Complex Mixture Analysis Using Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS) and Computational Techniques

Hossein Maleki
Complex Mixture Analysis Using Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS) and Computational Techniques

Hossein Maleki

Dissertation submitted to the Eberly College of Arts and Sciences at West Virginia University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Chemistry

Stephen J. Valentine, Ph.D., Committee Chairperson
Brian Popp, Ph.D.
Jonathan Boyd, Ph.D.
Harry O. Finklea, Ph.D.
H. Ilkin Bilgesu, Ph.D.

C. Eugene Bennett Dept. of Chemistry
Morgantown, West Virginia
2017

Key Terms: ion mobility, mass spectrometry, hydrogen deuterium exchange (HDX), ‘omics analysis, isotope scrambling, collision induced dissociation, predictive tools for HDX, petroleomics, metabolomics, lipidomics

Copyright 2017 Hossein Maleki
Compound characterization in complex mixtures has applications in diverse areas such as the pharmaceutical, health, and petroleum industries. Measurement problems arising from complex mixtures include the overwhelming of the analytical peak capacity, the wide concentration range of the different analytes, and the inter-sample concentration variability of given molecular species. Therefore advances in the rapid identification of compounds such as improvements in peak capacity and analyte selectivity are an active area of scientific research. This dissertation presents research that is directed at improving these measurement areas. For petroleum samples, silver cationization coupled with ion mobility spectrometry – tandem mass spectrometry (IMS-MS/MS) is presented as a means to improve compound selectivity (sulfur-containing molecules) and species identification. The approach showed that silver-sulfur ions demonstrate higher mobilities compared with protonated nitrogen-containing compounds. Furthermore, at lower m/z values, mostly elongated structures (with lower DBE values) were favored. Conversely, more compact structures (with higher DBE values) were favored at higher m/z values. Additionally, cyclic structures were favored over acyclic structures and S₁, S₂, and S₃ molecular classes were favored over other molecular classes in sulfur-containing petroleum compounds. In separate studies, hydrogen/deuterium exchange (HDX), hydrogen deuterium scrambling (HD scrambling), and collision-induced dissociation (CID) were coupled with IMS-MS measurements to assist in the characterization of metabolites, lipids, and peptides. Unique mass spectral patterns were obtained for different classes of biological molecules upon HDX. The use of these HDX patterns in combination with CID and collision cross section (CCS) data were demonstrated as an aid for the identification efforts for dataset features from a bovine heart extract sample. Lastly, HDX kinetics modeling was used to estimate the relative abundances of sucrose conformers contributing to experimental HDX profiles. These advances in compound identification have implications for the health and energy industries. An example for the former is the discovery of important biomarkers associated with disease progression. For the latter, such characterizations may help to optimize the refining process for different petroleum samples. Future trends to improve these new measurements are presented and discussed.
**Dedication**

I dedicate this work to my brother, Mohammad Hassan, and to my parents Kobra and Yahya. You were great inspirations for me to finish this path.

**Acknowledgements**

I would like to thank my research mentor Dr. Stephen J. Valentine for all the knowledge and skills he taught me. I learned a lot about science and life from him. I am greatly honored to be his student. I would like to express my thanks to all my committee members: Dr. Stephen J. Valentine, Dr. Jonathan Boyd, Dr. Brian V. Popp, Dr. Harry O. Finklea, and Dr. H. Ilkin Bilgesu for serving on my committee and helping me. I would like to thank my peers, Halle Edwards, Ahmad Kiani Karanji, Greg C. Donohoe, Megan M. Maurer, Kushani Attanayake, Sandra N. Majuta, Jim Arndt, Mahdiar Khakinejad, and Samaneh Kondalaji.
## TABLE OF CONTENTS

Abstract ii

Dedication iii

Acknowledgements iii

List of Contents v

List of Figures viii

List of Tables xii

List of Symbols and Abbreviations xiv
# List of Contents

## 1. Compound Characterization in Complex Mixtures Using Hydrogen Deuterium/Exchange (HDX) and Isotope Scrambling Coupled to Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS) and Collision-Induced Dissociation (CID)

1.1. Complex Mixture Analysis .............................. 1

1.2. Methods Used for Complex Mixture Analysis ...... 3

1.3. IMS-MS Basics ........................................... 5

1.3.1. IMS-MS Instrumentation: Linear Drift Tube Ion mobility Spectrometry Coupled to Linear Ion Trap Mass Spectrometer .............. 7

1.3.2. Application of IMS-MS for Complex Mixture Analysis .... 8

1.4. Applications of HDX Combined with IMS-MS ....... 10

1.4.1. Hydrogen Deuterium Scrambling (HD Scrambling) as an Additional Tool to Assist Compound Identification ......... 12

1.5. Application of Fragmentation Techniques for Complex Mixture Analysis .................. 13

1.5.1. Collision-Induced Dissociation ............... 15

1.6. HDX Kinetics Modeling ................................. 16

1.7. References .............................................. 19

## 2. Structural Assignments of Sulfur-containing Compounds in Crude Oil Using Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS)

2.1. Introduction ............................................. 27

2.2. Experimental Section .................................. 32

2.2.1. Sample Preparation ................................ 32

2.2.2. Instrumentation ................................... 32

2.2.3. Experimental and Theoretical CCS Calculations .......... 34
2.3. Results and Discussion

2.3.1. Comparison of Mass Spectra and 2D Plots for Two Samples

2.3.2. Criteria for Proposing Structures for Matching Molecular Formulas

2.3.3. Proposed Structures for Ions Having Nominal $m/z$ Values of 291 and 293

2.3.4. Proposed Structures for Ions Having Nominal $m/z$ Values of 353 and 355

2.4. Conclusion

2.5. References

3. Compound Characterization Using Hydrogen/Deuterium Exchange (HDX) and Isotope Scrambling Coupled to IMS-MS and CID

3.1. Introduction

3.2. Experimental

3.2.1. Materials

3.2.2. Sample Preparation

3.2.3. Data Collection Method

3.3. Results and Discussion

3.3.1. HDX and HD Scrambling for Distinguishing Molecules

3.3.2. Application of HDX and Scrambling Experiments for Metabolite Analysis

3.3.3. Drift Time Reproducibility after HDX Experiments

3.3.4. Isotopic Distribution Reproducibility in HDX Experiments

3.3.5. Mass Spectral Reproducibility after Isotope Scrambling Followed by HDX

3.3.6. HDX Behavior to Aid Compound Identification
### 4. Ion Mobility Spectrometry-Mass Spectrometry Coupled with Gas-phase Hydrogen/Deuterium Exchange for Metabolomics Analyses

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1. Introduction</td>
<td>95</td>
</tr>
<tr>
<td>4.2. Experimental Section</td>
<td>98</td>
</tr>
<tr>
<td>4.2.1 Sample Preparation</td>
<td>98</td>
</tr>
<tr>
<td>4.2.2. Instrumentation</td>
<td>101</td>
</tr>
<tr>
<td>4.3. Simulated Annealing and HDX Kinetics Modeling</td>
<td>102</td>
</tr>
<tr>
<td>4.4. Results and Discussion</td>
<td>104</td>
</tr>
<tr>
<td>4.4.1. Compound Specific HDX Behavior</td>
<td>104</td>
</tr>
<tr>
<td>4.4.2. Distinguishing Isobaric Ions</td>
<td>109</td>
</tr>
<tr>
<td>4.4.3. HDX Behavior of Lipid Ions</td>
<td>113</td>
</tr>
<tr>
<td>4.4.4. Structural Explanations for HDX Behavior</td>
<td>116</td>
</tr>
<tr>
<td>4.4.5. Proof-of-principle Application of HDX Data for Compound Identification</td>
<td>116</td>
</tr>
<tr>
<td>4.5. Conclusions</td>
<td>125</td>
</tr>
<tr>
<td>4.6. References</td>
<td>126</td>
</tr>
</tbody>
</table>

### 5. Future Directions: Developing New Methods to Improve Compound Identifications in the Complex Mixtures

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1. Construct a Library of CCS Values and HDX Distributions</td>
<td>131</td>
</tr>
<tr>
<td>5.2. Prediction of CCS Values for Biomolecules Using Intrinsic Size Parameters</td>
<td>132</td>
</tr>
<tr>
<td>5.3. Conformational Analysis of Biomolecules in the Gas Phase: Matching Theoretical CCS Values with Experimental CCS Profiles</td>
<td>133</td>
</tr>
</tbody>
</table>
5.4. Prediction of HDX Patterns for Metabolites and Lipids

5.5. Application of HDX Profiles to Assist Locating the Cis and Trans Double Bonds in Lipids at Elevated Temperatures

5.6. Application of New Fragmentation Techniques Followed by HDX

5.7. Application of Shift Reagents to Improve Characterization of Specific Classes of Biomolecules in 2D Plots of $m/z$ vs Drift Time

5.8. References

List of Figures

1.1. IMS-MS configuration in which a linear drift tube outfitted with ESI capabilities is coupled to a linear ion trap mass spectrometer

2.1. Schematic diagram of the instrument used in these studies. The home-built drift tube and linear ion trap mass spectrometer are shown. Funnels, activation regions, and gates are labeled.

2.2. Two-dimensional (2D) plot of $m/z$ and $t_D$ values obtained upon direct electrospray of Saudi crude oil co-incubated with silver ions. Intensities are represented on a color scale where red and black are the highest and lowest intensities, respectively. Three white lines are associated with molecular structures with different degrees of compactness. The theoretical locations of three of these structures within the 2D plot are shown. B) $t_D$ distributions of silver-containing samples and protonated crude oil samples for two $m/z$ ranges (650-660 and 550-560 for top and bottom traces, respectively). Solid and dashed lines are associated with silver-containing and protonated samples, respectively. Intensity values have been scaled to the tallest dataset feature representing a peak height of 1.

2.3. Expanded regions in 2D IMS-MS plots for A) ions with nominal $m/z$ values of 665 and 667 and, B) 683 and 685. Respective $t_D$ distributions for
panels A and B are shown in panels C and D. Green and blue traces represent lighter and heavier silver isotopes, respectively. Intensities in panels C and D have been normalized by dividing the y-value at each drift time by the sum of all y-values.

2.4. A) \( t_D \) profile for ions having nominal \( m/z \) values of 291 and 293. \( t_D \) values for each number are correlated to structures (Table 2.6) as shown with arrows. B) CID spectrum for the dataset feature of nominal \( m/z \) value of 291 obtained with the LTQ device of the FT-ICR instrument. Several fragments are labeled.

2.5. A) \( t_D \) profile for ions having nominal \( m/z \) values of 353 and 355. \( t_D \) values for each number are correlated to structures (Table 2.7) as shown with arrows. B) CID spectrum for the dataset feature of nominal \( m/z \) 353 obtained with the LTQ device of the FT-ICR instrument. Several fragments are labeled.

3.1. Different deuterium uptake patterns for model peptide ions. A) \([M+3H]^{3+}\) bradykinin ions. B) \([M+3H]^{3+}\) substance P ions. C) \([M+2H]^{2+}\) KKD peptide ions. The black line is the isotopic distribution in the absence of D\(_2\)O and the blue line is the isotopic distribution in the presence of D\(_2\)O.

3.2. Different deuterium uptake patterns for model peptides after HDX and HD scrambling with D\(_2\)O added to the drift tube after ion activation. A) \([M+3H]^{3+}\) bradykinin ions. B) \([M+3H]^{3+}\) substance P ions. C) \([M+2H]^{2+}\) KKD peptide ions. The black line is the isotopic distribution in the presence of D\(_2\)O and the blue line is the isotopic distribution in the presence of D\(_2\)O and ion activation.

3.3. Isotopic distributions of precursor ions of A) alanine, C) guanosine, E) dopamine, and G) acetaminophen. Distributions for mobility selected ions in the presence of He only and He and D\(_2\)O are shown as blue and orange traces, respectively. Panels B, D, F, and H show the isotopic distributions (green traces) for the respective molecular ions obtained after collisional activation in the drift tube.

3.4. Two dimensional (2D) drift time, \( m/z \) plots for triplicate measurements of a ubiquitin digest mixture. The data were recorded for conditions employing He buffer gas and random mobility selection. Several dataset features are indicated by numbers and correspond to the assigned ions in Table 3.2.

3.5. Two dimensional (2D) drift time, \( m/z \) plots for triplicate measurements of a ubiquitin digest mixture. The data were recorded for conditions employing He buffer gas and sequential mobility selection. Several dataset
features are indicated by numbers and correspond to the assigned ions in Table 3.3.

3.6. IMS-MS profile of two ubiquitin digest trials (A) demonstrating the run-to-run reproducibility of compounds in a complex mixture after addition of D₂O to the drift tube. Panel B shows the CV for select peptide ions. m/z and average drift time values are provided for the peptide ions. CV values for dataset features tentatively identified as 1- QRLIFAGKQ+3H⁺; 2- [NIQ+H⁺]; 3- [HLVLRL+2H]²⁺; 4- [GIPPDQQ+2H]²⁺; 5- [KIQDKE+2H]²⁺; 6- [VLRLRGG+2H]²⁺; 7- [IFAGKQL+2H]²⁺; 8- [DKE+H⁺]; 9- [VKTLTGKTITL+3H]³⁺ ions as labeled in (B).

3.7. Isotopic distributions for [M+2H]²⁺ KKD (A), [M+2H]²⁺ substance P (B), and [M+3H]³⁺ bradykinin (C) peptide ions. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.

3.8. Isotopic distributions for [M+3H]³⁺ KKD (A), [M+3H]³⁺ substance P (B), and [M+4H]⁴⁺ KKD (C) peptide ions. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.

3.9. Isotopic distributions for [M+2H]²⁺ KKD (A), [M+2H]²⁺ substance P (B), and [M+3H]³⁺ bradykinin (C) peptide ions. These data were recorded under conditions in which the ions were collisionally activated to induce increased deuterium incorporation. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.

3.10. Isotopic distributions for [M+3H]³⁺ KKD (A), [M+3H]³⁺ substance P (B), and [M+4H]⁴⁺ KKD (C) peptide ions. These data were recorded under conditions in which the ions were collisionally activated to induce increased deuterium incorporation. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.

3.11. Isotopic distributions of precursor ions having nominal m/z 90 obtained from a bovine heart extract. Panel A shows distributions for ions in the presence of He only and He and D₂O as blue and orange traces, respectively. Panel B shows the isotopic distribution (green trace) for these ions obtained after collisional activation in the drift tube.

3.12. CID spectra for alanine at A) 15 NCE and B) 25 NCE. CID spectra for a feature with nominal m/z value of 90 within the bovine heart metabolite extract sample at C) 15 NCE and D) 25 NCE.
4.1. [M-H]- sucrose ion conformers obtained from simulated annealing (see manuscript for details). A) Conformer with a deprotonated oxygen that is surrounded by four –OH groups. B) Conformer with a deprotonated oxygen that is surrounded by only one –OH group.

4.2. Shows 2D plots of m/z vs t\text{D} for a mixture of lysine, acetaminophen, and dopamine A) in the absence of D\text{2}O reagent gas and B) in the presence of D\text{2}O reagent gas. Structures of ions associated to peaks shown with white arrows.

4.3. Isotopic distributions obtained using a D\text{2}O partial pressure of ~0.02 Torr for A) [M+H]+ hydroxyproline ions, C) M+ 18:1 Lyso PC ions, E) [M+H]+ 14:1 PE ions, and G) M+ 18:1 SM ions. Isotopic distributions obtained using a D\text{2}O partial pressure of ~0.04 Torr for B) [M+H]+ hydroxyproline ions, D) M+ 18:1 Lyso PC ions, F) [M+H]+ 14:1 PE ions, and H) M+ 18:1 SM ions. Solid lines show the distribution obtained in the absence of D\text{2}O and the dashed lines show the isotopic distribution with the addition of the respective partial pressure of D\text{2}O to the drift tube buffer gas.

4.4. Drift time distribution for [M+H]+ 14:1 PE ions. Two conformers associated with these ions are partially resolved. See Table 1 in the manuscript for CCS values for the conformers.

4.5. Isotopic distributions for [M+H]+ A) adenine and B) homocysteine ions and for [M+H]+ C) lysine and D) glutamine ions. Solid lines show the distribution obtained in the absence of D\text{2}O and the dashed lines show the isotopic distribution with the addition of ~0.04 Torr D\text{2}O to the drift tube buffer gas.

4.6. Isotopic distributions for a mixture of two isobaric lipids (15:0 Lyso PC and 07:0 PC (DHPC)) at t\text{D} selections of A) 10.8-11.0, B) 11.2-11.4, and C) 11.4-11.6. The solid lines demonstrates isotopic patterns after deuterium addition.

4.7. Isotopic distributions obtained for A) 07:0 PC (DHPC) and B) 15:0 Lyso PC. Solid lines show isotopic distribution in the absence of D\text{2}O and dashed lines shows isotopic patterns in the presence of D\text{2}O.

4.8. Isotopic distributions obtained for A) Leucine and B) Isoleucine.

4.9. Isotopic distributions obtained for A) (Δ9-Cis) PC (14:1/14:1) and B) (Δ9-Trans) PC (14:1/14:1).

Solid lines show the distribution obtained in the absence of D_2O and the dashed lines show the isotopic distribution with the addition of ~0.04 Torr D_2O to the drift tube buffer gas.

4.11. Isotopic distributions obtained for the sodiated lipids including: A) 18:1 Lyso PC; B) 18:1 Ceramide; C) PC (14:1/14:1); D) 18:1 SM; and, E) 18:1 TG lipids. Solid lines show the isotopic distribution in the absence of D_2O and dashed lines show the isotopic distribution with the addition of ~0.04 Torr D_2O.

4.12. Bar graph representing the isotopic distribution for [M-H]^− sucrose ions upon addition of ~0.04 Torr D_2O showing the relative intensities of the M+0, M+1, M+2, M+3, M+4, and M+5 peaks (red bars). The blue bars show the best-fit isotopic distribution obtained from the modeled isotopic distributions for in-silico candidate structures.

4.13. Isotopic distributions obtained using a D_2O partial pressure of ~0.04 Torr for A) selected ions having a nominal m/z value of 132 in the bovine cardiac extract and B) [M+H]^+ hydroxyproline. Fragmentation patterns produced by CID of C) selected ions having a nominal m/z value of 132 in the bovine cardiac extract and D) [M+H]^+ hydroxyproline ions.

4.14. Isotopic distributions obtained using a D_2O partial pressure of ~0.04 Torr for A) selected ions having a nominal m/z value of 147 in the bovine cardiac extract and B) [M+H]^+ lysine ions. Fragmentation patterns produced by CID of C) selected ions having a nominal m/z value of 147 in the bovine cardiac extract and D) [M+H]^+ lysine ions.

List of Tables

2.1. Structures used for drawing trend lines in Figure 2.2

2.2. Selected structures for studying the tendency of silver to interact at different locations on sulfur-containing compounds

2.3. Molecular formulas observed within nominal m/z values of 291 and 293

2.4. Molecular formulas observed within nominal m/z values of 353 and 353

2.5. Tentative assignments of sulfur-containing ions from S1 class compounds
2.6. Proposed structures for ions having nominal \( m/z \) values of 291 and 293

2.7. Proposed structures for ions having nominal \( m/z \) values of 353 or 355

3.1. Structures of metabolite compounds used for the IMS-HDX-MS experiments and their reduced mobilities.

3.2. Drift times for ubiquitin peptide ions recorded using a random drift time selection process. The buffer gas was He.

3.3. Drift times for ubiquitin peptide ions recorded using a sequential drift time selection process. The buffer gas was He.

4.1. Structures of small-molecule standards.

4.2. Structures of lipid standards.

4.3. Experimental \( DT_{CCS}^{He} \) values calculated for the molecular standards.

**LIST OF SYMBOLS AND ABBREVIATIONS**

- \( v_d \) Drift velocity
- \( K \) Mobility
- \( E \) Electric field
- \( L \) Length of drift tube
- \( t_d \) Drift time
- \( N \) The number density of the buffer gas
- \( m \) Ion mass
- \( M \) Mass of the buffer gas
- \( e \) Elemental charge
- \( k_B \) Boltzmann distribution constant
- \( T \) Temperature
- \( z \) Ion charge
- \( \Omega \) Collision cross section
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>Pressure</td>
</tr>
<tr>
<td>IMS-MS</td>
<td>Ion mobility spectrometry-mass spectrometry</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem mass spectrometry</td>
</tr>
<tr>
<td>MS(^n)</td>
<td>Multistage tandem mass spectrometry</td>
</tr>
<tr>
<td>CID</td>
<td>Collision-induced dissociation</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionization</td>
</tr>
<tr>
<td>LIT</td>
<td>Linear ion trap</td>
</tr>
<tr>
<td>LTQ</td>
<td>Linear trap quadrupole (same meaning as LIT)</td>
</tr>
<tr>
<td>CCS</td>
<td>Collision cross section</td>
</tr>
<tr>
<td>F1</td>
<td>Funnel 1</td>
</tr>
<tr>
<td>G1</td>
<td>Gate 1</td>
</tr>
<tr>
<td>F2</td>
<td>Funnel 2</td>
</tr>
<tr>
<td>G2</td>
<td>Gate 2 (selection gate)</td>
</tr>
<tr>
<td>IA</td>
<td>Ion activation region</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>D(_2)O</td>
<td>Deuterium oxide</td>
</tr>
<tr>
<td>FA</td>
<td>Formic acid</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>NH(_4)OH</td>
<td>Ammonium hydroxide</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance spectroscopy</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>GC</td>
<td>Gas chromatography</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Gas chromatography-mass spectrometry</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid chromatography</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean square of the deviation</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of variation</td>
</tr>
</tbody>
</table>
1. Introduction: Tools Used to Enhance Compound Identification in Complex Mixtures: Collision-Induced Dissociation (CID), Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS), Hydrogen/Deuterium Exchange, and Isotope Scrambling

1.1. Complex Mixture Analysis

Complex mixture analysis usually refers to mixtures that present enormous challenges to analytical chemists for compound identification and characterization. These mixtures may include metabolites, lipids, proteins, and petroleum samples. The fields of study which deal with the analysis of these types of compounds are called Metabolomics, Lipidomics, Proteomics, and Petroleomics, respectively. One of the greatest challenges in ‘omics analysis is the characterization of isomeric and isobaric compounds.¹ ²

Petroleum is considered one of the most complex mixtures on earth.³ ⁴ Substantial efforts have been dedicated to the characterization of petroleum mixtures because they have significant roles in the pollution of air and contamination of water sources.⁵-¹² One of the challenges in petroleum analysis is the presence of numerous isomeric and isobaric compounds.¹ Having a knowledge of the structural composition of these compounds can shed light on understanding the best mechanism of catalysis during refining processes, thereby reducing the economic cost of petroleum usage.
Another challenging area for analytical chemists is metabolite characterization and quantification.\textsuperscript{13-15} Metabolomics encompasses the comprehensive characterization of metabolites in biological mixtures. This field faces many challenges due to the great variation of metabolites in terms of functional groups, molecular structures, molecular masses, and concentrations. For example, based on the Human Metabolome Database (HMD), human plasma includes numerous metabolites which contain a diverse range of biological classes including organic acids, carbohydrates, lipids, and peptides with different concentration levels.\textsuperscript{16, 17} Furthermore, the complexity of matrices necessitates the application of different extraction and separation techniques. Therefore, a combination of different extraction methods and instrumentation techniques is required to address this complexity. Metabolite analysis also has implications for biomarker discovery.\textsuperscript{18} Biomarkers are molecules which vary significantly in concentration during the development and progression of diseases. For example, studies have demonstrated that variations in metabolite profiles are associated with cancer.\textsuperscript{19, 20} Therefore, biomarker discovery can impact the elucidation of the mechanisms of disease development.\textsuperscript{21, 22} This may result in earlier disease diagnosis and the proposing of new treatments for diseases. Thus, proposing new extraction and instrumentation techniques which generate results with high reproducibility is an active area of research.
Lipids are a class of metabolites that also pose significant challenges for analytical chemistry. Lipidomics is an emerging field which includes the cumulative methods and techniques used for lipid analysis and characterization.\textsuperscript{23-25} Lipids have several biological functions including primary roles in cell structure, cell organization, signaling events, and trafficking.\textsuperscript{26, 27} Analysis of lipids is highly complex due to the combinations of polar headgroups, different positions of double bonds, and variations in the length of lipid chains.\textsuperscript{23, 24, 28-30}

1.2. Methods Used for Complex Mixture Analysis

Complementary approaches can be used to address the complexity of different mixtures. Mass spectrometry (MS) provides invaluable information about the molecular weight of compounds and is one of the most prevalent techniques for complex mixture analysis.\textsuperscript{14, 31} Generally separation methods in both the liquid and gas-phase are required to be coupled with MS to address the complexity of samples (e.g. due to the presence of numerous isomeric and isobaric species). The most common separation methods include gas chromatography (GC), liquid chromatography (LC), and ion mobility spectrometry (IMS). Therefore, instrumentation methods such as liquid chromatography-mass spectrometry (LC-MS)\textsuperscript{32, 33}, gas chromatography-mass spectrometry (GC-MS)\textsuperscript{34}, and ion mobility spectrometry-mass spectrometry (IMS-MS)\textsuperscript{1, 35} are frequently employed for this purpose.
A variety of techniques are used for petroleum analysis including GC, GC-MS, LC, near-infrared spectroscopy (NIR), fluorescence spectroscopy, nuclear magnetic resonance (NMR), and X-ray scattering techniques. One of the most frequently used techniques to decode the chemical mixture of petroleum is Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometry. This technique provides high resolving power and high mass accuracy required for the elucidation of different classes of compounds in petroleum. Although FT-ICR provides m/z data with high accuracy, it does not provide information regarding isomeric and some isobaric species. For this purpose, methods such as LC-MS, GC-MS, and IMS-MS are used frequently. IMS-MS is used extensively for structural characterization of petroleum compounds.

Different instrumentation techniques are used for the analysis of compounds in biological mixtures including metabolites and lipids. Coupling MS to GC, LC, and supercritical fluid chromatography (SFC) presents significant advantages for compound identification. Presently, the most common method for metabolite analysis is LC-MS as it provides comprehensive compound identification for polar species. However, this technique has some limitations including long analysis times and limited utility of data-dependent fragmentation techniques. Moreover, biological mixtures contain a great number of compounds with the same molecular weight (isomeric and
isobaric species) and thorough analysis is not possible using these chromatographic methods. For example, the determination of the number and location of double bonds in lipids poses a tremendous challenge.\textsuperscript{29, 30} To address these limitations, IMS-MS is employed to enhance analytical peak capacity and enable isobaric and isomeric characterization.\textsuperscript{49} Moreover, IMS-MS provides high-reproducible data and short-analysis times.\textsuperscript{50} Lastly, it addresses matrix effects using mobility selection. This has the potential for compound assignments through reducing the number of false positive and false negative identifications. Additionally, tandem MS (MS\textsuperscript{n}) can be employed in a data-independent manner with IMS-MS to improve structural studies of compounds.\textsuperscript{51} LC can also be coupled with IMS-MS to further improve compound characterization (i.e., LC-IMS-MS).\textsuperscript{52}

1.3. IMS-MS Basics

Ion mobility spectrometry (IMS) is a gas-phase separation technique in which gaseous ions travelling in a drift tube under the influence of a constant electric field are separated based on differences in the numbers of collisions and interaction with the buffer gas.\textsuperscript{53} Overall, ions are separated based on their size and charge. Helium, being inert, is used as a buffer gas in traditional IMS separations.\textsuperscript{51} IMS can be coupled with MS to enhance the peak capacity of measurements.\textsuperscript{49} IMS-MS provides a two-dimensional (2D) dataset of mass-to-charge (m/z) versus drift time (t\textsubscript{D}). t\textsubscript{D} is the time it
takes for ions to traverse the drift tube. Drift time can be translated to orientationally-averaged collision cross section (CCS).\textsuperscript{53} CCS values are a specific physicochemical property of the molecular ions and are very reproducible. Therefore, they can be used as an additional tool for the characterization of compounds in complex mixtures.\textsuperscript{54, 55} This work will demonstrate that both experimental and theoretical CCS values can be used for this purpose. The experimental CCS values of the ions can be calculated using Equation 1:\textsuperscript{53}

\[
\Omega = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_B T)^{1/2}} \left[ \frac{1}{m_i} + \frac{1}{m_b} \right] t_D E \frac{760}{L} \frac{T}{P} \frac{1}{273.2 N}
\]

In Equation 1, \(z\) is the charge of the ion, \(e\) is the charge of an electron, \(t_D\) is the drift time, \(E\) is the electric field strength, \(T\) is the temperature of the buffer gas, \(k_B\) is Boltzmann’s constant, \(m_i\) and \(m_B\) are the masses of the ion and the buffer gas respectively, \(L\) is the length of the drift tube, \(P\) is the pressure of the buffer gas, and \(N\) is the number density of the buffer gas at STP.

For theoretical CCS calculations, first in-silico structures are obtained using molecular dynamics simulations (MDS), and then CCS values are calculated using Mobcal software suite.\textsuperscript{56, 57} Comparison of theoretical CCS values of ion conformers belonging to biological molecules with experimental CCS profiles has implications for conformational studies.\textsuperscript{56} Additionally, matching theoretical CCS values of standard
compounds with experimental CCS values of unknown compounds in complex mixtures can be used for compound identification.\textsuperscript{54} Therefore, it is beneficial to create CCS databases for biological compounds using both experimental and theoretical methods.

1.3.1. IMS-MS Instrumentation: Linear Drift Tube Ion Mobility Spectrometry Coupled to a Linear Ion Trap Mass Spectrometer

Coupling a traditional linear drift tube with the trapping capabilities of linear ion trap mass spectrometers brings some advantages for complex mixture analysis.\textsuperscript{51} The schematic of the instrument used in the studies presented throughout this work is shown in Figure 1.1.

![Fig 1.1. IMS-MS configuration in which a linear drift tube outfitted with ESI capabilities is coupled to a linear ion trap mass spectrometer](image)

The instrument shown here uses a nano-ESI source for the generation of gaseous ions. Nano-ESI prevents contamination of the ion source while enhancing ionization, which is critical for complex mixture analysis. The drift tube contains He buffer gas and a constant electric field is applied to lenses while ions are travelling through the drift tube. Initially ions are introduced to the desolvation region and then
trapped in the funnel 1 (F1) region (Figure 1.1). From these, pulses of ions are released into the drift tube with a voltage drop at G1. Consequently, ions travel through the drift tube and reach the second ion funnel (F2) region which refocuses the ions (Figure 1.1). Ion funnels are essential to improve the focusing and duty cycle of the measurements thereby improving the sensitivity of the instrument. Finally, ions enter the mass spectrometer and are mass analyzed. The presence of a dual gate design containing first and second gates (G1 and G2 in Figure 1.1) enables the collection of drift time distributions by scanning the timing between the opening of G1 and G2.

1.3.2. Application of IMS-MS for Complex Mixture Analysis

IMS-MS is able to detect molecular features that are not detectable by MS alone. IMS-MS has been frequently used for the characterization of compounds in complex mixtures. Increased peak capacity, decreased matrix effects, and the acquisition of structural and conformational information are some advantages that have made IMS-MS a popular technique for complex mixture analysis. For example, it has been shown from 2D-plots of m/z vs tD, that different classes of biomolecules fell into specific regions in the analytical space. Notably, it is possible to use different fragmentation methods with IMS measurements. For example, there is the possibility for mobility selection at the back of the drift tube and fragmentation of selected ions. Therefore, it is possible to generate unique fragmentation patterns for specific isobars or isomers. As
a result, this technique complements the use of CCS in the elucidation of structural information. Moreover, coupling experimental and theoretical CCS calculations significantly enhances conformational studies of molecules. Additionally, matching experimental and theoretical CCS values has been used for compound characterization in complex mixtures. Experimental CCS values have been reported for many biomolecules in the literature.

Numerous work has focused on developing new methods to predict CCS values for peptides and metabolites. These attempts have implications for generating a library of CCS values to aid in compound identification. For example, for peptides, it was realized that the identity and sequence of peptides are significant factors affecting CCS values. A recent attempt has been performed to predict CCS values for metabolites. In this work, different chemical properties of metabolites including polarity, functional groups, presence of aromatic rings, and polarizability of molecules were considered as prediction descriptors. Another example of the application of IMS-MS for complex mixture analysis is the use of shift reagents to isolate specific classes of molecules within 2D-plots of $m/z$ vs $t_D$ which makes it possible to address a unique class of molecules. In separate works, the separation of enantiomeric and diasteromeric compounds was reported. Here, a chiral drift gas was employed in the drift tube. Separation was possible due to different interactions of two enantiomeric species with
the chiral gas. IMS-MS has been used successfully for the characterization of metabolites in human blood,\textsuperscript{16} Escherichia coli,\textsuperscript{78} rat lymph,\textsuperscript{79} and saliva\textsuperscript{80}.

The application of IMS–MS in the area of lipidomics has significantly increased over the past few years.\textsuperscript{81-83} IMS–MS presents advantages in terms of class identification and elucidation of structural information for complex lipid samples.\textsuperscript{54} Differences in the chain length, number and location of double bounds, identity of attached cation, and identity of polar headgroups are factors that contribute to the separation of lipid ions using IMS-MS.

1.4. Applications of HDX Combined with IMS-MS

HDX-MS as a standalone method has applications for the structural study of molecules including hetero-atomic compounds in petroleum, natural organic matter, carbohydrates, nucleotides, peptides, proteins, etc.\textsuperscript{84-93} For example, Uppal et. al implemented HDX to differentiate isomeric carbohydrate ions.\textsuperscript{94} In another HDX-MS work, Robinson studied exchange levels for 5’- and 3’-nucleotide monophosphates and the 3’,5’-cyclic nucleotides in negative ESI mode using FT-ICR MS.\textsuperscript{89} Additionally, Acter and coworkers have demonstrated the utility of HDX for the speciation of nitrogen-containing compounds in petroleum samples.\textsuperscript{95} That said, to date most of the reported HDX applications have focused on peptides and proteins.\textsuperscript{84-86} Seminal studies have used HDX-MS to study the folding mechanisms of proteins.\textsuperscript{96, 97} Furthermore, the
approach has applications in studying the interactions of proteins with other molecules.\textsuperscript{98, 99}

Gas-phase HDX coupled with IMS-MS measurements holds significant potential for the study of conformational forms of biomolecules (mostly peptides and proteins).\textsuperscript{100, 101} Early IMS-HDX-MS studies demonstrated the HDX levels for protein ion conformers.\textsuperscript{100} These studies exhibited lower HDX uptakes for compact structures relative to more elongated structures. Since that time, Rand and coworkers demonstrated the application of HDX combined with traveling-wave ion mobility (TWIM) to study proteins.\textsuperscript{102} More recently HDX combined with a linear drift tube and ion trap mass spectrometer was used to extract site-specific deuterium incorporation at labile sites.\textsuperscript{56}

To implement HDX, a deuterium-containing gas is introduced into the drift tube.\textsuperscript{50} Upon deuterium incorporation, the mass of the ion increases by 1 Dalton. Studying the isotopic patterns of molecular ions provides information regarding the number of sites that are more exposed to deuterium-containing gas. That is, more accessible sites incorporate more deuterium. D\textsubscript{2}O is one of the most prevalent gases used for HDX studies. For experiments employing D\textsubscript{2}O, exchange occurs via a relay mechanism.\textsuperscript{103} For the relay mechanism to occur, an intermediate forms between the charge site, another basic site, and the D\textsubscript{2}O molecule. Therefore, to a first approximation, the
propensity for deuterium incorporation is a function of the proximity between the charge site and the exchange site.\textsuperscript{103} The HDX level depends on the rate of exchange at different sites. In other words, more accessible hydrogens exchange faster than less accessible hydrogens.

1.4.1. Hydrogen Deuterium Scrambling (HD Scrambling) as an Additional Tool to Assist Compound Identification

Hydrogen deuterium scrambling (HD scrambling) is the rearrangement of labile hydrogens and deuteriums after increasing the internal energy of the ions. This phenomenon can be explained by "mobile" proton theory.\textsuperscript{104} HD scrambling is undesirable during structural studies of biomolecules especially those of proteins.\textsuperscript{105, 106} Therefore, in order to resolve this problem for structural studies of biomolecules, new fragmentation techniques such as electron transfer dissociation (ETD) have been employed.\textsuperscript{51, 107} During HD scrambling, deuteriums which are more exposed to the reagent gas D\textsubscript{2}O exchange with inner hydrogens that are less accessible to the D\textsubscript{2}O. Thus, newly generated labile hydrogen sites can exchange again.\textsuperscript{50} As a result, the overall HDX level increases after scrambling of ions. In general, HD scrambling depends on factors such as charge location on the ions and the structure of the molecules. Therefore, HD scrambling is a unique property of the chemical makeup and
the structure of the ion that is produced. As a result, it can be used as an additional tool for the characterization of compounds in complex mixtures.

1.5. Application of Fragmentation Techniques for Complex Mixture Analysis

Ion trap mass spectrometry has unique characteristics such as mass analysis and tandem mass spectrometry (MS/MS) capabilities, which makes it suitable for complex mixture analysis.\textsuperscript{108,109} MS/MS is a very common method for ‘omics analysis that provides useful information.\textsuperscript{110} Numerous MS/MS techniques can be implemented in ion traps including collision-induced dissociation (CID)\textsuperscript{111,112}, electron transfer dissociation (ETD)\textsuperscript{113,114}, electron capture dissociation (ECD)\textsuperscript{115}, electron detachment dissociation (EDD)\textsuperscript{116}, blackbody infrared radiative dissociation (BIRD)\textsuperscript{117}, and others. Generally, to identify compounds in complex mixtures, fragmentation methods are coupled with IMS-MS\textsuperscript{51} and LC-MS\textsuperscript{118} analyses. CID is the most common fragmentation method for the analysis of metabolites and lipids.\textsuperscript{50,119} For CID operation, mass-selected precursor ions encounter energizing collisions with a buffer gas (He, CO\textsubscript{2}). Consequently, the vibrational energy of the ion increases. Eventually, molecules dissociate at bonds that have the lowest dissociation energies. For example, single bonds cleave more readily than double bonds. Such considerations have promoted the generation of CID spectra for metabolite standards in order to aid metabolite identifications.\textsuperscript{120} Comparisons of the CID spectra of unknown compounds in a complex
mixture with the CID spectra of known metabolites has been used for this purpose. For peptide characterization, CID usually generates b- and y-type fragment ions. For fragmentation techniques used in LC-MS/MS analyses, fragment ions can be correlated to precursor ions at specific retention time values. However, variations in pressure and mobile phase composition can introduce significant errors for LC-MS analysis. To mitigate against such challenges, fragmentation techniques inside the drift tube (IMS-CID-MS) have been previously implemented for complex mixture analysis.

With regard to HDX studies, CID usually causes isotope scrambling which is not desired for structural studies of biomolecules especially proteins. To address this problem, other fragmentation techniques have been proposed. These techniques include ECD and ETD. McLafferty and coworkers first employed FT-ICR MS to implement ECD for fragmentation of peptides and proteins. For ECD experiments using FT-ICR MS, voltage used on the cathode generates thermal electrons that can react with peptide cations after trapping in the ICR cell. Consequently, electron capture by such ions causes dissociation. Dissimilar to CID techniques that generate b- and y-type ions for peptides, ECD produces c- and z-type ions upon dissociation of peptide ions. Unlike ECD that uses the low pressure conditions of an ICR cell, ETD can be implemented with high-pressure mass analyzers such as linear ion traps. For ETD to
occur, electrons are transferred from reagent anions and result in the generation of the characteristic c- and z-type ions.

1.5.1. CID

CID is the most frequently used method for tandem mass spectrometry analysis. For operation of CID, ions are accelerated in the presence of a neutral gas (for example He), and consequently undergo numerous collisions with the buffer gas. A portion of the high kinetic energy gained by ion collisions is converted into the internal energy of the ions. The net result is an increase in the vibrational energy of the ion and eventually the cleavage of the weakest chemical bonds. The resulting fragmentation pattern depends on the chemical structure (functional groups) of the molecules and can be used as a fingerprint for compound identification. Bond dissociation in CID may occur through both homolytic and heterolytic cleavages. In homolytic fragmentation each fragment ion receives one electron from a covalent bond. In heterolytic fragmentation, both electrons from a covalent bond transfer to just one of the fragment ions after severing of the bond.

CID has been used frequently for the sequencing of peptides in proteomics applications. Furthermore, it has been employed in combination with other techniques to aid in the characterization of small molecules including metabolites and lipids. Interpretation of fragmentation patterns for small molecules is more challenging than for peptides. Small molecules are more diverse than peptides in terms of functional
groups and structural motifs. Structural rearrangements during dissociation in small molecules makes interpretation of CID spectra even more challenging.\textsuperscript{126} For example, McLafferty rearrangement is frequently observed in the CID spectra of small molecules.\textsuperscript{127} In this rearrangement, a molecule which contains a keto-group experiences β-cleavage, with the acceptance of a γ-hydrogen atom. The internal energy of the molecular ion and the kinetics of dissociation are two important factors that influence CID fragmentation patterns. Fast CID processes are desirable. This prevents fragmentation of metastable ions during traveling to the detector. Because of this, low and high-energy CID are used which may result in distinct fragmentation patterns.

1.6. HDX Kinetics Modeling

Over the last two decades, new analytical prediction tools have been established based on molecular ion structures. These efforts include the development of methods to predict CCS values and the gas phase HDX behavior of ions.\textsuperscript{62, 68} Having the HDX patterns for different classes of metabolites facilitates identification of these compounds in complex mixtures.\textsuperscript{50} Providing a library of the experimental HDX patterns is time intensive and expensive. Therefore, efforts to discover new HDX prediction methods are worthwhile. Valentine et.al have reported new algorithm to predict the HDX behavior of peptides.\textsuperscript{128} They used hydrogen accessibility scoring (HAS) which relies on an ion-neutral collision model to predict HDX patterns. In this work, efforts were directed at
gaining insights into the peptide ion conformers mostly contribute the most to the observed HDX behavior. In the workflow, first molecular dynamics (MD) simulations were performed on select peptide ions. Second, conformers which best match with experimental CCS values were selected to use for HDX kinetics modeling. For the HAS algorithm, the surface accessibility of exchange sites on ion conformers is a crucial parameter to consider for accurate HDX prediction. That is, more exposed labile hydrogens exchange more rapidly than buried hetero-atom hydrogens. Another factor considered by the algorithm was the inverse distance between charge sites and less basic sites (e.g. carbonyl). As mentioned above, for HDX experiments which employ D$_2$O as a reagent gas, exchange proceeds through a relay mechanism.$^{103}$ Therefore, the proximity of the charge site to the exchange site is critical for HDX to proceed. Using this method, the HDX contributions from each site for each of conformers were determined. Then the overall exchange was computed for each conformer. Finally, non-equilibrium linear regression approach was employed to find the best-fit HDX distribution based on combinations of different conformers.

For the HDX kinetics modeling reported in this dissertation, the modeling was performed on sucrose anions. Notably, sucrose has different deprotonation locations. Two deprotonation sites were considered for HDX kinetics modeling. Simulated Annealing was first employed to find conformers which best match with experimental
CCS profiles. Typically, 1000 energy-minimized structures were obtained from Simulated Annealing and their CCS values were calculated. A modified HAS scoring method was used to perform HDX kinetics modeling for sucrose. In this method, the distance between the charge site and the basic site was still considered an important factor contributing to HDX kinetics. However, the surface accessibility was not considered due to the fact that all –OH groups in sucrose are equally exposed to D₂O. To narrow down the number of selected conformers, a new accessibility factor was introduced. The factor takes into account the propensity of hydrogen bond formation between the charge site and the exchange site and also the possibility that intra-molecular rotation locates the charge site in proximity to the exchange site.
1.7. References


2. Structural Assignments of Sulfur-containing Compounds in Crude Oil Using Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS)


2.1. Introduction

Petroleum is one of the most complex mixtures encountered in analytical chemistry. Despite its challenge, petroleum characterization is paramount and its full analysis is outstanding.\textsuperscript{1, 2} Many studies in recent years have focused on molecular characterization in order to better understand the chemical makeup of petroleum. One objective is to catalog the types of functional groups and structural motifs in petroleum. Such characterization is enticing as it may find applications for indicating the origin of particular petroleum samples.\textsuperscript{3} It also provides information for improving petroleum refining and transportation.\textsuperscript{4, 5}

The geographical location of a petroleum well has a decisive effect on the chemical and physical nature of the petroleum sample.\textsuperscript{6} The molecular mixture in petroleum is mostly comprised of 3 categories of hydrocarbons which include alkanes, naphthenes, and aromatic compounds (conjugated five- or six-carbon member rings).\textsuperscript{7} Hetero-atom compounds in petroleum are mostly composed of nitrogen, oxygen, and sulfur-containing species.\textsuperscript{8} Petroleum samples also contain trace amounts of metals such as nickel, vanadium, and iron.\textsuperscript{9} Acids are the main constituents of oxygen containing molecules.\textsuperscript{10} This class of compounds is especially important because it plays a main role in the
corrosion of transportation tubes of petroleum. Nitrogen containing compounds are important to study because of their role in air pollution.

In general, crude oil with a high percentage of polar compounds is less desirable because it creates greater pollution upon consumption, more readily disturbs catalytic processes, and increases the corrosion of transportation tubes. In this regard, lighter crude oil is more desirable. With that being said, analysis of heavy petroleum mixtures is increasingly important because the petroleum resources of the world are moving toward heavy petroleum. Therefore, analysis of polar compounds in petroleum is of high demand. Having noted the connection between petroleum composition and issues associated with refining and transportation, it is difficult to overestimate the implications of extensive compound structure characterization. For example, structural analyses could help alleviate environmental challenges by providing information about the chemical reactivity of petroleum mixtures. Such information could be helpful to identify petroleum wells that have more desirable molecular characteristics and thereby cost reduction could be achieved for refining and also for transportation processes. For these reasons, the hetero-atom content of petroleum has been the subject of detailed structural characterization.

There are a number of important reasons for characterizing sulfur compounds in petroleum. One significant reason is that these compounds are the most abundant hetero-atom constituents in petroleum. The content of sulfur in crude oil can be as high as 10%. Sulfur-containing compounds cause corrosion in petroleum pipelines and interfere in catalysis processes which has been termed catalysis poisoning. Specifically, thiols and sulfides have implications with regard to environmental and human health presenting high
exposure risk factors. Furthermore, petroleum analysis is difficult as sulfur-containing compounds are very absorptive, adsorptive, and reactive.\textsuperscript{30} Therefore, a knowledge of the constituent sulfur-containing compounds could aid other measurement strategies to obtain fuller characterization of petroleum samples. Different methods have been proposed for the depletion of sulfur compounds from petroleum. Hydrodesulfurization (HDS) is the main method used for this purpose \textsuperscript{31, 32} because these compounds disturb catalytic processes. Fuller structural characterization of sulfur-containing compounds in petroleum may provide information about the required types of catalytic processes. That is, the distribution of different sulfur containing species is helpful when selecting specific catalysts. For example, the efficiency of a catalyst depends on the match between the surface of the catalyst and the interacting molecules.\textsuperscript{33} In this manner a thorough knowledge about the size and structural distribution of sulfur-containing species can have a pronounced effect on the optimization of catalytic processes. The same characterization has implications for guiding the HDS processes.

A variety of instrumentational and analytical techniques have been employed for the determination of the overall chemical nature of crude oil.\textsuperscript{34-41} Gas chromatography (GC),\textsuperscript{42} gas chromatography-mass spectrometry (GC-MS),\textsuperscript{43, 44} liquid chromatography (LC)\textsuperscript{45}, near-infrared spectroscopy (NIR)\textsuperscript{46}, fluorescence spectroscopy,\textsuperscript{47} nuclear magnetic resonance (NMR),\textsuperscript{48} and X-ray scattering techniques\textsuperscript{49} have been used to study petroleum compounds. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is a technique that is frequently used for petroleum compound characterization.\textsuperscript{16} This technique provides extraordinary resolving power and high mass accuracy required to identify compound classes within a complex mixture of petroleum.
That said, often a more comprehensive analysis is required for petroleum characterization; seminal work has shown that FTMS can be combined with other analytical techniques as an integrative approach for improving compound characterization.\textsuperscript{50, 51}

The polarities of petroleum compounds are varied and thus selecting the appropriate ionization method for mass spectrometry is important.\textsuperscript{52} Soft ionization techniques are desirable in many experiments because the proportion of ion fragmentation is significantly reduced. Electrospray ionization (ESI) is able to readily generate cations and anions from polar compounds but is less suitable to ionize mid-polar compounds such as sulfur-containing species in petroleum.\textsuperscript{53} To address this problem, different methods have been proposed. As an example, methylation of sulfur atoms has been used for this purpose.\textsuperscript{54, 55} Another technique that has been used frequently is the addition of a metal that forms a complex with petroleum compounds.\textsuperscript{56} Silver cationization has recently been used with MS to study sulfur containing compounds in petroleum.\textsuperscript{57}

FT-ICR MS is very efficient for producing molecular fingerprints of petroleum samples.\textsuperscript{58} However, this technique cannot provide structural information associated with isomers (and some isobars) required for more complete structural characterization. Ion mobility spectrometry-mass spectrometry (IMS-MS) has been used to address this problem and has become an increasingly popular choice for studying petroleum mixtures in recent years.\textsuperscript{28, 59-64} IMS is a gas-phase separation technique in which ions traveling in a pressurized drift tube under the influence of a constant electric field are separated based on their shape and charge.\textsuperscript{66-68} IMS-MS provides additional information in the form of the mass-to-charge ($m/z$) values for molecular ions. Additionally, IMS in combination with MS
provides improved peak capacities, as well as enhancing molecule identification.\(^{69}\)

Finally, drift times \((t_D)\) measured by IMS-MS can be converted to experimental collision cross section (CCS) values that can be used for structural assignments.

Experimental \(t_D\) profiles and theoretical CCS values have long been employed for structural estimation of proteins and peptides.\(^{70-88}\) Another powerful method for the elucidation of structural information is collision-induced dissociation (CID).\(^{89-97}\) FT-ICR MS in combination with CID has been used for structural and molecular characterization of petroleum compounds.\(^{19,22}\) This approach provides high mass accuracy and fragmentation information about precursor ions that may be used to complement CCS values for structural assignments.

This research presents complementary IMS-MS and MS and tandem MS (MS/MS) methods for structural determination of sulfur-containing compounds in petroleum. This is accomplished by comparison of experimental and theoretical CCS values of silver ion complexes combined with the application of high mass accuracy and CID data. Silver contains two different isotopes of nearly equal abundance and, therefore, ion doublets in two-dimensional (2D) plots of \(t_D\) and \(m/z\) values can be used for the tracking and characterization of metalated compounds.\(^{57}\) Here, the selection of sulfur-containing compounds combined with IMS-MS and FT-MS analysis is presented as proof-of-principle method for characterizing this important class of compounds. As such, only limited examples of CCS-based structure comparison and MS/MS comparisons are presented. Future studies will extend the usage of these techniques for more complete characterization.
2.2. Experimental Section

2.2.1. Sample Preparation

Methanol and toluene [both HPLC grade], concentrated acetic acid (HCOOH), and silver triflate (AgOTf) were purchased from Sigma-Aldrich (St. Louis, MO). Saudi crude oil samples were purchased from ONTA, Inc. (Canada). Two solutions were prepared to test the efficiency of silver complexation with sulfur containing compounds. The first solution contained a crude oil sample with 1% acetic acid. The second solution contained a crude oil sample with silver cation. Saudi crude oil samples were dissolved in methanol/toluene in a 1:1 (v/v) ratio to a final concentration of 1 mg/mL and AgOTf was added to a concentration of 0.1 mg/mL.

2.2.2. Instrumentation

For mobility experiments, an instrument consisting of a home built drift tube coupled to an LTQ Velos mass spectrometer was used as shown in Figure 2.1. The IMS-MS instrument used here employed a nano-ESI source for the generation of gaseous ions and their introduction into the drift tube. In Figure 2.1, the locations of the ion desolvation region, ion gates, ion funnels, and drift tube are displayed.

The two ion gates in the drift tube (Figure 2.1) enabled the measurement of the $t_D$ distributions. Initially, ions were electrosprayed into the desolvation region and then are trapped in the first funnel (F1 region in Figure 2.1). From the F1 region, ions were released into the drift tube in a pulsed manner. Consequently, ions passed into a home built, 1-meter-long drift tube, which was operated at 10 V.cm$^{-1}$ and filled with 2.72 Torr He (300 K) buffer gas. Upon reaching the second ion funnel (F2 region in Figure 2.1), ions were focused for extraction into the mass spectrometer. The ions subsequently entered the
linear ion trap mass spectrometer (LTQ Velos, Thermo Electron, San Jose, CA, USA) for ion detection and $m/z$ determination.

![Figure 2.1](image)

**Figure 2.1.** Schematic diagram of the instrument used in these studies. The home-built drift tube and linear ion trap mass spectrometer are shown. Funnels, activation regions, and gates are labeled.

distributions were created by scanning the delay time associated with the opening of the second ion gate (G2 in Figure 2.1) relative to the first ion gate (G1 in Figure 2.1). This time delay was stepped by 200 µs (the G2 open time) across the entire $t_d$ distribution. In this manner, mobility-resolved $m/z$ datasets were collected. The DPlot software (DPlot Graph Software, HydeSoft Computing, 2001-2010, Vicksburg, Mississippi) was used for the generation of 2D plots of $m/z$ and $t_d$ values. Accurate mass measurements were performed with an FT-ICR instrument (Thermo Electron, San Jose, CA) equipped with a 7 T horizontal superconducting solenoid magnet. The resolving power of instrument is determined to be ~120,000 on average. The experiments were performed in positive ESI mode. The CID experiments were performed using the LTQ device of FT-ICR MS (Thermo Electron, San Jose, CA). Both instruments were operated in positive ESI mode. The CID analysis was conducted by isolating a precursor ion with a $m/z$ isolation window of ±1.0 Da relative to the centroid $m/z$ value. Selected ions were fragmented using a normalized activation energy of 35.0%.
2.2.3. Experimental and Theoretical CCS Calculations

Experimental CCS values were calculated according to Equation 1.\(^99\)

\[
\Omega = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_B T)^{1/2}} \left[ \frac{1}{m_i} + \frac{1}{m_b} \right] \frac{t_D E}{L} \frac{T}{P} \frac{1}{273.2 N} \tag{1}
\]

In Equation 1, \(z\) is the charge of the ion, \(e\) is the charge of an electron, \(E\) is the electric field strength, \(T\) is the temperature of the buffer gas, \(k_b\) is Boltzman’s constant, \(m_i\) and \(m_b\) are the masses of the ion and the buffer gas respectively, \(L\) is the length of the drift tube, \(P\) is the pressure of the buffer gas, and \(N\) is the number density of the buffer gas at STP. Notably, Equation 1, relates CCS values and \(t_D\) values. Demonstratively, elongated ions (larger \(\Omega\)) require more time to reach the detector, and compact ions reach the detector at earlier times.

Experimental CCS values were determined for standard ions in order to check the accuracy of the measurements. Tetra-butyl ammonium salt and reserpine were used for this purpose. The error is estimated by comparing the CCS values of tetra-butyl ammonium salt (113.1 Å\(^2\)) and reserpine (180.9 Å\(^2\)) with those obtained from the literature. The measured CCS values were within 1.7% and 1.17% of reported values \(^{100}\) for tetra-butyl ammonium salt and reserpine ions, respectively. This is similar to values obtained previously for large protein ions.\(^{101}\) The precision of our measurement is markedly better (typically <1% as measured by triplicate measurements for some of these compounds). This difference was taken into consideration during association of theoretical structures to \(t_D\) profiles.

Theoretical CCS values were obtained by first generating initial ion structures with the Avogadro software suite. Second, sulfur-silver structures were optimized with
Second-Order Møller-Plesset Perturbation Theory (MP2) using a LanL2MB basis set.\textsuperscript{102} Calculations were performed using Gaussian.\textsuperscript{103} Finally, the Mobcal software employing the trajectory method (TM) was used for calculation of CCS values for compounds with optimized geometries.\textsuperscript{104, 105} Theoretical CCS values were determined for five randomly oriented structures for each compound. Average theoretical CCS values are reported.

2.3. Results and Discussion

2.3.1. Comparison of Mass Spectra and 2D Plots for Two Samples

Figure 2.2A shows a 2D $t_D$, $m/z$ plot for the silver-containing sample. Both protonated and metalated datasets are comprised primarily of singly-charged ions. As reported previously, the addition of silver to crude oil samples should mostly generate [M+Ag]$^+$ and [2M+Ag]$^+$ ions.\textsuperscript{56} For this study, only [M+Ag]$^+$ ions are considered because they reveal more structural information related to isomers and isobars. The data encompasses a $t_D$ range of $\sim$4.5 to $\sim$12 ms. Additionally, the data are comprised of ions that predominantly have $m/z$ values in two separate regions of the mass spectrum. The first extends from $\sim$300 $m/z$ to 400 $m/z$ and the second region extends from $\sim$500 $m/z$ to 700 $m/z$. In Figure 2.2A, white lines are associated with structural classes exhibiting different degrees of compactness. Here the lines represent $t_D$ regions of similar structures.\textsuperscript{106} The selected structures differ from each other by their benzene ring orientations. Trend lines 1 and 3 represent the most compact and least compact structures among these three molecular types. The calculated locations for a number of these structures are indicated in the 2D plot.
Figure 2.2. A) Two-dimensional (2D) plot of m/z and tD values obtained upon direct electrospray of Saudi crude oil co-incubated with silver ions. Intensities are represented on a color scale where red and black are the highest and lowest intensities, respectively. Three white lines are associated with molecular structures with different degrees of compactness. The theoretical locations of three of these structures within the 2D plot are shown. B) tD distributions of silver-containing samples and protonated crude oil samples for two m/z ranges (650-660 and 550-560 for top and bottom traces, respectively). Solid and dashed lines are associated with silver-containing and protonated samples, respectively. Intensity values have been scaled to the tallest data file feature representing a peak height of 1.
The trend lines are generated using nine ions comprising three different structural classes of molecules which have been selected for comparison to the experimental data. Theoretical CCS values have been determined for the nine ions and plots of $m/z$ versus theoretical $t_D$ values have been generated. For each separate molecular class, linear regression is used to generate the trend lines. The molecular structures and their nominal $m/z$ values are provided in Table 2.1.

In Figure 2.2, the structures representing those associated with lines 1 and 2 have roughly the same $t_D$ values but different $m/z$ values. This is based on the fact that the structures associated with line 1 are the most compact among the three types requiring them to have higher $m/z$ values for ions of similar mobility. Conversely, structures associated within lines 2 and 3 have the same $m/z$ but different $t_D$ values. Such comparisons are instructive for considering the predominant types of structures observed in the silver-containing sample at different $m/z$ values. In general, for individual features, as the $m/z$ value increases, the $t_D$ profile widens. This could result from the fact that the number of isomeric and isobaric species increases at higher masses. Figure 2.2A also shows a decrease in $t_D$ increment with $m/z$ values (shift toward lines 1 and 2). As demonstrated in previous studies, as the number of carbon atoms for petroleum compounds increases, the DBE range becomes wider and shifts toward higher DBE values. Higher DBE values correlate to more compact structures, which would explain a decrease in $t_D$ change with higher $m/z$ values. Notably, the same trends have been observed for asphaltenes in petroleum.
Table 2.1. Structures used for drawing trend lines in Figure 2.2.

<table>
<thead>
<tr>
<th>Structurea</th>
<th>Nominal m/z b</th>
<th>Trend Linec</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Structure" /></td>
<td>365, 367</td>
<td>1</td>
</tr>
<tr>
<td><img src="image2" alt="Structure" /></td>
<td>537, 539</td>
<td>1</td>
</tr>
<tr>
<td><img src="image3" alt="Structure" /></td>
<td>659, 661</td>
<td>1</td>
</tr>
<tr>
<td><img src="image4" alt="Structure" /></td>
<td>341, 343</td>
<td>2</td>
</tr>
<tr>
<td><img src="image5" alt="Structure" /></td>
<td>491, 493</td>
<td>2</td>
</tr>
<tr>
<td><img src="image6" alt="Structure" /></td>
<td>691, 693</td>
<td>2</td>
</tr>
<tr>
<td><img src="image7" alt="Structure" /></td>
<td>341, 343</td>
<td>3</td>
</tr>
<tr>
<td><img src="image8" alt="Structure" /></td>
<td>491, 493</td>
<td>3</td>
</tr>
<tr>
<td><img src="image9" alt="Structure" /></td>
<td>641, 643</td>
<td>3</td>
</tr>
</tbody>
</table>

a Structures used to draw trend lines  
b Nominal m/z values associated with each structure  
c Trend line number from Figure 2.2
Comparisons of the IMS-MS data between the two samples (acid and silver-containing crude oil samples) show significant differences in the \( t_D \) values for predominant ions of select \( m/z \) values. Based on previous studies, \(^{57, 109}\) it may be expected that the silver-treated samples favor mostly ionization of sulfur-containing compounds, and the acid-treated samples favor mostly ionization of amine-containing compounds. Silver attachment to sulfur-containing compounds generate higher mobility ions relative to protonated compounds of the same \( m/z \). Consequently, there is a slight shift toward lower \( t_D \) values for the silver-treated crude oil sample compared to the protonated sample over the same \( m/z \) range. Figure 2.2B compares \( t_D \) distributions of the silver-treated sample with the acid-treated sample for two \( m/z \) ranges (650-660 and 550-560). These two \( m/z \) ranges were selected because they show a great degree of mobility difference between acid-treated and silver-treated analytical peaks. As may be expected, compounds in the protonated samples (dashed line in Figure 2.2B) have higher \( t_D \) values than the silver-treated samples (a solid line in Figure 2.2B). Overall, compounds in the acid-treated samples have higher \( t_D \) values than the silver-treated samples. This can generally be explained by the more massive charge carrier (silver-treated samples), where covalent structure is essentially replaced by the silver cation in comparison to protonated ions. That is, silver has a smaller size/mass ratio compared with other isobaric functional groups (in the case of protonated ions). Consequently, in isobaric ion comparisons (protonated versus metalated), when silver is attached to any sulfur-containing molecule, the molecular core must necessarily be smaller in size resulting in higher mobilities at given \( m/z \) ratios. Such a shift in \( t_D \) values has been demonstrated in several studies by
Another observation is that less structural diversity is observed for the silver-treated samples. This may indicate selective silver complexation. Mclean and coworkers.  

There are two isotopes of silver ($^{107}\text{Ag}$, $^{109}\text{Ag}$) with respective abundances of 51.84% and 48.16%. Therefore, upon silver attachment, a doublet pattern would be expected in the 2D IMS-MS plot in which the same $t_D$ profiles would be observed separated by 2 m/z units. Silver may attach to either sulfur atoms or aromatic rings. Notably, the interaction of silver with sulfur is stronger. To check the preferred location of attachment for gas-phase ions, silver was attached in silico to different positions on the molecule including sulfur atoms and aromatic rings. The structures were optimized using Gaussian and the energy of these structures were compared. This is similar to methods employed to assign protonation sites for in silico structures in which those sites with the highest apparent gas-phase basicity (most stable) are used to represent the ions of interest. For four selected molecules, the location of silver is shown in Table 2.2. As observed in the table, the attachment of silver to sulfur is generally more energetically favorable. These observations have been considered in the selection of representative structures for other ions.
Table 2.2. Selected structures for studying the tendency of silver to interact at different locations on sulfur-containing compounds.

<table>
<thead>
<tr>
<th>Structure&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Energy&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>-608.074760</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>-608.008270</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>-608.002246</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>-508.071911</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>-508.071913</td>
</tr>
<tr>
<td><img src="image6.png" alt="Structure 6" /></td>
<td>-508.071916</td>
</tr>
<tr>
<td><img src="image7.png" alt="Structure 7" /></td>
<td>-508.052750</td>
</tr>
<tr>
<td><img src="image8.png" alt="Structure 8" /></td>
<td>-596.384416</td>
</tr>
<tr>
<td><img src="image9.png" alt="Structure 9" /></td>
<td>-596.384412</td>
</tr>
<tr>
<td><img src="image10.png" alt="Structure 10" /></td>
<td>-596.381320</td>
</tr>
<tr>
<td><img src="image11.png" alt="Structure 11" /></td>
<td>-596.384414</td>
</tr>
</tbody>
</table>

<sup>a</sup> Structure representing the location of attachment of silver to the molecule
<sup>b</sup> Energy obtained after energy optimization with Gaussian
Figure 2.3A and 2.3B show expanded dataset regions encompassing ions of nominal m/z values of 665 and 667 (A) and 683 and 685 (B), respectively. The $t_D$ distributions associated with Figure 2.3A and 2.3B are shown in Figure 2.3C and 2.3D, respectively. An interesting observation of the $t_D$ profiles in Figure 2.3A and 2.3B is that, for all four nominal m/z values within a specific $t_D$ profile, there is a slight increase in m/z values as $t_D$ increases. It should be noted that, for a given nominal m/z value, different classes of sulfur-containing compounds having different molecular formulas can be observed using the silver cationization technique. These compounds with the same nominal m/z value are isobars. Such isobaric classes include compounds with S, S$_2$, S$_3$, NS, NS$_2$, OS, OS$_2$, NOS, and O$_2$S functional groups. The abundances of these classes can vary significantly based on crude oil types and origins. For example, in Figure 2.3, nominal m/z values of 665 and 667 can correspond to molecular formulas of C$_{39}$H$_{58}$S$\text{Ag}$, C$_{37}$H$_{50}$S$_2$$\text{Ag}$, C$_{34}$H$_{54}$S$_3$$\text{Ag}$, C$_{38}$H$_{56}$NS$\text{Ag}$, C$_{35}$H$_{60}$NS$_2$$\text{Ag}$, C$_{38}$H$_{54}$OS$\text{Ag}$, C$_{36}$H$_{46}$OS$_2$$\text{Ag}$, C$_{37}$H$_{52}$OSN$\text{Ag}$, and C$_{37}$H$_{50}$O$_2$$\text{SAg}$. A similar analysis can be performed for ions of nominal m/z values of 683 and 685. Another observation is that the isotopologue intensities of the $t_D$ profiles for the two silver complex compounds are slightly different. As expected, and shown in Figures 2.3C and 2.3D, the $t_D$ profiles at nominal m/z values of 665 and 683 are slightly more intense than the $t_D$ profiles at nominal m/z values of 667 and 685.

In these structural studies of sulfur-containing compounds in petroleum, only S$_1$, S$_2$, and S$_3$ classes of compounds have been considered. One reason for this preference is that other studies have shown the predominance of these types of compounds by silver adduction. These data are in agreement with information extracted from FT-ICR data of silver-treated Saudi sample used in this work (Table 2.3 and 2.4). In such studies, library
databases have been used to assign sulfur containing compounds. Table 2.5 provides some of these types of assignments belonging to the S1 class of compounds and their relative mass accuracies as measured by FT-ICR MS.

Figure 2.3. Expanded regions in 2D IMS-MS plots for A) ions with nominal m/z values of 665 and 667 and, B) 683 and 685. Respective t_D distributions for panels A and B are shown in panels C and D. Green and blue traces represent lighter and heavier silver isotopes, respectively. Intensities in panels C and D have been normalized by dividing the y-value at each drift time by the sum of all y-values.

Table 2.3. Molecular formulas observed within nominal m/z values of 291 and 293

<table>
<thead>
<tr>
<th>Molecular Formula</th>
<th>DBE^b</th>
<th>Theoretical mass^c</th>
<th>Observed Mass^d</th>
<th>Mass Accuracy^e</th>
<th>Relative Intensity^f</th>
</tr>
</thead>
<tbody>
<tr>
<td>C_{12}H_{8}S_{107}Ag</td>
<td>9</td>
<td>290.9398</td>
<td>290.9400</td>
<td>1.0</td>
<td>100</td>
</tr>
<tr>
<td>C_{11}H_{20}S_{107}Ag</td>
<td>2</td>
<td>291.0337</td>
<td>291.0341</td>
<td>1.6</td>
<td>24.99</td>
</tr>
<tr>
<td>C_{9}H_{12}S_{2}S_{107}Ag</td>
<td>4</td>
<td>290.9431</td>
<td>290.9429</td>
<td>-0.9</td>
<td>35.32</td>
</tr>
<tr>
<td>C_{8}H_{8}S_{2}O_{107}Ag</td>
<td>5</td>
<td>290.9067</td>
<td>290.9064</td>
<td>-1.1</td>
<td>10.83</td>
</tr>
<tr>
<td>C_{10}H_{16}SO_{107}Ag</td>
<td>3</td>
<td>290.9973</td>
<td>290.9978</td>
<td>1.7</td>
<td>8.33</td>
</tr>
</tbody>
</table>

^a Molecular formula obtained using mMass software

^b Double Bond Equivalence

^c Exact mass of the ion obtained from the molecular formula

^d Observed mass of the ion obtained by FT-ICR-MS analysis

^e Mass accuracy obtained with the FT-ICR instrument. Values are reported in parts in million

^f Relative intensity of peaks observed using FT-ICR MS
Table 2.4. Molecular formulas observed within nominal $m/z$ values of 353 and 353

<table>
<thead>
<tr>
<th>Molecular Formula $^a$</th>
<th>DBE$^b$</th>
<th>Theoretical mass$^c$</th>
<th>Observed Mass$^d$</th>
<th>Mass Accuracy$^e$</th>
<th>Relative Intensity$^f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$<em>{14}$H$</em>{14}$S$_2$O$^{107}$Ag</td>
<td>8</td>
<td>352.9591</td>
<td>352.9588</td>
<td>1.0</td>
<td>100</td>
</tr>
<tr>
<td>C$<em>{11}$H$</em>{18}$S$_3$O$^{107}$Ag</td>
<td>3</td>
<td>352.9615</td>
<td>352.9622</td>
<td>-1.9</td>
<td>90.73</td>
</tr>
<tr>
<td>C$<em>{17}$H$</em>{20}$S$_2$O$^{107}$Ag</td>
<td>13</td>
<td>352.9558</td>
<td>352.9554</td>
<td>0.9</td>
<td>50.21</td>
</tr>
<tr>
<td>C$<em>{16}$H$</em>{22}$S$_2$O$^{107}$Ag</td>
<td>6</td>
<td>353.0498</td>
<td>353.0493</td>
<td>1.3</td>
<td>5.3</td>
</tr>
<tr>
<td>C$<em>{13}$H$</em>{10}$S$_2$O$^{107}$Ag</td>
<td>9</td>
<td>352.9228</td>
<td>352.9224</td>
<td>1.1</td>
<td>8.9</td>
</tr>
<tr>
<td>C$<em>{11}$H$</em>{18}$S$_2$O$^{107}$Ag</td>
<td>3</td>
<td>352.9805</td>
<td>352.9800</td>
<td>1.7</td>
<td>5.2</td>
</tr>
<tr>
<td>C$<em>{15}$H$</em>{18}$SO$^{107}$Ag</td>
<td>7</td>
<td>353.0122</td>
<td>353.0129</td>
<td>-1.9</td>
<td>3.2</td>
</tr>
<tr>
<td>C$<em>{12}$H$</em>{22}$S$_2$O$^{107}$Ag</td>
<td>2</td>
<td>353.0166</td>
<td>353.0163</td>
<td>0.9</td>
<td>1.1</td>
</tr>
</tbody>
</table>

$^a$ Molecular formula obtained using mMass software  
$^b$ Double Bond Equivalence  
$^c$ Exact mass of the ion obtained from the molecular formula  
$^d$ Observed mass of the ion obtained by FT-ICR-MS analysis  
$^e$ Mass accuracy obtained with the FT-ICR instrument. Values are reported in parts per million  
$^f$ Relative intensity of peaks observed using FT-ICR MS

Table 2.5. Tentative assignments of sulfur-containing ions from S$_1$ class compounds

<table>
<thead>
<tr>
<th>Molecular Formula$^a$</th>
<th>DBE$^b$</th>
<th>Exact Mass$^c$</th>
<th>Observed Mass$^d$</th>
<th>Mass Accuracy$^e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$<em>{27}$H$</em>{48}$S$^{107}$Ag</td>
<td>4</td>
<td>511.2527</td>
<td>511.2532</td>
<td>1.0</td>
</tr>
<tr>
<td>C$<em>{27}$H$</em>{48}$S$^{109}$Ag</td>
<td>4</td>
<td>513.2524</td>
<td>513.2517</td>
<td>-1.4</td>
</tr>
<tr>
<td>C$<em>{27}$H$</em>{50}$S$^{107}$Ag</td>
<td>3</td>
<td>513.2684</td>
<td>513.2692</td>
<td>1.5</td>
</tr>
<tr>
<td>C$<em>{27}$H$</em>{50}$S$^{109}$Ag</td>
<td>3</td>
<td>515.2680</td>
<td>515.2675</td>
<td>-1.0</td>
</tr>
<tr>
<td>C$<em>{30}$H$</em>{54}$S$^{107}$Ag</td>
<td>4</td>
<td>553.2997</td>
<td>553.3006</td>
<td>1.6</td>
</tr>
<tr>
<td>C$<em>{30}$H$</em>{54}$S$^{109}$Ag</td>
<td>4</td>
<td>555.2993</td>
<td>555.2984</td>
<td>-1.6</td>
</tr>
<tr>
<td>C$<em>{31}$H$</em>{54}$S$^{107}$Ag</td>
<td>5</td>
<td>565.2997</td>
<td>565.3003</td>
<td>1.1</td>
</tr>
<tr>
<td>C$<em>{31}$H$</em>{54}$S$^{109}$Ag</td>
<td>5</td>
<td>567.2994</td>
<td>567.2987</td>
<td>-1.2</td>
</tr>
<tr>
<td>C$<em>{29}$H$</em>{58}$S$^{107}$Ag</td>
<td>1</td>
<td>577.3030</td>
<td>577.3038</td>
<td>1.4</td>
</tr>
<tr>
<td>C$<em>{29}$H$</em>{58}$S$^{109}$Ag</td>
<td>1</td>
<td>579.3026</td>
<td>579.3034</td>
<td>1.4</td>
</tr>
<tr>
<td>C$<em>{30}$H$</em>{60}$S$^{107}$Ag</td>
<td>1</td>
<td>591.3186</td>
<td>591.3192</td>
<td>1.0</td>
</tr>
<tr>
<td>C$<em>{30}$H$</em>{60}$S$^{109}$Ag</td>
<td>1</td>
<td>593.3183</td>
<td>593.3175</td>
<td>-1.3</td>
</tr>
<tr>
<td>C$<em>{33}$H$</em>{68}$S$^{107}$Ag</td>
<td>5</td>
<td>593.3310</td>
<td>593.3302</td>
<td>-1.3</td>
</tr>
<tr>
<td>C$<em>{33}$H$</em>{68}$S$^{109}$Ag</td>
<td>5</td>
<td>595.3307</td>
<td>595.3314</td>
<td>1.2</td>
</tr>
<tr>
<td>C$<em>{34}$H$</em>{60}$S$^{107}$Ag</td>
<td>5</td>
<td>607.3467</td>
<td>607.3461</td>
<td>-1.0</td>
</tr>
<tr>
<td>C$<em>{34}$H$</em>{60}$S$^{109}$Ag</td>
<td>5</td>
<td>609.3463</td>
<td>609.3470</td>
<td>1.1</td>
</tr>
</tbody>
</table>

$^a$ Molecular formula obtained using mMass software  
$^b$ Double Bond Equivalence  
$^c$ Exact mass of the ion obtained from the molecular formula  
$^d$ Observed mass of the ion obtained by FT-ICR-MS analysis  
$^e$ Mass accuracy obtained with the FT-ICR instrument. Values are reported as parts per million
Four ions have been considered for structural elucidation of sulfur compounds (nominal isotopologue m/z values of 291 and 293 and 353 and 355) (Table 2.3 and 2.4). Here it is noted that because of the large number of different structures that can match with collision cross section data, the exclusion of certain structures and the inclusion of others is, at present, tentative. The $t_D$ distribution for isotopologue ions of m/z 291 and 293 is $\sim 0.5$ ms wide (Figure 2.4A) and the $t_D$ distribution for isotopologue ions of m/z 353 and 355 is 1.4 wide (Figure 2.5A). As mentioned above, this may demonstrate that the number of isomeric and isobaric species increases with increasing mass. For ions with nominal m/z values of 291 and 293, the $t_D$ profiles are mostly centered at $\sim 5$ ms. For ions with nominal m/z values of 353 and 355, the $t_D$ profiles are mostly centered at $\sim 5.6$ ms, but some relatively low-intensity features can also be observed at $\sim 6.2$ ms. This reveals additional information about structural heterogeneity for ions of these nominal m/z values. That is, these ions have greater structural variability and may belong to different classes of sulfur-containing ions.

It is important to note that the $t_D$ distribution is essentially identical for both isotopic doublets as the electronic structure of the two silver isotopes is the same. Accurate mass measurements suggest that the ions with nominal m/z values of 291 and 293 belong to the S$1$, S$2$, SO, and S$_2$O classes of sulfur-containing compounds as shown in Table 2.3. Overall, the most abundant species belong to the S$1$ and S$2$ classes (Table 2.6). Notably, for these m/z values, no nitrogen-containing species are observed. Therefore, in the discussion below only compounds containing sulfur as the sole heteroatom are considered. For the S$1$ class, ions with molecular formulas of C$_{12}$H$_8$S$^{107}$Ag or C$_{12}$H$_8$S$^{109}$Ag [which have a double bond equivalence (DBE) of 9] or those with molecular formulas of
C\textsubscript{11}H\textsubscript{20}S\textsuperscript{107}Ag or C\textsubscript{11}H\textsubscript{20}S\textsuperscript{109}Ag (which have DBE values of 2) are possible. For the S\textsubscript{2} class, ions with molecular formulas of C\textsubscript{9}H\textsubscript{12}S\textsubscript{2}\textsuperscript{107}Ag or C\textsubscript{9}H\textsubscript{12}S\textsubscript{2}\textsuperscript{109}Ag (which have DBE values of 4) are possible.

For the ions at nominal \(m/z\) values of 353 and 355 exhibiting different dominant structures centered at drift times of \(~5.6\) ms and \(~6.2\) ms (Figure 2.5A), exact mass measurements suggest the ions comprise S\textsubscript{1}, S\textsubscript{2}, S\textsubscript{3}, S\textsubscript{2}O, SO, and S\textsubscript{2}O\textsubscript{2} classes of sulfur-containing compounds. Abundance information for the ions having these two nominal \(m/z\) values is presented in Table 2.4. In general, the most abundant species belong to the S\textsubscript{1}, S\textsubscript{2}, and S\textsubscript{3} classes. For the S\textsubscript{1} class, molecules with molecular formulas of C\textsubscript{17}H\textsubscript{10}S\textsuperscript{107}Ag or C\textsubscript{17}H\textsubscript{10}S\textsuperscript{109}Ag (which have DBE values of 13) or C\textsubscript{16}H\textsubscript{22}S\textsuperscript{107}Ag or C\textsubscript{16}H\textsubscript{22}S\textsuperscript{109}Ag (which have DBE values of 6) are possible. For the S\textsubscript{2} class, molecules with molecular formulas of C\textsubscript{14}H\textsubscript{14}S\textsubscript{2}\textsuperscript{107}Ag or C\textsubscript{14}H\textsubscript{14}S\textsubscript{2}\textsuperscript{109}Ag (which have DBE values of 8) are possible. For the S\textsubscript{3} class, molecules with molecular formulas of C\textsubscript{11}H\textsubscript{18}S\textsubscript{3}\textsuperscript{107}Ag or C\textsubscript{11}H\textsubscript{18}S\textsubscript{3}\textsuperscript{109}Ag (which have DBE values of 3) are possible.

### 2.3.2. Criteria for Proposing Structures for Matching Molecular Formulas

Molecular structures are assigned to \(t_d\) profiles based on the following three criteria. First, CID spectra have been used to select the best structures based on plausible fragmentation patterns for ions with the four nominal \(m/z\) values of 291, 293, 353, and 355. Here the intensities of peaks in the CID spectra are considered for the estimation of the relative abundances of specific molecules. Second, structures, which have already been reported in petroleum mixtures analysis receive higher priority. For example, cyclic structures (saturated and aromatic rings) are preferred over linear, extended structures\textsuperscript{24}, \textsuperscript{111-113}. Third, comparisons of theoretical CCS values with the experimental CCS profiles
have been used as criteria for structural selection. Consequently, compounds containing theoretical CCS values that do not match the $t_D$ profiles have not been considered further.

### 2.3.3. Proposed Structures for Ions Having Nominal $m/z$ Values of 291 and 293

Table 2.6 provides a number of different molecular structures that are considered for the molecular formulas associated with ions of nominal $m/z$ values of 291 and 293. The molecular formulas, DBE values, theoretical CCS and $t_D$ values, and relevant peaks found in the CID spectrum of the ions are also reported in this table. The CID spectrum for ions having a nominal $m/z$ value of 291 is shown in Figure 2.4B. Notably, in the CID spectrum, many peaks are observed. This is likely due to the many isobaric and isomeric species that can be associated with a specific nominal $m/z$ value (e.g., see Figure 2.3). In this work, efforts were directed in assigning the most intense peaks observed in the CID spectrum and correlating them to $S_1$, $S_2$, and $S_3$ classes which are mostly observed in FT-ICR spectrum. Therefore, not all of the observed CID fragment peaks are considered in structural comparisons. The dominant fragmentation pattern in this CID spectrum can easily be explained by the type of chemical bond cleaved in the proposed structures of the sulfur-containing ions. That is, bonds having lower dissociation energies break more readily than other bonds. Arguably, the detachment of silver cation ($^{107}\text{Ag}^+$ and $^{109}\text{Ag}^+$) is the most energetically favorable fragmentation pathway resulting in large peaks at $m/z$ 107 and 109. These two fragment ions are not shown in Figure 2.4 because of their relatively high abundances. The next dominant fragmentation trend is de-alkylation. Comparatively, the dissociation of C=C bonds is not a favorable pathway. Consequently, there is no observable high intensity peak that can be correlated to double bond cleavage. Notably, others have observed this behavior in studies of petroleum
compounds. In the CID spectrum in Figure 2.4B, for every single fragment peak, multiple neighboring peaks are observed. This can be explained by the fact that, the CID process has been performed on a select \( m/z \) value that can belong to different petroleum classes such as NS, OS, OS\(_2\), or N\(_2\)S. Therefore, the fragment ions from all of these classes are expected to be observed. That said, in the CID spectrum, most fragment ions result from precursor dissociation while silver is still attached.

Other molecular structures may also be considered to represent these molecular formulas based on fragmentation pattern analysis. These have not been considered for the theoretical CCS calculation comparison. Such compounds vary from the proposed structures shown in Table 2.6 only in terms of the location of double bonds, the location of the S-containing functional group on the molecule (only if it does not change the fragmentation pattern), the size of aromatic or aliphatic cycles, and the orientation of the two cycles relative to each other.
Table 2.6. Proposed structures for ions having nominal m/z values of 291 and 293

<table>
<thead>
<tr>
<th>Number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Structure&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Molecule Formula&lt;sup&gt;c&lt;/sup&gt;</th>
<th>DBE&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Fragment Ions&lt;sup&gt;e&lt;/sup&gt;</th>
<th>Theoretical CCS (t&lt;sub&gt;D&lt;/sub&gt;)&lt;sup&gt;f&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1" alt="Structure" /></td>
<td>C&lt;sub&gt;12&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;Ag</td>
<td>9</td>
<td>(107) (109)</td>
<td>77.7 (4.76)</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2" alt="Structure" /></td>
<td>C&lt;sub&gt;12&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;Ag</td>
<td>9</td>
<td>(107) (109)</td>
<td>80.6 (4.93)</td>
</tr>
<tr>
<td>3</td>
<td><img src="image3" alt="Structure" /></td>
<td>C&lt;sub&gt;12&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;Ag</td>
<td>9</td>
<td>(107) (109)</td>
<td>79.2 (4.85)</td>
</tr>
<tr>
<td>4</td>
<td><img src="image4" alt="Structure" /></td>
<td>C&lt;sub&gt;12&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;Ag</td>
<td>9</td>
<td>(107) (109)</td>
<td>79.5 (4.87)</td>
</tr>
<tr>
<td>5</td>
<td><img src="image5" alt="Structure" /></td>
<td>C&lt;sub&gt;12&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;Ag</td>
<td>9</td>
<td>(107) (109)</td>
<td>79.9 (4.89)</td>
</tr>
<tr>
<td>6</td>
<td><img src="image6" alt="Structure" /></td>
<td>C&lt;sub&gt;12&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;Ag</td>
<td>9</td>
<td>(107) (109)</td>
<td>78.3 (4.80)</td>
</tr>
<tr>
<td>7</td>
<td><img src="image7" alt="Structure" /></td>
<td>C&lt;sub&gt;9&lt;/sub&gt;H&lt;sub&gt;12&lt;/sub&gt;Ag</td>
<td>4</td>
<td>(M-14) (M-28) (M-42) (107) (109)</td>
<td>76.1 (4.66)</td>
</tr>
<tr>
<td>8</td>
<td><img src="image8" alt="Structure" /></td>
<td>C&lt;sub&gt;9&lt;/sub&gt;H&lt;sub&gt;12&lt;/sub&gt;Ag</td>
<td>4</td>
<td>(M-15) (M-28) (107) (109)</td>
<td>80.6 (4.93)</td>
</tr>
<tr>
<td>9</td>
<td><img src="image9" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;Ag</td>
<td>2</td>
<td>(M-15) (M-29) (M-43) (107) (109)</td>
<td>83.3 (5.09)</td>
</tr>
<tr>
<td>10</td>
<td><img src="image10" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;Ag</td>
<td>2</td>
<td>(M-15) (M-30) (M-45) (107) (109)</td>
<td>82.2 (5.03)</td>
</tr>
<tr>
<td>11</td>
<td><img src="image11" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;Ag</td>
<td>2</td>
<td>(M-15) (M-30) (M-44) (107) (109)</td>
<td>83.3 (5.09)</td>
</tr>
<tr>
<td>12</td>
<td><img src="image12" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;Ag</td>
<td>2</td>
<td>(M-15) (M-29) (M-42) (107) (109)</td>
<td>82.0 (5.01)</td>
</tr>
</tbody>
</table>
Table 2.6. To be continued

<table>
<thead>
<tr>
<th>Number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Structure&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Molecule Formula&lt;sup&gt;c&lt;/sup&gt;</th>
<th>DBE&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Fragment Ions&lt;sup&gt;e&lt;/sup&gt;</th>
<th>Theoretical CCS (t&lt;sub&gt;D&lt;/sub&gt;)&lt;sup&gt;f&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td><img src="image1" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;SAg</td>
<td>2</td>
<td>(M-15) (M-29) (M-42) (107) (109)</td>
<td>81.3 (4.97)</td>
</tr>
<tr>
<td>14*</td>
<td><img src="image2" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;SAg</td>
<td>2</td>
<td>(M-15) (M-29) (107) (109)</td>
<td>91.1 (5.55)</td>
</tr>
<tr>
<td>15*</td>
<td><img src="image3" alt="Structure" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;20&lt;/sub&gt;SAg</td>
<td>2</td>
<td>(M-15) (M-29) (107) (109)</td>
<td>89.0 (5.43)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Number assigned to each molecule  
<sup>b</sup> Molecular structure drawn by ChemDraw freeware  
<sup>c</sup> Molecular formula obtained using the mMass freeware  
<sup>d</sup> Double Bond Equivalence  
<sup>e</sup> Fragment ions from CID spectrum obtained from the Orbitrap mass spectrometer. M shows the m/z of the molecular ion.  
<sup>f</sup> Theoretical CCS and t<sub>D</sub> values calculated using the Mobcal software  
<sup>*</sup> Numbers represented with star are associated with structures which their theoretical CCS values were not fitted in the experimental t<sub>D</sub> profiles

Theoretical CCS values for select compound structures have been compared with experimental t<sub>D</sub> profiles for ions having nominal m/z values of 291 and 293. Figure 2.4A shows the t<sub>D</sub> distribution for these ions and the t<sub>D</sub> values for select compound structures obtained from CCS calculations (Table 2.6). In Figure 2.4A, the matching cross sections do not require that the ions are observed. Rather, the ions have matching CCS values, accurate masses, and confirming CID fragmentation patterns. From Table 2.6, structures 1, 2, 3, 4, 5, 6 all have the molecular formula of C<sub>12</sub>H<sub>6</sub>SAg and a DBE value of 9. Theoretical CCS values for these structures fall in the experimental t<sub>D</sub> profile. Structures 7 and 8 correspond to a molecule having the molecular formula of C<sub>9</sub>H<sub>12</sub>S<sub>2</sub>Ag and a DBE value of 4. The theoretical CCS value of these two
structures also match with the experimental $t_D$ profile. Structures 9, 10, 11, 12, and 13 (Table 2.6) all have the same molecular formula, C$_{11}$H$_{20}$SAg, and a DBE value of 2. Each of these structures have a theoretical CCS value falling within the experimental $t_D$ profile. Structures 14 and 15 have theoretical CCS values that do not match experimental $t_D$ profiles. Consequently, there were not considered as possible structures. Furthermore, the results from Table 2.6 show that S$_1$ structures with DBE values of 9 should be more prevalent than S$_1$ structures with DBE values of 2 although this cannot be observed from the CID data. In addition, S$_1$ structures should be more prevalent than S$_2$ structures regardless of DBE value.

Figure 2.4. A) $t_D$ profile for ions having nominal m/z values of 291 and 293. $t_D$ values for each number are correlated to structures (Table 2.6) as shown with arrows. B) CID spectrum for the dataset feature of nominal m/z value of 291 obtained with the LTQ device of the FT-ICR instrument. Several fragments are labeled.
2.3.4. Proposed Structures for Ions Having Nominal m/z Values of 353 and 355

Table 2.7 lists a number of structures that are proposed for the four molecular formulas considered for ions having a nominal m/z value 353 or 355. The molecular formula, DBE values, theoretical CCS and $t_D$ values, and potential ion matches from the CID spectrum for these structures are also shown in Table 2.7. The CID spectrum of ions exhibiting nominal m/z values of 353 is shown in Figure 2.5B. The same methodology used for the interpretation of the CID spectrum for lower mass ions is also used here. In the CID spectrum, two peaks (m/z 107 and 109) are observed that belong to the detachment of silver cation from the sulfur-containing compounds (not shown in Figure 2.5B). As before, significant de-alkylation is observed in the CID spectrum. Notably, others have observed this behavior in studies of petroleum compounds.\textsuperscript{19} Lastly, evidence for cleavages of carbon-sulfur bonds (C-S) is prevalent. Peaks observed at (M-60), (M-100), (M-114), and (M-128) can be correlated to M-(CH$_2$-CH$_2$-S), M-(CH$_2$-CH$_2$-S-CH-CH-CH-$\text{CH}_2$), M-(CH$_2$-CH$_2$-S-CH-CH-CH$_2$-CH$_2$), and M-(CH$_2$-CH$_2$-S-CH-CH-CH$_2$-CH-$\text{CH}_2$), respectively. These fragments would arise from compounds that belong to S$_2$ and S$_3$ classes of petroleum compounds. Specifically, for molecules that belong to the S$_3$ class, (M-92) and (M-118) fragments that correspond to M-(S-CH$_2$-CH$_2$-S) and M-(S-CH$_2$-CH$_2$-S-CH-CH) can be observed. It is important to note that many other structures may also be considered for these molecular formulas based on fragmentation patterns alone. These structures are ultimately not considered for theoretical CCS calculations.
Table 2.7. Proposed structures for ions having nominal m/z values of 353 or 355

<table>
<thead>
<tr>
<th>Number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Structure&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Molecule Formula&lt;sup&gt;c&lt;/sup&gt;</th>
<th>DBE&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Fragment Ions&lt;sup&gt;e&lt;/sup&gt;</th>
<th>Theoretical CCS (td)&lt;sup&gt;f&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td><img src="image1" alt="Structure" /></td>
<td>C₁₁H₁₈S₃Ag</td>
<td>3</td>
<td>(M-14) (M-28)</td>
<td>90.9 (5.54)</td>
</tr>
<tr>
<td>17</td>
<td><img src="image2" alt="Structure" /></td>
<td>C₁₄H₁₄S₂Ag</td>
<td>8</td>
<td>(M-60) (107) (109)</td>
<td>91.4 (5.57)</td>
</tr>
<tr>
<td>18</td>
<td><img src="image3" alt="Structure" /></td>
<td>C₁₄H₁₄S₂Ag</td>
<td>8</td>
<td>(M-14) (107) (109)</td>
<td>95.8 (5.83)</td>
</tr>
<tr>
<td>19</td>
<td><img src="image4" alt="Structure" /></td>
<td>C₁₄H₁₄S₂Ag</td>
<td>8</td>
<td>(M-60) (107) (109)</td>
<td>93.4 (5.69)</td>
</tr>
<tr>
<td>20</td>
<td><img src="image5" alt="Structure" /></td>
<td>C₁₄H₁₄S₂Ag</td>
<td>8</td>
<td>(M-15) (107) (109)</td>
<td>99.6 (6.05)</td>
</tr>
<tr>
<td>21</td>
<td><img src="image6" alt="Structure" /></td>
<td>C₁₄H₁₄S₂Ag</td>
<td>8</td>
<td>(M-60) (107) (109)</td>
<td>96.0 (5.84)</td>
</tr>
<tr>
<td>22</td>
<td><img src="image7" alt="Structure" /></td>
<td>C₁₄H₁₄S₂Ag</td>
<td>8</td>
<td>(M-14) (M-28)</td>
<td>91.7 (5.59)</td>
</tr>
<tr>
<td>23</td>
<td><img src="image8" alt="Structure" /></td>
<td>C₁₆H₂₂SAg</td>
<td>6</td>
<td>(M-15) (107) (109)</td>
<td>102.0 (6.19)</td>
</tr>
<tr>
<td>24</td>
<td><img src="image9" alt="Structure" /></td>
<td>C₁₆H₂₂SAg</td>
<td>6</td>
<td>(M-15) (107) (109)</td>
<td>106.0 (6.43)</td>
</tr>
</tbody>
</table>
Table 2.7. To be continued

<table>
<thead>
<tr>
<th>Number&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Structure&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Molecule Formula&lt;sup&gt;c&lt;/sup&gt;</th>
<th>DBE&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Fragment Ions&lt;sup&gt;e&lt;/sup&gt;</th>
<th>Theoretical CCS (t&lt;sub&gt;D&lt;/sub&gt;)&lt;sup&gt;f&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td><img src="image" alt="Structure 25" /></td>
<td>C&lt;sub&gt;17&lt;/sub&gt;H&lt;sub&gt;10&lt;/sub&gt;SAg</td>
<td>13</td>
<td>(107) (109)</td>
<td>94.5 (5.75)</td>
</tr>
<tr>
<td>26</td>
<td><img src="image" alt="Structure 26" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;18&lt;/sub&gt;S&lt;sub&gt;3&lt;/sub&gt;Ag</td>
<td>3</td>
<td>(M-60) (M-92) (M-118) (107) (109)</td>
<td>99.0 (6.02)</td>
</tr>
<tr>
<td>27</td>
<td><img src="image" alt="Structure 27" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;18&lt;/sub&gt;S&lt;sub&gt;3&lt;/sub&gt;Ag</td>
<td>3</td>
<td>(M-60) (M-92) (M-118) (107) (109)</td>
<td>92.1 (5.61)</td>
</tr>
<tr>
<td>28</td>
<td><img src="image" alt="Structure 28" /></td>
<td>C&lt;sub&gt;11&lt;/sub&gt;H&lt;sub&gt;18&lt;/sub&gt;S&lt;sub&gt;3&lt;/sub&gt;Ag</td>
<td>3</td>
<td>(M-60) (M-92) (M-118) (107) (109)</td>
<td>92.4 (5.63)</td>
</tr>
<tr>
<td>29*</td>
<td><img src="image" alt="Structure 29" /></td>
<td>C&lt;sub&gt;16&lt;/sub&gt;H&lt;sub&gt;22&lt;/sub&gt;SAg</td>
<td>6</td>
<td>(M-15) (107) (109)</td>
<td>114.9 (6.95)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Number assigned to each molecule

<sup>b</sup> Molecular structure drawn by ChemDraw freeware

<sup>c</sup> Molecular formula obtained using the mMass freeware

<sup>d</sup> Double Bond Equivalence

<sup>e</sup> Fragment ions from CID spectrum obtained from LTQ mode of FT-ICR instrument. M shows the m/z of the molecular ion.

<sup>f</sup> Theoretical CCS and t<sub>D</sub> values calculated using the Mobcal software.

* Numbers represented with star are associated with structures which their theoretical CCS values were not fitted in the experimental t<sub>D</sub> profiles.

**t<sub>D</sub> values associated with theoretical CCS values have been compared with the experimental t<sub>D</sub> profiles. Structure 16 (Table 2.7) would be expected to generate fragment ions that have the highest intensities such as those in the CID spectrum (Figure 2.5B). This structure corresponds to a molecule with a molecular formula of C<sub>11</sub>H<sub>18</sub>S<sub>3</sub>Ag and**
DBE value of 3. Structure 17 would also generate relatively high intensity fragment ion peaks such as those in the CID spectrum. Additionally, its theoretical CCS value matches more closely with the most intense feature in the experimental $t_D$ profile. Structures 19, 20, 21, and 22 (Table 2.7) show the potential to produce fragments that are observed experimentally and their theoretical CCS values fall within $t_D$ regions. Structure 23 matched within experimental $t_D$ profile. Structure 24 appears in a region with much lower intensities within the experimental $t_D$ profile. Structure 25 would not be expected to display the various fragments observed in the CID spectrum due to the strength of the double bonds in the highly aromatic structure. However, its theoretical CCS value matches with the most intense feature in the experimental $t_D$ profile, and it is therefore presented for comparison in Figure 2.5A. Structures 26, 27, and 28 belong to the $S_3$ class with DBE value of 3. All these compounds show fragments with relative higher intensities than other potential structures in the CID spectrum. Structures 27 and 28 show relative higher intensities than structure 26 in the $t_D$ profile. Structure 29 has theoretical CCS values that did not match experimental $t_D$ profiles. Consequently, it was not considered as possible match. This compound belongs to $S_1$ class which includes two rings in its structure.
Figure 2.5. A) $t_D$ profile for ions having nominal m/z values of 353 and 355. $t_D$ values for each number are correlated to structures (Table 2.7) as shown with arrows. B) CID spectrum for the dataset feature of nominal m/z 353 obtained with the LTQ device of the FT-ICR instrument. Several fragments are labeled.

In conclusion, the FT-ICR spectra, CID spectra, and experimental $t_D$ profiles show that structures containing 2 and 3 sulfur atoms are more abundant than structures that have only 1 sulfur atom for 353 and 355 nominal m/z values.
2.4. Conclusion

The experiments reported here highlight the power of a combined IMS-MS, tandem MS, and high mass accuracy measurement approach applied to sulfur containing species from petroleum samples. Overall, it is possible to better characterize compound structures for sulfur-containing species present in petroleum mixtures. Generally, compounds of lower \( m/z \) values exhibit more elongated structures and structures with higher DBE values (more compact structures) are favored at higher \( m/z \) values. The method provides information about prevalent functional groups, structural motifs, and ion abundance contributions from plausible compounds within specific \( m/z \), \( t_D \), and ion fragment ranges. Comparisons with precursor ions of two different nominal \( m/z \) values reveal that molecule type (\( S_1, S_2, S_3 \)) and DBE value can be associated with different dataset features. The functional group information may be valuable for predicting the relative chemical reactivity of petroleum. Furthermore, structural information may provide insight regarding the selection of appropriate catalytic processes during petroleum refining. It should be noted that similar studies may be employed to elucidate the structure of other hetero-atom containing compounds in petroleum.

Although the work presented here is proof-of-principle, several improvements would significantly enhance the characterization capabilities. For example, tuning the CID energy and applying sets of standards with different structural motifs can improve the interpretation of CID data for petroleum ions. Additionally, novel approaches employing \( t_D \) shift reagents can be utilized. One example involves energy-resolved fragmentation at the back of the drift tube for differentiation of isomeric species.\(^{69}\)
2.5. References


19. Qian, K. In Determination of structural building blocks in heavy petroleum systems by collision-induced dissociation Fourier transform ion cyclotron resonance mass spectrometry, American Chemical Society: 2012; pp ENFL-398.


30. Song, C.; Ma, X. L., New design approaches to ultra-clean diesel fuels by deep desulfurization and deep dearomatization. Applied Catalysis B-Environmental 2003, 41 (1-2), 207-238.


86. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V., Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics. *Journal of the American Chemical Society* 2012, 134 (7), 3429-3438.
111. Snyder, L. R., NITROGEN AND OXYGEN COMPOUND TYPES IN PETROLEUM - TOTAL ANALYSIS OF A 400-700 DEGREES F DISTILLATE FROM A CALIFORNIA CRUDE OIL. *Analytical Chemistry* 1969, 41 (2), 314-
112. Snyder, L. R., NITROGEN AND OXYGEN COMPOUND TYPES IN PETROLEUM - TOTAL ANALYSIS OF AN 850-1000 DEGREES F DISTILLATE FROM A CALIFORNIA CRUDE OIL. *Analytical Chemistry* 1969, 41 (8), 1084-
3. Compound Characterization Using Hydrogen Deuterium/Exchange (HDX) and Isotope Scrambling Coupled to IMS-MS and CID


3.1. Introduction

Analysis of compounds in complex mixtures encounter numerous problems resulting from different molecular masses, concentrations, polarities, and molecular structures.1-3 Analysis techniques used to characterize compounds in complex mixtures include a diverse range of techniques.4, 5 Thorough analyses often require a combination of different methods which can compromise factors such as speed, selectivity, sensitivity, and cost. To a great extent the sample complexity is determinant in the instrumentation method utilized for sample analysis. For example, nuclear magnetic resonance (NMR) spectroscopy is best suited for the analysis of concentrated species.6-11 Mass spectrometry (MS) is employed mostly for the characterization of a wide range of compounds with different polarities and size as well as concentration.2 Another advantage of MS analysis is the ability to fragment precursor ions. Collision-induced dissociation (CID) is mostly used for fragmentation of small molecules.12 The fragmentation patterns are considered fingerprints for the identification of compounds.2, 13, 14

Despite having advantages in terms of sensitivity and compound identification, direct MS suffers from the competitive ionization phenomenon. Furthermore, it has difficulty in the characterization of isomeric and isobaric species. Therefore, coupling MS with separation techniques is required to improve compound characterization. Liquid
chromatography (LC), gas chromatography (GC), and ion mobility spectrometry (IMS) are mostly used in hyphenated methods for this purpose.\textsuperscript{15-17} Each of these hyphenated methods have advantages and disadvantages for compound characterization in complex mixtures. Although LC-MS has potential for the separation of isomeric and isobaric species with different polarities, it suffers from long analysis times and relatively poor reproducibility. GC-MS is a technique for the analysis of volatile species and therefore a derivatization step is required for the analysis of polar compounds before analysis. IMS-MS provides significant structural and conformation information.\textsuperscript{18, 19} Furthermore, it has been used for the analysis of isomeric and isobaric species.\textsuperscript{19} That said, IMS-MS is a post-ionization method and incurs competitive ionization problems.

An IMS drift tube is composed of metal rings separated by insulators.\textsuperscript{20, 21} The drift tube is filled with He buffer gas. Ions are guided in the drift tube using a constant electric field. Separation of ions occurs based on different collision interactions with the buffer gas. Ions that have more compact structures reach the detector sooner than ions that have more elongated structures. Coupling IMS to MS provides many advantages including increased peak capacity.\textsuperscript{22} IMS-MS enables the separation of some isomeric and isobaric species.\textsuperscript{19} The ion transit time or drift time ($t_D$) of ions obtained from IMS-MS measurements can be employed to calculate experimental collision cross section (CCS) values using the Mason-Schamp equation (Equation 1 in the Introduction).\textsuperscript{20} CCS values are unique physicochemical characteristics of ions and can be used to aid compound characterization in complex mixtures. Moreover, in two-dimensional (2D) plots of $m/z$ vs $t_D$, different classes of compounds fall into distinct regions.\textsuperscript{18} For example, classes of lipids fall into specific regions in 2D plots based on their distinct headgroups.
Gas-phase hydrogen/deuterium (HDX) coupled to MS (HDX-MS) is usually used for conformational analysis of proteins but it is also used for the analysis of small molecule structure.\textsuperscript{23-28} Gas-phase HDX can be coupled to IMS-MS to enhance compound characterization in complex mixtures.\textsuperscript{29} During the reaction, deuterium atoms can exchange with labile hydrogen atoms resulting in an increase in the $m/z$ of the ions. Exchange proceeds via a relay mechanism in which a hydrogen-bonded intermediate forms between the charge site, another basic site, and D$_2$O.\textsuperscript{30,31} Thus, HDX profiles are dependent on the proximity of charge sites to basic sites resulting from different conformational forms of ions. It has been previously demonstrated that HDX profiles for peptides depends to some degree on their amino acid sequence.\textsuperscript{32} Another important factor affecting HDX profiles is the partial pressure of the reagent gas. Therefore, isotopic distributions (HDX profiles) of ions can be very distinctive due to differences in reaction kinetics. This demonstrates that although IMS-HDX-MS has been employed for conformational studies of peptides and proteins\textsuperscript{29}, the technique has potential for small molecule characterization in complex mixtures. The idea is to compare HDX profiles and CCS values of unknown compounds in such mixtures with the equivalent parameters of standard compounds. In order to compare datasets, the robustness of the measurements must be established. Pressure control may be the most challenging part for generating reproducible data for IMS-HDX-MS experiments.

Hydrogen/deuterium (HD) scrambling can occur upon increasing the internal energy of ions through energizing collisions.\textsuperscript{33,34} Thus, hydrogen sites that are less exposed to D$_2$O can exchange for deuteriums. Ultimately, this generates more exchangeable sites where, the overall deuterium uptake level can increase after HD
scrambling experiments if reagent gas is present. Therefore, collisional activation in the drift tube can be employed as another analytical method for compound characterization in complex mixtures. Notably, studies suggest that HD scrambling on peptide ions depends on the amino acid sequence\textsuperscript{33, 35, 36} and the identity of charge carrier.\textsuperscript{35, 37, 38}

In the current work, the combination of gas-phase HDX and HD scrambling with IMS-MS is used to investigate the potential for characterizing of compounds in complex mixtures. Different classes of compounds were studied and their HDX behavior is reported. For application of these techniques in the routine analysis of compounds, the reproducibility of separate measurements must be high. Numerous reproducibility measurements have been performed to evaluate the analytical power of the method. Lastly, the power of this technique is evaluated for compound identification efforts in a bovine heart extract sample.

3.2. Experimental

3.2.1. Materials

Dopamine hydrochloride (99%, AK Scientific), guanosine (99%, Acros Organics), acetaminophen (98%, Sigma Aldrich), alanine (98%, Alfa Aesar) were purchased and used without further purification. LC-MS grade water, acetonitrile, formic acid, acetic acid, D\textsubscript{2}O (99.9% D), ammonium bicarbonate (≥ 99.5%), urea (ACS reagent grade), ubiquitin (bovine, 98.0%), pepsin (porcine, ≥2,500 units/mg protein) were purchased from Sigma-Aldrich (St. Louis, MO). The peptides bradykinin (≥95.0%) and substance P (≥95.0%) were purchased from GenScript (Piscataway, NJ). Synthetic peptide KKDDDDDDIIKKII
(KKD peptide, 94.1%) was also purchased from GenScript. Bovine heart extract was purchased from Lipid Avanti, Inc.

### 3.2.2. Sample Preparation

A 1.0 mg/mL solution of Ubiquitin was made by dissolving the protein in ultra pure water (18 megohm) followed by subsequent acidification and then was acidified using acetic acid (0.1 M) to a PH of ~2.8. Pepsin was added to the protein solution to a final concentration of 0.33 mg/mL. Protein digestion was performed at 37 °C for 60 minutes. Acetonitrile and acetic acid were then added to the protein solution such that the final ESI solution composition was 1:1 water:acetonitrile with 1% acetic acid. The sample was loaded into a gas tight syringe. The syringe was placed into a syringe pump and the sample was infused at 300 nL/min through the electrospray needle which is biased at +2000 V relative to the desolvation region entrance. A mixture of the peptide standards was made in 1:1 water: acetonitrile with 0.1% (v/v) acetic acid. The concentration of peptides were 45 µg/mL (bradykinin and substance P) and 450 µg/mL (KKD). Dopamine, guanosine, acetaminophen, and alanine standard samples were made by dissolving 10 mg of each metabolite in methanol and ACN (1:1 v/v) to a final volume of 100 mL (0.1 mg/mL). Acetic acid (1%) was used to enhance the ionization of these compounds. A lipid working solution was prepared by mixing 250 mL of methylene chloride, 250 mL of methanol, and 5 mL of ammonium acetate solution. The bovine heart extract was prepared by diluting 400 µL of bovine heart extract with 3600 µL of lipid working solution.
3.2.3. Data Collection Method

A home-built, dual-gate IMS device coupled to a thermo LTQ Velos mass spectrometer (Thermo Fisher, San Jose, CA) was used for data collection. A total of 3 sample sets were collected for sequential and random drift time analyses. For random drift time data collection, the delay times between initiation and selection gates were randomized (For example, 6.2 ms, 5.2 ms, 4.8 ms, etc.). For sequential collection datasets, the delay times were stepped through entire drift time range. The collection of random and sequential datasets was also randomized resulting in triplicated analysis for both types. Random selections were achieved using the pseudo random number generator in the Excel software (Microsoft, Washington). For reproducibility experiments, ~2.60 Torr of the buffer gas was used.

For all HDX experiments partial pressures of ~2.50 and ~0.02 Torr were used for He and D$_2$O, respectively. A constant temperature of 35°C was applied to tube transferring D$_2$O to the drift tube to avoid condensation. This prevents errors in measurements resulting from pressure fluctuations. Very little D$_2$O is used under these conditions. Drift time reproducibility using the ubiquitin digest samples subjected to HDX was obtained using the random delay selection process. Three trials were collected. No drift time selection was used in the HDX reproducibility studies of the model peptides. In all, for each HDX and HD scrambling experiments a total of 6 replicates were performed. To perform the reproducibility experiments, the voltage on the drift tube was turned off and drift tube gases were removed between each replicate. The drift tube partial pressures were returned to the same values using constant leak valve settings (Granville Phillips, MKS Instruments, Andover, MA). The buffer gas temperature was determined
using a thermocouple attached to the stainless steel plate through which the buffer gas is introduced; notably, gentle heating of D$_2$O introduction lines to prevent condensation did not measurably change the temperature. The buffer gas pressure was measured using a capacitance manometer.

IMS-MS measurements provide drift time ($t_D$), m/z, and intensity values for different ions. Two-dimensional (2D) contour plots of $m/z$ vs $t_D$ were generated using the DPlot software (v. 2.3.5.3, HydeSoft Computing, LLC; Vicksburg, MS, USA). $t_D$ and m/z datasets corresponding to ubiquitin digest peptide ions were compared to examine the reproducibility in drift time. Coefficients of variation (CVs) associated with $t_D$ values were calculated for several dataset features associated with ubiquitin digest peptide ions.

HDX reproducibility in the absence and in the presence of HD scrambling was determined for the peptide mixture containing three model peptides without $t_D$ selection. HD scrambling experiments were performed by accelerating ions in the last drift tube region (~0.1 m from drift tube exit). Consequently, HDX was carried out in this drift tube region. In all, six replicates were collected for these model peptide mixture. Isotopic distribution reproducibility determinations were made for digest peptides presented above and small-molecule metabolites.

3.3. Results and Discussion

3.3.1. HDX and HD Scrambling for Distinguishing Molecules

As discussed previously, implementation of new analytical methods to improve the characterization of isomeric and isobaric compounds is in high demand. In order to use an additional method for the characterization of compounds in complex mixtures, it should
probe a specific but separate physicochemical property of the molecule or ion. Regarding
gas-phase HDX, the physicochemical property is structure as related to the positions of
charge sites and exchange sites as well as their surface accessibility. Therefore,
molecules with different orientations of charge sites and basic sites are expected to
demonstrate different HDX patterns. Figure 3.1 demonstrates HDX patterns of 3 peptide
ions present in the model peptide mixture. For example, for [M+3H]^{3+} bradykinin ions, ~6
deuteriums are incorporated (Figure 3.1A). In comparison, for the [M+3H]^{3+} ions of
substance P, little to no deuterium (~1) uptake is observed as shown in Figure 3.1B.
Lastly, for [M+2H]^{2+} KKD peptide ions, deuterium uptake is observed to be ~13 (Figure
3.1C).

At the partial pressure setting used in these experiments, it is expected that all
fast-exchanging hydrogens (easily accessible) reach. In order to understand the origin of
differences in deuterium uptake, the numbers of exchangeable hydrogens for these
peptides were considered. The number of exchangeable hydrogens for triply-charged
bradykinin, triply-charged substance P, and doubly-charged KKD peptide ions are 20, 25,
and 30, respectively.
Figure 3.1. Different deuterium uptake patterns for model peptide ions. A) [M+3H]^{3+} bradykinin ions. B) [M+3H]^{3+} substance P ions. C) [M+2H]^{2+} KKD peptide ions. The black line is the isotopic distribution in the absence of D_2O and the blue line is the isotopic distribution in the presence of D_2O.

Even though these ions have nearly the same number of exchangeable hydrogens, they demonstrate significant differences in deuterium incorporation levels.
Presumably this is due to differences in exchange site accessibility to basic sites and the accessibility to collisions with D$_2$O. In summary, doubly-charged KKD peptide ions demonstrate the much higher HDX efficiency than triply-charged substance P and bradykinin ions. This results despite the fact that the former ions have fewer reaction centers.

Another method that can be used to further characterize molecules in complex mixtures is induced HD scrambling followed by HDX. As described above HD scrambling can be explained in the context of "mobile proton" model.$^{32}$ Briefly, collisional activation of ions leads to intramolecular proton transfer events and the deuteriums can also be mobilized. Therefore, less accessible hydrogens can exchange with deuteriums during the scrambling process thereby generating newly exchangeable hydrogen for further exchange. Therefore, it should be possible to observe more deuterium uptake after HD scrambling. Figure 3.2 shows the HDX profiles for the same peptides represented in Figure 3.1 after HD scrambling and further HDX in the drift tube. As shown in Figure 3.2, all peptide ions gain mass after HD scrambling. For doubly-charged KKD peptide ions, more deuterium uptake is observed than for the triply-charged substance P and bradykinin ions upon HD scrambling.
Figure 3.2. Different deuterium uptake patterns for model peptides after HDX and HD scrambling with D₂O added to the drift tube after ion activation. A) [M+3H]⁺⁺ bradykinin ions. B) [M+3H]⁺⁺ substance P ions. C) [M+2H]⁺⁺ KKD peptide ions. The black line is the isotopic distribution in the presence of D₂O and the blue line is the isotopic distribution in the presence of D₂O and ion activation.
3.3.2. Application of HDX and Scrambling Experiments for Metabolite Analysis

The application of HDX and scrambling experiments were extended to small-molecule metabolite samples in order to demonstrate the universality of these techniques for complex mixture analysis. Figure 3.3A, 3.3C, 3.3E and 3.3G present the isotopic distributions for alanine, guanosine, dopamine, and acetaminophen, respectively after being subjected to gas-phase HDX in the drift tube. Figure 3.3B, 3.3D, 3.3F, and 3.3H show the isotopic distributions after collisional activation of the same ions in the drift tube. The molecular structures of these ions and recorded reduced mobilities are presented in Table 3.1.

Table 3.1. Structures of metabolite compounds used for the IMS-HDX-MS experiments and their reduced mobilities.

<table>
<thead>
<tr>
<th>Name</th>
<th>m/z(^a)</th>
<th>Structure(^b)</th>
<th>Reduced Mobility(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>90.1</td>
<td><img src="image" alt="Alanine Structure" /></td>
<td>8.84 X 10(^{-4})</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>152.2</td>
<td><img src="image" alt="Acetaminophen Structure" /></td>
<td>7.8 X 10(^{-4})</td>
</tr>
<tr>
<td>Dopamine</td>
<td>154.2</td>
<td><img src="image" alt="Dopamine Structure" /></td>
<td>7.99 X 10(^{-4})</td>
</tr>
<tr>
<td>Guanosine</td>
<td>284.2</td>
<td><img src="image" alt="Guanosine Structure" /></td>
<td>5.76 X 10(^{-4})</td>
</tr>
</tbody>
</table>

\(^a\) Nominal m/z associated with the monoisotopic peak for the respective compound  
\(^b\) Molecular structure obtained from ChemDraw freeware
Figure 3.3. Isotopic distributions of precursor ions of A) alanine, C) guanosine, E) dopamine, and G) acetaminophen. Distributions for mobility selected ions in the presence of He only and He and D$_2$O are shown as blue and orange traces, respectively. Panels B, D, F, and H show the isotopic distributions (green traces) for the respective molecular ions obtained after collisional activation in the drift tube.

[M+H]$^+$ alanine ions have a nominal $m/z$ value of 90. As demonstrated in Figure 3.3A, these ions incorporate ~1.3 deuteriums, on average. This is evident from the fact the ions with a nominal $m/z$ of 91 are the most abundant ions. Upon collisional activation
of the ions in the drift tube, increased deuterium incorporation is evident by the fact that ions of nominal $m/z$ 92 are in equal abundance to those of $m/z$ 91. The average deuterium incorporation increases to 1.7 deuteriums. As with the peptide ions, the isotopic patterns are extremely reproducible using the evaporative introduction system; average RMSD values are 1.1% and 1.6% for the precursor and activated ions, respectively.

The behavior of alanine ions can be contrasted with that of [M+H]$^+$ guanosine ions. Guanosine has 7 exchangeable hydrogens. Protonated guanosine ions have a nominal $m/z$ value of 284. Upon introduction of D$_2$O in the drift tube, the isotopic distribution of these ions shows a high intensity peak at $m/z$ 285 as shown in Figure 3.3B. Figure 3.3D, shows that after activation of the ions, the peak at $m/z$ 285 remains the most abundant peak. Indeed the isotopic distributions for the mobility-selected ions are nearly identical. On average, the deuterium incorporation is ~1.0 for the precursor and the collisionally-activated ions. As with the other ions presented in this study, the reproducibility of the HDX measurements is high; average RMSD values for the isotopic distribution comparisons of the precursor and activated ions are 0.7% and 1.5%, respectively.

[M+H]$^+$ dopamine ions have a nominal $m/z$ value of 154. As demonstrated in Figure 3.3E, these ions do not undergo appreciable exchange upon introduction of D$_2$O reagent gas to the drift tube. Additionally, after collisional activation in the drift tube, the dopamine ions do not significantly increase the level of deuterium incorporation (Figure 3.3F). Indeed, using the average mass difference computed from the isotopic distributions, no deuteriums are incorporated for the precursor ions and the collisionally activated ions. The average RMSD value for the isotopic distribution comparisons are 1.1% and 0.5% for the precursor and activated ions, respectively.
Finally, Figure 3.3G and 3.3H present the HDX behavior of [M+H]^+ acetaminophen ions. The ions have a nominal m/z value of 152 and contain up to 3 exchangeable hydrogens. As shown in Figure 3.3G, with the addition of D_2O to the drift tube, the dominant feature has a nominal m/z value of 153 with an increase in features of m/z 154 and 155. After activation of the ions, the higher m/z peaks slightly increase in relative intensity within the isotopic distribution. For isotopic distribution comparisons, average RMSD values of 1.2% and 1.0% are obtained for the precursor and activated ions, respectively. The data indicate that acetaminophen ions uptake ~2.3 deuteriums with the introduction of D_2O and that the uptake increases with ion activation. Overall, Figure 3.3 shows different, distinguishing HDX behavior for the various ions.

3.3.3. Drift Time Reproducibility after HDX Experiments

The matching of ion drift times in He and in He/D_2O buffer mixtures can be used as another analytical tool for ion assignments in ‘omics mixtures. Therefore, high reproducibility of these measurements is desired. One question that arises is whether gating order affects the reproducibility of collected data. In order to answer this question, drift times were collected by sequentially or randomly stepping through He buffer gas as described in the Methods Section. Three replicates of each type of acquisition were obtained for a peptide digest of ubiquitin. The average CVs for sixteen peptide ion peaks is determined to be 0.52 ± 0.11% and 0.33 ± 0.23% for random and sequential data collection, respectively. Two-dimensional t_d vs m/z contour plots for the 16 peptide ions are shown in Figure 3.4 and Figure 3.5 for random and sequential data collections, respectively. The identities of some of peptide ions and their respective drift times are
also provided in Table 3.2 and Table 3.3 for random and sequential data collections, respectively.

Figure 3.4. Two dimensional (2D) drift time, \( m/z \) plots for triplicate measurements of a ubiquitin digest mixture. The data were recorded for conditions employing He buffer gas and random mobility selection. Several dataset features are indicated by numbers and correspond to the assigned ions in Table 3.2.
Table 3.2. Drift times for assigned peptide ions from experimental runs employing random mobility selection.

<table>
<thead>
<tr>
<th>Peptide Ion a</th>
<th>m/z b</th>
<th>Run 1 c</th>
<th>Run 2 c</th>
<th>Run 3 c</th>
<th>Average d</th>
<th>CV e</th>
</tr>
</thead>
<tbody>
<tr>
<td>[QRLIFAGKQ+3H]3+</td>
<td>354.58</td>
<td>6.96</td>
<td>7.01</td>
<td>6.93</td>
<td>6.97</td>
<td>0.58</td>
</tr>
<tr>
<td>[NIQ+H]1+</td>
<td>374.27</td>
<td>8.25</td>
<td>8.30</td>
<td>8.22</td>
<td>8.26</td>
<td>0.49</td>
</tr>
<tr>
<td>[HLVRL+2H]2+</td>
<td>375.77</td>
<td>6.91</td>
<td>6.97</td>
<td>6.90</td>
<td>6.93</td>
<td>0.55</td>
</tr>
<tr>
<td>[GIPPDQQ+2H]2+</td>
<td>377.75</td>
<td>6.33</td>
<td>6.39</td>
<td>6.35</td>
<td>6.36</td>
<td>0.48</td>
</tr>
<tr>
<td>[KIQDKE+2H]2+</td>
<td>380.72</td>
<td>6.20</td>
<td>6.25</td>
<td>6.20</td>
<td>6.22</td>
<td>0.46</td>
</tr>
<tr>
<td>[VLRLRG+2H]2+</td>
<td>385.77</td>
<td>6.94</td>
<td>6.98</td>
<td>6.92</td>
<td>6.95</td>
<td>0.44</td>
</tr>
<tr>
<td>[IFAGQL+2H]2+</td>
<td>388.75</td>
<td>6.07</td>
<td>6.11</td>
<td>6.10</td>
<td>6.09</td>
<td>0.34</td>
</tr>
<tr>
<td>[DKE+H]1+</td>
<td>391.24</td>
<td>8.46</td>
<td>8.55</td>
<td>8.43</td>
<td>8.48</td>
<td>0.74</td>
</tr>
<tr>
<td>[VKLTGKTITL+3H]3+</td>
<td>392.28</td>
<td>7.11</td>
<td>7.07</td>
<td>7.05</td>
<td>7.08</td>
<td>0.43</td>
</tr>
<tr>
<td>[NIQKESTLHL+3H]3+</td>
<td>394.92</td>
<td>6.46</td>
<td>6.50</td>
<td>6.46</td>
<td>6.47</td>
<td>0.36</td>
</tr>
<tr>
<td>[KESTLHL+2H]2+</td>
<td>414.25</td>
<td>6.49</td>
<td>6.56</td>
<td>6.52</td>
<td>6.52</td>
<td>0.54</td>
</tr>
<tr>
<td>[QRL+H]1+</td>
<td>416.26</td>
<td>8.51</td>
<td>8.56</td>
<td>8.47</td>
<td>8.51</td>
<td>0.53</td>
</tr>
<tr>
<td>[NVKAA]+</td>
<td>431.33</td>
<td>9.07</td>
<td>9.16</td>
<td>9.08</td>
<td>9.10</td>
<td>0.54</td>
</tr>
<tr>
<td>[NIQKESTL+2H]2+</td>
<td>466.79</td>
<td>7.61</td>
<td>7.70</td>
<td>7.60</td>
<td>7.64</td>
<td>0.72</td>
</tr>
<tr>
<td>[HLVL]+</td>
<td>481.34</td>
<td>9.42</td>
<td>9.49</td>
<td>9.40</td>
<td>9.44</td>
<td>0.50</td>
</tr>
<tr>
<td>[VRL]+</td>
<td>500.40</td>
<td>10.62</td>
<td>10.73</td>
<td>10.60</td>
<td>10.65</td>
<td>0.66</td>
</tr>
</tbody>
</table>

aPeptide ion assignments based on m/z matches to monoisotopic peaks and charge (obtained from isotopic distribution)
bExperimental m/z values for monoisotopic peaks from the 2D datasets
cIon drift times in ms as obtained from 2D dataset centroids for the assigned ions
dAverage ion drift times in ms for the three experimental runs
eCoefficients of variation for the drift times expressed as a percentage
Figure 3.5. Two dimensional (2D) drift time, m/z plots for triplicate measurements of a ubiquitin digest mixture. The data were recorded for conditions employing He buffer gas and sequential mobility selection. Several dataset features are indicated by numbers and correspond to the assigned ions in Table 3.3.
Table 3.3. Drift times for ubiquitin peptide ions recorded using a sequential drift time selection process. The buffer gas was He.

<table>
<thead>
<tr>
<th>Peptide Iona</th>
<th>m/zb</th>
<th>Run 1c</th>
<th>Run 2c</th>
<th>Run 3c</th>
<th>Average d</th>
<th>CVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>[QRLIFAGKQ+3H]3+</td>
<td>354.58</td>
<td>7.00</td>
<td>6.96</td>
<td>6.99</td>
<td>6.98</td>
<td>0.30</td>
</tr>
<tr>
<td>[NIQ+H]+</td>
<td>374.27</td>
<td>8.26</td>
<td>8.23</td>
<td>8.20</td>
<td>8.22</td>
<td>0.26</td>
</tr>
<tr>
<td>[HLVLR+2H]2+</td>
<td>375.77</td>
<td>6.93</td>
<td>6.91</td>
<td>6.91</td>
<td>6.91</td>
<td>0.00</td>
</tr>
<tr>
<td>[GIPPDQ+2H]2+</td>
<td>377.75</td>
<td>6.31</td>
<td>6.35</td>
<td>6.29</td>
<td>6.32</td>
<td>0.67</td>
</tr>
<tr>
<td>[KIQDE+2H]2+</td>
<td>380.72</td>
<td>6.23</td>
<td>6.20</td>
<td>6.18</td>
<td>6.19</td>
<td>0.23</td>
</tr>
<tr>
<td>[VLRLRG+2H]2+</td>
<td>385.77</td>
<td>6.98</td>
<td>6.93</td>
<td>6.91</td>
<td>6.92</td>
<td>0.20</td>
</tr>
<tr>
<td>[IFAGK+2H]2+</td>
<td>388.75</td>
<td>6.07</td>
<td>6.00</td>
<td>6.06</td>
<td>6.03</td>
<td>0.70</td>
</tr>
<tr>
<td>[DKE+H]+</td>
<td>391.24</td>
<td>8.47</td>
<td>8.40</td>
<td>8.50</td>
<td>8.45</td>
<td>0.84</td>
</tr>
<tr>
<td>[VKLTGKT+3H]3+</td>
<td>392.28</td>
<td>7.08</td>
<td>7.07</td>
<td>7.08</td>
<td>7.08</td>
<td>0.10</td>
</tr>
<tr>
<td>[NIQESTHL+3H]3+</td>
<td>394.92</td>
<td>6.50</td>
<td>6.48</td>
<td>6.46</td>
<td>6.47</td>
<td>0.22</td>
</tr>
<tr>
<td>[KESTL+2H]2+</td>
<td>414.25</td>
<td>6.51</td>
<td>6.50</td>
<td>6.47</td>
<td>6.49</td>
<td>0.33</td>
</tr>
<tr>
<td>[QRL+H]+</td>
<td>416.26</td>
<td>8.53</td>
<td>8.46</td>
<td>8.51</td>
<td>8.49</td>
<td>0.42</td>
</tr>
<tr>
<td>[NVKAA]+</td>
<td>431.33</td>
<td>9.08</td>
<td>9.03</td>
<td>9.08</td>
<td>9.06</td>
<td>0.39</td>
</tr>
<tr>
<td>[NIQESTL+2H]2+</td>
<td>466.79</td>
<td>7.68</td>
<td>7.65</td>
<td>7.67</td>
<td>7.66</td>
<td>0.18</td>
</tr>
<tr>
<td>[HLVL]+</td>
<td>481.34</td>
<td>9.44</td>
<td>9.40</td>
<td>9.43</td>
<td>9.42</td>
<td>0.23</td>
</tr>
<tr>
<td>[VLRL]+</td>
<td>500.40</td>
<td>10.68</td>
<td>10.70</td>
<td>10.66</td>
<td>10.68</td>
<td>0.26</td>
</tr>
</tbody>
</table>

aPeptide ion assignments based on m/z matches to monoisotopic peaks and charge (obtained from isotopic distribution)
bExperimental m/z values for monoisotopic peaks from the 2D datasets
cIon drift times in ms as obtained from 2D dataset centroids for the assigned ions
dAverage ion drift times in ms for the three experimental runs
eCoefficients of variation for the drift times expressed as a percentage

It can be argued that variation in buffer gas pressure for the mixture of He and D₂O is more prone to error due to the relatively small amount of D₂O used in these experiments (compared to the He buffer gas). Therefore, small fluctuations in partial pressure of D₂O gas may lead to significant errors for drift time measurements. The following experiments were performed in order to study reproducibility of drift time values in a mixture of He and D₂O buffer gas. Figure 3.6 shows the drift time run-to-run variability for the ubiquitin digest peptide ions after experiencing HDX. Dataset features for nine [M+H]+, [M+2H]2+, and [M+3H]3+ ions show very little difference in drift time (Figure 3.6A). Consider the CV for
peak 9 (Figure 3.6A) which has been assigned to the peptide ion \([\text{VKTLTGKTITAL}+3\text{H}]^{3+}\). The CV value is 0.36% across all three replicates (Figure 3.6B).

![Figure 3.6. IMS-MS plots of two ubiquitin digest trials (A) demonstrating the run-to-run reproducibility for ion from a complex mixture upon addition of \(\text{D}_2\text{O}\) to reagent gas. Panel B shows CV values for assigned peptide ions. CV values are provided for dataset features tentatively identified as 1- QRLIFAGKQ+3H]^{3+}; 2- [NIQ+H]^+; 3- [HLVRL+2H]^{2+}; 4- [GIPPDQQ+2H^]^{2+}; 5- [KIQDKE+2H]^{2+}; 6- [VRLRGG+2H]^{2+}; 7- [IFAGKL+2H]^{2+}; 8- [DKE+H]^+; 9- [VKTLTGKTITAL+3H]^{3+} ions as labeled in (B).

Labeled dataset features examined in the drift time distributions. An average CV of 0.55 ± 0.20% (Figure 3.6) was found. Overall, for the reproducibility of IMS experiments which uses a dopant gas still a high reproducibility range was achieved associated with mobility measurements.

3.3.4. Isotopic Distribution Reproducibility in HDX Experiments

The reproducibility of HDX patterns can be tested via comparisons of RMSD values for combined spectral features. To perform these experiments, special attention was given to maintaining the partial pressures of constituent buffer gases. To challenge the reproducibility of each measurement, \(\text{D}_2\text{O}\) gas was removed for each trial and then
leaked back into the drift tube to reach to the same partial pressure. This was performed for six replicates. For all observable ions the RMSDs were all found to be less than 1% (see Figures 3.7 and 3.8).

Figure 3.7. Isotopic distributions for [M+2H]^{2+} KKD (A), [M+2H]^{2+} substance P (B), and [M+3H]^{3+} bradykinin (C) peptide ions. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.
Figure 3.8. Isotopic distributions for \([M+3H]^3+\) KKD (A), \([M+3H]^3+\) substance P (B), and \([M+4H]^4+\) KKD (C) peptide ions. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.

To consider achievable reproducibility for a peptide mixture upon using mobility selection, isotopic RMSDs have been calculated for the ubiquitin digest ions (Figure 3.4 and 3.5). For example, mobility extraction of the features associated with \([DKE+H]^+\), \([VLRLRGG+2H]^2+\), and \([QRLIFAGKQ+3H]^3+\) peptide ions, provides isotopic distribution RMSD values that are all below 1.5%. Thus high reproducibility is still achieved even for mobility-selected features with the lower intensity. That said, it is instructive to study the RMSD changes at different concentrations because the concentration profile for compounds in complex mixtures varies significantly. Notably, the RMSD increases for lower intensity species such as the \([HLVLRL+2H]^2+\) peptide ions yielding a value of ~2.5%; these ions have an abundance that is 20 fold less than the \([VLRLRGG+2H]^2+\) ions. However, this data also shows that the isotopic comparisons can be used even at
lower concentrations for compound identification due to the unique shift in molecular masses of ions after deuterium incorporation. Finally, the degradation of RMSD values with drift window location is reported. For the [VLRLRG+2H]^{2+} ions, at 10% height of the drift time peak apex, the RMSD value increases by a factor of 3.

3.3.5. Mass Spectral Reproducibility after Isotope Scrambling Followed by HDX

Experiments were performed on the model peptide mixture in order to study the reproducibility of HDX experiments after HD scrambling. As before, significant care was given to maintain a constant partial pressure for D_{2}O reagent gas. The isotopic distribution RMSDs for the 6 HD scrambling trials for model peptide ions were less than 1%. These distributions are provided in Figures 3.9 and 3.10.

![Isotopic distributions for [M+2H]^{2+} KKD (A), [M+2H]^{2+} substance P (B), and [M+3H]^{3+} bradykinin (C) peptide ions. These data were recorded under conditions in which the ions were collisionally activated to induce increased deuterium incorporation. Six replicate analyses were performed for the RMSD calculations. RMSD values were less than 1% for these isotopic distributions.](image)

---

87
3.3.6. HDX Behavior to Aid Compound Identification

As a final proof-of-principle demonstration of the use of HDX behavior in compound identification efforts, the isotopic distributions of [M+H]$^+$ alanine ions and bovine heart metabolite extract ions of the same nominal $m/z$ value (90) and similar mobility ($9.1 \times 10^{-4} \text{ m}^2\cdot\text{V}^{-1}\cdot\text{s}^{-1}$, see also Table 3.1) are compared. Figure 3.11 shows the isotopic distributions obtained for the ions having nominal $m/z$ 90 from the bovine heart extract. Upon introduction of D$_2$O to the drift tube, similar isotopic distributions for these ions and the alanine ions are observed. A possible exception is the presence of a feature at $m/z$ 90 (ions that do not undergo HDX) for the unknown compound in the metabolite extract.
The most noticeable difference in HDX behavior is observed for the activated ions for which alanine shows an increase in deuterium incorporation (Figure 3.11B) while no similar increase is observed for the heart extract ions (Figure 3.11B).

![Figure 3.11](image_url1)

**Figure 3.11.** Isotopic distributions of precursor ions having nominal m/z 90 obtained from a bovine heart extract. Panel A shows distributions for ions in the presence of He only and He and D₂O as blue and orange traces, respectively. Panel B shows the isotopic distribution (green trace) for these ions obtained after collisional activation in the drift tube.

Considering the reproducibility of the measurements as described above, this would be an indication that such ions from the metabolite extract could not be assigned to the amino acid alanine. This assessment is supported by ion fragmentation experiments showing different fragmentation patterns for the mobility-selected ions (Figure 3.12). Although the above example shows an approach in which a false positive
assignment is identified, with a database of sufficient size (i.e., one containing the metabolite), assignment confirmation may be achieved by matching all of the new measurement parameters.

Figure 3.12. CID spectra for alanine at A) 15 NCE and B) 25 NCE. CID spectra for a feature with nominal m/z value of 90 within the bovine heart metabolite extract sample at C) 15 NCE and D) 25 NCE.
3.4. Conclusions

These experiments provide a proof-of-concept method for fuller characterization of compounds within complex mixtures. The method encompasses the usage of IMS measurements coupled with HDX and HD scrambling. To employ this for routine analysis, a library of IMS and HDX data must be generated for standard compounds. The promise is that comparisons of HDX profiles of unknown compounds in complex mixtures with HDX profiles of standard compounds can provide information about the identities of mixture components. Furthermore, it is instructive to briefly examine a potentially significant advantage of the IMS-HDX-MS approach. Over the years, significant effort has been exerted to develop analytical region prediction tools based on biomolecular ion structure. These include methods for predicting ion-neutral collision cross sections,\textsuperscript{39-41} and the gas-phase HDX behavior of specific ions.\textsuperscript{30, 42-45} In the context of comparative ‘omics, the ability to predict ion-neutral collision cross sections for small molecules could be beneficial for species for which no measurements are available\textsuperscript{41} such as newly emerging synthetic drugs and their metabolites or newly discovered neuropeptides. Adding the ability to predict isotopic distributions upon HDX may significantly enhance the ion distinguishing capabilities of such an approach.\textsuperscript{43} Notably, significant work to build out and utilize ion-neutral collision cross section and HDX databases as well as to construct predictive algorithms would be required to avail such an approach and this is a focus of future research efforts. Because of the extremely high reproducibility of the approach, it may be possible to incorporate HDX database searching into ‘omics workflows in much the same way that GC-MS employs small-molecule databases.
3.5. References


4. Ion Mobility Spectrometry-Mass Spectrometry Coupled with Gas-phase Hydrogen/Deuterium Exchange for Metabolomics Analyses

Reprinted in part with permission from Journal of American Society for Mass Spectrometry: Ion mobility spectrometry-mass spectrometry coupled with gas-phase hydrogen/deuterium exchange for metabolomics analyses. Hossein Maleki, Ahmad K. Karanji, Sandra Majuta, Megan M. Maurer, Stephen J. Valentine. JASMS.

4.1. Introduction

Comparative metabolomics experiments can provide researchers with vital information regarding disease processes. A goal of metabolomics investigations is the identification of new biomarkers for early disease diagnosis and the monitoring of disease progression. Biomarker discovery is accomplished through the comparison of identified metabolite abundances within biological samples (e.g., biofluids and tissues). As an example of the power of comparative metabolomics, consider that different types of coronary artery disease (CAD) can be differentiated by unique molecular profiles. Understanding the bio-pathways implicated by changes in molecular concentrations as a function of physiological state may play a crucial role for developing effective treatments. Therefore, one suggested consequence of the effective metabolomics studies noted above would be the tailoring of therapeutic strategies for CAD.

Characterization of metabolomics mixtures presents a number of challenges for analytical chemists resulting from the physicochemical diversity of metabolites (vis-à-vis functional groups, molecular structure, and molecular mass), the existence of numerous isomers and isobars, and the wide range in concentration of endogenous
metabolites\textsuperscript{10, 11}. Related to the latter characteristic, molecular concentration ranges in biofluids representing specific physiological states can span more than an order of magnitude.\textsuperscript{10, 11} Together these properties present obstacles to high-throughput comparative metabolomics studies and invite the development of new technologies and techniques for complex mixture studies.

Ion mobility spectrometry (IMS) is a gas-phase separation technique in which ions traveling in a pressurized drift tube under the influence of a constant electric field are separated based on their shape and size.\textsuperscript{12, 13} IMS-MS offers advantages in sensitivity by removing signals from low-intensity ions from the interfering signals of higher intensity species thereby increasing the overall measurement peak capacity.\textsuperscript{14} This technique also enables the separation of isomeric and isobaric species realizing some of the advantages of condensed-phase separations.\textsuperscript{15-19} It has been proposed that the addition of collision cross section (CCS) data to metabolomics databases increases the identification accuracy compared to traditional analytical methods alone.\textsuperscript{20} Furthermore, two-dimensional (2D) plots of $m/z$ vs $t_D$ have shown that different classes of compounds can be relegated to distinct regions of analytical measurement space.\textsuperscript{21} For example, classes of lipids can be easily distinguished from each other based on their occupation of such regions. Overall, IMS-MS has become an increasingly popular combination for studying metabolite mixtures.\textsuperscript{22-31} That said, IMS-MS analysis has two significant limitations for mixture characterization. IMS is a post-ionization separation step and thus still suffers from effects associated with competitive ionization. Additionally, ion mobilities are strongly correlated with $m/z$ values limiting the overall separation capacity of the IMS-MS combination.
More than two decades ago, gas-phase hydrogen/deuterium exchange (HDX) coupled with MS (HDX-MS) was introduced as a structural analysis tool for protein and peptide ions.\(^{32-34}\) This work was extended to include the study a wide array of different molecules.\(^{35-41}\) Shortly later, the approach was combined with IMS to provide shape and reactivity information for biological ions.\(^{42,43}\) More recently, with the development of non-ergodic ion fragmentation techniques, IMS-HDX-MS was used to determine the relative deuterium uptake in a site-specific manner for peptide ions.\(^{44,45}\) From such studies, a question that has arisen is whether or not HDX profiles can be availed to aid in the characterization of complex mixtures of small molecules.

In a recent article, the usage of gas-phase HDX in combination with ion mobility techniques was proposed as a means for compound identification in ‘omics mixtures.\(^{46}\) The work described various IMS-HDX measurements that could serve as unique identifiers for different molecular species and demonstrated the necessary robustness to be implemented in ‘omics workflows. In order to be useful as a predictive tool, the physicochemical properties of small molecules affecting HDX must be clearly elucidated. This work presents initial efforts to map such properties. Additionally, an example of how the approach would be used to aid identification of compounds in ‘omics mixtures is presented. Through these proof-of-principle studies, the utility of IMS-HDX-MS for metabolite identification is further explored.
4.2. Experimental Section

4.2.1. Sample Preparation

Hydroxyproline, adenine, leucine, isoleucine homocysteine, glutamine, lysine, dopamine, sucrose and acetaminophen were purchased from Sigma Aldrich (St. Louis, MO) and used without further purification. The Avanti lipid mixture and the bovine heart extract were purchased from Lipid Avanti Inc. Methanol, acetonitrile (ACN), methylene chloride, ammonium acetate and glacial acetic acid (HCOOH) were also purchased from Sigma-Aldrich (St. Louis, MO). Tables 4.1 and 4.2 show the molecular structures of the compounds used in this study.

A number of single standard and mixture sample solutions were prepared to assess the HDX reactivities of a wide range of sample types both with respect to size and composition. Small-molecule mixture consisted of one containing lysine, acetaminophen, and dopamine while a second solution contained all metabolites. For the mixtures and the single standards mentioned above, the concentration of the molecules was 0.01 mg/mL in methanol:acetonitrile (1:1). Acetic acid (1%) was added to enhance compound ionization. For preparation of the sucrose standard solution, 10 mg of standard was dissolved in 10 mL of water:acetonitrile (50:50) solution. Ammonium hydroxide (1%) was added to aid compound ionization.

In the study of lipids, two lipid mixture solutions were analyzed as well as the bovine heart extract. A lipid working solution was first prepared by mixing 250 mL of methylene chloride, 250 mL of methanol, and 5 mL of ammonium acetate solution (1 M solution prepared by dissolving 1.54 grams of ammonium acetate in 20 mL of methanol).
Table 4.1. Structures of small-molecule standards.

<table>
<thead>
<tr>
<th>Name</th>
<th>$m/z^a$</th>
<th>Structure$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxyproline</td>
<td>132.1</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Leucine</td>
<td>132.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Isoleucine</td>
<td>132.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Adenine</td>
<td>136.1</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Homocysteine</td>
<td>136.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Glutamine</td>
<td>147.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Lysine</td>
<td>147.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>152.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Dopamine</td>
<td>154.2</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>Sucrose</td>
<td>341.3</td>
<td><img src="image" alt="Structure" /></td>
</tr>
</tbody>
</table>

$^a m/z$ recorded for each ion

$^b$ Molecular structure drawn by the ChemDraw freeware
Table 4.2. Structures of lipid standards.

<table>
<thead>
<tr>
<th>Name</th>
<th>(m/z^a)</th>
<th>Adduct(^b)</th>
<th>Structure(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:0 Lyso PC</td>
<td>482.3</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>07:0 PC (DHPC)</td>
<td>482.3</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>18:1 Lyso PC</td>
<td>522.4</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td></td>
<td>544.4</td>
<td>(M+Na)</td>
<td></td>
</tr>
<tr>
<td>18:1 Ceramide</td>
<td>564.6</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td></td>
<td>586.6</td>
<td>(M+Na)</td>
<td></td>
</tr>
<tr>
<td>14:1 PE</td>
<td>632.5</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>((\Delta9^-\text{Cis})) PC (14:1/14:1)</td>
<td>674.6</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>((\Delta9^-\text{Trans})) PC (14:1/14:1)</td>
<td>674.6</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td>18:1 SM</td>
<td>729.7</td>
<td>(M+H)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
<tr>
<td></td>
<td>751.7</td>
<td>(M+Na)</td>
<td></td>
</tr>
<tr>
<td>18:1 TG</td>
<td>907.9</td>
<td>(M+Na)</td>
<td><img src="image" alt="Structure" /></td>
</tr>
</tbody>
</table>

\(^a\) \(m/z\) recorded for each ion  
\(^b\) Adduct observed for each ion  
\(^c\) Molecular structure drawn by the ChemDraw freeware

The first lipid standard mixture was prepared by transferring 400 \(\mu\)L of Avanti lipid mix solution to a 20 mL glass vial and adding 3.6 mL of lipid working solution. The second lipid standard mixture contained the isobars 15:0 Lyso PC and 07:0 PC (DHPC). For this
sample, 0.1 mg/mL of each standard solution was prepared by dissolving 1 mg of each standard at 10 mL of lipid working solution. Isobaric lipid mixture was prepared by adding 1000 μL of two lipid standard solutions. The bovine heart extract was prepared by using 400 μL of bovine heart extract and adding 3.6 mL of lipid working solution.

4.2.2. Instrumentation

For mobility measurements, an instrument consisting of a home-built drift tube coupled to a linear ion trap (LTQ Velos) mass spectrometer was used. The instrumentation has been described in detail elsewhere.\textsuperscript{47} Briefly, samples were infused (400 nL/min) through a pulled-tip capillary which was biased at ±2000 V relative to the entrance of a desolvation region. The desolvation region houses an ion funnel trap\textsuperscript{48,49} where ions are stored and periodically (frequency of 50 Hz) released into the drift tube. Here ions separate under the influence of a constant electric field (10 V/cm). The drift tube contains He (300 K) at ~2.5 Torr. For HDX experiments, partial pressures of ~0.02 and ~0.04 Torr of D\textsubscript{2}O were added to the He buffer gas via evaporation through a variable leak valve. The experiments for the metabolite standard, the lipid standard, and the bovine heart extract sample were performed in positive ESI mode. For the sucrose sample, experiments were performed in negative ion mode. CID experiments were performed by isolating a precursor ion with a m/z isolation window of ±1.0 Da relative to the centroid m/z value. Selected ions were fragmented using a normalized activation energy of 35.

Drift time (t\textsubscript{D})-resolved mass spectra were created with the application of a gating pulse to a Tyndall gate at the back of the drift tube. This gating pulse was delayed with respect to the ion introduction pulse setting the t\textsubscript{D} of each mass spectrum. The delay was
scanned across the entire \( t_D \) range using a step size of 0.2 ms. Each mass spectrum was converted to a text file and in-house software was used to assemble a 3-column array file containing \( t_D, m/z \), and intensity \((i)\) values. The DPlot software suite (HydeSoft Computing, Vicksburg, Mississippi) was used for the generation of 2D plots of \( m/z \) and \( t_D \) values. Using 2D \( t_D, m/z \) plots as a guide, in house software was used to extract \( t_D \)-resolved isotopic distributions for individual dataset features.

Experimental collision cross sections (\( \Omega \)) were calculated using the following expression:\(^{13}\)

\[
\Omega = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_BT)^{1/2}} \left[ \frac{1}{m_i} + \frac{1}{m_b} \right] \frac{t_D E}{L P} \frac{760}{273.2} \frac{T}{N}
\]  

(1)

where \( ze \) is the charge of the ion, \( E \) is the electric field strength, \( T \) is the temperature of the buffer gas, \( k_B \) is Boltzmann's constant, and \( m_i \) and \( m_b \) are the masses of the ion and the buffer gas, respectively. \( L \) is the length of the drift tube, \( P \) is the pressure of the buffer gas, and \( N \) is the number density of the buffer gas at STP.

4.3. Simulated Annealing and HDX Kinetics Modeling

Sucrose ions having two different deprotonation sites were generated using the Avogadro software suite. Sites of deprotonation were -OH3 and -OH6 moieties (on the fructose monomer) as shown in Figure 4.1. Both ions were subjected to simulated annealing using the NAMD software suite. To accomplish this, the ions were heated instantaneously to 1000 K, maintained at 1000 K for 40 ps, and cooled to 50 K over 40 ps. The 50 K structures were then energy-minimized and the resulting structures was saved for CCS calculation. Here we note that all experimental measurements were recorded using a constant-field drift tube (DT) while employing helium (He) as a buffer gas. Therefore, using the standard nomenclature, cross section values (experimental
and theoretical) are referred to as $^{\text{DT}}\text{CCS}_{\text{He}}$. This same structure served as the starting point for the next round of simulated annealing. This was performed for 1000 cycles to generate 1000 candidate structures. $^{\text{DT}}\text{CCS}_{\text{He}}$ values were computed for all structures using the trajectory method in the MOBCAL software.\textsuperscript{50,51} One structure from each charge site configuration having a $^{\text{DT}}\text{CCS}_{\text{He}}$ value near that observed experimentally was selected for HDX kinetics modeling.

**Figure 4.1.** [M-H]⁻ sucrose ion conformers obtained from simulated annealing (see manuscript for details). A) Conformer with a deprotonated oxygen that is surrounded by four $–\text{OH}$ groups. B) Conformer with a deprotonated oxygen that is surrounded by only one $–\text{OH}$ group.

HDX kinetics were modeled using a modified hydrogen accessibility scoring (HAS) approach.\textsuperscript{52} Briefly, the contribution of each heteroatom site to the ion’s measured HDX rate was scaled according to two factors. The first factor was the distance of the site to the charge site in which values scaled according to inverse distance. Using these values, all heteroatom sites were then normalized to provide fractional contribution to the overall HDX rate. Next, because distance alone does not describe accessibility, a blocking factor was used to scale each of the HDX rate contribution scores. The blocking factor was considered based on whether or not the charge site could access the exchange site via molecular motion as well as whether or not the exchange site was involved in significant
intramolecular hydrogen bonding. After scaling the individual contributions to HDX rate, the HDX process was modeled for each conformer using an approach described previously.\textsuperscript{52} The theoretical isotopic distributions were then calculated for the two different ion conformer types.

The best-fit isotopic distribution obtained from contributions by both conformer types was computed using a pseudo Monte Carlo approach. Here, random, fractional contributions from both theoretical isotopic distributions were evaluated in an iterative manner. The fit did not improve after \(1 \times 10^5\) iterations. The in house software then provided the contribution by each conformer type to the best-fit isotopic distribution.

4.4. Results and Discussion

4.4.1. Compound Specific HDX Behavior

To serve as a unique compound identifier, an analytical technique must yield a measurement that is orthogonal to any combined measurements. That is, the new measurement must not exhibit significant correlation to other measurements. To demonstrate such a feature for HDX with respect to ion mobility measurements consider the deuterium uptake of \([M+H]^+\) acetaminophen, dopamine, and lysine ions as shown in Figure 4.2. These ions were selected to demonstrate variable reactivities for species of relatively similar molecular weight. Additionally, because isotopic distributions tend to broaden upon HDX, acetaminophen and dopamine being of more similar molecular weight offer a test case for resolution of ion signatures. \(\text{DTCCS}_{\text{He}}\) values for these ions are provided in Table 4.3. Upon introduction of \(\sim 0.02\) Torr of \(\text{D}_2\text{O}\) into the drift tube, the major
isotopologue peaks demonstrate an incorporation of 6, 1, and 0 deuterions for lysine, acetaminophen, and dopamine, respectively.

Figure 4.2. Shows 2D plots of m/z vs t_D for a mixture of lysine, acetaminophen, and dopamine A) in the absence of D_2O reagent gas and B) in the presence of D_2O reagent gas. Structures of ions associated to peaks shown with white arrows.

Considering the limited number of heteroatom hydrogens for many small-molecule metabolites, a question of utility arises. That is, if many compounds of similar mass have only 2 exchangeable hydrogens, how can such an approach be useful? To show utility requires a brief discussion of ion populations and HDX kinetics. The mechanism of HDX requires the formation of a long-lived reaction intermediate involving the charge site and a less basic site on the molecule. Therefore, the ion structure and the atomic composition dictate the rate of HDX of the ions and thus the amount of incorporated deuterium. It might be expected then that the rates of exchange are nearly as diverse as the compound types encountered in the mixture. Indeed such variability in HDX behavior is demonstrated in Figure 4.2. For these ions as well as others examined in this study,
different physical attributes result in unique isotopic distributions with characteristic intensities for the M+0, M+1, M+2, etc. isotopologue peaks.

Table 4.3. Experimental DTCCS_{He} values calculated for the molecular standards.

<table>
<thead>
<tr>
<th>Name</th>
<th>m/z</th>
<th>Adduct</th>
<th>Experimental DTCCS_{He}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxyproline</td>
<td>132.1</td>
<td>(M+H)</td>
<td>58.2</td>
</tr>
<tr>
<td>Leucine</td>
<td>132.2</td>
<td>(M+H)</td>
<td>64.3</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>132.2</td>
<td>(M+H)</td>
<td>64.1</td>
</tr>
<tr>
<td>Adenine</td>
<td>136.1</td>
<td>(M+H)</td>
<td>57.4</td>
</tr>
<tr>
<td>Homocysteine</td>
<td>136.2</td>
<td>(M+H)</td>
<td>61.2</td>
</tr>
<tr>
<td>Glutamine</td>
<td>147.2</td>
<td>(M+H)</td>
<td>67.6</td>
</tr>
<tr>
<td>Lysine</td>
<td>147.2</td>
<td>(M+H)</td>
<td>64.8</td>
</tr>
<tr>
<td>Acetaminophen</td>
<td>152.2</td>
<td>(M+H)</td>
<td>69.2</td>
</tr>
<tr>
<td>Dopamine</td>
<td>154.2</td>
<td>(M+H)</td>
<td>70.8</td>
</tr>
<tr>
<td>Sucrose</td>
<td>341.3</td>
<td>(M-H)</td>
<td>112.4</td>
</tr>
<tr>
<td>15:0 Lyso PC</td>
<td>482.3</td>
<td>(M+H)</td>
<td>156.8</td>
</tr>
<tr>
<td>07:0 PC (DHPC)</td>
<td>482.3</td>
<td>(M+H)</td>
<td>155.1</td>
</tr>
<tr>
<td>18:1 Lyso PC</td>
<td>522.4</td>
<td>(M+H)</td>
<td>170.3</td>
</tr>
<tr>
<td></td>
<td>544.4</td>
<td>(M+Na)</td>
<td>172.0</td>
</tr>
<tr>
<td>18:1 Ceramide</td>
<td>564.6</td>
<td>(M+H)</td>
<td>194.2</td>
</tr>
<tr>
<td></td>
<td>586.6</td>
<td>(M+Na)</td>
<td>196.5</td>
</tr>
<tr>
<td>14:1 PE</td>
<td>632.5</td>
<td>(M+H)</td>
<td>185.4/193.0</td>
</tr>
<tr>
<td>(Δ9-Cis) PC (14:1/14:1)</td>
<td>674.6</td>
<td>(M+H)</td>
<td>196.1</td>
</tr>
<tr>
<td>(Δ9-Trans) PC (14:1/14:1)</td>
<td>674.6</td>
<td>(M+H)</td>
<td>195.6</td>
</tr>
<tr>
<td>18:1 SM</td>
<td>729.7</td>
<td>(M+H)</td>
<td>220.9</td>
</tr>
<tr>
<td></td>
<td>751.7</td>
<td>(M+Na)</td>
<td>222.4</td>
</tr>
<tr>
<td>18:1 TG</td>
<td>907.9</td>
<td>(M+Na)</td>
<td>269.7</td>
</tr>
</tbody>
</table>

a m/z recorded for each ion  
b Adduct observed for each ion  
c Experimental DTCCS_{He} values calculated from the measured tD values  
d DTCCS_{He} values reported for the two observed conformers belonging to ions of m/z 632.5

Because of differences in rates of exchange, the ions may also exhibit unique HDX behavior at different partial pressure settings of D₂O. Figure 4.3 shows the HDX behavior of mobility-selected ions at D₂O partial pressures of ~0.02 and ~0.04 Torr. Upon addition of ~0.02 Torr D₂O, mobility-selected [M+H]⁺ hydroxyproline ions exhibit an isotopic
distribution in which the M+1 peak is the dominant feature and the M+0 peak is largely absent (Figure 4.3A). With the same partial pressure of D₂O, [M+H]+ 18:1 Lyso PC ions produce an isotopic distribution in which the M+0 peak is yet the dominant feature with a nearly equivalent amount of the M+1 peak (Figure 4.3C). The M+3 peak is ~30% of the height of the M+1 peak. Upon increasing the D₂O partial pressure to ~0.04 Torr, the isotopic features for the singly-charged hydroxyproline do not change in relative intensity (Figure 4.3B). At this same partial pressure setting, the singly-charged ions of 18:1 Lyso PC produce an isotopic distribution in which the most intense feature is the M+2 peak and the M+3 peak is also a major (~30%) constituent (Figure 4.3D).

Other ions show isotopic distributions that are more similar at the two different partial pressure settings. For example, [M+H]+ 14:1 PE and M+ 18:1 SM ions show similar isotopic patterns upon introduction of 0.02 Torr of D₂O. For each, the major peak is the M+0 peak; however, for both ion types, there is a noticeable – but similar – increase in the M+1 peak (Figures 4.3E and 4.3G). After increasing the partial pressure of D₂O to ~0.04 Torr, both ion types show isotopic distributions in which the M+1, M+2, and M+3 peaks are major constituents. Uniquely for the 14:1 PE ions, the M+4 peak is the dominant feature. Such a bimodal distribution could be indicative of the HDX behavior of different conformer types. Indeed, in the He buffer gas, two conformers are resolved (Figure 4.4). However, these conformers are not resolved in the He:D₂O mix. That said, even for slowly-exchanging conformer types, the 14:1 PE and the 18:1 SM ions show some difference. For example the M+3 peak shows a greater relative intensity for the former ions compared to the latter ions.
Figure 4.3. Isotopic distributions obtained using a D$_2$O partial pressure of ~0.02 Torr for A) [M+H]$^+$ hydroxyproline ions, C) M$^+$ 18:1 Lyso PC ions, E) [M+H]$^+$ 14:1 PE ions, and G) M$^+$ 18:1 SM ions. Isotopic distributions obtained using a D$_2$O partial pressure of ~0.04 Torr for B) [M+H]$^+$ hydroxyproline ions, D) M$^+$ 18:1 Lyso PC ions, F) [M+H]$^+$ 14:1 PE ions, and H) M$^+$ 18:1 SM ions. Solid lines show the distribution obtained in the absence of D$_2$O and the dashed lines show the isotopic distribution with the addition of the respective partial pressure of D$_2$O to the drift tube buffer gas.
Figure 4.4. Drift time distribution for [M+H]+ 14:1 PE ions. Two conformers associated with these ions are partially resolved. See Table 1 in the manuscript for CCS values for the conformers.

4.4.2. Distinguishing Isobaric Ions

Considering the implementation of IMS-HDX with a low-resolution mass spectrometer, it is necessary that isobaric ions exhibit differences in HDX uptake. To provide examples of differences in HDX behavior for isobaric ions, data have been recorded for [M+H]+ adenine, homocysteine, lysine, and glutamine ions as shown in Figure 4.5. The adenine and homocysteine ions have the same nominal m/z value of 136. Upon addition of D$_2$O (~0.04 Torr) to the drift tube, a portion of the adenine ions undergo HDX producing a bimodal distribution in which the dominant peaks in the isotopic distribution become the M+0 and M+2 peaks (Figure 4.5A). In contrast, for homocysteine ions the dominant peak in the isotopic distribution becomes the M+1 peak (Figure 4.5B). As mentioned above, lysine ions undergo significant HDX with the M+6 peak being the dominant feature in the isotopic distribution (Figure 4.5C). In contrast, for glutamine ions the dominant peak in the isotopic distribution becomes the M+1 peak (Figure 4.5D).
Figure 4.5. Isotopic distributions for [M+H]+ A) adenine and B) homocysteine ions and for [M+H]+ C) lysine and D) glutamine ions. Solid lines show the distribution obtained in the absence of D$_2$O and the dashed lines show the isotopic distribution with the addition of ~0.04 Torr D$_2$O to the drift tube buffer gas.

Figures 2, 3, and 5 present results for which the major isotopic peaks at specific $t_\text{D}$ values could be used to distinguish different ions. Higher molecular weight species can present a greater challenge especially as the isotopic distribution becomes broader. Figure 4.6 shows mobility-selected isotopic distributions obtained upon HDX experiments for a two-component lipid mixture. The mixture contains the lipid isobars having a nominal m/z of 482. From the compound structures in Table 4.2 it is evident that the maximum number of exchangeable hydrogens for 15:0 Lyso PC and 07:0 PC (DHPC) are 2 and 1, respectively. For comparison, the isotopic distributions obtained from HDX experiments of the separate isobaric lipids are shown in Figure 4.7. The [M+H]+ ions yield $^{12}$CCS$_{\text{He}}$ values of 155.3 Å$^2$ and 156.9 Å$^2$ (Table 4.3) and are not resolved in the mobility
measurement in which He or He and D₂O buffer gas is used. However, for HDX experiments, clear differences in isotopic distributions are observed at select drift times.

**Figure 4.6.** Isotopic distributions for a mixture of two isobaric lipids (15:0 Lyso PC and 07:0 PC (DHPC)) at t₀ selections of A) 10.8-11.0, B) 11.2-11.4, and C) 11.4-11.6. The solid lines demonstrates isotopic patterns after deuterium addition.
Figure 4.7. Isotopic distributions obtained for A) 07:0 PC (DHPC) and B) 15:0 Lyso PC.

As shown in Figure 4.6, the M+1 isotopologue peak is the major feature observed at a drift time selection of 10.8-11.0 ms. At 11.2-11.4 ms, the M+2 isotopologue peak is of nearly equal abundance while at 11.4-11.6 ms it becomes the dominant feature in the distribution. This change in the distribution is consistent with a mobility selection that is dominated by 07:0 PC (DHPC) ions at shorter $t_D$ values and then becomes dominated by the 15:0 Lyso PC at longer times. Admittedly, the HDX approach is of limited value for isomeric lipid ions (as well as small-molecule isomers) having the same cross section such as [M+H]$^+$ ($\Delta$9-Trans) PC (14:1/14:1) and ($\Delta$9-Cis) PC (14:1/14:1) ions as shown in
Figure 4.8 and Figure 4.9. This is somewhat expected as the head group is identical for this species and the ions under HDX for the same amount of time.

Figure 4.8. Isotopic distributions obtained for A) Leucine and B) Isoleucine.

Figure 4.9. Isotopic distributions obtained for A) (Δ9-Cis) PC (14:1/14:1) and B) (Δ9-Trans) PC (14:1/14:1).

4.4.3. HDX Behavior of Lipid Ions

The M⁺ and [M+H]⁺ lipid ions (Table 4.2) also display unique HDX behavior as shown in Figure 4.10. In summary, ions produce isotopic distributions in which M+1, M+2, or M+3 peaks are the dominant species. For example, the 18:1 Lyso PC ions produce an isotopic distribution in which the M+2 peak is the dominant peak upon introduction of ~0.04 Torr of D₂O (Figure 4.10A). The M+3 peak is also present with a relative intensity
of ~30%. In comparison, the 18:1 Ceramide ions yield a broader isotopic distribution where the M+3 peak is the dominant feature (Figure 4.10B). M+1, M+2, and M+4 peaks also comprise the distribution having relative intensities of 16%, 45%, and 35%, respectively.

**Figure 4.10.** Isotopic distributions for the A) M⁺ 18:1 Lyso PC, B) [M+H]⁺ 18:1 Ceramide, C) [M+H]⁺ 14:1 PE, D) M⁺ 14:1 PC, and E) M⁺ 18:1 SM ions. Solid lines show the distribution obtained in the absence of D₂O and the dashed lines show the isotopic distribution with the addition of ~0.04 Torr D₂O to the drift tube buffer gas.
With the addition of ~0.04 Torr D₂O, the 14:1 PE, PC (14:1/14:1), and 18:1 SM ions all display isotopic distributions having dominant M+1 peaks. As mentioned above, the 14:1 PE ions also produce a large M+4 peak and the bimodal distribution may result from two different ion conformer types that are mobility resolved in He buffer gas (Figure 4.4). For the purposes of comparison to the other lipids, the discussion here is focused on the lower m/z features. A comparison of the isotopic distributions of these ions shows differences in the relative intensities of the M+2 and M+3 peaks. For the former ions, the relative intensities are 58% (Figure 4.10C), 42% (Figure 4.10D), and 70% (Figure 4.10E) for the 14:1 PE, PC (14:1/14:1), and 18:1 SM ions, respectively. For the M+3 peaks the respective intensities are 37%, 9%, and 22%. A question that arises is whether or not such differences are significant; that is, can such differences be used to obtain reliable identifications? Extensive experiments have shown that such disparity in isotopic distributions is significant as run-to-run comparisons yield root mean square displacement values of <1.5%.⁴⁶ Finally, it is noted that no deuterium uptake is observed for the sodiated form of the lipids as shown in Figure 4.11. This can be explained by the fact that the sodium cannot participate in the hydrogen-bonded intermediate of the relay mechanism.⁵⁴
Figure 4.11. Isotopic distributions obtained for the sodiated lipids including: A) 18:1 Lyso PC; B) 18:1 Ceramide; C) PC (14:1/14:1); D) 18:1 SM; and, E) 18:1 TG lipids. Solid lines show the isotopic distribution in the absence of D$_2$O and dashed lines show the isotopic distribution with the addition of ~0.04 Torr D$_2$O.

4.4.4. Structural Explanations for HDX Behavior

Significant effort has been exerted to develop collision cross section databases for small molecules.$^{31,55-58}$ It has been proposed that such databases could find utility in ion identification either through direct matching of ion mobilities or through the locating of an ion’s mobility to an analytical region that is associated with a family of compounds.$^{21,59}$ Previously, it was proposed that matching of isotopic distributions could be used in this fashion where comparisons would be made to database compounds.$^{46}$ One potential
advantage of ion mobility techniques is that, to some degree, measurement predictions are accessible for unknown compounds. Such approaches range from the consideration of substituent contributions to mobility to the use of powerful quantum mechanics and molecular mechanics approaches to obtain structures for cross section calculation check.\textsuperscript{60-62} Adding to the capability of HDX prediction to an IMS prediction method could provide a powerful tool for the identification of unknown compounds.

In order to develop predictive tools for HDX behavior, it is important to elucidate the factors governing the reaction process. Because the HDX mechanism requires the formation of a long-lived reaction intermediate involving the charge site and a less basic site that are both accessible to collisions with D\textsubscript{2}O\textsuperscript{43, 53}, it is important to first consider elements of compound structure that favor or disfavor intermediate formation to explain ion reaction behavior. As described above, [M+H]\textsuperscript{+} dopamine ions do not undergo HDX even at the highest partial pressure settings of D\textsubscript{2}O (Figure 4.2). This is observed despite the fact that such ions have 5 heteroatom hydrogens. From the molecular structure (Table 4.1), it can be argued that exchange does not occur because D\textsubscript{2}O cannot hydrogen bond at the charge site and a less basic hydroxyl as these substituents are located distantly across the aromatic ring from the charge site. Thus, the relay mechanism is not operable and HDX is not achieved. In contrast, the flexibility of the side chain of the lysine ions permits the positioning of the charge in relative proximity of a less basic site (e.g., carbonyl) such that HDX proceeds efficiently (Figure 4.2B). The incorporation of 1 deuterium for the [M+H]\textsuperscript{+} acetaminophen ions would then require increased access of the less basic site to the charge site despite being located across the aromatic ring. In this case the charge site is closer to the less basic site compared with the dopamine ions.
For singly-charged adenine ions, there are 4 heteroatom hydrogens. Upon addition of ~0.04 Torr of D$_2$O, a bimodal distribution is produced where one dominant feature in the isotopic distribution is the M+0 peak and the other is the M+2 peak (Figure 4.5A). It is possible that the bimodal distribution results from two different charge arrangements - one in which the charge resides on the primary amine and the other in which it resides on the aromatic ring shown in Table 4.1. The lack of exchange of all 4 hydrogens in both isotopic distributions can be ascribed to the fact that the heteroatom hydrogens are relatively distant to one another. Singly-charged glutamine ions have 6 heteroatom hydrogens. These ions exhibit a relatively small degree of HDX as shown in Figure 4.5D. Because the heteroatom hydrogens reside at the end of the side chain and the N- and C-termini, the data suggest that these ions are less flexible in that the side chain does not approach the other heteroatom sites as efficiently as is observed for lysine.

Providing a rationale for HDX behavior of homocysteine ions based on compound structure (Table 4.1) is more challenging. These ions have 5 heteroatom hydrogens and the dominant feature in the isotopic distribution is the M+1 peak. Considering that the charge would reside at the amino terminus and thus present significant redundancy in exchangeable hydrogens, the low level of exchange suggests accessibility of the side chain or C-terminal hydrogen alone.

For the lipids, the 18:1 Lyso PC ions have 2 heteroatom sites. The isotopic distribution shows nearly complete exchange with the M+2 peak being the dominant feature. This can be reasoned by the fact that both sites are near less basic sites onto which the deuterium can be shuttled via the relay mechanism. The 18:1 Ceramide ions have 4 heteroatom hydrogens and the dominant feature in the isotopic distribution is the
M+3 peak. Additionally, the presence of significant M+1 and M+2 peaks indicates lower HDX efficiency. Assuming that the charge resides at the secondary amine, the HDX behavior can be explained by less accessibility of the terminal hydroxyl group to the less basic carbonyl site for example. Here the carbonyl site would serve as the less basic site onto which the deuterium is originally shuttled via the relay mechanism. For the 14:1 PE ions, there are 4 heteroatom hydrogens. The isotopic distribution is bimodal with dominant features corresponding to the M+1 and M+4 peaks. The latter must arise from an ion conformer type in which the protonated primary amine has access to a less basic site allowing transfer of the incorporated deuterium back to the charge site. The conformer type leading to the M+1 peak being the dominant feature must necessarily allow less access of the terminal amine to a less basic site. Again it is noted that two conformer types are resolved by IMS alone. (Figure 4.4) The 14:1 PC ions have only 1 heteroatom hydrogen which exchanges efficiently; the M+1 peak is the dominant feature. This is similar to the high-efficiency exchange of the phosphate hydrogen for the 18:1 Lyso PC ions suggesting facile formation of the reaction intermediate at the site of the phosphate group. Finally, the 18:1 SM ions have 4 heteroatom hydrogens. Overall, the HDX efficiency of some sites is low as the isotopic distribution is dominated by the M+1 peak with the M+2 peak contributing to the distribution as well. In this case it is possible that the secondary amine – which is not protonated for these ions – and the hydroxyl group are less accessible to either the formation of the reaction intermediate or to the less basic site serving as the original site of deuterium incorporation.

The HDX behavior of the ions shown in here has been described in terms of the ability of different heteroatom sites to participate in the reaction intermediate involved in
the relay mechanism or their relative accessibility to the original site of deuterium incorporation. Thus a cursory examination of the molecular structure represents the beginning of the development of HDX prediction capability for compounds for which measurements have not been obtained. That said, although a cursory examination of compound structure may help to describe the observed reactivity, for certain ions, this may be more difficult. For example, bimodal distributions coupled with considerations of exchange site redundancy is used above to suggest the presence of different conformer types (including charge site configuration). For some species a clear distinction in multiple conformer types may not be drawn easily from the isotopic distribution or the mobility distribution. For example, the differences in m/z range of the isotopic distributions could suggest that some ion populations are comprised of greater structural heterogeneity (conformer types). That is, as with the bimodal distributions discussed above, it is possible that different conformer types could give rise to the broad isotopic distribution observed for the 18:1 Ceramide ions (Figure 4.10).

To consider the effect of multiple, unresolved conformers, [M-H]- sucrose ions have been studied. Even though these ions exhibit a single conformer type in the mobility distribution (Table 4.3), a broad isotopic distribution is observed in which the dominant peaks are the M+0, M+1, M+2, M+3, and M+4 species as illustrated in Figure 4.12. The presence of lower m/z peaks suggests that at least one sucrose ion conformer type shows limited access to exchange sites. One explanation could be that different charge site configurations result in conformers that exhibit varying degrees of accessibility to exchange sites. Simulated annealing of deprotonated ions produces several conformers with collision cross sections that match that observed experimentally. Two conformer
types with different charge sites were selected for HDX modeling (see Methods section above). Conformer 1 (Figure 4.1A), being deprotonated at the –OH3 site, is charge solvated by hydroxyl groups at the –OH1, –OH2, –OH4, and –OH5 moieties. Conformer 2 (Figure 4.1B), being deprotonated at the –OH6 site, is solvated by the –OH7 hydroxyl moiety. Other sites are not easily accessed by the charge site for this conformer as determined by modified HAS scoring (see experimental section). Using the simulated isotopic distributions from the two conformers and a pseudo Monte Carlo best-fit approach, the optimal populations of the two candidate conformers have been computed. The best-fit isotopic distribution as shown in Figure 4.12 suggests that Conformer types 1 and 2 would contribute best to the overall HDX behavior with a respective 0.63:1 ion population ratio. That said, the best-fit approach does not accurately capture the intensity of the M+4 peak. It is possible that other conformer types of lower abundance are present allowing for increased HDX.

![Figure 4.12. Bar graph representing the isotopic distribution for [M-H]- sucrose ions upon addition of ~0.04 Torr D2O showing the relative intensities of the M+0, M+1, M+2, M+3, M+4, and M+5 peaks (red bars). The blue bars show the best-fit isotopic distribution obtained from the modeled isotopic distributions for in-silico candidate structures.](image-url)
From the analysis above, the range of HDX behavior across metabolite space is expected to be as diverse as the differences in the types of compound substituents and their locations in ion structures. It is stressed that even though the same isotopologues may be the dominant features in the different HDX patterns, their relative intensities can be significantly different (e.g., Figure 4.10). It is this diversity that offers the opportunity for using HDX in conjunction with IMS and MS for compound identification. That said, the magnitude of the challenge in developing predictive HDX tools becomes readily apparent with this level of diversity as well. Herculean effort will be required to populate experimental databases and utilize high-quality in-silico structures to develop such tools.

4.4.5. Proof-of-Principle Application of HDX Data for Compound Identification

It is instructive to consider how the IMS-HDX approach might be used for 'omics mixture characterization. Figure 4.13A shows the isotopic distribution after HDX of an unknown compound from the bovine heart extract sample. Here the dominant species are the M+1 ions. There are also significant contributions to the distribution from the M+0 and M+2 ions. This is very similar to the isotopic distribution obtained for the precursor ions of hydroxyproline. Additionally, the DTCCSHe value for the unknown ions is 58.6 which is near that determined for hydroxyproline (Table 4.3). Therefore, based on IMS-HDX-MS data alone, a match would be obtained for hydroxyproline. However, a comparison of the CID data for the unknown ions and hydroxyproline ions reveals a major difference as shown in Figure 13C and 13D, respectively. Notably, an intense ion at m/z 90 is present in the fragmentation spectrum for unknown compound in bovine sample. This ion is completely absent in the fragmentation spectrum of hydroxyproline. Notably, parallel CID$^{63}$ with HDX and deuterium scrambling with HDX have been presented
previously as a means to provide other unique identifiers. These methods can also be implemented without adding significant experimental time to create an IMS-HDX-MS/MS approach. In that manner, high-throughput analyses could be availed to avoid false positive assignments.

![Figure 4.13](image)

**Figure 4.13.** Isotopic distributions obtained using a D$_2$O partial pressure of ~0.04 Torr for A) selected ions having a nominal m/z value of 132 in the bovine cardiac extract and B) [M+H]$^+$ hydroxyproline. Fragmentation patterns produced by CID of C) selected ions having a nominal m/z value of 132 in the bovine cardiac extract and D) [M+H]$^+$ hydroxyproline ions.

In a separate example, a dataset feature from the bovine heart extract is observed to have a DTCCS$_{He}$ value of 64.1 Å$^2$. Upon adding ~0.04 Torr of D$_2$O, the ion produces an isotopic distribution in which the M+6 peak is the dominant feature (Figure 4.14A). Indeed the distribution is very similar to that observed for lysine ions (Figure 4.14B). Additionally, the DTCCS$_{He}$ value matches that determined for singly-charged lysine ions (Table 4.3). To support such an assignment, the precursor ions from the bovine heart extract have been isolated and subjected to CID. The dominant fragments are produced
by the neutral loss of -NH$_3$ and -H$_2$O. Low-intensity fragment ions are observed at $m/z$ 82 and 103. The same dominant and low-intensity fragment ions are produced upon CID of mobility-selected lysine ions. This analysis would support the assignment of the unknown ions as lysine ions in the heart extract sample. One advantage to the assignment of the ions based on the IMS-HDX-MS information alone is that such information can be obtained from dispersive measurements on very short timescales (seconds). That is, no ion selection or $m/z$ or mobility scanning is required with currently available commercial instrumentation.

**Figure 4.14.** Isotopic distributions obtained using a D$_2$O partial pressure of ~0.04 Torr for A) selected ions having a nominal $m/z$ value of 147 in the bovine cardiac extract and B) [M+H]$^+$ lysine ions. Fragmentation patterns produced by CID of C) selected ions having a nominal $m/z$ value of 147 in the bovine cardiac extract and D) [M+H]$^+$ lysine ions.
4.5. Conclusions

Proof-of-principle experiments have been conducted to demonstrate the utility of IMS combined with gas-phase HDX and MS analysis. The work demonstrates that, despite limited numbers of heteroatom hydrogens on many small molecules, unique isotopic distributions are produced. Such selectivity results from differences in HDX kinetics arising from the three-dimensional arrangement of the heteroatoms and charge sites within the various compounds. Although the wide range of different structural arrangements offers great opportunity with regard to selectivity, it presents significant challenges with regard to deciphering the rules governing HDX. Elucidating such rules is necessary to develop accurate tools for predicting HDX behavior. The value of the ability to accurately predict HDX behavior cannot be overstated. Such an approach can be combined with IMS prediction methods \textsuperscript{61, 64} to suggest assignments of heretofore unmeasured compounds. These could include species such as newly emerging neuropeptides or drugs of abuse. Therefore, the careful study of the HDX behavior of numerous database compounds should be a significant pursuit in the future.
4.6. References


5. Future Directions: Developing New Methods to Improve Compound Identification in Complex Mixtures

5.1. Construct a Library of CCS Values and HDX Distributions

CCS values can be determined for metabolites and lipids.\textsuperscript{1,2} These CCS values can be employed for the identification of unknown features in 'omics analyses.\textsuperscript{3} In this manner, the CCS values of unknown features in complex mixtures can be compared with CCS values of standards in the library. A CCS is a physicochemical property of different molecules that depends on numerous factors including the polarity of the molecule, the number of rings and double bonds the molecule contains, and the structure of the ion. Therefore, similarity of CCS values can be used as unique identifiers for unknown features in complex mixtures. The HDX behavior of ions depends on the atomic composition and structure of the ions. From this work it is shown that isotopic distributions after HDX can also serve as unique identifier.\textsuperscript{4}

For traditional drift tube IMS-MS instruments, experimental CCS values can be calculated directly by using drift time values and the Mason-Schamp equation.\textsuperscript{5} Therefore, a library of experimental CCS values can be constructed for different classes of biomolecules.\textsuperscript{1} Although experimental CCS values have been measured for classes of metabolites, such as lipids, peptides, carbohydrates. Due to the diversity of compounds in 'omics mixtures, more measurements are required to institute robust compound search capabilities. Admittedly, the construction of a database of sufficient size containing both CCS and HDX information is a herculean effort. Additionally a question arises as to the universal usage of HDX data. For example, the isotopic distributions are depended on
reaction kinetics for which some instrumentation may not be able to provide similar measurement parameters. This is a tremendous challenge. That said, if site-specific kinetics information is included in the database, HDX behavior can be extrapolated for different experimental conditions.

5.2. Prediction of CCS Values for Biomolecules Using Intrinsic Size Parameters

The measurement of experimental CCS values is time intensive and expensive. Although IMS-MS is a powerful method for compound identification using CCS values, its power is limited by the number of available CCS values. Additionally, many metabolite compounds are not commercially available which narrows the number of experimental CCS measurements. Lastly, theoretical CCS determinations using the Mobcal software Suite is time consuming. Because of these challenges and the potential to identify novel compounds, many research groups have implemented the application of intrinsic size parameters to predict CCS values.\textsuperscript{6-9} The rationale is to find parameters which could potentially sum to the CCS values of molecules. Valentine and coworkers proposed a method to predict CCS values of peptides in He buffer gas.\textsuperscript{6} They considered each amino acid as having an intrinsic size parameter contributing to the overall ion CCS values. Zhou and coworkers tried to expand the application of CCS predictions to metabolites.\textsuperscript{9} In their work, 14 common descriptors of metabolites were employed to predict CCS values in N\textsubscript{2} buffer gas. A number of these descriptors include accurate mass, the formal charge, molecular polarizability, the octanol/water partition coefficient, etc. Using their algorithm, CCS values were predicted for 35,203 metabolites in the Human Metabolome Database (HMD). Additionally, CCS values for ion adducts in both positive and negative ESI mode were included. Overall, hypothetical CCS values for 176,015 metabolites were reported.
Furthermore, a few researchers have tried to implement artificial neural network (ANN) analysis for prediction of CCS values.\textsuperscript{10, 11}

Although the work above shows that CCS values were successfully predicted for peptides in the He buffer gas and for metabolites in the N\textsubscript{2} buffer gas, however, more extensive work is required to predict CCS values in both He and N\textsubscript{2} buffer gas for such a great number of metabolites and lipids. This is necessary to implement a database search approach. Furthermore, due to the diversity and complexity of metabolite structures, modified descriptors are required to improve the accuracy of CCS calculations. One approach that will be pursued is to major factors governing CCS values. For example, for larger species, conformer “folding” is expected to play a greater role. In contrast, for smaller ions, their overall density should be considered. Therefore, factors that scale CCS according to density (e.g. number of hetero-atoms) could be incorporated and weighted relations to folding parameters.

5.3. Conformational Analysis of Biomolecules in the Gas-phase: Matching Theoretical CCS Values with Experimental CCS Profiles

Because specific conformer types of biomolecules may play main role in the progression of different disease states,\textsuperscript{12, 13} conformational analysis of biomolecules is a significant endeavor. Conformational analysis has been reported for peptides and proteins\textsuperscript{14, 15, 16} However, little has been accomplished for conformational analysis of metabolites and lipids. These types of analyses are in high demand because lipids have a major role in the evolution and progression of many diseases.\textsuperscript{17, 18} The traditional method for conformational analysis is to perform MD simulations on biomolecules and then find conformations which match with experimental CCS profiles.\textsuperscript{19} That said, little
has been achieved with regard to important efforts such as force field optimization for MD simulations. For example, HDX can be coupled to CCS measurements to aid conformational analysis of biomolecules.\textsuperscript{15} It will be possible to test and optimize MD simulations. The construction of a large database of CCS values and HDX information will advance efforts to improve MD simulations.

5.4. Prediction of HDX Patterns for Metabolites and Lipids

The ability to predict isotopic patterns upon HDX would significantly enhance compound identification in complex mixtures. As demonstrated in this work, HDX profiles can be generated at different pressures which adds to the power of this method. That said, generation of experimental HDX profiles for metabolites and lipids is very time-consuming and expensive. Furthermore, some metabolites and lipids are not commercially available. Therefore, developing predictive methods for HDX prediction would be rewarding. Extensive research has been conducted to study factors contributing to HDX behavior of peptides.\textsuperscript{16, 20} However, little work has been performed with regard to the prediction of HDX behavior of metabolites and lipids. For peptides, a hydrogen accessibility factor (HAS) is employed. In this method, the surface accessibility of labile hydrogens and the proximity of the charge site to the basic site were considered. HDX modeling of metabolites and lipids is more challenging because all the hydrogens in the molecule are exposed to D\textsubscript{2}O reagent gas. Therefore, surface accessibility cannot be taken into account for HDX modeling in this case. However, new factors would be proposed in order to narrow the number of possible conformations. To overcome this issue, an accessibility factor based on propensity to encounter a charge site due to dynamics was introduced in this dissertation. This parameter can be explored by
extensive MD simulations. In summary, finding new factors that can predict HDX behaviors of metabolites and lipids would be of demand. Additionally, proposing new parameters can shed light on the conformational analysis of metabolites and lipids.

5.5. Application of HDX Profiles to Assist in Locating the Cis and Trans Double Bonds in Lipids at Elevated Temperatures

Lipids have major roles in biological processes such as cell membranes and metabolic pathways. Additionally, they are involved in the progression of diseases such as obesity, diabetes, and cancer. Similar to proteins, the biological function of lipids is dependent on molecular structure. Furthermore, isomeric lipids differing only in the positioning of C=C on the fatty acyl chain may have unique functional characteristics. Therefore, finding new methods to localize the positions of double bonds, number of double bonds, and cis/trans isomers of lipids is significant. Mass spectrometry combined with fragmentation methods is the most common technique for lipid analysis. However, in many cases, MS/MS data does not provide information about isomeric species with different positions of double bonds. Different hyphenated methods such as GC-MS have been used to tackle this problem. GC-MS is usually used to provide assignment of the double bond locations in mono-saturated lipids. However, cis/trans information cannot be gained for GC-MS analysis. Thomas and coworkers used ozonolysis for localization of double bonds. This method provides significant information about the position of double bonds. However, the procedure is difficult to implement. Furthermore, the fragmentation patterns are complex. It is therefore constructive to study HDX behavior of isomeric lipids with distinctive positioning of double bonds with a goal of developing methods for distinguishing ions. The rationale is that cis/trans isomeric
lipid molecules may generate different HDX profiles. This can be used as a fingerprint for assignments of cis/trans isomers of lipid number of double bonds in lipids. Future work will examine HDX behavior at elevated drift tube temperatures as this has been shown to cause exchange of aliphatic hydrogens.29

5.6. Application of New Fragmentation Techniques Followed by HDX

Fragmentation patterns for molecules that have undergone HDX can provide invaluable information about the locations of deuterium incorporation. Therefore, information about functional groups and relative orientation of the charge site to the basic site can be gained. In the present work, CID fragmentation patterns were studied after HDX. However, CID can result in ion scrambling as discussed earlier.3 Ion scrambling can be explained by the theory of the “mobile” proton in which protons transfer between different sites before ion fragmentation.30 As a result, valuable information about the location of exchangeable sites may be lost. To alleviate this issue, other fragmentation techniques such as ETD or charge transfer dissociation (CTD) can be implemented.28, 31-33

5.7. Application of Shift Reagents to Improve Characterization of Specific Classes of Biomolecules in 2D Plots of m/z vs Drift Time

To further improve the separation of metabolites and lipids in complex mixtures, ion mobility shift reagents may be used.34 A shift reagent is a molecule which forms a gas-phase complex with an analyte ion. This complex can be formed by non-covalent interactions. This complex helps improve mixture component separations by altering the mobilities of the ions arising from a change in their structures.35 The use of shift reagents can also enhance the overall peak capacity of the characterization method. Increased
peak capacity means more isomer and isobar compound determinations can be made in petroleum. Finally, for some analytes there will be a remarkable increase in ionization efficiency of the analyte/shift reagent complex compared with the analyte alone resulting in an increase in signal-to-noise (S/N) ratio.\textsuperscript{36, 37} The stabilization of charge by the analyte/shift reagent complex is the main factor for increasing S/N in this case.

In 2D IMS-MS data plots several molecular classes can be observed in specific areas called “conformation space”.\textsuperscript{36} The space in 2D plots of IMS-MS datasets that a specific class of molecules like peptides and lipids occupy is called a correlation band.\textsuperscript{1} After adding shift reagents to analytes, specific classes of molecules can be attached to these shift reagents and consequently shift to areas outside of the correlation band. Based on increasing or decreasing the mobility of these complexes, shift reagents are classified as low-density shift reagents or high-density shift reagents. Crown ethers are an example of low density shift reagents that shift the mobility above the correlation band.\textsuperscript{35} Lanthanide-based shift reagents are an example of high-density shift reagents that shift the mobility below the correlation band.\textsuperscript{36} The parameters that should be considered in choosing a shift reagent for a specific class of molecules involve: the size of the cavity, soft- and hard-interaction considerations, the number of hydrogen bonds, and the gas-phase basicity of the hetero-atomic sites in the shift reagent and the analyte ion, and the pH of the solution. Cryptands are another class of compounds that have the potential to be used as shift reagents. An interesting property of these molecules is that they are composed of two rings as opposed to a single ring. This leads to increases in shift reagent rigidity and enhancement in their selectivity. By changing the cavity size
(internal portion of the ring structure), it is possible to capture specific molecules. In the future, Cryptands will be implemented as part of the CCS database search approach.
5.8. References


