to bromoanilines. The isomeric separation changed between minimum and moderate field. The increase in field did lead to an overall increase in maximum E_C position from <-200 V/cm at 63 % helium in minimum field to >-200 V/cm at 63% helium in moderate field. In minimum field (Figure 25a) all three isomers of MBA are completely separated at helium percentages 40-60. In

moderate field (Figure

demonstrates the massi

positions of a compoun

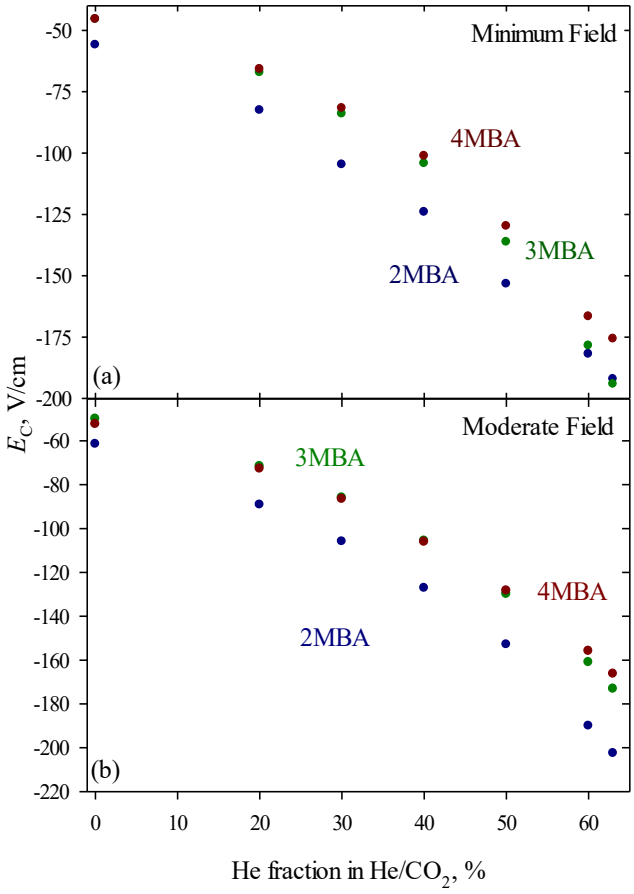

Figure 25: Isomeric positions in various helium/carbon dioxide gas compositions at minimum field (a) and moderate field (b).

Minimum field isotopologic shifts in CO_2 showed a slightly different pattern than that for MCA. When looking at the +1 Da isotopologic shift of the MBA isomers the first striking difference is the order of shift intensities. With the monobromoanilines 2MBA<4MBA<3MBA, unlike the MCA isomers where 2MCA was the largest shift. Otherwise the ability to differentiate the isomers amongst the isotopologic shifts is nearly identical to the MCAs. The +1 Da shifts all easily distinguish between the different isomers (Figure 26a,b,e), while the +2 Da

shifts all fall nearly on top of one another (Figure 26c,f). The intensities of each of the shifts is less than those of MCA at equivalent field, which is expected since bromine is more than two times the size of chlorine. With increasing amounts of helium nearly all shifts increased, though some by minute amounts.

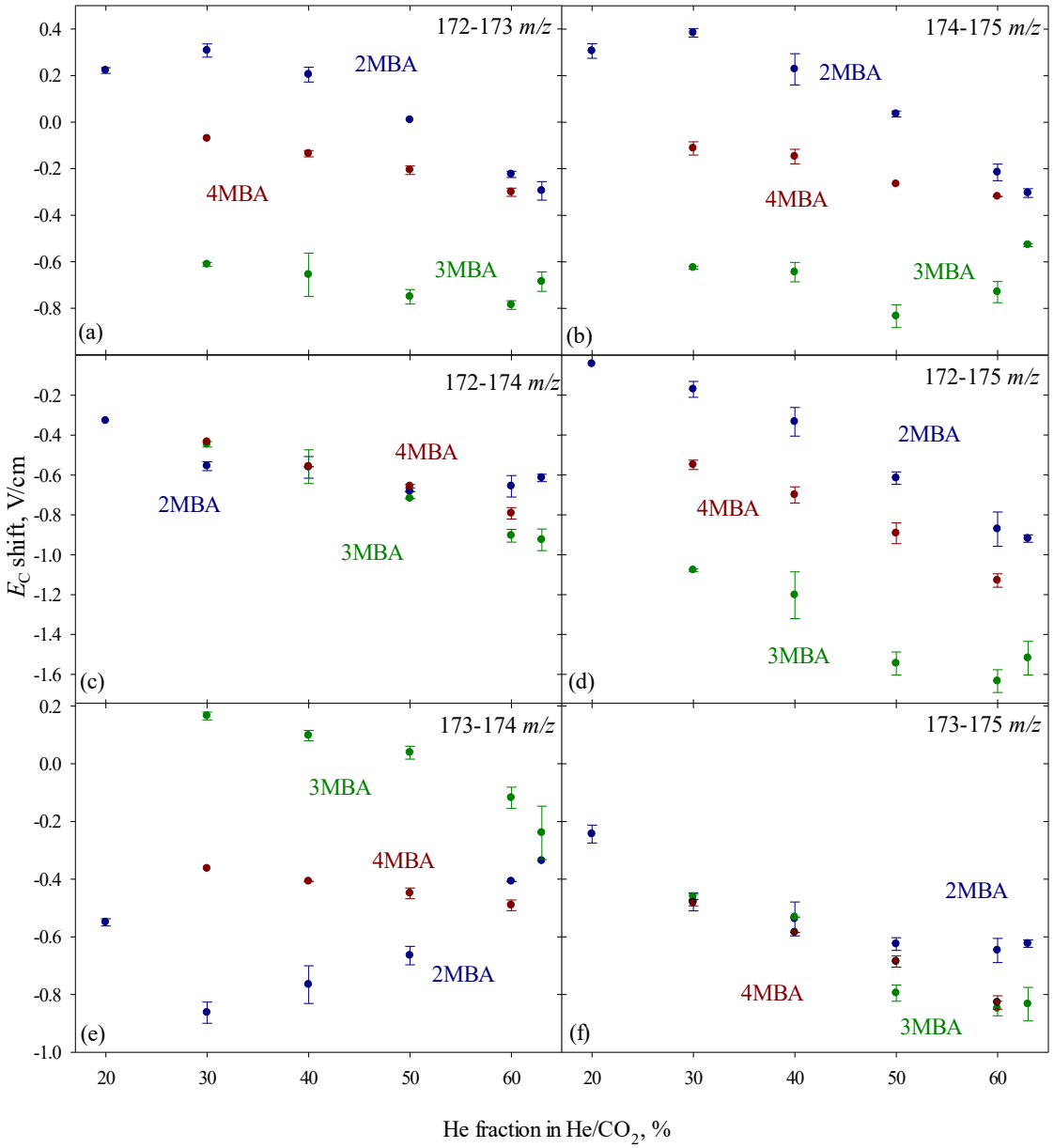

Figure 26: Assorted isotopologic shifts at minimum field for isomers of MBA in He/CO_2 mixtures.

At moderate field the shifts for MBA altered slightly. For the internal shifts between isotopologues base peak and +1 Da (172-173 *m/z*) , 2 MBA has a positive E_C shift from 0.2 V/cm at 20% up to 0.21 V/cm at 30-40% He then decreasing until -0.25 V/cm at 60-63% He. 3MBA remains consistently around -0.3- -0.4 V/cm throughout all amounts of helium. 4MBA starts at -0.25 V/cm at 20% helium decreasing in an absolute sense to -0.15 V/cm at 50% then increasing again to -0.3 V/cm at 60-63% (Figure 27a). The other halogen to +1 Da shift (174-175 *m/z*) has a very similar layout. 2MBA follows nearly an identical pattern, while 3MBA delves slightly deeper to -0.5 V/cm at 60% helium. 4MBA stays right around -0.1 V/cm until the upper helium percentages where it increases closer to -0.2 V/cm (Figure 27b).

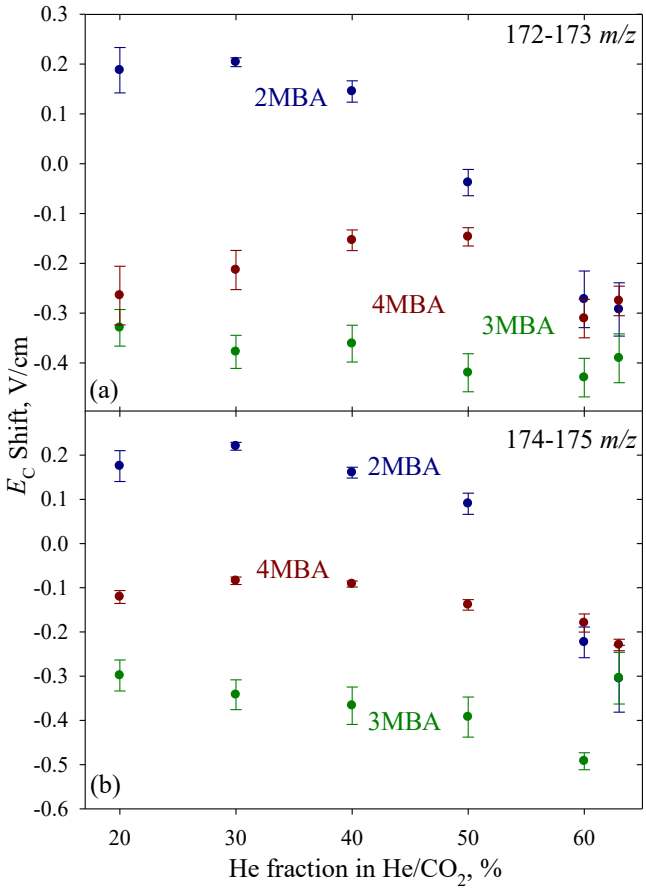

Figure 27: +1 Da isotopologue shifts for MBA with ^{79}Br (a) and ^{81}Br (b) at moderate field in He/CO_2 gas buffer.

For the halogen shifts (172-174 *m/z*) the isomers overlap throughout all helium percentages. Each isomer has a shift increase over 0.5 V/cm from increasing the helium percentage form 20-63% (Figure 28a). For the 173-175 *m/z* shifts of all isomers again increase from 20% to 63%, and they increase uniformly preventing differences between isomers (Figure 28b). For both sets at 60% and 63% there starts to be some separation but not satisfactory amounts. The only separation is on a magnitude similar to the nitrogen buffer gas with MCAs at minimum field which was deemed not suitable for continuation.

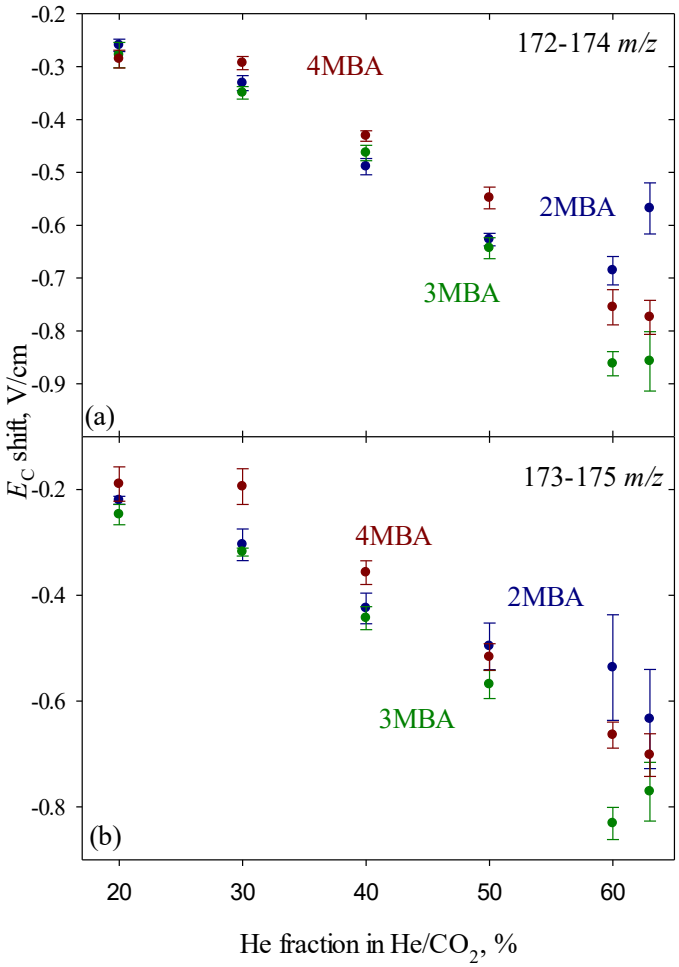

Figure 28: At moderate field in He/CO_2 gas buffer, +2 Da isotopologue shifts between the base ^{79}Br and ^{81}Br (a) and their +1 Da between ^{35}Br + ^{13}C and ^{81}Br + ^{13}C (b).

The last set of +1 Da shifts 173-174 *m/z,* have full separations at most helium percentages. 2MBA starts at -0.4 V/cm and descends to -0.6 V/cm at 40% helium then proceeding to jump back up to -0.3 V/cm at 63% helium. 4MBA starts at -0.1 V/cm increasing steadily to -0.5 V/cm by 63% He. 3MBA starts at 0.1 V/cm descending consistently to -0.45 V/cm at 63% helium. This puts the isomers in the opposite order from the other +1 Da shifts, which enforces peak additivity. The last shift of +3 Da from base peak to the heavy halogen with a heavy carbon (172-175 *m/z*) has differentiation between isomers at all percentages of helium with a steady absolute increase of all shifts during the helium increases.

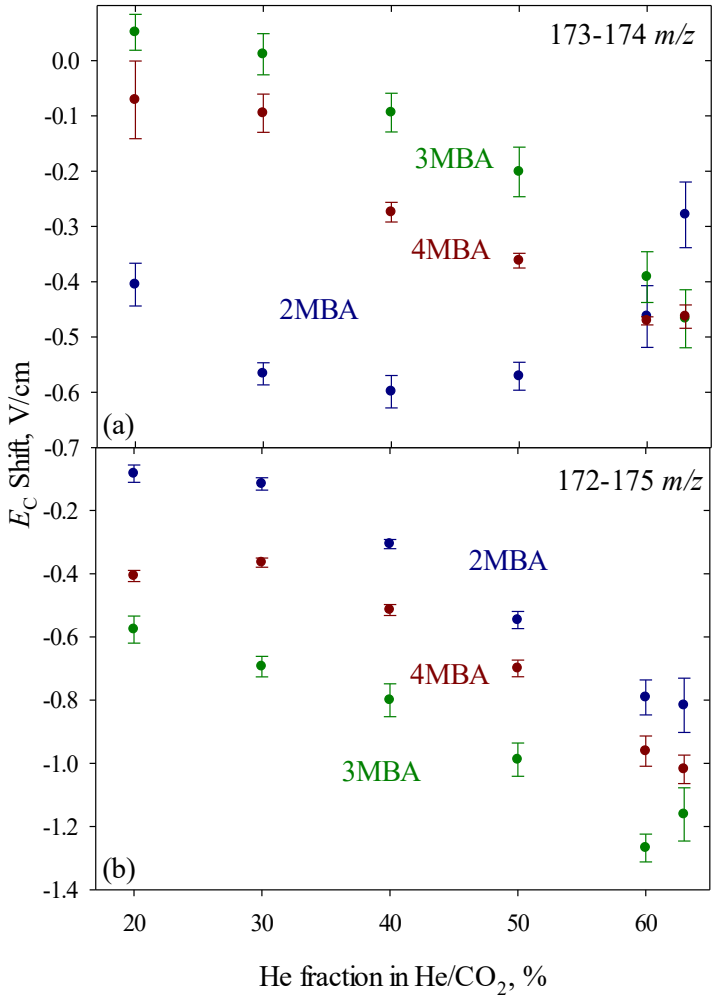

Figure 29: Isotopologue shifts +1 Da between the base $^{79}Br^{+13}C$ and ^{81}Br (a) and +3 Da between $^{35}Br$and $^{81}Br^{+13}C$ (b) at moderate field in He/CO_2 gas buffer.

Peak additivity can be seen in the monobromoanilines as well. The +2 Da shifts are the sum of the individual +1 Da shifts despite some variance in where the heavy carbon atom is added. Table 5 has the shift additions from the graphs at moderate field demonstrating the idea of adding different shifts to determine another shift.

Table 5

At moderate fields in He/CO2, the peak additivity for MBA +2 Da shifts (first) and +3 Da shifts (second).

+2 Da Shifts		Shifts Added		Measured Shift			Shifts Added		Measured Shift	
Isomer	%He	172-173 173-174	STD	172-174	STD		173-174 174-175	STD	173-175	STD
2MBA	20	-0.2174	0.0597	-0.2598	0.0116		-0.2301	0.0521	-0.2204	0.0076
2MBA	30	-0.3626	0.0219	-0.3312	0.0142		-0.3465	0.0218	-0.3044	0.0299
2MBA	40	-0.4537	0.0364	-0.4891	0.0154		-0.4384	0.0318	-0.4247	0.0290
2MBA	50	-0.6088	0.0366	-0.6270	0.0118		-0.4808	0.0348	-0.4965	0.0440
2MBA	60	-0.7349	0.0798	-0.6859	0.0271		-0.6862	0.0658	-0.5368	0.0999
2MBA	63	-0.5714	0.0799	-0.5681	0.0483		-0.5847	0.0960	-0.6339	0.0939
3MBA	0	-0.2204	0.0439	-0.1987	0.0101		-0.2795	0.0405	-0.2162	0.0225
3MBA	20	-0.2783	0.0490	-0.2783	0.0239		-0.2472	0.0479	-0.2472	0.0192
3MBA	30	-0.3661	0.0498	-0.3497	0.0118		-0.3303	0.0501	-0.3182	0.0075
3MBA	40	-0.4552	0.0510	-0.4633	0.0145		-0.4607	0.0549	-0.4433	0.0218
3MBA	50	-0.6210	0.0588	-0.6432	0.0198		-0.5936	0.0638	-0.5687	0.0266
3MBA	60	-0.8209	0.0600	-0.8620	0.0227		-0.8837	0.0498	-0.8314	0.0304
3MBA	63	-0.8575	0.0718	-0.8575	0.0562		-0.7713	0.0785	-0.7713	0.0556
4MBA	20	-0.3355	0.0917	-0.2857	0.0168		-0.1918	0.0717	-0.1897	0.0325
4MBA	30	-0.3086	0.0524	-0.2932	0.0126		-0.1794	0.0356	-0.1944	0.0337
4MBA	40	-0.4279	0.0274	-0.4312	0.0099		-0.3658	0.0190	-0.3571	0.0225
4MBA	50	-0.5087	0.0227	-0.5483	0.0205		-0.5005	0.0179	-0.5167	0.0250
4MBA	60	-0.7814	0.0393	-0.7552	0.0332		-0.6504	0.0218	-0.6643	0.0248
4MBA	63	-0.7385	0.0365	-0.7742	0.0321		-0.6925	0.0248	-0.7021	0.0404

Table 5 (continued)

+3 Da Shifts		Shifts Added		Shifts Added		Shifts Added		Measured Shift	
Isomer	%He	172-173 173-175	STD	172-174 174-175	STD	172-173 173-174 174-175	STD	172-175	STD
2MBA	20	-0.0326	0.0461	-0.0847	0.0368	-0.0423	0.0692	-0.0835	0.0275
2MBA	30	-0.1005	0.0312	-0.1112	0.0167	-0.1426	0.0236	-0.1160	0.0197
2MBA	40	-0.2797	0.0361	-0.3287	0.0197	-0.2934	0.0384	-0.3063	0.0146
2MBA	50	-0.5343	0.0513	-0.5369	0.0266	-0.5187	0.0437	-0.5467	0.0270
2MBA	60	-0.8089	0.1150	-0.9094	0.0441	-0.9583	0.0870	-0.7917	0.0554
2MBA	63	-0.9264	0.1081	-0.8740	0.0897	-0.8773	0.1100	-0.8164	0.0858
3MBA	0	-0.4797	0.0410	-0.5213	0.0314	-0.5430	0.0530	-0.5006	0.0401
3MBA	20	-0.5767	0.0414	-0.5767	0.0425	-0.5767	0.0603	-0.5767	0.0429
3MBA	30	-0.6959	0.0340	-0.6917	0.0357	-0.7081	0.0601	-0.6938	0.0325
3MBA	40	-0.8045	0.0430	-0.8301	0.0446	-0.8220	0.0662	-0.8005	0.0521
3MBA	50	-0.9885	0.0465	-1.0356	0.0495	-1.0134	0.0743	-0.9885	0.0525
3MBA	60	-1.2608	0.0491	-1.3543	0.0298	-1.3132	0.0630	-1.2677	0.0439
3MBA	63	-1.1618	0.0741	-1.1618	0.0812	-1.1618	0.0926	-1.1618	0.0842
4MBA	20	-0.4543	0.0673	-0.4065	0.0223	-0.4564	0.0928	-0.4071	0.0176
4MBA	30	-0.4078	0.0518	-0.3775	0.0152	-0.3928	0.0531	-0.3650	0.0145
4MBA	40	-0.5108	0.0306	-0.5228	0.0121	-0.5195	0.0282	-0.5149	0.0173
4MBA	50	-0.6635	0.0311	-0.6870	0.0238	-0.6474	0.0257	-0.6997	0.0263
4MBA	60	-0.9751	0.0459	-0.9350	0.0390	-0.9613	0.0443	-0.9613	0.0479
4MBA	63	-0.9775	0.0501	-1.0036	0.0346	-0.9679	0.0387	-1.0189	0.0452

3.4 Center of Mass

Noted earlier was the difference in shift patterns between monochloroanilines and monobromoanilines. These differences will be discussed as the differences between the base peak and the +1 Da peak so 128-129 *m/z* for monochloroaniline and 172-173 *m/*z for monobromoaniline. Each of these isotopologue shifts is essentially caused by the addition of ^{13}C at any 1 of 6 positions in the molecule. This change in shift order stems from the difference in mass between the halogens and how it shifts the center of mass. While traveling in a gas the molecules are spinning. Mainly the spin is determined by the center of mass of the molecule and

for the unhalogenated aniline this is dictated by the amine group at 17 Da (because of the required protonation for MS we are treating it as NH_3^+ for this thought process). Chlorine at a mass of 35 Da is only ~2 times heavier than the amine group shifting the center of mass only small amounts for each isomers (Figure 30). Bromine at a mass of 79 Da is ~4.6 times heavier than the amine, greatly shifting the center of mass to past the edge of the inner ring. Exact coordinates for each center of mass can be found in Table 6.

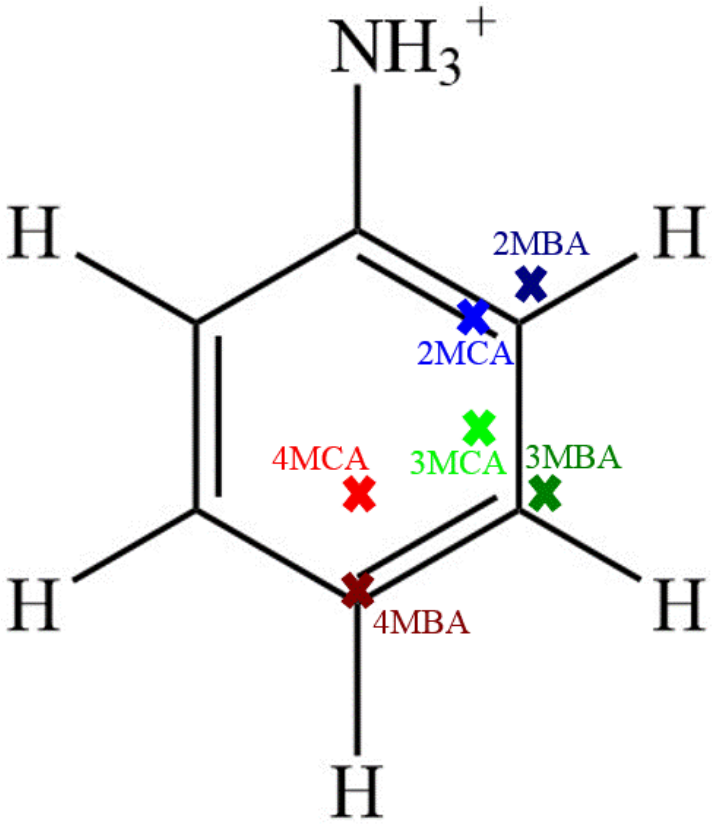

Figure 30: Illustration of the center of mass for the halogenated anilines on a 2D plane.

First, the shifts for MCA are greater in intensity than MBA because bromine is so much heavier than chlorine. This ~2.3 times mass difference means that the addition of +1 Da on the different carbons affects the center of mass far less for bromine than chlorine. With the center of mass shifts each of the isomers is affected in a similar pattern. 4 monohalogenatedaniline (MHA) has the smallest shifts since the addition of ^{13}C appears either in line with the other 2

heavy groups or on either side. These side positions do not move the center of mass much due to the trapped position between the heaviest parts of the aniline, so it is either near 1 heavy group or the other. The halogen addition to the 3 position shifts the center of mass along 2 axes rather than just 1 axis in 4 MHA. This creates 3 positions that add away from the heavy parts of the molecule, with 1 position having the most effect. The addition of a halogen to the 2 position creates a weight right next to the aniline shifting the center of mass along 2 axes. This does create a difference of having 4 positions for the addition of ^{13}C that are away from the heavy groups, but with 2 positions that have the most manipulation of the center of mass. This in turn is reflected by the intensities of shifts in MCA where 4MCA<3MCA<2MCA. 3 and 4 monohalogenated isomers retain similar patterns for both bromine and chlorine, but 2MCA and 2MBA act entirely different. This indicates that there is another factor in play for 2MBA.

Table 6

Center of mass coordinates of the monohalogenatedanilines relative to the center of the benzene ring:

Compound	X Coordinate (Å)	Y Coordinate (Å)	Z Coordinate (Å)
2MCA	0.717598	0.734113	0.000021
3MCA	0.709576	-0.072672	0.000004
4MCA	0.004570	-0.499677	0.000043
2MBA	1.291884	0.960699	-0.000047
3MBA	1.274612	-0.484172	0.000026
4MBA	0.003445	-1.245445	-0.000372

3.5 Conclusion

MCA and MBA had complete isomer separation with FAIMS. They also exhibited unique isotopologue shift patterns for the different isomers for certain shifts. The shifts that showed differences between isomers was any shift from a base halogen to an isotopologue caused by a ^{13}C. MCA presented 4MCA as the smallest shifts with 2MCA as the greatest shifts

between 128-129 *m/z*. For MBA, 2MBA had the positive shift, while 3MBA had about the same shift, but negative in value. 4MBA had a shift between 2 and 3 MBA. The shifts between halogen isotopes was consistently indifferent between isomers. The shifts between the +2 Da and +4 Da peaks also did not vary between isomers for either halogen.

These identifying positions for MCA and MBA isomers can be used to identify unknown isomers. Systematically through the combination of the shifts, even with adding nitrogen as a buffer gas for added individuality of shifts, any ion can potentially be identified through its isotopologic shifts in a way that is reminiscent of NMR. Through this, an alternative to NMR is approachable.

CHAPTER 4

INTRINSIC SIZE PARAMETER AND POST TRANSLATIONAL MODIFICATIONS

4.1 Methodology

ISP is, as stated in Chapter 1, proportional to the ratio of the cross sectional area to the mass. The cross sectional area is approximated to be a sphere making the cross section area into the area of a circle. Thus, the ISP of a moiety was approximated as proportional to the sum of projection areas for all atoms divided by the total mass:

$$ISP \propto \sum \pi r_i^2 / \sum m_i \tag{4.1}$$

where r_i and m_i are the radius and the mass of i-th atom. Two sets of radii were assumed. Set 1 is comprised of van der Waals radii combined with the extended Bondii radii [43,44] [1.1 Å (H), 1.7 Å (C), 1.55 Å (N), 1.52 Å (O), 1.80 Å (S), 1.47 Å (F), 1.80 Å (P), 1.75 Å (Cl), 1.85 Å (Br), 1.98 Å (I)]. Set 2 stems from a previous a priori ISP for amino acids with most radii being identical to set 1 while the exceptions are [1.60 Å (C,N,O), 2.00 Å (S)], values stemming from the original a priori ISP calculations for amino acids [35]. With both sets, the ISPs are scaled to the arithmetic mean of 1.0 for natural amino acids. The consequent ISP range for each PTM conveys the error margin resulting from reasonable uncertainties of atomic radii. We assumed the monoisotopic masses for all moieties. The change from average masses is minute because of small mass differences and the above ISP normalization.

Most PTMs are a mixture of addition and subtraction of atoms to a peptide. For example, citrullination (cit) replaces NH in an arginine by O. Thus Equation 4.1 can be generalized for a PTM as:

$$ISP \propto \frac{\Delta P}{\Delta m} = \sum \pi (r_i^2 - r_j^2) / \sum (m_i - m_j) \tag{4.2}$$

Where ΔP and Δm are the net changes of sum projection area and mass, and i and j count the added and removed atoms respectively. This gives the possibility of a negative ISP, such as cit where the cumulative projection area for N and H exceeds that for O so the numerator is negative and the mass increase is positive.

ISP controls the contribution of a moiety to cross section per unit mass, the visible effect is the variance from the mean Ω/m trend: a PTM with small change in mass would affect the cross section very little even if its ISP grossly deviates from the mean. To account for this, we introduce the cross section impact score.

$$\Omega_{IMP} = (ISP - 1) \times \Delta m \tag{4.3}$$

This impact quantity can be positive or negative depending on the signs of the factors. This becomes most important when considering a peptide with multiple PTMs. The Ω_{IMP} value for multiple (same or different) PTMs is the sum of their Ω_{IMP}. The concept of impact scores is crucial for PTMs where Δm values range over four orders of magnitude (from 1 Da for cit to >10 kDa for SUMOs), whereas amino acids only vary as much as ~3 times (from 57 Da for glycine to 186 Da for tryptophan).

4.2 Results

The ISPs calculated for p (0.53-0.58), cam (0.92-0.94), and pal (1.29-1.33) essentially match the measurements, confirming that the a priori model works for PTMs. The ensuing domain delineantion is partial for p [Ω_{IMP} = - (33-38)], none for cam (Ω_{IMP} = -4), and strong to perfect for pal (Ω_{IMP} = 69-78) or *bis*-phosphorylation [Ω_{IMP} = - (67-75)]. This allows tentatively projecting little, partial, and strong domain delineations for absolute Ω_{IMP} of <20, 20-60, >60, respectively. Within this framework, we have evaluated ISPs and Ω_{IMP} for 100 more common PTMs found in Table 7.

Table 7

PTMs and their ISPs and Ω_{IMP} for radii sets 1 and 2. [a-r]See appendix 1 for PTM sources.

	PTM Name	Δm	Stoichiometry change	ISP1	ISP2	Ω_{IMP1}	Ω_{IMP2}
1	Amidation[a]	-1	H1;N1;O-1	-4.84	-4.65	5.9	5.7
2	Citrullination[a]	1	H-1;N-1;O1	-4.84	-4.65	-5.9	-5.7
3	Deamidation[b]	1	O1;N-1;H-1	-4.84	-4.65	-5.9	-5.7
4	Iodination[c]	126	I1;H-1	0.0801	0.0828	-116	-116
5	Bromination[c]	78	Br1;H-1	0.102	0.105	-70	-70
6	Nitrosylation[a]	29	H-1;N1;O1	0.450	0.519	-16	-14
7	Pyrophosphorylation[d]	157	P2;H-1;O6	0.473	0.505	-86	-78
8	Sulfation[a]	80	S1;O3	0.473	0.562	-42	-35
9	Nitration[a]	45	N1;O2;H-1	0.481	0.553	-23	-20
10	Trifluoroacetyl[e]	96	C2;O1;F3;H-1	0.518	0.519	-46	-46
11	Phosphorylation[a]	80	H1;O3;P1	0.530	0.583	-38	-33
12	Hydroxylation[a]	16	O1	0.538	0.616	-7.4	-6.2
13	Methionine Sulfone[a]	32	O2	0.538	0.616	-15	-12
14	Oxidation[f]	32	O2	0.538	0.616	-15	-12
15	2,4,5 trichlorophenyl[g]	178	C6;H1;Cl3	0.580	0.557	-75	-79
16	Pentafluorophenyl (Pfp)[g]	166	C6;F5;H-1	0.604	0.578	-66	-70
17	Carboxyglutamic Acid[a]	44	C1;O2	0.636	0.671	-16	-14
18	2Bromobenzyloxycarbonyl[h]	212	C8;H5;O2;Br1	0.652	0.635	-74	-77
19	Formylation[a]	28	C1;O1	0.692	0.703	-8.6	-8.3
20	2Nitrophenylsulphenyl (Nps)[h]	153	C6;H3;N1;O2;S1	0.760	0.771	-37	-35
21	Carbamylation[i]	43	N1;H1;C1;O1	0.763	0.795	-10	-8.8
22	Malonylation[j]	86	C3;O3;H2	0.780	0.795	-19	-18
23	Cysteinylation[g]	119	C3;H5;N1;O2;S1	0.782	0.821	-26	-21
24	ADP-ribosylation[g]	541	C15;H21;N5;O13;P2	0.807	0.827	-104	-93
25	Pyridoxal phosphate[g]	229	C8;H8;N1;O5;P1	0.813	0.819	-43	-41
26	2chlorobenzyloxycarbonyl (Cl-Z)[g]	168	C8;H5;O2;Cl1	0.817	0.795	-31	-34
27	5'-adenosylation[g]	329	C10;H12;N5;O6;P1	0.821	0.836	-59	-54
28	Carboxymethylcystine (Cme)[i]	58	C2;O2;H2	0.823	0.840	-10	-9.3
29	2Nitrobenzoyl (NBz)[k]	149	C7;H3;O3;N1	0.829	0.821	-25	-27

Table 7 (continued)

	PTM Name	Δm	Stoichiometry change	ISP1	ISP2	Ω_{IMP1}	Ω_{IMP2}
30	Pyruvoyl[a]	70	C3;H2;O2	0.836	0.836	-11	-11
31	FAD[g]	783	C27;H31;P2; N9;O15	0.848	0.858	-119	-111
32	Hydroxysuccinimide[k]	97	C4;H3;O2;N1	0.853	0.855	-14	-14
33	4Toluenesulphonyl (Tosyl,Tos)[k]	154	C7;H6;S1;O2	0.855	0.857	-22	-22
34	4Nitrophenyl (Onp)[g]	121	C6;H3;N1;O2	0.861	0.848	-17	-18
35	FMN(On Histidine)[g]	454	C17;H19:N4; O9;P1	0.867	0.873	-60	-58
36	Succinylation[j]	100	C4;H4;O3	0.869	0.876	-13	-12
37	Gluthathionylation[a]	305	C10;H15;N3; O6;S1	0.871	0.893	-39	-33
38	FMN(On Threonine)[l]	438	C17;H19;N4; O8;P1	0.879	0.882	-53	-51
39	FMN(On Cystine)[m]	456	C17;H21:N4; O9;P1	0.883	0.890	-53	-50
40	4'Phosphopantetheine[g]	339	C11;H20;N2; O6;P1;S1	0.891	0.909	-37	-31
41	N-Glycolylneuraminic acid[h]	307	C11;H17;N1;O 9	0.916	0.931	-26	-21
42	Lipoylation[g]	188	C8;H12;N0; O1;S2	0.920	0.932	-15	-13
43	Carboxyamidomethylcystine[k]	57	C2;O1;N1;H3	0.923	0.936	-4.4	-3.6
44	Dansylation[k]	233	C12;H11;N1; O2;S1	0.931	0.920	-16	-19
45	Acetylation[a]	42	C2;H2;O1	0.931	0.925	-2.9	-3.1
46	N-acetylneuraminic acid[g]	291	C11;H17;N1;O 8	0.937	0.949	-18	-15
47	Pentoses (Xlyose)[g]	132	C5;H8;O4	0.941	0.954	-7.7	-6.1
48	Hexoses(Mannose)[a]	162	C6;H10;O5	0.942	0.956	-9.4	-7.1
49	Aldohexosylation[g]	162	C6;H10;O5	0.942	0.956	-9.4	-7.1
50	Glc3Mana9GlcNAc2[n]	2368	C88;H148; O71;N2	0.947	0.960	-126	-94
51	BG B Trisaccharide[o]	470	C18;H30;O14	0.956	0.968	-21	-15
52	Biotinylation[g]	226	C10;H14;N2; O2;S1	0.964	0.966	-8.2	-7.6
53	BG A Pentasaccharide[o]	819	C32;H53;N1; O23	0.965	0.975	-29	-21
54	BG H type 1[o]	511	C20;H33;N1; O14	0.965	0.975	-18	-13

Table 7 (continued)

	PTM Name	Δ*m*	Stoichiometry change	ISP1	ISP2	Ω_{IMP1}	Ω_{IMP2}
55	BG A Trisaccharide[o]	511	C20;H33;N1; O14	0.965	0.975	-18	-13
56	Acetylgalactosamine[a]	203	C8;H13;O5;N1	0.969	0.977	-6.4	-4.6
57	Benzyloxycarbonyl (Z)[g]	134	C8;H6;O2	0.973	0.943	-3.7	-7.6
58	Glucosamine[g]	161	C6;H11;O4;N1	0.978	0.991	-3.5	-1.5
59	Benzoyl (Bz)[g]	104	C7;H4;O1	0.980	0.937	-2.1	-6.6
60	Deoxyhexoses (fucose)[a]	146	C6;H10;O4	0.986	0.993	-2	-1
61	Pupylation (PUP_MYCTO)[c]	6938	C285;H456; N85; O115;S1	0.992	0.997	-55	-24
62	Thioanisyl[k]	106	C7;H6;S;O-1	0.998	0.966	-0.2	-3.6
63	9-Fluorenylmethyloxycarbonyl[k]	222	C15;H10;O2	1.01	0.964	1.7	-8
64	SUMO3 Human Isoform 1[c]	11623	C502;H799; N144;O164;S5	1.01	1.01	138	150
65	Pentamethylchroman6sulphonyl[k]	266	C14;H18;O3;S1	1.01	1.00	3.6	0.6
66	SUMO1 Human Isoform 1[c]	11543	C501;H799; N138;O165;S5	1.01	1.01	162	171
67	SUMO4 Human Isoform 1[c]	10672	C465;H746; N131;O147;S5	1.02	1.02	190	194
68	Crotonylation[g]	68	C4;H4;O1	1.02	0.998	1.7	-0.1
69	Ubiquitination[p]	8556	C378;H630; N105;O118;S1	1.04	1.04	316	310
70	Glypiation (GPI anchor)[q]	2179	C96;H175;P4; N4;O42	1.04	1.04	88	84
71	t-Butyloxycarbonyl (Boc)[g]	100	C5;H8;O2	1.07	1.06	7.1	6.2
72	Benzyloxymethyl (Bom)[g]	120	C8;H8;O1	1.09	1.05	11	5.9
73	Triphenylmethyl[g]	242	C19;H14	1.10	1.04	25	10
74	Diphenylmethyl (Dpm)[g]	166	C13;H10	1.11	1.05	19	8.6
75	Anisyl[g]	90	C7;H6	1.14	1.08	12	6.9
76	Benzyl (Bzl)[g]	90	C7;H6	1.14	1.08	12	6.9
77	Lipoprotein to D[a]	294	C17;H26;O4	1.14	1.12	40	34
78	S-archaeolyl[a]	316	C20;H28;O3	1.16	1.13	51	41
79	Hypusine[g]	87	C4;H9;N1;O1	1.16	1.16	14	14
80	4-methylbenzyl (Meb)[k]	104	C8;H8	1.17	1.12	18	12
81	Phosphatidylethanolamine amidated glycine[a]	699	C39;H74;O7; N1;P1	1.19	1.17	135	121
82	O-octanoylserine[a]	126	C8;H14;O1	1.25	1.22	32	28
83	O-decanoyl serine[a]	154	C10;H18;O1	1.28	1.25	43	38
84	S-diacylglycerol[a]	577	C37;H69;O4	1.29	1.26	166	148

Table 7 (continued)

	PTM Name	Δm	Stoichiometry change	ISP1	ISP2	Ω_{IMP1}	Ω_{IMP2}
85	Myristoylation[g]	210	C14;H26;O1	1.32	1.28	66	59
86	Farnesylation[g]	204	C15;H24	1.32	1.27	65	55
87	Pyrrolidone carboxylic acid[a]	-17	H-3;N-1	1.32	1.40	-5.5	-6.8
88	Geranyl[g]	136	C10;H16	1.32	1.27	44	37
89	Pyroglutamic acid from Gln[g]	-17	N-1:H-3	1.32	1.40	-5.5	-6.8
90	S-Geranyl-geranylation[g]	272	C20;H32	1.32	1.27	87	74
91	Palmitoylation[a]	238	C16;H30;O1	1.33	1.29	78	69
92	Stearoylation[a]	266	C18;H34;O1	1.34	1.30	89	79
93	Cholesterylation[a]	369	C27;H45	1.34	1.29	124	106
94	Methyl Esterification[r]	14	C1;H2	1.41	1.37	5.8	5.2
95	Methylation[a]	14	C1;H2	1.41	1.37	5.8	5.2
96	Ethylation[g]	28	C2;H4	1.41	1.37	11	10
97	t-Butylation[k]	56	C4;H8	1.41	1.37	23	21
98	Dimethylation[a]	28	C2:H4	1.41	1.37	11	10
99	Trimethylation[a]	42	C3:H6	1.41	1.37	17	15
100	Disulfide Bridge[g]	-2	H-2	4.50	4.65	-7	-7.3

Three PTMs have negative ISPs of – (4.6-4.8) citrullination, amidation, and deamidation. Despite these extreme ISPs, the minimum absolute Δm = 1 Da means $\left| \Omega_{IMP} \right| = 6$ so they should cause no significant deviation from the benchmark. At the other end is the ISP for the disulfide bridge, but $\left| \Omega_{IMP} \right| = 7$ is similarly minor because of the small Δm. The ISPs for other PTMs vary from 0.08 for iodination to 1.4 for methylation. As anticipated, ISP<<<1 are found for PTMs rich in heavy atoms-first of all halogens, then O, S, and P. Iodination and bromination combine very low ISPs (~0.1) with substantial mass, leading to Ω_{IMP} of -116 (I) and -70 (Br) that ought to make for (near) completely demarcated domains below the benchmarks. The next large group of PTMs (including sulfation, nitration, trifluroacetylation, (pyro) phosphorylation, oxidation, and hydroxylation) have essentially identical ISPs of 0.5-0.6. Their Ω_{IMP} accordingly scale with mass, amounting to – (6-15) for oxidation and hydroxylation (likely no significant domain delineation), - (20-46) for nitration, sulfation, phosphorylation, and trifluoroacetylation

(partial delineation), and – (80-90) for pyrophosphorylation (full delineation). Then IMS/MS would not broadly distinguish the nominally isobaric sulfation and nitration with near-equal ISPs, or pyrophosphorylation from double phosphorylation (an objective of some analyses). Higher but still low ISPs are found for others such formylation (~0.7), malonylation and cysteinylation (~0.8). Their Ω_{IMP} values of – (8-26) are not promising for major domain delineation.

Again, the PTM mass often matters more than the magnitude of (ISP-1). For example, the coincident Ω/m graphs for +1 oligonucleotides and carbohydrates lie below that for unmodified peptides [45,46]. Indeed, our calculated ISPs are 0.81-0.83 for ADP-ribose (a sugar derivative, 0.85-0.89 for FAD (flavin adenine dinucleotide), and FMN (flavin monocucleotide). Despite these only moderately low ISPs, the large masses of ribose and FAD lead to Ω_{IMP} of circa – (100-120), which should differentiate the peptides from ribosylated [47] and flavo [48] proteins. The Ω_{IMP} of FMN is only – (50-60), but should still cause notable domain separation. Succinylation and glutathionylation also have ISP ~ 0.87-0.89, which leads to a low $\left| \Omega_{IMP} \right|$<10 for the former but substantial $\left| \Omega_{IMP} \right|$ ~30-40 for the latter. Most glycans [pentoses, hexoses, deoxyhexoses, sialic acid, and hexosamines] have ISPs slightly below 1 (0.94-0.99) with $\left| \Omega_{IMP} \right|$<10 (except ~15-18 for sialic acid), suggesting no real shift from the benchmark as seen for cam [39] and, thus, no domain separation for glycopeptides. These observations also tell that the ISP approach can predict the ordering of IMS/MS trend lines for different biomolecular classes such as sugars, peptides, and nucleotides.

Some PTMs with ISP close to 1 are acetylation (0.93), biotinylation (0.96-0.97), and crotonylation (1.00-1.02): all with $\left| \Omega_{IMP} \right|$ ~ - 10-0. Protein PTMs also have ISP ~1, namely 1.01-1.02 (SUMO) and 1.04 (ubiquitin): many different amino acids in typical proteins average

to ~1.0. With the mass of ~9- 12 kDa, the theoretical Ω_{IMP} values of ~ 140-190 (SUMO) and ~ 310 (ubiquitin) are large but imprecise because of outsized impacts of small ISP variations. Although we include these and several other very large PTMs for completeness, we recognize that their size would place the adducts outside of the standard peptide mass range, where the baseline mass/mobility correlation loosens with the growing role of protein conformation.

A prominent PTM group is lipids. The Ω/m line for +1 lipids lies above that for unmodified peptides [45, 46], and our ISPs for lipid PTMs (e.g., farnesylation, geranylation, myristoylation, pal, stearoylation, cholesterylation) are ~1.3. The associated Ω_{IMP} values are, therefore, proportional to mass, and those for myristoylation and farnesylation (~60), geranylgeranylation and stearoylation (~80-90), and cholesterylation (~110-120) are expected, like pal [39], to produce stronger or fully delineated domains above the peptide benchmark. The Ω_{IMP} of ~40 for geranylation could still be enough for substantial separation. The alkyl PTMs (mostly methylation or methylesterification, but also ethylation and butylation) have even higher ISP ~1.4. The Ω_{IMP} of just 5-6 for light methylation would not afford material trend separation with monomethylation, but small size of Me common di- and trimethylation of the same residues with Ω_{IMP} of ~10-11 and ~15-17. A serious task in proteomics is distinguishing trimethylation from acetylation that is only 36 mDa lighter [48]. With the swing of ~20 between Ω_{IMP} of trimethylation and acetylation, the domain separation between acetylated and trimethylated proteomes should be noticeable although perhaps insufficient with single substitution, but substantial to strong with multiple substitutions.

Many peptides which are multiply modified with the same or different PTMs scale Ω_{IMP} by their number, which may substantially augment the domain differentiation as observed for multiple phosphorylation [37, 49] ubiquitous in biology (e.g. for tau proteins relevant to

Alzheimer's) [50, 51]. The same should occur with multiple nitrations, sulfations, oxidations, methylations, etc. For example, 0.3% of the H3 histone tails (characterized in middle-down proteomics) feature eight or nine methylations [52]. Their total computed Ω_{IMP} is ~40-50, potentially inducing substantial domain separation that may help detect those rare proteoforms. Superposition of different PTMs may increase or decrease the effect, depending on the ISPs. For example, histones often feature [53] acetylation and/or p ($\Omega_{IMP} < 0$) and methylation ($\Omega_{IMP} > 0$), which may cancel. For instance, same H3 tails have forms including diacetylation and methylation with Ω_{IMP} ~0 [52]. The same may happen within a PTM, e.g., glypiation that anchors proteins to cell membranes [54] consists of a glycan core (ISP <1) and lipid tail (ISP > 1) for the overall ISP of 1.04.

The mean ISP slightly increase for heavier PTMs, but scatter swamps the trend (Figure 31). The values converge to ~1.0 for heavier PTMs: large moieties tend to comprise a diversity of light and heavy atoms with ISP contributions that average close to 1.0, and the largest PTMs are peptides with ISP ~ 1.0 by default. The mean ISP of all PTMs analyzed is 0.81, reflecting that all biomolecules except lipids are denser than peptides [45, 46] In particular, heavy sulfur atoms lower the ISP, and no sulfur-containing PTM has ISPs much above 1.0. Overall, 21 and 26 PTMs have $\left| \Omega_{IMP} \right|$ over 60 and from 20-60 for supposed strong and substantial domain delineation, respectively. To compare, the values for amino acids lie between -18 (cysteine) and 19 (leucine or isoleucine): by the same criteria, no single amino acid would substantially move a peptide from the mean trend (Figure 31). The greater effect of PTMs is due to both their wider mass range and richer compositional diversity that result in a wider ISP range.

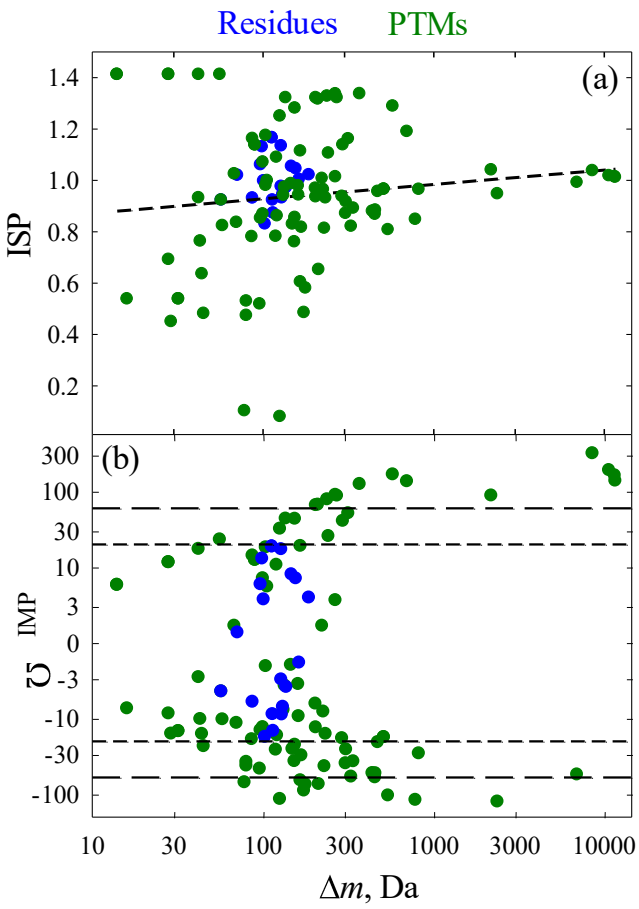

Figure 31: Calculated ISPs (a) and impact scores (b) of 94 PTMs from Table 1 (radii set 1) as a function of net peptide mass change (six PTMs with $\Delta m < 0$ or ISP outside of the 0-1.5 range are not shown). The values for amino acid residues are added for comparison. The mass and Ω_{IMP} axes are on the log scale. The dashed line (a) is the first-order regression for ISPs of PTMs. The short and long dashed lines (b) mark the absolute Ω_{IMP} of 20-60.

4.3 Conclusions

Besides raising the utility of IMS in proteomics, ISPs convey information about the global influence of an amino acid or PTM on the tertiary peptide structure. That is, in experiment entities compacting the peptide by attracting and tightly packing the surrounding groups (through intramolecular solvation) would have low ISPs, whereas entities repelling those groups and thus loosening the 3-D structure would have high ISPs. The issue is that one must compare the measured ISPs to theoretical values (Table 7) rather than the average ISP ~ 1.0;

otherwise a correlation of density would be mistaken for structural effects. Specifically polar and aromatic residues that have small ISPs were believed to contract peptides by charge-dipole and dipole-dipole interactions (with other residues or the backbone) or aromatic ring stacking [38, 55-57], while residues with long aliphatic side chains were rationalized to possess freer conformations that expand the peptide geometries. A close match between the measured and calculated ISPs for both categories has proven those effects, while plausible and potentially pertinent for individual species, to be inoperative on average [35]. Similarly, small ISPs of cam and especially p were argued to be a manifestation of polar groups attracted to the charged sites [39] or intramolecular interactions that lead to structural compaction [38], and a large ISP for pal was thought to reflect its length and hydrophobicity [39]. Present findings clarify that these PTMs do not affect the average tightness of peptide folding.

Extending a priori calculations of intrinsic size parameters from residues [35] to PTMs permits predicting the mobilities of modified peptides. The resulting ISPs match those for PTMs measured so far (phosphorylation, carboxyamidomethylation, and palmitoylation). Along with the agreement for amino acid residues [35], this validates the approach and shows that the ISPs for both amino acids and PTMs reflect primarily their density rather than cooperative intramolecular interactions. The agreement of ISPs for PTMs formed from nucleotides, sugars, peptides, and lipids with the established arrangement of IMS/MS domains from that for unmodified peptides, and some others may do so with multiple modification. These estimates ought to help plan IMS/MS analyses of modified proteomes and improve the quality and speed of identifications. They should be utilized only statistically, though, as for any individual peptide the effect of specific geometry may outweigh that of the density of constituents.

CHAPTER 5

BUILDING INTO THE FUTURE

5.1 Growing into Larger Species

Di- and tri- alanines are not the end of the road. The extension of this project into 4 and 5 amino acids with individually labeled amino acids (similar to previously used) provides two things. First it would give a chance to try predicting shift changes, if they follow the trend presented with the closer to the amino end the farther the shift. Second, it would explore how large the peptide must be before the size of the peptide outweighs the specificity of a fully labeled amino acid.

Another option, switching to test viability with modern techniques, would include sticking a label on the large peptide fragment and determining how large a label is needed before it affects the FAIMS spectra. Other work including PTM localization on larger peptide sequences [58] indicates this is a viable option. Combining this localization of a label on a peptide with the separation of isobaric species has an interesting consequence, leading down the path to another way to determine in what manner proteins are made. Order of amino acid incorporation can be determined with proper amino acid identification and clever amino acid insertions. The knowledge from peptide creation can help track down problems about why proteins become faulty.

5.2 Cycling Forward

Monohalogenated anilines are only the tip of the iceberg. Next in line would be the dihalogenated anilines, then maybe trihalogenated anilines (where available). These molecules could be extended as far as rigid multiringed molecules. These would assist in exploring how far identification of isomers can be pushed. For starters sticking to rigid cyclic molecules limits the

number of enantiomers and other such assorted isomers from interfering. Limiting the number of isomers is important for the first attempts at identification based on shifts, though it should be carefully introduced as confidence in the approach grows.

5.3 Understanding FAIMS

The difference in shifts (not just shift intensities) between 2MCA and 2MBA indicate that something has changed, probably beyond the center of mass, for such a drastic change. The factor that overpowers the center of mass could be attributed to several things. There could be solvent interactions from improper desolvation that change how one of the halogens reacts. Maybe chlorine is small enough to wedge a methanol molecule between it and the amine group while bromine is too big to allow that, or chlorine is too small to trap the methanol but bromine was large enough. Though due to the high voltage of the ESI needle and the internal heating through FAIMS desolvation is usually considered complete for small molecules. This clustering is also not detected in the mass spectrometer since the major peaks only matched the monohalogenated aniline without a +32 Da from methanol. Running more samples with different solvents such as isopropanol that would be too large to become trapped in the molecule could be a path for future insight.

There could be other causes, it could be interactions with the carrier gas due to the polarizability of carbon dioxide. 2MBA starts in a very different spot and with added helium starts to drop until it starts to resemble the expected pattern from 2MCA near the breakdown threshold with He/CO_2. Since this effect is in place when the buffer gas is mostly CO_2 it probably stems from the polarizability of CO_2 in a high electric field and its interaction with the less electronegative bromine and with interactions to the amine group. This would explain the drastic difference in shifts between 2MBA and 2MCA isomers. The chlorine is too small to

cause the CO_2 to interact with both the chlorine and the amine group, and its electronegativity is closer to that of oxygen than to that of bromine, making it less likely to interact as well. In the 3- and 4- isomers the halogen atoms are so far away that they do not have the same potential for those interactions coinciding, and so are not affected. Having an electron density map of each molecule may provide some insight regarding that idea. Running the MBA isomers in nitrogen (and the MCA ones in more detail) does not further the ability to identify isomers by their shifts, but could demonstrate how the gas is affecting the mobilities.

5.4 ISPs, Innovations, and Inorganic PTMs

PTMs are not limited to the list analyzed in chapter 4. There remain various other PTMs, including inorganic PTMs including metalation. Metals complicate PTM calculation. First, since many metals involved in PTMs have multiple charge states, the ionic radii become important. Van der Waals radii may not be applicable after rising above the +1 charge, then ionic radii could become more important. Second, the added localized charges could affect the structural makeup of the peptides, more than the density represented by the ISP. These might combine to say that the +1 metals affect the ISPs different than the +2 or +3 metals, but without exploration by both a priori and experimental this remains uncertain.

Although ISPs provide more accurate peptide mobilities than the mass alone, there are major limitations. The IMS resolution of sequence [59] and PTM localization [38, 60] isomers means that the positions of both amino acids and PTMs (not captured by current ISPs) matter. The influence of sequence could possibly be emulated to some extent by sequence specific ISPs for pairs of adjacent residues [61]. However, this escalates the number of ISPs by at least an order of magnitude, and eliciting their statistically significant values requires a much larger and more diverse experimental set. Approaches considering the environment can also be devised for

PTMs, although no direct equivalent exists as the neighboring PTMs on the backbone are often too far apart to interact. One enhancement may be to treat same PTMs on different amino acids separately, although phosphorylations of serine, tyrosine, and threonine appear to have the same effect as stated. Eventually, ISPs would be supplanted by sophisticated artificial neural networks—the machine learning algorithms that integrate numerous (frequently non-obvious) structural descriptors to predict chemical properties. Such models that incorporate the peptide sequence and size have succeeded for chromatographic retention [62, 63] and should work here, but require massive training sets.

5.5 Overall conclusions

All three projects have succeeded. The labeled isotopomers were separated completely for dialanine and partly for trialanine. With both monochloroaniline and monobromoaniline, full separation of isomers with identification based on the +1 Da isotopologic shifts in FAIMS was achieved. Lastly, my calculated ISPs for PTMs have matched the published measurements, validating our a priori approach to predicting the mobilities of modified peptides based on the ISP calculated from first principles and the new "impact score" quantity that we introduce.

REFERENCES

REFERENCES

[1] Watson, T. J.; Sparkman, D. O.: *Introduction to Mass Spectrometry: Instrumentation, Applications and Strategies for Data Interpretation*. 4th ed. John Wiley and Sons, Hoboken New Jersey, (2008).

[2] Thompson, J. J.: Rays of positive electricity, *Proc. R. Soc. A* **89**, 1-20 (1913).

[3] Hiltzik, M.: *Big science: Ernest Lawrence and the invention that launched the military-industrial complex.* Simon and Schuster, New York (2015).

[4] Shilo, Y.; Aebersold, R.: Quantitative proteome analysis using isotope coded affinity tags and mass spectrometry. *Nat. Protoc.* **1**, 139-145 (2006).

[5] Ong, S. E.; Mann, M.: Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods Mol. Biol.* **359**, 37-52 (2007).

[6] Phanstiel, D.; Zhang, Y.; Marto, J. A.; Coon, J. J.: Peptide and protein quantification using iTRAQ with electron transfer dissociation. *J. Am. Soc. Mass Spectrom.* **19**, 1255-1262 (2008).

[7] Wiechert, W.: ^{13}C Metabolic flux analysis. *Metab. Eng.* **3** 195-206 (2001).

[8] Koubaa, M.; Mghaieth, S.; Thomasset, B.; Roscher, A.: Gas chromatography-mass spectrometry analysis of ^{13}C labeling in sugars for metabolic flux analysis. *Anal. Biochem***. 425,** 183-188 (2012).

[9] Allen, K. D.; Bates, P. D.; Tjellstrom, H.: Tracking the metabolic pulse of plant lipid production with isotopic labeling and flux analyses: past, present, and future. *Prog. Lipid Res.* **58,** 97-120 (2015).

[10] Murrell, J. C., Whiteley, A. S.: *Stable isotope probing and related technologies*. ASM Press, Washington DC (2010).

[11] Silverstein, R. M.; Webster, F. X.; Kiemle, D. J.: *Spectrometric Identification of Organic Compounds*. 7th ed. John Wiley and Sons Hoboken, NJ (2005).

[12] Eiceman, G. A.; Karpas, Z.; Hill, H. H.: *Ion mobility spectrometry* CRC Press, Boca Raton (2013).

[13] Valentine, S. J.; Clemmer, D. E.: Treatise on the Measurement of Molecular Masses with Ion Mobility Spectrometry. *Anal. Chem*. **81**, 5876-5880 (2009).

[14] Shvartsburg, A. A.: *Differential Ion Mobility Spectrometry: Nonlinear Ion Transport and Fundamentals of FAIMS.* Boca Raton, FL (2009).

[15] Lalli, P. M.; Corilo, Y. E.; Fasciotti, M.; Riccio, M. F.; de Sa, G. F.; Daroda, R. J.; Souza, G. H. M. F.; McCullagh, M.; Bartberger, M. D.; Eberlin, M. N.; Campuzano, I. D. G.: Baseline resolution of isomers by traveling wave ion mobility mass spectrometry: investigating the effects of polarizable drift gases and ionic charge distribution. *J. Mass Spectrom.* **48**, 989-997 (2013).

[16] Shvartsburg, A. A.; Isaac, G.; Leveque, N.; Smith, R. D.; Metz, T.O.: Separation and classification of lipids using differential ion mobility spectrometry. *J. Am. Soc. Mass Spectrom*. **22**, 1146-1155 (2011).

[17] Shvartsburg, A. A.; Creese, A. J.; Smith, R. D.; Cooper, H. J.: Separation of peptide isomers with variant modified sites by high-resolution differential ion mobility spectrometer. *Anal. Chem*. **82**, 8327-8334 (2010).

[18] Shvartsburg, A. A.; Zheng, Y.; Smith, R. D.; Kelleher, N.: Ion mobility separations of variant histone tails extending to the middle-down range. *Anal. Chem.* **84**, 7271-7276 (2012).

[19] Shvartsburg, A. A.; Seim, T. A.; Danielson, W. F.; Norheim, R.; Moore, R. J.; Anderson, G. A.; Smith, R. D.: High-definition differential ion mobility spectrometry with resolving power up to 500. *J. Am. Soc. Mass Spectrom*. **24**, 109-114 (2014).

[20] Shvartsburg, A. A.; Smith, R. D.: High resolution differential ion mobility spectrometry of a protein. *Anal. Chem.* **85,** 10-13 (2013).

[21] Shvartsburg, A. A.: Ultra high-resolution differential ion mobility separations of conformers for proteins above 10 kDa--onset of dipole alignment? *Anal. Chem.* **86**, 10608-10615 (2014).

[22] Shvartsburg, A. A.; Tang, K.; Smith, R. D.: Understanding and designing field asymmetric waveform ion mobility spectrometry separations in gas mixtures. *Anal. Chem.* **76,** 7366-7374 (2004).

[23] Shi, S. D. H.; Hendrickson, C. L.; Marshall, A. G.: Counting individual sulfur atoms in a protein by ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry: experimental resolution of isotopic fine structure in proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 11532-11537 (1998).

[24] Winter, D.; Pipkorn, R.; Lehmann, W. D: Separation of peptide isomers and conformers by ultra-performance liquid chromatography. *J. Sep. Sci.* **32**, 1111-1119 (2009).

[25] Szarka, S.; Prokai-Tatrai, K.; Prokai, L.: Application of screening experimental designs to assess chromatographic isotope effect upon isotope coded derivatization of quantitative liquid chromatography-mass spectrometry. *Anal. Chem.* **86,** 7033-7040 (2014).

[26] Keller, J.: *Understanding NMR spectroscopy*. Wiley, NT (2010).

[27] Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Clemmer, D. E.: Resolving and Assigning N-linked Glycan Structural Isomers from Ovalbumin by IMS-MS. *J. Am. Soc. Mass Spectrom.* **19**, 1706-1715 (2008).

[28] Shu, M.; Bendiak, B.; Clowers, B.; Hill, H. H. Jr.: Ion mobility-mass spectrometry analysis of isomeric carbohydrate precursor ions. *Anal. Bioanal. Chem*. **394**, 1853-1867 (2009).

[29] Shvartsburg, A. A.; Clemmer, D. E.; Smith, R. D.: Isotopic effect on ion mobility and separation of isotopomers by high-field ion mobility spectrometry. *Anal. Chem*. **82**, 8074-8051 (2010).

[30] Guevremont, R.; Barnett, D. A.; Purves, R. W.; Vandermey, J.: Analysis of a tryptic digest of pig hemoglobin using ESI-FAIMS-MS, *Anal. Chem.* **72**, 4577-4584 (2000).

[31] Shvartsburg, A.A.; Liu, B.; Jarrold, M.F.; Ho, K.M.: Modeling ionic mobilities by scattering on electronic density isosurfaces: application to silicon cluster anions. *J. Chem. Phys.* **112**, 4517–4526 (2000).

[32] Alexeev, Y.; Fedorov, D.G.; Shvartsburg, A.A.: Effective ion mobility calculations for macromolecules by scattering on electron clouds. *J. Phys. Chem. A* **118**, 6763–6772 (2014).

[33] Shvartsburg, A.A.; Liu, B.; Siu, K.W.M.; Ho, K.M.: Evaluation of ionic mobilities by coupling the scattering on atoms and on electron density. *J. Phys. Chem*. **104**, 6152–6157 (2000).

[34] Valentine, S.J.; Ewing, M.A.; Dilger, J.M.; Glover, M.S.; Geromanos, S.; Hughes, C.; Clemmer, D.E.: Using ion mobility data to improve peptide identification: cross section databases and intrinsic amino acid size parameters. *J. Proteome Res.* **10**, 2318–2329 (2011).

[35] Shvartsburg, A.A.; Siu, K.W.M.; Clemmer, D.E.: Prediction of peptide ion mobilities via a priori calculations from intrinsic size parameters of amino acid residues. *J. Am. Soc. Mass Spectrom.* **12**, 885–888 (2001).

[36] Ruotolo, B.T.; Verbeck IV, G.F.; Thomson, L.M.; Woods, A.S.; Gillig, K.J.; Russell, D.H.: Distinguishing between phosphorylated and nonphosphorylated peptides with ion mobility-mass spectrometry. *J. Proteome Res*. **1**, 303–306 (2002).

[37] Ruotolo, B.T.; Gillig, K.J.; Woods, A.S.; Egan, T.F.; Ugarov, M.V.; Schultz, J.A.; Russell, D.H.: Analysis of phosphorylated peptides by ion mobility-mass spectrometry. *Anal. Chem.* **76**, 6727–6733 (2004).

[38] Glover, M.S.; Dilger, J.M.; Acton, M.D.; Arnold, R.J.; Radivojac, P.; Clemmer, D.E.: Examining the influence of phosphorylation on peptide ion structure by ion mobility spectrometry–mass spectrometry. *J. Am. Soc. Mass Spectrom*. **27**, 786–794 (2016).

[39] Li, Z.; Dilger, J.M.; Pejaver, V.; Smiley, D.; Arnold, R.J.; Mooney, S.D.; Mukhopadhyay, S.; Radivojac, P.; Clemmer, D.E.: Intrinsic size parameters for palmitoylated and carboxyamidomethylated peptides. *Int. J. Mass Spectrom*. **368**, 6–14 (2014).

[40] Barnett, D. A.; Ells, B.; Guevremont, R.; Purves, R. W.; Vichland, L. A.: Evaluation of carrier gasses for use in high-field asymmetric waveform ion mobility spectrometry. *J. Am. Soc. Mass Spectrom.* **11,** 1125-1133 (2000).

[41] Shvartsburg, A. A.; Danielson, W.; Smith, R. D.: High-Resolution Differential Ion Mobility Separations Using Helium-Rich Gases. *Anal. Chem.* **82**, 2456-2462 (2010).

[42] Cimbala, J. M.: Outliers. *Penn State University*, Sept. 2011.

[43] Boni, A: van der Waals volumes and radii. *J. Phys. Chem.* **68**, 441-451 (1964).

[44] Mantina, M.; Chamberlin, A. C.; Valer, R.; Cramer, C. J.; Truhlar, D.G.: Consistent van der Waals radii for the whole main group. *J. Phys. Chem. A.* **113**, 5806-5812 (2009).

[45] Fenn, L. S.; Kliman, M.; Mahsut, A., Zhao, S. R.; McLean, J.A.: Characterizing ion mobility-mass spectrometry conformation space for the analysis of complex biological samples. *Anal. Bioanal. Chem.* **394**, 235-244 (2009).

[46] May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L., Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Staffford, G. C.; Fieldsted, J. C.; McLean, J. A.: Conformational ordering of biomolecules in the gas phase: nitrogen collision cross sections measured on a prototype high resolution drift tube ion mobility-mass spectrometer. *Anal. Chem.* **86**, 217-2116 (2014).

[47] Haag, F.; Buck, F.: Identification and analysis of ADP-ribosylated proteins. *Curr. Top. Microbiol*. **384**, 33-50 (2015).

[48] Morris, S. A.; Rao, B.; Garcia, B. A.; Hake, S. B.; Diax, R. L., Shabanowitz, J.; Hunt, D. F.; Allis, C. D.; Lieb, J. D.; Strahl, B. D.: Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *J. Biol. Chem.* **282**, 7632-7640 (2007).

[49] Thalassinos, K.; Grabenauer, M.; Slade, S. E.; Hilton, G. R.; Bowers, M.T.; Scrivens J. H.: Characterization of phosphorylated peptides using traveling wave-based and drift cell ion mobility mass spectrometry. *Anal. Chem*. **81**, 248-254 (2009).

[50] Shvartsburg, A. A.; Singer, D.; Smith, D. R.; Hoffmann, R.: Ion mobility separation of isomeric phosphopeptides from a protein with variant modification of adjacent residues. *Anal. Chem.* **83**, 5078-5085 (2011).

[51] Mair, W.; Muntel, J.; Tepper, K.; Tang, S.; Biernat, J.; Seeley, W. W.; Kosik, K. S.;Mandelkow, E.; Steen, H.; Steen, J. A.: FLEXITau: quantifying post-translational modifcations of tau protein in vitro and in human disease. *Anal. Chem*. **88**, 3704-3714 (2016).

[52] Sidoli, S.; Schwammle, V.; Ruminowicz, C.; Hansen, T. A.; Wu, X.; Helin, K.; Jensen, O. N.: Middle-down hybrid chromatography/tandem mass spectrometry workflow for characterization of combinatorial post-translational modifications in histones. *Proteomics* **14**, 2200-2211 (2014).

[53] Fraga, M. F.; Ballestar, E.; Villar-Garea, A.; Biox-Chornet, M.; Espada, J.; Schotta, G.; Bonaldi, T.; Haydon, C.; Ropero, S.; Petrie, K.; Iyer, N. G.; Perez-Rosado, A.; Calvo, E.; Lopez, J. A.; Cano, A.; Calasanz, M. J.; Colomer, D., Piris, M. A.; Ahn, N.; Imhof, A.; Caldas, C.; Jenuwein, T.; Esteller, M.: Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391-400 (2005).

[54] Paulick, M. G.; Bertozzi, C. R.: The glycosyphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. *Biochemistry* **47**, 6991-7000 (2008).

[55] Valentine, S. J.; Counterman, A. E.; Hoaglund-Hyzer, C. S.; Clemmer, D. E.: Intrinsic amino acid size parameters from a series of 113 lysine-terminated tryptic digest peptide ions. *J. Phys. Chem. B.* **103**, 1203-1207 (1999).

[56] Henderson, S. C.; Li, J.; Counterman, A. E.; Clemmer, D. E.: Intrinsic size parameters for Val, Ile, Leu, Gln Thr, Phe, and Trp residues from ion mobility measurements of polyamino acid ions. *J. Phys. Chem. B.* **103** 8780-8785 (1999).

[57] Valentine, S. J.; Counterman, A. E.; Clemmer, D. E.: A database of 660 peptide ion cross-sections: use of intrinsic size parameters for bona fide predictions of cross sections. *J. Am. Soc. Mass Spectrom*. **10**, 1188-1211 (1999).

[58] Baird, M. A.; Shvartsburg, A. A.: Localization of post-translational modifications in peptide mixtures via high-resolution differential ion mobility separations followed by electron transfer dissociation. *J. Am. Soc. Mass Spectrom.* **27**, 2064-2070 (2016).

[59] Wu, C.; Siems, W. F.; Klasmeier, J.; Hill, H. H.: Separation of isomeric peptides using electrospray ionization/high-resolution ion mobility mass spectrometry. *Anal. Chem.* **72**, 391-395 (2000).

[60] Ibrahim, Y.; Shvartsburg, A. A.; Smith, R. D.; Belov, M. E.: Ultrasensitive identification of localization variants of modified peptides using ion mobility spectrometry. *Anal. Chem*. **83**, 5617-5623 (2011).

[61] Hilderbrand, A. E.; Clemmer, D. E.: Determination of sequence-specific intrinsic size parameters from cross sections for 162 tri-peptides. *J. Phys. Chem. B*. **109**, 11802-11809 (2005).

[62] Krokhin, O. V.: Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-Å pore size C18 sorbents. *Anal. Chem.* **78**, 7785-7795 (2006).

[63] Petritis, K.; Kangas, L. J.; Yan, B.; Monroe, M. E.; Strittmatter, E. F.; Qian, W. J.; Adkins, J. N.; Moore, R. J.; Xu, Y.; Lipton, M. S.; Camp, D. G.; Smith, R. D.: Improved peptide elution time prediction for reversed-phase liquid chromatography-MS by incorporating peptide sequence information. *Anal. Chem.* **78**, 5026-5039 (2006).

APPENDIX

APPENDIX

LIST OF REFERENCES FOR PTMS

a. www.dbptm.mbc.nctu.edu.tw
b. Catack, S.; Mondard, G.; Aviyente, V.; Ruiz-Lopez, M. F. Deamidation of asparagine residues: direct hydrolysis versus succinimide-mediated deamidation mechanisms. J. Phys. Chem. **113**, 1111 - 1120 (2009).
c. www.uniprot.org
d. Marmelstein, A. M.; Yates, L. M.; Conway, J. H.; Fiedler, D. Chemical pyrophosphorylation of functionally diverse peptides. J. Am. Chem. Soc. **136**, 108 - 111 (2014).
e. Wolfrom, M. L.; Bhat, H. B. Trichloroacetyl and trifluoroacetyl as N-blocking groups in nucleoside synthesis with 2-amino sugars. J. Org. Chem. **32**, 1821 - 1823 (1967).
f. Larsen, M. R.; Trelle, M. B.; Thingholm, T. E.; Jensen, O. N. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. Biotechniques **40**, 790 - 798 (2006).
g. www.ebi.ac.uk/ols/index
h. www.chemicalbook.com
i. www.ionsource.com
j. Xie, Z.; Dai, J.; Dai, L.; Tan, M; Cheng, Z; Wu, Y; Boeke, J. D.; Zhao, Y. Lysine succinylation and lysine malonylation in histones. Mol. Cell. Proteomics. **11**, 100 - 107 (2012).
k. www.sigmaaldrich.com
l. www.nextprot.org
m. Christie, J. M. Phototropin blue-light receptors. Annu. Rev. Plant Biol. **58**, 21- 45 (2007).
n. Satoh, T.; Yamaguchi, T.; Kato, K. Emerging structural insight into glycoprotein quality control coupled with N-glycan processing in the endoplasmic reticulum. Molecules **20**, 2475 - 2491 (2015).
o. www.carbosynth.com
p. www.phosphosite.org/homeAction
q. Fraga, M. F.; Ballestar, E.; Villar-Garea, A.; Biox-Chornet, M.; Espada, J.; Schotta, G.; Bonaldi, T.; Haydon, C.; Ropero, S.; Petrie, K.; Iyer, N. G.; Perez-Rosado, A.; Calvo, E.; Lopez, J. A.; Cano, A.; Calasanz, M. J.; Colomer, D., Piris, M. A.; Ahn, N.; Imhof, A.; Caldas, C.; Jenuwein, T.; Esteller, M.: Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391-400 (2005).
r. Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.; Zagrovic, B. A systematic framework for molecular dynamics simulations of protein post-translational modifications. PLOS Comput. Biol. **9**, e1003154 (2013).