Forensic Applications of Gas Chromatography—Differential Mobility
Spectrometry, Gas Chromatography/Mass Spectrometry, and Ion Mobility Spectrometry with Chemometric Analysis

A dissertation presented to
the faculty of
the College of Arts and Sciences of Ohio University

In partial fulfillment
of the requirements for the degree
Doctor of Philosophy

Yao Lu
March 2010
© 2010 Yao Lu. All Rights Reserved.
This dissertation titled
Forensic Applications of Gas Chromatography—Differential Mobility
Spectrometry, Gas Chromatography/Mass Spectrometry, and Ion Mobility
Spectrometry with Chemometric Analysis

by

YAO LU

has been approved for
the Department of Chemistry and Biochemistry
and the College of Arts and Sciences by

______________________________
Peter B. Harrington
Professor of Chemistry and Biochemistry

______________________________
Benjamin M. Ogles
Dean, College of Arts and Sciences
Abstract

LU, YAO, Ph.D., March 2010, Chemistry and Biochemistry

Forensic Applications of Gas Chromatography—Differentail Mobility Spectrometry, Gas Chromatography/Mass Spectrometry, and Ion Mobility Spectrometry with Chemometric Analysis (148 pp.)

Director of Dissertation: Peter B. Harrington

Rapid, practical, and low-cost analytical methods are always desirable in forensic analysis. Using proper sample preparation techniques with the application of gas chromatography/mass spectrometry (GC/MS), gas chromatography—differential mobility spectrometry (GC–DMS) and ion mobility spectrometry (IMS) with chemometric analysis, analytical methods were developed for fast and practical identification and classification of analytes in complicated matrices.

GC–DMS was investigated as a tool for analysis of ignitable liquids from fire debris. The combined information afforded by gas chromatography and differential mobility spectrometry provided unique two-way patterns for each sample of ignitable liquid. Fuzzy rule-building expert system (FuRES) models constructed with the neat ignitable liquids identified the spiked samples from simulated fire debris with 99.07±0.04% accuracy. The performances of DMS as gas chromatographic detector was also compared with mass spectrometry (MS) using a chemometric tool, projected difference resolutions (PDRs). The PDR results show that one-way mass spectra data exhibit higher resolution than DMS data, while total ion chromatograms from
GC–DMS show higher resolution than that from GC/MS for differentiating seven kinds of ignitable liquids.

Direct methylation and solid phase microextraction (SPME) were used as a sample preparation technique for classification of bacteria based on fatty acid methyl esters (FAMEs) profiles. Compared with traditional chemical derivatization and liquid-liquid extraction (LLE), the method presented in this work avoids using inorganic and organic solvents and greatly decrease sample preparation time as well. The difference between Gram-positive and Gram-negative bacteria was clearly observed with the application of principal component analysis (PCA) of GC/MS data of bacterial FAMEs. The cross-validation study using ten bootstrap Latin partition (BLP) and fuzzy rule building expert system (FuRES) presented an 87±3% correct classification rate.

A comparatively rapid and reliable screening method for detection of cocaine and its metabolites, benzoylecgonine and cocaethylene in urine was demonstrated using solid phase extraction (SPE) coupled with IMS. Data analysis with alternating least squares (ALS) was used to model the IMS spectral datasets and separate the reactant ion peak from the product ion peaks. This method provides forensic chemists a viable approach for fast and simple drug screening.

Approved: ____________________________________________________

Peter B. Harrington
Professor of Chemistry and Biochemistry
Acknowledgments

First of all, I would like to give my deepest gratitude to my research advisor, Dr. Peter B. Harrington for giving me the opportunity of being a member in his research group. Not only instructing me in my research, he also generously spent a lot of his precious time reviewing and correcting my conference abstracts, papers and dissertation, helped me practice all my presentations. This dissertation could not be accomplished without his invaluable help and guidance. I would also like to thank my dissertation committee, Drs. Glen P. Jackson, Shiyong Wu, Hao Chen, and Xiaozhuo Chen, for their assistance in my academic and research progress.

I would like to thank the Department of Chemistry and Biochemistry at Ohio University for granting me the great opportunity of achieving my doctoral degree. The Center for Intelligent Chemical Instrumentation is thanked for supporting conference expenses. The Wright Patterson Air Force Base and United States Department of Agriculture, Agricultural Research Service are also thanked for their financial supports. Sionex Corporation is acknowledged for the use of differential mobility spectrometers and providing technical assistance. Amgen, Inc. is thanked for providing summer intern opportunity.

To all my group members, thank you for the companionship, helpful comments and suggestions, you make my stay in the laboratory enjoyable. To all my friends, no matter where you are, thank you for the friendship, this is always a precious asset in my life.
Last but not least, I thank my father Baosen Lu, my mother Chuanxin Liu and my brother Zhen Lu, without your constant support and encouragement I could not make this far. Although we are separated by distance, I can feel your love every day. Special thank to my husband Junfeng Huang, who took care of me during all these years, giving me courage and love and accompanying with me along this journey.
Table of Contents

Abstract .......................................................................................................................... 3

Acknowledgement ......................................................................................................... 5

List of Tables ............................................................................................................... 11

List of Figures ............................................................................................................. 13

List of Abbreviations .................................................................................................... 17

Chapter 1  Introduction .............................................................................................. 20
  1.1  Sample Preparation Techniques ........................................................................... 20
    1.1.1  Solid Phase Extraction .................................................................................. 20
    1.1.2  Solid Phase Microextraction ......................................................................... 24
  1.2  Ion Mobility Spectrometry .................................................................................. 28
  1.3  Differential Mobility Spectrometry ..................................................................... 32
  1.4  Data Processing and Analysis ............................................................................. 36
    1.4.1  Principal Component Analysis ...................................................................... 37
    1.4.2  Fuzzy Rule Building Expert System .............................................................. 39
    1.4.3  Projected Difference Resolution .................................................................... 40
    1.4.4  Alternating Least Squares ............................................................................ 42

Chapter 2  Forensic Application of GC–DMS with Two-Way Classification of Ignitable Liquids from Fire Debris ................................................. 47
  2.1  Introduction ......................................................................................................... 47
  2.2  Experimental Section .......................................................................................... 50
<table>
<thead>
<tr>
<th>2.2.1</th>
<th>Chemicals and Materials</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2</td>
<td>Instruments</td>
<td>51</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Procedure</td>
<td>52</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Data Processing</td>
<td>53</td>
</tr>
<tr>
<td>2.3</td>
<td>Results and Discussion</td>
<td>55</td>
</tr>
<tr>
<td>2.3.1</td>
<td>GC−DMS of Neat Ignitable Liquids</td>
<td>55</td>
</tr>
<tr>
<td>2.3.2</td>
<td>GC−DMS of Burned Ignitable Liquids</td>
<td>62</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Principal Component Analysis</td>
<td>63</td>
</tr>
<tr>
<td>2.3.4</td>
<td>FuRES Classifier</td>
<td>68</td>
</tr>
<tr>
<td>2.4</td>
<td>Conclusions</td>
<td>74</td>
</tr>
</tbody>
</table>

Chapter 3  Comparison of Differential Mobility Spectrometry and Mass Spectrometry for Gas Chromatographic Detection of Ignitable Liquids from Fire Debris Using Projected Difference Resolution

<table>
<thead>
<tr>
<th>3.1</th>
<th>Introduction</th>
<th>76</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Experimental Section</td>
<td>78</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Materials and Chemicals</td>
<td>78</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Instruments</td>
<td>78</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Neat Sample Collection</td>
<td>79</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Matrix Exposed Sample Preparation and Collection</td>
<td>80</td>
</tr>
<tr>
<td>3.3</td>
<td>Data Processing</td>
<td>80</td>
</tr>
<tr>
<td>3.4</td>
<td>Results and Discussion</td>
<td>81</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Two-way GC/MS and GC−DMS of Ignitable Liquids</td>
<td>81</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Projected Difference Resolutions</td>
<td>84</td>
</tr>
</tbody>
</table>
Chapter 4  Classification of Bacteria by Simultaneous Methylation−Solid Phase Microextraction and Gas Chromatography/Mass Spectrometry Analysis of Fatty Acid Methyl Esters

4.1  Introduction

4.2  Experimental Section

4.2.1  Bacterial Agents and Growth Conditions

4.2.2  Instrumentation and Apparatus

4.2.3  Sample Preparation

4.2.4  Data Analysis

4.3  Results and Discussion

4.3.1  Derivatization of Bacterial Fatty Acid

4.3.2  FAME GC/MS Profiles

4.3.3  Multivariate Analysis

4.4  Conclusions

Chapter 5  Detection of Cocaine and Its Metabolites in Urine Using Solid Phase Extraction−IMS with Alternating Least Square

5.1  Introduction

5.2  Experimental

5.2.1  Chemicals and Reagents

5.2.2  SPE Procedures

5.2.3  IMS Parameters
5.2.4 Data Acquisition and Processing ........................................ 115
5.3 Results and Discussion ....................................................... 116
  5.3.1 Ion Mobility Spectra .................................................. 116
  5.3.2 Reduced Mobilities .................................................... 118
  5.3.3 Solid Phase Extraction ............................................... 120
  5.3.4 Alternating Least Squares .......................................... 122
  5.3.5 Adulterants Influence ............................................... 123
5.4 Conclusions ...................................................................... 123

Chapter 6 Summary and Future Work .................................... 128

Appendix A Publications .......................................................... 147

Appendix B Presentations .......................................................... 148
List of Tables

Table 2-1. Confusion matrix with 95% confidence interval for FuRES analysis of ignitable liquids burned with nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP FuRES models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine................................................................. 72

Table 2-2. Confusion matrix with 95% confidence interval for optimized PLS analysis of ignitable liquids burned with nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP PLS models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine................................................................. 73

Table 2-3. Comparison of prediction results of TIC, DMS, and two-way data confusion matrices obtained from optimized PLS and FuRES using 10 bootstraps of four Latin partitions training-prediction pairs of the neat samples. The fire debris samples on the polyester and nylon carpets were classified with the 40 resulting models from the BLP evaluation.................. 75

Table 3-1. Minimum projected difference resolution values with 95% confidence interval, calculated from neat sample datasets...................... 85

Table 3-2. Confusion matrix with a 95% confidence interval for FuRES analysis of two-way GC/MS data of burned samples, nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP FuRES models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine................................................................. 88

Table 3-3. Confusion matrix with a 95% confidence interval for FuRES analysis of two-way GC–DMS data of burned samples, nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP FuRES models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine................................................................. 89

Table 3-4. Two-way prediction errors by FuRES analysis using 10 bootstraps of four Latin partitions training and prediction pairs of neat samples, spiked samples, and burned samples................................................................. 90

Table 4-1. Bacteria investigated in this study.................................................. 96
Table 4-2. Confusion matrix with 95% confidence interval for FuRES analysis of GC/MS data from bacterial fatty acid methyl esters. 10 bootstraps and 5 Latin partitions.

Table 5-1. The molecular masses, drift times, and measured reduced mobility values of cocaine and its metabolites investigated using SPE and IMS.
## List of Figures

| Figure 1-1. Schematic representation of solid phase extraction using cartridge. Adapted from http://www.sigmaaldrich.com/Graphics/Supelco/objects/4600/4538.pdf (Accessed in December, 2008). | 23 |
| Figure 1-2. Schematic diagram of a SPME device. | 25 |
| Figure 1-3. Schematic of Barringer Ionscan® Model 350 ion mobility spectrometer adapted from manufacture’s training manual. The circles indicate the formation of different ions, which are separated based on their size-to-charge ratio in the drift region. | 30 |
| Figure 1-4. Dependence of ion mobility on electric field strength \( (E) \). \( K(E_h) \) and \( K(E_l) \) stand for ion mobilities at high and low electric field, respectively. \( a \) is a characteristic parameter for a specific ion. | 35 |
| Figure 1-5. Operation schematic of a differential mobility spectrometer made by Sionex. | 35 |
| Figure 1-6. IMS spectral plot of 1 µL cocaethylene at concentration of 1 mg/mL in acetonitrile solution. The spectra were not modeled with ALS. | 45 |
| Figure 1-7. ALS-modeled IMS spectra (top) and ALS extracted concentration profiles (bottom) of 1 µL of cocaethylene at concentration of 1 mg/mL in acetonitrile analyzed in positive ion mode. Reactant ion peak (—); analyte ion peaks (—). | 46 |
| Figure 2-1. Topographic plot from GC–DMS characterization of neat gasoline. Sampling method: Headspace solid SPME. The top plot is the profile after data alignment and normalization. The bottom plot is the ion squared profile of the top plot. | 56 |
| Figure 2-2. Two-way GC-DMS positive ion data objects of SPME samples of gasoline, diesel, lighter fuel, and turpentine without data normalization. | 57 |
| Figure 2-3. Average ion chromatograms from GC–DMS analysis of neat sample liquids absorbed on SPME fibers. | 60 |
Figure 2-4. Average differential mobility spectra from GC-DMS analysis of neat sample volatiles absorbed on SPME fibers. ................................................. 61

Figure 2-5. Total ion chromatograms from GC–DMS analysis of neat gasoline (top) and burned gasoline with nylon carpet sample (bottom). ......................... 63

Figure 2-6. Principal component analysis score plot of two-way GC–DMS profiles of neat ignitable liquids. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine. Each letter represents a replicate of the ignitable liquid sample. A 95% confidence interval is drawn around the mean of each class. ........................................... 65

Figure 2-7. Principal component analysis score plot of two-way GC–DMS profiles of seven samples extracted from different matrices. A. Neat samples, B. Simulated fire debris (Nylon carpet), C. Simulated fire debris (Polyester carpet). A 95% confidence interval is drawn around the mean of each class. ........................................................................................................ 66

Figure 2-8. Principal component analysis score plot of two-way GC–DMS profiles of seven samples with different treatment methods. A. Neat samples, B. Spiked samples from simulated fire debris (burned nylon carpet), C. Burned samples (burned with nylon carpet). A 95% confidence interval is drawn around the mean of each class. ........................................................................... 67

Figure 2-9. FuRES classification tree for seven neat ignitable liquid samples. Nc is the number of samples and H is the absolute fuzzy entropy of classification. .................................................................................... 69

Figure 2-10. Image of the weights for the two-way rule #1 (middle) with negative features characteristic of classes to left of the tree (top) and positive features characteristic of classes to right of the tree (bottom). ..... 70

Figure 3-1. Two-way data profiles of GC/MS (top) and GC–DMS (bottom) analysis of neat gasoline adsorbed on SPME fibers. ............................................. 83

Figure 3-2. FuRES classification trees for seven neat ignitable liquid samples on two-way GC/MS (top) and GC–DMS (bottom) data with a 95% classification rate. Nc is the number of replicates and H is the absolute fuzzy entropy of classification. ................................................................. 86

Figure 4-1. Total ion chromatogram of Gram-negative bacterium Ecoli (ATCC 43895) FAMEs using SPME as the sampling method, without (top) and with (bottom) derivatizing bacterial cells using TMAH methanol solution. ....... 100
Figure 4-2. Influence of adsorption temperature on the performance of SPME. ........................................................................................................ 102

Figure 4-3. Total ion chromatogram of Gram-negative bacterium *Ecoli* (ATCC 43895) FAMEs with different sample preparation methods. Top: TMAH derivatization and headspace SPME; bottom: Traditional alkaline saponification and LLE. ........................................................................ 103

Figure 4-4. Total ion chromatogram of Gram-positive bacteria *Enterococcus* (ATCC 19433) FAMEs with different sample preparation methods. Top: TMAH derivatization and headspace SPME; bottom: Traditional alkaline saponification and LLE. ........................................................................ 104

Figure 4-5. Top: PCA score plot of six bacteria. The first two PCs account for 69% variance in total. The same letter corresponds to the replicates of each bacterium listed in Table 1. Bottom: The 1st PCA loading plot. ............... 107

Figure 4-6. FuRES classification tree of six species of bacteria. Nc is the number of replicates and H is the absolute fuzzy entropy of classification. 109

Figure 5-1. Chemical structures for (top) cocaine and its metabolites, (bottom left) benzoylecgonine, (bottom right) cocaethylene. .................. 112

Figure 5-2. Ion mobility spectra image (top), ALS spectra (bottom left), and concentration profile (bottom right) of urine sample containing 100 ng/mL benzoylecgonine with SPE treatment. Reactant ion (−−−−); analyte ion (−). ................................................................. 117

Figure 5-3. ALS spectra of 1 µg/mL cocaine (top), benzoylecgonine (middle) and cocaethylene (bottom) added on IMS disk directly. Reactant ion peak (-- ---); analyte ion peak (——). ............................................................... 119

Figure 5-4. ALS spectra of mixture of benzoylecgonine and cocaethylene. Reactant ion peak (−−−−); analyte ion peaks (—). ...................... 119

Figure 5-5. ALS spectra of (a) negative urine sample without SPE process (top); (b) urine sample containing 100 ng/mL benzoylecgonine without SPE process (middle); (c) urine sample containing 100 ng/mL benzoylecgonine with SPE process (bottom). Reactant ion peak (-----); analyte ion peaks (—). ........................................................................................................ 121
Figure 5-6. ALS spectra of urine sample containing 100ng/mL cocaethylene after SPE process. Number of components: two (top); three (bottom). Reactant ion peak (−−−−); analyte ion peak (−); interference (- — - —). ................................................................. 124

Figure 5-7. Comparison of average spectrum (top), maximum spectrum (middle) and ALS spectra (bottom) of urine sample containing 150 ng/mL of benzoylecgonine. ............................................................................. 125

Figure 5-8. Comparison of average spectrum (top), maximum spectrum (middle) and ALS spectra (bottom) of urine sample containing 150 ng/mL of cocaethylene. .................................................................................. 126

Figure 5-9. ALS spectra of urine sample containing 150 ng/mL cocaethylene and the manufacturer’s suggested amount of Add-it-ive® after the SPE process. Three components were chosen when applying ALS. Reactant ion peak (−−−−); analyte ion peak (−); interference (- − - −). ............... 127
Abbreviations

ALS ............................................................... alternating least squares
ANN ............................................................. artificial neural network
ANOVA ............................................................... analysis of variance
ANOVA-PCA ................................................. analysis of variance-principal component analysis
APCI .......................................................... atmospheric-pressure chemical ionization
ASTM .......................................................... American Society for Testing and Materials
ATCC ......................................................... American Type Culture Collection
BHIA .......................................................... brain-heart infusion agar
BHIB .......................................................... brain-heart infusion broth
BL2 ............................................................... biosafety level 2
BLP .............................................................. bootstrap Latin partition
BTX .............................................................. benzene, toluene and xylene
CDC .......................................................... Centers for Disease Control
CI ................................................................. confidence interval
CWAs .......................................................... chemical warfare agents
DMS .......................................................... differential mobility spectrometry
EI ............................................................... electron ionization
FAME .......................................................... fatty acid methyl ester
FAIMS .......................................................... field asymmetric ion mobility spectrometry
FTICR .......................................................... Fourier transform ion cyclotron resonance
FuRES .......................................................... fuzzy rule building expert system
GC–DMS .......... gas chromatography–differential mobility spectrometry
GC/MS ......................... gas chromatography/mass spectrometry
HPLC.......................... high performance liquid chromatography
HS .......................................................... headspace
IMS ....................................................... ion mobility spectrometry
IR ............................................................ infrared
K .............................................................. mobility
K₀ ........................................................... reduced mobility constant
MALDI ......................... matrix assisted laser desorption/ionization
MS ........................................................ mass spectrometry
MuRES ......................... multivariate rule building expert system
MCR ................................................. multivariate curve resolution
NA ........................................................ nutrient agar
NB ........................................................ nutrient broth
NIR ....................................................... near infrared
NSDUH ......................... National Survey on Drug Use and Health
ODS ...................................................... octadecylsilane
PC .............................................................. principal component
PCA ............................................... principal component analysis
PCS ................................................... principal component similarity
PCT ................................................... principal component transform
PDMS .................................................. polydimethylsiloxane
PDR .................................................... projected difference resolution
PLS...............................................................partial least squares
RIP.................................................................reactant ion peak
SAMHSA.....Substance Abuse and Mental Health Services Administration
SIMPLISMA......Simple-to-use interactive self-modeling mixture analysis
SPE .................................................................solid phase extraction
SPME .............................................................solid phase microextraction
STP ...............................................................standard temperature & pressure
SVD .................................................................singular value decomposition
TMAH.......................................................tetramethylammonium hydroxide
TMH.............................. thermally assisted hydrolysis and methylation
TSA .................................................................tryptic soy agar
TSB .................................................................tryptic soy broth
TIC.................................................................total ion chromatogram
VI.................................................................virtual instrument
Chapter 1  Introduction

This dissertation presents research focused on applications of different analytical instruments with multivariate data analysis to provide practical methods for forensic analysis. The general introduction to sample preparation techniques, instrumentation and data analysis methods are given in Chapter 1. Chapter 2 presents the forensic application of gas chromatography-differential mobility spectrometry with two-way classification of ignitable liquids from fire debris. In Chapter 3, the performances of differential mobility spectrometry and mass spectrometry for gas chromatographic detection of ignitable liquids are compared using a chemometric tool, projected difference resolution. Chapter 4 presents the classification of bacteria by simultaneous methylation–solid phase microextraction and gas chromatography/mass spectrometry analysis of fatty acid methyl esters. Detection of cocaine and its metabolites in urine using solid phase extraction–ion mobility spectrometry with alternating least squares is given in Chapter 5. Finally, the summary and future works are introduced in Chapter 6. The publications and selected presentations associated with this dissertation are shown in Appendix A and B, respectively.

1.1  Sample Preparation Techniques

1.1.1 Solid Phase Extraction

The development of a complete analytical method includes many steps from collecting sample to finally reporting the results. Compared with the
automation of the modern analytical instrument, sample preparation is often the most time-consuming step. The samples in forensic analysis typically are come in complicated matrices. To purify and concentrate samples for direct analysis using analytical instruments, sample preparation is a key step for forensic analysis.

According to the complexity of the sample, the matrix, and the available analytical instruments, there are many sample preparation techniques that can be applied individually or sequentially. Liquid-liquid extraction (LLE) is one of the most frequently used techniques for sample preparation, which consists of sample dissolution, purification, and extraction. However, the disadvantages with LLE include the use of large volumes of organic solvent, higher cost, possible creation of emulsions, time-consuming, and exposure to potentially hazardous solvents. Nowadays, solid phase extraction (SPE) is increasingly used as a sample preparation technique in forensic laboratories to isolate analytes of interest from a wide variety of matrices like urine, blood, hair, and food. Compared with LLE, SPE is simple and safe to use, allowing using lower sample quantities. The reduced-volume columns allow elution in smaller volumes of solvent that can be dried down quickly.

The principle of SPE is similar to that of LLE, involving the distribution of analytes between two phases. In SPE, the analytes to be extracted are adsorbed by a solid phase with a greater affinity for the analytes than for the sample matrix. Then the analytes retained on the solid phase can be eluted
by a solvent with a greater affinity for the analytes. Choosing appropriate solid phase can help increase the extraction efficiency. Basically, depending on how the solid phases behave during the process of SPE, retention mechanisms can be hydrophobic, polar, and ionic, which correspond to the application of reversed phase, normal phase, and ion exchange. A reverse phase octadecylsilane (ODS) (C\textsubscript{18}) solid phase was used in the experimental section described in Chapter 5 for extraction of cocaine metabolites in urine sample.

SPE is widely performed with a solid phase cartridge or disk. Figure 1-1 presents the schematic of SPE setup using solid phase cartridge with packed stationary phase inside. The experiment described in Chapter 5 was performed using the same setup. The solid phase in cartridges needs to be solvated to interact efficiently with aqueous samples. Organic solvents (such as methanol or acetonitrile) and buffer solutions are used to condition the solid phase before the extraction step. Then the aqueous samples are loaded onto the stationary phase, where either the desired analytes of interest or undesired impurities in the sample are retained. If the desired analytes are retained, an appropriate solvent is selected to pass through the stationary phase and the effluent is kept for further analysis. During this process, vacuum is maintained to suck the solvent through the column of solid phase.

SPE is an effective sample treatment technique in drug analysis\textsuperscript{2} as it gives high recoveries and clean extracts. The detection of cocaine and its metabolites in different matrices, such as urine,\textsuperscript{3} hair,\textsuperscript{4} oral fluid samples,\textsuperscript{5}
human plasma,\textsuperscript{6} amniotic fluid,\textsuperscript{7} and meconium,\textsuperscript{8} has already been investigated using SPE coupled with chromatographic analysis. In Chapter 5, cocaine metabolites in urine samples were extracted by SPE reversed-phase C\textsubscript{18} cartridges and then analyzed using a bench top ion mobility spectrometry (IMS).

![Figure 1-1. Schematic representation of solid phase extraction using cartridge. Adapted from http://www.sigmaaldrich.com/Graphics/Supelco/objects/4600/4538.pdf (Accessed in December, 2008).]
1.1.2 Solid Phase Microextraction

In addition to SPE, another sample preparation technique, solid phase microextraction (SPME), was also applied in this dissertation. SPME was first introduced by Pawliszyn in 1989, and then quickly developed and widely used in different fields of analytical chemistry in the 1990’s. SPME is a simple, efficient, and solvent free sample preparation technique. The steps of conventional LLE, such as extraction, concentration, and transfer to chromatography can be integrated in one step and one device, greatly simplifying the sample preparation procedure. In Chapter 2 and Chapter 3, SPME was used to extract volatile ignitable liquids from fire debris samples. In Chapter 4, SPME was also utilized to extract fatty acid methyl esters from bacterial cells.

The SPME apparatus looks like a modified syringe, as seen in Figure 1-2. It consists of a fiber holder, an adjustable depth gauge, a plunger, and a fiber assembly. The SPME fiber is a 1 cm long polymer coated fused silica, which is connected to the plunger. The needle can protect the fiber and also insert into the septum of the sampling vial when sampling. The fiber is retracted into the stainless steel needle after sampling, and then transferred to the chromatograph. Gas chromatography is one of the preferential techniques used with SPME. In this case, the analytes are thermally desorbed in the hot injector and carried by the carrier gas into capillary column directly. Splitless injection is usually used because there is no solvent involved in this process. In this dissertation, a 100 μm non-polar and
non-porous polydimethylsiloxane (PDMS) fiber was applied for all studies due to the properties of the analytes. In general, polar fibers are used for polar analytes and non-polar fibers for non-polar analytes. The SPME device and fibers used in this dissertation were purchased from Supelco Inc. (Bellefonte, PA).

Figure 1-2. Schematic diagram of a SPME device.
The SPME process is based on equilibrium between the sample and the polymer coated fiber. Each sampling is a two-step process including extraction and desorption. During extraction, the fiber is inserted directly into the sample or the headspace (HS) of the sample, where the analytes are concentrated. Then the fiber is exposed and the analyte is either thermally desorbed or eluted by the appropriate solvents. When the analytes of interest are present with complex matrices such as biological samples, some food products, and environmental samples, the SPME technique is mainly applied to extract the analytes from the HS in such cases.

Many factors will influence the desorption efficiency. When thermally desorbing the analytes from the SPME fiber, the analyte volatility, the thickness of the fiber coating, injection depth, injector temperature, and fiber exposure time may affect the desorption efficiency. In addition to selecting splitless injection mode when analyzing with GC, the optimal desorption temperature is approximately equal to the boiling point of the least volatile component in the analytes. The extraction efficiency can be increased by agitating the sample solution with a magnetic stirrer when direct SPME sampling is applied. Some other experimental factors, such as nature of the fiber, SPME mode, extraction temperature, and ultrasonic assistance may also affect the final results. To maintain a good reproducibility for each replicate measurement, all the experimental parameters have to be kept constant, especially when quantitative results are required.
Analyzing polar compounds is a challenge especially when they are present in biological and environmental matrices, because they are hydrophilic and difficult to be extracted with organic solvents and adsorbed without the interference from the matrices. Derivatization sometimes is necessary to chemically transform this kind of analytes into a form which is more suitable for analysis. As a result, extraction efficiency and selectivity can be improved by increasing the volatility and reducing the polarity of the analytes. This approach has been applied to phenol analysis by converting them to acetates with acetic anhydride.\textsuperscript{18} Analysis of amphetamine, methamphetamine, methylenedioxyamphetamine, and methylenedioxymethamphetamine in whole blood was performed by using in-matrix ethyl chloroformate derivatization and simultaneous HS-SPME followed by GC/MS.\textsuperscript{19} Vinas \textit{et al.} used SPME on-fiber derivatization method to analyze some polyphenols in wine and grapes, followed by GC/MS measurement.\textsuperscript{20} This method was also used to determine butyl and phenyltin compounds in human urine after derivatization with tetraethylborate.\textsuperscript{21} In Chapter 4, a derivatization reagent, tetramethylammonium hydroxide (TMAH), was used to derivatize bacterial fatty acids into more volatile fatty acid methyl esters for GC analysis. In this project, \textit{in situ} derivatization is applied by direct adding TMAH to the bacterial cells, then the derivatization takes place and the SPME fiber extracts the volatile fatty acid methyl esters from the headspace.
1.2 Ion Mobility Spectrometry

Ion mobility spectrometry (IMS), previously known as plasma chromatography, was developed in the early 1970’s to detect trace organic compounds based on the difference between their gas phase ion mobilities. IMS can characterize chemical substances based on velocity of gas-phase ions in an electric field $E$ (V/cm). The drift velocity $v_d$ (cm/s), which is the ion velocity through the electric field, is proportional to the electric field strength as in equation (1-1):

$$v_d = KE$$

for which $K$ is termed the mobility coefficient of the ion. This coefficient is usually normalized to a standard pressure ($P$, 760 Torr) and temperature ($T$, 273 K) to yield a reduced mobility ($K_0$). This process is expressed by equation (1-2):

$$K_0 = \left(\frac{v_d}{E}\right)\left(\frac{273}{T}\right)\left(\frac{P}{760}\right)$$

for which $v_d$ is obtained by dividing the drift length by the ion’s transit time, $T$ is temperature in Kelvin and $P$ is pressure in Torr of gas atmosphere through which the ions move. The last two factors correct the drift gas density to standard temperature and pressure (STP) conditions.

In the experimental section described in Chapter 5, a Barringer Ionscan® Model 350 ion mobility spectrometer was used to detect cocaine and its metabolites extracted from urine sample combined with chemometric analysis. The schematic of this spectrometer is given in Figure 1-3. The
analyte is first placed on a glass fiber disk and inserted into a sample cartridge. After fixed on a sample cartridge assembly, the cartridge is then slid to the top of the desorber, where the sample material is heated by the desorber heater and inlet heater and vaporized to gas phase. The vaporized sample is blown by the carrier gas to the ionization region where the ions are generated by atmospheric-pressure chemical ionization (APCI). The ions are then injected by ion shutter into a drift region where they are separated and reach the detector at different drift times according to their sizes and charges. In this IMS model, the ionization process is initiated by beta particles emitted from a $^{63}$Ni source. The positive ions can be formed through proton transfer reactions between the analyte and the reactant ion. Nicotinamide, C$_5$H$_4$NCONH$_2$, is the internal calibrant when using the Ionscan® Model 350 in the positive ion mode.

For the production of positive ions, proton transfer in the equation (1-3) is the charge transfer mechanism:

$$A + H^+(C_5H_4NCONH_2) \rightarrow H^+ + C_5H_4NCONH_2$$

In the above reaction, A is the analyte under investigation. It is ionized by collision with the reactant ions H$^+(C_5H_4NCONH_2)$ under atmospheric pressure. In the positive detection mode, the ion shutter grid and counter-current drift gas prevent negatively charged ions and neutral molecules from entering the drift region of the spectrometer. When the ions travel through the drift tube, they will be separated based on their differences in ion mobility, which primarily depends on the ion’s size-to-charge ratio. As the ions are collected
at the detector plate, the resultant currents are measured. An ion mobility spectrum is depicted by the ion intensity as a function of drift time.

![Schematic of Barringer Ionscan® Model 350 ion mobility spectrometer](image)

**Figure 1-3.** Schematic of Barringer Ionscan® Model 350 ion mobility spectrometer adapted from manufacture’s training manual. The circles indicate the formation of different ions, which are separated based on their size-to-charge ratio in the drift region.

In IMS, drift times are not appropriate for compound identification because of the variations among instruments, such as drift potential, temperature, and pressure, which may result in variation of the drift times among the measurements. Reduced mobilities ($K_0$) are often used instead for identification purposes to correct for the instrumental variations, such as drift potential and drift tube length as well as pressure and temperature of the drift region. Reduced mobilities provide qualitative results that may be
used for identification and help compare results between instruments, laboratories, and different days. The drift times of a reference ion with known reduced mobility can be used to calculate the reduced mobility of the analyte of interest, according to equation (1-4):

\[
K_0 \text{unknown} = \frac{K_0 \text{calibrant}}{K_0 \text{calibrant}}
\]

for which \(K_0 \text{unknown}\) and \(K_0 \text{calibrant}\) are the reduced mobilities of the analyte ion and calibrant ion, respectively, and \(t\) refers to drift time. In Chapter 5, nicotinamide was used as the internal calibrant under positive detection mode, having a reduced mobility of 1.86 ± 0.04 cm\(^2\)V\(^{-1}\)s\(^{-1}\). In the negative detection mode of Ionscan® Model 350 IMS, methyl salicylate (\(K_0\) of 1.54 cm\(^2\)V\(^{-1}\)s\(^{-1}\)) is used as the internal calibrant.\(^{22}\)

IMS is sensitive for compounds with high electronegativity and proton affinity. It has been implemented for chemical warfare agents (CWAs) detection,\(^{23, 24}\) medical diagnostics,\(^{25}\) biomolecules and biopolymers detection,\(^{26, 27}\) and industrial and environmental monitoring.\(^{28-31}\) It is also employed by pharmaceutical companies for cleaning verification of manufacturing equipment and direct analysis of formulations.\(^{32}\) IMS offers the advantages of atmospheric ionization, fast screening, high sensitivity, simple operation, low cost, and portability,\(^{22}\) compared with GC and high performance liquid chromatography (HPLC) when quantitative results are not crucial. These characteristics make IMS amenable not only to laboratories or the industry, but also to field deployment.
1.3 Differential Mobility Spectrometry

In Chapter 2 and 3, differential mobility spectrometry (DMS) was applied as a GC detector and accomplished with a Sionex microfabricated analyzer that characterizes analytes based on the difference between ion mobilities under high and low electric fields at ambient pressure. Different from the previously introduced IMS, ions separated in DMS are conveyed by a drift gas that is orthogonal to the applied electric field, which is an AC voltage. Also, DMS instruments do not need ion-shutter, drift rings or aperture grids which make the miniaturization possible and increase ion transmission efficiency. The principle of separating ions by their changes in mobility at different electric field strengths was proposed in 1993\textsuperscript{33}. Many publications have introduced the principle of DMS in detail\textsuperscript{34-36}. This technique has also been referred to as field asymmetric ion mobility spectrometry (FAIMS)\textsuperscript{37}, although these analyzers have cylindrical electrodes and DMS has planar electrodes. The planar electrode geometry of DMS offers an advantage over FAIMS in that both positive and negative ions can be separated simultaneously.

The dependence of ion mobility on electric field strength ($E$) is given in Figure 1-4. As can be seen from this figure, different ions may exhibit similar ion mobility in low electric fields ($E_L$). While under high electric field strength ($E_H$), basically higher than 10 kV/cm, the ion mobility may diverge from the similar mobility that occurs at low field strengths. Ions that do not change mobility with respect to field strength will traverse the electrodes and...
be detected. Ions having different mobilities under the high frequency asymmetric electric field will migrate to one of the two electrodes where they will be neutralized on contact and no longer detected. The relationship for the electric field dependence with the mobility coefficient is given in the following equation:

\[ K(E_h) = K(E_L)[1 + \alpha \left( \frac{E}{N} \right)] \]  

for which \( K(E_h) \) and \( K(E_L) \) are the ion mobilities at high and low electric field, respectively. The \( \alpha \) values are characteristic parameters for different ions, and \( N \) is the gas number density. The \( \alpha \left( \frac{E}{N} \right) \) characterizes the electric field dependence of a specific ion. Many other factors also affect the ion mobility, such as the chemical structure of the ion, its local environment, and temperature. The operation schematic of a Sionex differential mobility spectrometer is given in Figure 1-5. An analyte enters the spectrometer where it is photoionized. Then the ions pass the drift area between two electrodes where a high frequency asymmetric voltage is applied. To prevent the ions with \( \alpha \) value not equal to zero from being neutralized at the electrodes, they are selectively filtered by varying an applied compensation voltage that is a low dc voltage superimposed on the high frequency asymmetric field. The compensation voltage corrects the migration of the ion caused by the differential mobility so that the ion can pass between the electrodes and be detected. A differential mobility spectrum is the current at the detection electrode with respect to the compensation voltage.
The differential mobility spectrometer used in this study contains two microplanar electrodes that are 25 mm long × 5 mm wide × 0.5 mm apart. The miniaturized model has the potential to be used for onsite sensors or detectors for gas chromatography. The small size and the ratio of performance to cost compared with other bench top instruments make DMS a promising technique. A variety of applications have been investigated by DMS. For example, DMS was used as a gas chromatographic detector to analyze jet fuels.\textsuperscript{38} The determination of nitro-organic explosives using DMS was achieved by modifying the drift gas with a small amount of methylene chloride.\textsuperscript{39} DMS also demonstrated ability in detecting environmental contaminants, such as chlorocarbons in ground water\textsuperscript{40}, volatile organic compounds in ambient air\textsuperscript{41}, and DMS coupled with SPME was exploited to detect benzene, toluene, and xylene (BTX) in water.\textsuperscript{42} In view of the growing concern about biological terrorism, DMS also exhibited potential as a sensor for detecting trace amounts of bioorganisms.\textsuperscript{43, 44}
Figure 1-4. Dependence of ion mobility on electric field strength \(E\). \(K(E_H)\) and \(K(E_L)\) stand for ion mobilities at high and low electric field, respectively. \(\alpha\) is a characteristic parameter for a specific ion.

Figure 1-5. Operation schematic of a differential mobility spectrometer made by Sionex.
1.4 Data Processing and Analysis

Chemometrics is a subject that uses mathematical and statistical methods to design experiments, analyze chemical measurements, and obtain useful information. It offers powerful tools for interpreting complex data. When investigating chemical data from different instruments, some chemometric methods can act as a complementary technique of analytical separation, mathematically resolving mixtures by determining the response profile (spectra) and concentration profiles. In case of sample identification and classification, optimal chromatographic separation with baseline resolution is difficult to achieve with many complex mixtures, data preprocessing like baseline correction and peak alignment is of significance in determining accurate results. Chemometric methods are powerful in data interpretation, qualitative analysis, quantitative analysis, and pattern recognition.\textsuperscript{45, 46}

In this dissertation, bold uppercase symbols designate matrices and bold lowercase symbols designate vectors. The transpose of matrices or vectors is indicated by a superscript T. For all the experimental data matrices present in this dissertation, each row contains one measurement and the measurement resolution elements are saved in the columns. Introduction of the chemometric methods applied in this dissertation will be described in this section.
1.4.1 **Principal Component Analysis**

To visualize the differences in GC–DMS profiles of ignitable liquids and GC/MS profiles of bacterial fatty acid methyl esters, principal component analysis (PCA), was applied to the whole datasets in each project. PCA is an efficient multivariate statistical tool for rendering overwhelming data in a reduced form that is amenable to visual interpretation and efficient for further computation. PCA is a mathematical procedure that transforms a larger number of correlated variables into a smaller number of uncorrelated variables called principal components (PCs). During this process, the main variations in the dataset are modeled with fewer variables. The PCs are mutually orthogonal to each other. The PCs are ordered by the variance that they span, thus the first PC accounts for the largest variance, and the second PC accounts for the second largest amount of variance and so on. In this dissertation, PCA observation score plots were generated using the first two PCs that span the largest proportion of variance. The PCs are calculated using singular value decomposition (SVD) defined in the following equation (1-6):

\[ \mathbf{D} = \mathbf{USV}^T \]  

(1-6)

for which \( \mathbf{D} \) is a data matrix, \( \mathbf{V}^T \) is the transpose of a matrix comprising the PCs or variable loadings as columns, \( \mathbf{S} \) is a diagonal matrix containing the singular values, and \( \mathbf{U} \) comprises the normalized scores. The singular values have the same units as \( \mathbf{D} \) and are proportional to the standard deviations that are spanned by the corresponding principal components. Observation scores can be obtained by the product \( \mathbf{DV} \) or \( \mathbf{US} \), which should be apparent
from equation (1-6). The relative and absolute variances are given for each component. A 95% confidence interval can also be plotted around the mean of each sample class to visualize the statistical significance.

Analysis of variance (ANOVA) is a statistical method that separates sources of variation from properly designed experiments and tests the significance of the variance for each experimental factor by comparing it to the residual error. Combined with ANOVA, PCA score plots can provide a clear picture of the significance as well as the relationship among the levels of each experimental factor. ANOVA-PCA was developed to simplify PCA of complex data. The original dataset is decomposed into additive matrices that characterize different experimental factors and the residual error, and all the matrices have the same dimensionality.\textsuperscript{47} PCA is then applied to the combination of each factor matrix and residual error matrix. Thus, interference between different experimental factors can be avoided but the experiment must be well designed so the factors are not confounded. For the significant factor, its variation will dominate over the pure error and the first PC in the score plot will characterize this variation if it is significant.

Recently it has been widely used as a powerful chemometric tool in a variety of fields such as discriminating between cultivars and treatment of broccoli using mass spectral fingerprint,\textsuperscript{48} investigating the reproducibility of matrix assisted laser desorption/ionization (MALDI) mass spectrometry measurements of bacteria,\textsuperscript{49} studying mid-infrared spectra of carraghenan gels,\textsuperscript{50} and characterizing sources of variance in plant materials using UV
spectra. This method was applied to evaluate the two-way GC–DMS of ignitable liquids data in Chapter 2.

1.4.2 Fuzzy Rule Building Expert System

A fuzzy rule building expert system (FuRES) is applied as the classification method for differentiating GC–DMS and GC/MS ignitable liquids data and GC/MS bacteria data. Rule building expert systems use machine learning to construct logical rules from exemplary sets of data. FuRES is a multivariate classification algorithm that uses fuzzy logic instead of classic logic to classify data. Fuzzy logic has been used in investigating chromatograms, classification and feature selection of mass spectra, identification of plant materials using IR spectra, selecting band and identifying pigments in Raman spectra. The application of FuRES builds a collection of membership functions in the form of an inductive classification tree, for which each branch is a multivariate fuzzy rule that minimize the entropy of classification $H(C|A)$ which is obtained from:

$$H(C|A) = \sum_{j=1}^{m} p(a_j)H(C|a_j)$$  \hspace{1cm} (1-7)

$$H(C|a_j) = -\sum_{i=1}^{g} p(c_i|a_j) \ln p(c_i|a_j)$$  \hspace{1cm} (1-8)

where $H(C|a_j)$ is the entropy for a given character $a_j$ and can be obtained by summing the entropy for each class $c_i$. The entropies are totaled for all classes $g$. Because the classification model is fuzzy and soft, overlapping data can be accommodated without forcing the classifier into ill-conditioned
solutions. The structure of the FuRES classification tree reveals the inductive logic of the classifier. General rules are found at the root of the tree and precise rules are found at the leaves. Thus, the samples separated at the bottom branches are more similar compared to the samples separated by the other branches. The advantage of fuzzy classification trees over network classifiers, such as neural networks, is that they afford a simple inductive structure that is amenable to interpretation.

By testing the experimental data in the FuRES classification tree model, a confusion matrix containing the classification results can be constructed. In this matrix, each row designates the class membership, while the predicted classes are located in the columns. Therefore, the numbers in the matrix diagonal indicate correct classification, and the other matrix elements indicate misclassification. A 95% confidence interval is also calculated with each classification result by applying bootstrap Latin partitions (BLP). In the following chapters, confusion matrices with 95% confidence interval will be presented with all classification models.

1.4.3 Projected Difference Resolution

The projected difference resolution (PDR) metric was developed for quantifying the separation of classes in multivariate data space that is an approach easy to understand. It was used as a chemometric tool to compare the performances of GC/MS and GC–DMS in classifying ignitable liquids in Chapter 3. PDR has been successfully used to characterize and optimize the classification of bacteria from MALDI-MS data. The PDR
measure of class separation is obtained by first calculating the difference of two class means as given below:

\[ \overline{x}_{a,b} = \overline{x}_a - \overline{x}_b \]  

(1-9)
for which \( \overline{x} \) is a data object. For two-way data matrices such as GC/MS or GC–DMS, the objects are unfolded by taking each row of the matrix and concatenating the rows, so that a vector is obtained. This equation calculates a difference vector between the mean of two classes \( a \) and \( b \).

Next the objects in the two classes \( a \) and \( b \) are projected onto the difference vector according to equation (1-10):

\[ s_i = x_i \overline{x}_{a,b}^T \]  

(1-10)
for which \( s_i \) is the score of the \( i \)th object onto the difference vector. The resolution \( R_s \) is calculated from these projections, so for classes \( a \) and \( b \), \( R_s \) is calculated by equation (1-11):

\[ R_{s_{a,b}} = \frac{|s_a - s_b|}{2(\sigma_a + \sigma_b)} \]  

(1-11)
for which \( R_{s_{a,b}} \) is calculated by dividing the absolute difference of the averages of the projections \( s_a \) and \( s_b \) with twice the sum of the standard deviations \( \sigma_a \) and \( \sigma_b \) of the projections. This resolution is an analog to chromatographic resolution and values greater than 1.5 indicate that the classes are well resolved along the direction defined by the mean difference of the two classes. Resolution values are calculated for all combinations of pairs of classes. PDR can also be expressed as geometric averages, which are calculated using equation (1-12):
for which \( g \) is the number of classes. A resolution value between two classes indicates their difference in the data space. Classes with smaller resolution are more difficult to classify. Higher classification efficiency is expected from data sets with larger resolution values. In Chapter 3, the minimum resolution values calculated from all pairs of classes will be reported.

1.4.4 Alternating Least Squares

Alternating least squares (ALS) is a multivariate curve resolution method which is useful for modeling complex mixtures data because of the mathematical separation of overlapping peaks.\(^62-64\) ALS has been successfully applied in analyzing ion mobility spectra of chemical warfare agent stimulants,\(^65\) investigating the near infrared (NIR) spectra from the reaction of curing epoxy resins,\(^66\) and resolving infrared (IR) spectra.\(^67\) In Chapter 5, ALS was applied to model all IMS spectral datasets of drug metabolites and separate the reactant ion peaks from the product ion peaks.

Before the application of ALS to IMS dataset, the initial estimation of concentration profiles \( \mathbf{C} \) from a data matrix \( \mathbf{D} \) was obtained by using SIMPlE-to-use interactive self-modeling mixture analysis (SIMPLISMA), a self-modeling curve resolution method developed by Windig and Guilment,\(^68-73\) to find the pure variables in a dataset by modeling the data matrix \( \mathbf{D} \) as a product of the concentration profiles \( \mathbf{C} \) and the spectral intensity \( \mathbf{S} \) as in equation (1-13):

\[
Rs_i^n = g(g-1)/2 \sqrt{\sum_{j=1}^{g} R_{i,j}^{2}}
\]
\[
\mathbf{D} = \mathbf{CS}^T + \mathbf{E} \tag{1-13}
\]

Assuming each spectrum has \( n \) data points and there are \( p \) pure components in the sample, an \( m \times n \) data matrix \( \mathbf{D} \) is obtained after acquiring \( m \) spectra. Thus, \( \mathbf{C} \) is an \( m \times p \) matrix, and \( \mathbf{S} \) is a \( n \times p \) matrix. \( \mathbf{E} \) is an unmodeled residual error matrix with the same dimensions as \( \mathbf{D} \).

In IMS, a drift time is a pure variable in which the intensity is due to only one of the components of the sample mixture. Every pure variable has a unique concentration profile and a relative large variance with respect to intensity in a dynamic system. The corresponding concentrations must vary during the measurement to resolve different components. To find the pure variables, the first step of SIMPLISMA is to calculate the purity of each variable. One variable with higher purity indicates that it has a unique variance profile and large relative standard deviation with respect to intensity. The purity of the variable can be calculated from

\[
P_{ij} = \frac{\sigma_j}{\mu_j + \alpha} \times w_{ij} \tag{1-14}
\]

for which \( i \) and \( j \) represent the component and original variable (drift times), respectively. The mean and the standard deviation of the \( j \)th candidate variable are indicated by \( \mu_j \) and \( \sigma_j \). To avoid selecting variables that characterize noise, a noise correction term \( \alpha \) is used. Generally, \( \alpha \) is given a value at 5% of the maximum peak intensity of the mean of the data matrix. The \( w_{ij} \) is a weighted term that characterizes the linear independence of the \( j \)th candidate variable with respect to pre-determined variables. After the pure variables are found, the initial concentration matrix \( \mathbf{C} \) can be found.
directly from the data matrix $D$, and the spectral matrix $S$ can be calculated by a least square method using the following equation

$$S = (C'C)^{-1}C'D'$$

(1-15)

for which $S$ is then normalized to unit vector length to remove the model ambiguity. Similarly, new concentration profiles $C$ are also generated by a least squares method using equation (1-16):

$$C = DS(S'S)^{-1}$$

(1-16)

As a result of least squares method, negative artifacts may arise in the concentration and spectral profile in SIMPLISMA solution, even though the data contains no negative features. In the following application of ALS algorithm, a non-negative constraint is applied to the model and profiles $C$ and spectra $S$ can be optimized iteratively until convergence or the maximum number of iteration is reached.

In Chapter 5, the goal of applying SIMPLISMA and ALS is to iteratively improve the least square fit of the model to the IMS dataset. The number of components in concentration profile can be adjusted depending on the complexity of the sample with each component having a corresponding spectrum. Therefore, the IMS spectrum of the target compound will be mathematically extracted. In Figure 1-6, an image of IMS spectral of cocaethylene is presented. Being processed with SIMPLISMA and ALS, the spectral profile and concentration profile corresponding to reactant ion and product ion were resolved in the model, as given in Figure 1-7. Qualitative information was indicated in the spectral profiles, from which the reduced ion
mobility can be calculated according to the drift times of reactant ion and analyte ion. The concentration profiles provided quantitative information. An obvious increase of the concentration of analyte ion could be observed when desorption started. As a result of extracting the profiles of target compound, the sensitivity and selectivity of the analytical method can be further enhanced.

Figure 1-6. IMS spectral plot of 1 µL cocaethylene at concentration of 1 mg/mL in acetonitrile solution. The spectra were not modeled with ALS.
Figure 1-7. ALS-modeled IMS spectra (top) and ALS extracted concentration profiles (bottom) of 1 µL of cocaethylene at concentration of 1 mg/mL in acetonitrile analyzed in positive ion mode. Reactant ion peak (—); analyte ion peaks (—).
Chapter 2  Forensic Application of GC–DMS with Two-Way Classification of Ignitable Liquids from Fire Debris

Adapted with permission from Lu, Y. and Harrington, P.B.; Analytical Chemistry, 2007, 79, 6752-6759.

Copyright 2007 American Chemical Society

2.1 Introduction

According to the statistics from the U.S. Fire Administration, arson is the leading cause of fires and second leading cause of deaths and injuries, which results in lost lives and property. As a result, arson investigation is of forensic significance for federal government, the fire service, and the criminal justice system. A variety of ignitable liquids may be used as accelerants by criminals to start a fire. Ignitable liquids are typically commercially available fuels or solvents that are mixtures consisting of hundreds of components. Identification is difficult because ignitable liquids may comprise the same or similar components but with different distributions of concentration. Identification is further complicated by combustion because volatile components may have evaporated from the liquid and additional pyrolysis products may be produced that can alter the composition of the residual ignitable liquids detected in fire debris.

Gas chromatography–mass spectrometry (GC/MS) has proven to be an effective method for the characterization of ignitable liquids from suspected arson scenes. GC/MS is applied as the standard American Society for Testing and Materials (ASTM) method for ignitable liquid analysis. A large
database of ignitable liquids GC/MS data is available on the internet. In addition to GC/MS, a variety of methods have also been explored for ignitable liquids analysis to improve the resolution, such as heartcut multidimensional gas chromatography, tandem mass spectrometry (GC/MS/MS), Fourier transform ion cyclotron resonance (FT-ICR) high-resolution mass spectrometry, and two-dimensional gas chromatography (GC × GC). Although the modern analytical technology would provide the specificity needed for analyses or measurements of complex mixtures of hydrocarbon compounds, cost and maintenance have to be considered.

Differential mobility spectrometry (DMS) was accomplished with a Sionex microfabricated analyzer that characterizes analytes based on the difference between ion mobilities under high and low electric fields at ambient pressure. Detailed descriptions on DMS have been clearly presented in Chapter 1. In this applied DMS model, the asymmetric waveform is applied to two microplanar electrodes that are 25 mm long × 5 mm wide × 0.5 mm apart. The planar electrode geometry of DMS can separate both positive and negative ions simultaneously. Also, DMS has the potential to be used for onsite sensors or detectors for gas chromatography. The small size and the ratio of performance to cost compared with other bench top instruments make DMS a promising technique. A variety of applications have been investigated by DMS. The determination of nitro-organic explosives using DMS was achieved by modifying the drift gas with a small amount of methylene chloride. DMS also demonstrated ability in detecting
environmental contaminants, such as chlorocarbons in ground water\textsuperscript{40}, volatile organic compounds in ambient air\textsuperscript{41}, and DMS coupled with solid phase microextraction (SPME) was exploited to detect benzene, toluene, and xylene (BTX) in water.\textsuperscript{42} In view of the growing concern about biological terrorism, DMS also exhibited potential as a sensor for detecting trace amounts of bioorganisms.\textsuperscript{43, 44}

Multivariate statistical methods offer powerful tools for interpreting complex data. Generally, optimal chromatographic separation with baseline resolution is difficult to achieve for arson analysis because the large number of chemical components exceeds the peak capacities of gas chromatographic columns. Visual comparison of analytes with reference chromatograms is also limited by the skill and experience of the analyst. Chemometric methods are helpful in data interpretation, especially in the pattern recognition of unknown complex mixtures. Tan et al. performed soft independent model classification analogy and principal components analysis (PCA) on binned GC/MS chromatograms of five classes of petroleum-based ignitable liquids.\textsuperscript{76} Artificial neural networks (ANNs) are another powerful tool for pattern recognition. Doble et al. trained an ANN to recognize premium and regular gasoline from GC/MS total ion chromatograms and achieved 97\% classification accuracy of their samples.\textsuperscript{84}

In this work, DMS was utilized for classification of ignitable liquids from fire debris with FuRES. As an alternative ASTM method for the separation and concentration of ignitable liquids from fire debris samples,\textsuperscript{85-88} headspace
solid phase microextraction (SPME) was also applied as sampling method in this study. GC–DMS provides unique profiles for different ignitable liquids. Subject to FuRES classifier built on the entire two-way GC–DMS data objects, 99.07±0.04% of spiked ignitable liquids extracted from simulated fire debris could be correctly classified. Better prediction results can be obtained by using two-way GC–DMS data than only using one-way total ion chromatograms or integrated differential mobility spectra. Another more prevalent method partial least squares (PLS) was also used as a reference method for comparison with the FuRES classifier. The results demonstrated that the application of GC–DMS with FuRES on two-way data objects could be successfully used for forensic analysis of ignitable liquids from fire debris.

2.2 Experimental Section

2.2.1 Chemicals and Materials

Ignitable liquid samples were gasoline (BP), diesel (BP), mineral spirits (Klean Strip), paint thinner (Klean Strip), paint remover (OOPS!), lighter fuel (Ronsonol), and turpentine (Sunnyside). Two types of carpet with different fibers, nylon (Mannington) and polyester (Carpet One) were used as material matrices. Both samples and material matrices were obtained from local grocery store, carpet supply house, and gas station. SPME devices, 100 µm polydimethylsiloxane (PDMS) fiber, and the vials (4 mL) for neat sample extraction were purchased from Supelco Inc. (Bellefonte, PA). The septa-jar wide-mouth containers (125 mL) for extracting ignitable liquids from fire debris were obtained from Fisher Scientific (Fair Lawn, NJ).
2.2.2 Instruments

Experiments were performed using a HP 5890A gas chromatograph (Agilent Technologies, Palo Alto, CA, USA) interfaced to a differential mobility spectrometer (DMS) (Model SDP-1, Sionex Corporation, Bedford, MA). The GC was equipped with a HP splitless injector, which was operated in the splitless mode with a purge delay of 2 min. The helium carrier gas (99.99%) was maintained at a constant flow rate of 2.0 mL/min. GC was carried out on a 30 m × 0.25 mm × 0.25 µm capillary column, with a stationary phase of 5% diphenyl and 95% dimethyl polysiloxane cross-linked (Rtx®-5MS; Restek Corporation, Bellefonte, PA). The initial oven temperature was 40 °C which was held for 5 min. The temperature was increased from 40 °C to 120 °C at the rate of 3 °C/min, then from 120 °C to 250 °C at the rate of 30 °C/min and held at 250 °C for 1 min. GC and DMS were interfaced using stainless steel tubing and a Swagelok® (Solon, OH) tee union. The total transfer line was 14 cm. The GC capillary column penetrated 2.5 cm into the DMS inlet. The DMS detector was equipped with a UV photoionization source (10.6 eV). The drift gas was dry air that was heated in the stainless steel tubing to 80 °C at a flow rate of 250 mL/min. The DMS was operated at a separation voltage of 1300 V. Compensation voltages were scanned from -15 V to +6 V with an interval of 0.084 V. In this work, only positive ion mode data were reported because no signals were observed in the negative ion mode, which is expected for photoionization of hydrocarbons.
2.2.3 Procedure

New SPME fibers were thermally preconditioned for at least 30 min by insertion into the GC injector at 270 °C. A 20 µL aliquot of neat ignitable liquid sample was placed in a 4 mL amber vial and allowed to equilibrate for 30 min. The SPME fiber was then exposed to the headspace for 30 seconds. After extraction the fiber was immediately transferred to the GC injector and chromatographic analysis by GC–DMS was carried out. Desorption at 270 °C for 3 min occurred in the GC injector. The SPME fibers were conditioned between runs for 10 min at 270 °C. Five replicates of each sample were collected using a random block design.

In another experiment, 5.5 cm × 5.5 cm squares of carpet that had 200 µL of neat ignitable liquid added to them were burned for 3 different periods of time (30 s, 60 s, and 90 s). The burning carpets were extinguished by placing under a 500 mL inverted beaker. After a 5 min cooling time, the burned carpets with residual ignitable liquid were transferred to 125 mL septa-jar wide-mouth containers for further sampling, in which the SPME fiber was exposed to the headspace for 3 min at room temperature. Simulated fire debris was also obtained by adding 50 µL neat ignitable liquid samples to the burned nylon and polyester carpet squares before placement in the containers. Both spiked samples and burned samples were investigated in this work.
2.2.4 Data Processing

After GC–DMS analysis, a LabVIEW Version 7.0 (National Instruments, Austin, TX) virtual instrument software program was used to collect the data, which was saved as comma separated text. These files were converted from formatted text to Excel worksheets. The data was read by MATLAB 2006B Version 7.3 (The MathWorks Inc., Natick, MA) and stored as a data matrix in which the rows correspond to the GC retention times and the columns correspond to the DMS compensation voltages. Each component of the matrix was the positive ion intensity in volts for a given retention time and compensation voltage. Because the values for retention time and compensation voltage vary from run to run, linear interpolation was used to convert the data objects to the same retention time and compensation voltage values. The point spacings were 1.6 s for the retention time way and 0.084 V for the compensation voltage way during collection. After standardization using linear interpolation, the point spacings were 1.0 s for the retention time way and 0.10 V for the compensation voltage way.

Each DMS scan was baseline corrected by subtracting the average of points obtained from the 5 lowest (from -15.0 to -14.6 V) and 5 greatest compensation voltages (from 5.6 to 6.0 V) of all the points in each scan. Retention time alignment was implemented that used a second order polynomial to fit each two-way data object so that it maximized the correlation with the average of all the two-way data objects in the training data set, which consisted of neat ignitable liquid GC–DMS data. After
retention time alignment, each two-way data object was normalized to unit Euclidean length. To emphasize the characteristic peaks and sharpen peaks in the two-way data, the intensities of the two-way data were squared before baseline correction, normalization, alignment, visualization, and classification. This preprocessing step improved both the mathematical resolution and classification performance of the objects. The fire debris and prediction samples were always aligned to the average of the aligned neat liquids in the model building set.

The two-way data objects comprised 2,221 retention time and 211 compensation voltage measurements to furnish 468,631 points per object. Each object was unfolded into a vector representation. The principal component transform (PCT) was used to reduce the dimensionality of the vector from 468,631 data points to the number of objects before implementing the FuRES algorithm. This transformation does not affect the FuRES model but significantly reduces the time to build the model.

The PLS-2 algorithm was applied directly to the untransformed unfolded data objects. The PLS-2 models were favorably biased and present a best case for prediction. The models were constructed for the entire set of latent variables and the model with the lowest prediction error was always used. In both FuRES and PLS algorithms, the object classes were presented as binary encoded matrices (i.e., the rows correspond to the object and the columns to the class). The performance of the classifiers of both FuRES and PLS were validated using bootstrap Latin partitions (BLPs).
All evaluations were run with MATLAB without any specialized toolboxes, except that the FuRES algorithm used the “fminunc” function from the Optimization Toolbox 3.1 for rule construction. All evaluations were run on a home-built AMD Opteron 150 computer with 4 GB of 250 MHz DRAM. The CPU operated at 2.8 GHz (256 MHz × 11) and 1.55 V for the core. The DRAM was operated in dual channel model and synchronized (1:1) with the front side bus. The computer was operated under Microsoft Windows XP Pro x64 Edition SP1 (Redmond, WA).

2.3 Results and Discussion

2.3.1 GC–DMS of Neat Ignitable Liquids

Figure 2-1 (top) is a typical image of a GC–DMS object, in which the more intense peaks are shown with darker colors. The abscissa is retention time and the ordinate is the compensation voltage. The image from the squared preprocessing step is given below. One reason that this preprocessing step was successful is that the peaks from the burned samples tend to give less intense peaks. The squaring of the intensities is a nonlinear transform that will accentuate the more intense peaks in the data objects. Henceforth, the squared intensities of the two-way data are presented.
Figure 2-1. Topographic plot from GC–DMS characterization of neat gasoline. Sampling method: Headspace solid SPME. The top plot is the profile after data alignment and normalization. The bottom plot is the ion squared profile of the top plot.
Figure 2-2. Two-way GC-DMS positive ion data objects of SPME samples of gasoline, diesel, lighter fuel, and turpentine without data normalization.
Images of the two-way data objects from GC–DMS analysis of four ignitable liquids, gasoline, diesel, lighter fuel, and turpentine, are given in Figure 2-2. Diesel exhibits the lowest intensity range because it contains more compounds. Components eluting at the same time may correspond to different compensation voltages, which cannot be observed from chromatograms.

Figure 2-3 and Figure 2-4 give gas chromatographic and differential mobility spectral profiles of the four different ignitable liquids obtained by averaging across the retention time way or the compensation voltage way. This approach is used to reduce the two-way GC-DMS image into two one-way representations; an average chromatogram and an average spectrum.

For Figure 2-3 most of the chromatographic peaks occur between 5 and 30 min. For highly volatile ignitable liquids, like gasoline, many of the components elute before 10 min. As a heavy distillate product, diesel differs from gasoline with respect to the GC profile in that the retention times of the components are much longer than those for gasoline. For lighter fuel, the components of the mixture are distributed in a narrower molecular weight range than gasoline and diesel. Turpentine is an example of a distillate with fewer high concentration components than the other three liquids used to generate this figure.

Peaks in the average DMS spectra for the four examples in Figure 2-4 occur between -10 and 5 V. Most of the components in ignitable liquids are hydrocarbon compounds, which make their DMS profiles very similar.
Because the DMS spectra presented here are averaged spectra of all retention times, they do not correspond to a particular retention time. Differences in the patterns of the differential mobility spectra can be observed. The average DMS spectra of gasoline and turpentine look similar at first glance, but the most intensive position for gasoline is at -1 V compensation voltage and turpentine at 0 V. Considering the presence of some nonvolatile components in ignitable liquids, SPME sampling at room temperature may not obtain all components in the ignitable liquids, thus the results may not be an inclusive sample characterization. However, for the purpose of analysis under relatively simple conditions, characteristic profiles of each sample can be acquired from room temperature SPME GC–DMS.
Figure 2-3. Average ion chromatograms from GC–DMS analysis of neat sample liquids absorbed on SPME fibers.
Figure 2-4. Average differential mobility spectra from GC-DMS analysis of neat sample volatiles absorbed on SPME fibers.
2.3.2 GC–DMS of Burned Ignitable Liquids

Nylon and polyester carpets were selected as matrices in this work. Pyrolysates from these petroleum-based materials may contain identical chemical components in ignitable liquids, which will make the data analysis difficult. Another factor that will affect ignitable liquids classification is evaporation and combustion. Chromatographic peaks from highly volatile components may decrease after combustion or being present in high temperatures that arise in fires. Figure 2-5 compares the total ion gas chromatograms from GC–DMS of neat gasoline and burned gasoline on a nylon carpet sample. It can be seen that all the peaks before 10 min retention time are not apparent except for a peak at 8 min, which may be caused by incomplete combustion of this component of gasoline. The intensities of peaks between 13-15 min were also decreased, compared with the less volatile components eluting out afterwards. After 15 min the chromatograms are remarkably similar.
2.3.3 Principal Component Analysis

PCA was conducted on the training set, which was preprocessed from GC–DMS data of neat ignitable liquids extracted by SPME. The goal of performing PCA is to get a series of principal components, which account for the variance of dataset. Figure 2-6 is a PCA score plot of neat ignitable liquid samples, in which the first two principal components account for 41% of the
total variance of the original dataset. Gasoline (sample A) has the largest confidence interval in this figure, which suggests relatively lower reproducibility in the two-way GC–DMS profiles. This lack of reproducibility may be expected because gasoline comprises components with high volatility that may be lost during sampling. The training set of neat ignitable liquids was reevaluated after three months. Almost identical PCA results were obtained after three replicates of each sample were added to the original dataset. All samples collected from different matrices and treated with different methods were also visualized by PCA plot.

The effect of fire debris matrices can be seen from Figure 2-7. Ignitable liquids extracted from both nylon and polyester carpets deviated from those extracted from the neat samples. There is also an obvious difference between the two matrices themselves. Figure 2-8 gives the PCA score plot of different ignitable liquids extracted from three conditions: the neat samples, the spiked samples, and the burned samples. Different treatment methods, mainly varying in temperature and duration in a fire, can result in changes in comparative amount of volatile components, which makes the identification of ignitable liquids more difficult.
Figure 2-6. Principal component analysis score plot of two-way GC–DMS profiles of neat ignitable liquids. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine. Each letter represents a replicate of the ignitable liquid sample. A 95% confidence interval is drawn around the mean of each class.
Figure 2-7. Principal component analysis score plot of two-way GC–DMS profiles of seven samples extracted from different matrices. A. Neat samples, B. Simulated fire debris (Nylon carpet), C. Simulated fire debris (Polyester carpet). A 95% confidence interval is drawn around the mean of each class.
Figure 2-8. Principal component analysis score plot of two-way GC–DMS profiles of seven samples with different treatment methods. A. Neat samples, B. Spiked samples from simulated fire debris (burned nylon carpet), C. Burned samples (burned with nylon carpet). A 95% confidence interval is drawn around the mean of each class.
2.3.4 FuRES Classifier

The FuRES classification tree of seven neat samples is given in Figure 2-9. All samples were separated using this 6-rule tree with no splitting of the samples in the leaf nodes. Nonlinearities are accommodated by dividing the objects at each branch of the classification tree, which is a multivariate fuzzy rule. The depth of the tree encodes the similarities of the objects in the class, so that two classes separated at the bottom of the tree will be more similar than classes separated earlier. Besides being useful for classification, FuRES multivariate discriminants are also amenable for interpretation. Figure 2-10 (middle) gives the image of the two-way discriminant for rule one. Two-way averages were calculated from all the objects separated to the left and right of the tree by the #1 rule. This rule separating the ignitable liquids is primarily based on volatility. Paint remover and lighter fuel are classified to the right of the rule because most of their components are highly volatile with chromatographic peaks before 16 min retention time, as can be seen from their two-way image (Figure 2-10 bottom). The top image is the average GC–DMS profile of the other five ignitable liquids, which contain peaks eluting after 16 min. Features in the FuRES discriminant plots may also be used to find selective and specific peaks that can be identified and exploited as signature peaks for a specific ignitable liquid.
Figure 2-9. FuRES classification tree for seven neat ignitable liquid samples. Nc is the number of samples and H is the absolute fuzzy entropy of classification.
Figure 2-10. Image of the weights for the two-way rule #1 (middle) with negative features characteristic of classes to left of the tree (top) and positive features characteristic of classes to right of the tree (bottom).
To evaluate the performance of FuRES classifier, ten BLPs were applied so that the neat sample data sets were split into 75% training and 25% prediction. The Latin partitions are randomly divided to furnish 4 training-prediction set pairs so that each sample is used only once for prediction and the proportions of the class distributions remain fixed (i.e., 1/7 for training and prediction).

Confusion matrices with 95% confidence interval obtained from the FuRES classification of ignitable liquids burned with nylon carpet are given in Table 2-1. The confidence intervals were calculated from the 40 BLP results for the fire debris samples. Each column in the matrix is the predicted class of ignitable liquid and each row is the actual class of the liquid that was added before igniting the nylon carpet. Therefore, correct classification occurs along the diagonal of the confusion matrix, while the off-diagonal elements correspond to misclassifications.

PLS was also applied to the Latin partitions in parallel with FuRES. The numbers of latent variables were selected to yield the minimum error for the prediction sets. The confusion matrices obtained from this optimized PLS are given in Table 2-2. Burned lighter fuel is always misclassified for PLS and FuRES because most of its components are highly volatile and easily lost during combustion. When applying classifier constructed from the neat liquids, dissimilarity in the two-way profiles would cause misclassification.
Table 2-1. Confusion matrix with 95% confidence interval for FuRES analysis of ignitable liquids burned with nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP FuRES models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0.1±0.1</td>
<td>1.0±0.3</td>
<td>1.5±0.5</td>
<td>0</td>
<td>0</td>
<td>0.4±0.3</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 2-2. Confusion matrix with 95% confidence interval for optimized PLS analysis of ignitable liquids burned with nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP PLS models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2.98±0.05</td>
<td>0</td>
<td>0</td>
<td>0.03±0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>1.1±0.2</td>
<td>0</td>
<td>1.9±0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0.1±0.1</td>
<td>0</td>
<td>0</td>
<td>2.9±0.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1.1±0.1</td>
<td>0</td>
<td>1.9±0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>2.9±0.1</td>
<td>0.1±0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

The prediction results obtained from optimized PLS and FuRES are summarized in Table 2-3. For PLS and FuRES analysis of the neat samples, a classification accuracy of 100% was consistently achieved. The two-way classification accuracies for samples from simulated fire debris were 99.07±0.04% (FuRES) and 98.00±0.04% (PLS), which were 10% and 7% higher than the prediction results for one-way TIC data; 47% and 53% higher than the prediction results for one-way DMS data. Because of the great loss of volatile components during combustion, FuRES analysis of ignitable liquids extracted from the burned samples showed a lower classification accuracy of 82.3±0.1% for the two-way data, which showed
3% and 27% improvements compared with the prediction results for TIC and DMS data, respectively. It is also observed that the prediction results from FuRES were not statistically better than those from PLS. However, PLS was positively biased in that the number of latent variables was selected for each model that yielded the lowest prediction error. FuRES is self-optimizing, so the prediction data is never used to adjust the FuRES model, thus its predictions were unbiased.

2.4 Conclusions

GC–DMS can provide characteristic two-way data profiles for different petroleum-based ignitable liquids. Squaring the intensities of the GC–DMS data sharpens peaks and simplifies the images, which helps improve the resolution among the classes of ignitable liquids. FuRES is a powerful chemometric tool for classifying intricate and overwhelming datasets. FuRES models constructed with the two-way GC neat ignitable liquids data identified ignitable liquids spiked to burnt carpet with 99.07±0.04% accuracy. The results were also encouraging for burned samples with 82.3±0.1% classification accuracy. Compared with one-way total ion chromatograms, two-way GC–DMS data provides more information for classification and better prediction results can be obtained than using either chromatograms or DMS spectra. Combining the application of the smaller size and lower cost GC–DMS with robust FuRES, this study presented a promising onsite
detection method for arson investigation. Future work will investigate an integrated micro GC–DMS system.

Table 2-3. Comparison of prediction results of TIC, DMS, and two-way data confusion matrices obtained from optimized PLS and FuRES using 10 bootstraps of four Latin partitions training-prediction pairs of the neat samples. The fire debris samples on the polyester and nylon carpets were classified with the 40 resulting models from the BLP evaluation.

<table>
<thead>
<tr>
<th>Experiment Number of Evaluations</th>
<th>Neat Sample 10</th>
<th>Spiked Sample 40</th>
<th>Burned Sample 40</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PLS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIC</td>
<td>99.7±0.6%</td>
<td>91.1±0.1%</td>
<td>75.2±0.1%</td>
</tr>
<tr>
<td>DMS</td>
<td>99.7±0.6%</td>
<td>44.93±0.04%</td>
<td>50.5±0.2%</td>
</tr>
<tr>
<td>Two-way Data</td>
<td>100%</td>
<td>98.00±0.04%</td>
<td>78.33±0.04%</td>
</tr>
<tr>
<td><strong>FuRES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIC</td>
<td>100%</td>
<td>89.1±0.1%</td>
<td>79.0±0.1%</td>
</tr>
<tr>
<td>DMS</td>
<td>96±1%</td>
<td>51.71±0.06%</td>
<td>45.30±0.01%</td>
</tr>
<tr>
<td>Two-way Data</td>
<td>100%</td>
<td>99.07±0.04%</td>
<td>82.3±0.1%</td>
</tr>
</tbody>
</table>
Chapter 3  Comparison of Differential Mobility Spectrometry and Mass Spectrometry for Gas Chromatographic Detection of Ignitable Liquids from Fire Debris Using Projected Difference Resolution

Adapted with permission from Lu, Y., Chen, P. and Harrington, P.B.; Analytical and Bioanalytical Chemistry, 2009, 394, 2061-2067.

Copyright 2009 Springer-Verlag

3.1  Introduction

Identification of ignitable liquids from fire debris is of forensic importance for arson investigation. As a standard method of the American Society for Testing and Materials (ASTM) for ignitable liquid analysis, gas chromatography/mass spectrometry (GC/MS) has been widely used by forensic chemists in arson analysis. With the development of new analytical technologies, a variety of methods have also been explored for ignitable liquid analysis. In Chapter 2, the application of GC–DMS for classification of ignitable liquids from fire debris was investigated. DMS characterizes an analyte based on its change in ion mobility under high and low electric fields. As an alternative detector for GC, DMS shows promise for on-site analysis because it is small, has low cost to purchase and operate, and is sensitive. In addition to analyzing ignitable liquids, it has also been applied in many other areas, such as the detection of explosives, environmental contaminants, and bioorganisms.

To statistically compare the classification efficiency of differential mobility spectrometry with that of mass spectrometry for chromatographic
detection of ignitable liquids from fire debris, projected difference resolution (PDR)\textsuperscript{60} is applied as the criteria to evaluate the separation for all pairs of classes of the ignitable liquids. The performances of two-way and one-way data in classifying ignitable liquids from fire debris were also predicted using the same criterion. PDR has been successfully used to characterize and optimize the classification of bacteria from MALDI-MS data.\textsuperscript{61} The esolutions between two classes indicate the class separation. Small values for resolution indicate overlapping classes. The minimum resolution values calculated from different datasets among all pairs of classes were used to compare the performances of GC/MS and GC–DMS in classifying ignitable liquids.

A set of seven ignitable liquids extracted by solid phase microextraction (SPME) from fire debris were analyzed by both GC/MS and GC–DMS. The performances of GC/MS and GC–DMS were statistically compared using PDR, a novel chemometric tool for calculating resolutions among groups of multivariate data objects. The prediction accuracies based on FuRES\textsuperscript{52} models exhibited consistent results with PDR, which demonstrated the feasibility of predicting the classification results and comparing the performances of different analysis methods for pattern recognition by calculating PDR values among several classes of analytes.
3.2 Experimental Section

3.2.1 Materials and Chemicals

Solid-phase microextraction (SPME) devices, 100 µm polydimethylsiloxane (PDMS) fiber, and the vials (4 mL) for neat sample extraction were purchased from Supelco Inc. (Bellefonte, PA). The fiber was conditioned as recommended by the manufacturer in the GC injector at 270 °C for 1 h. The septa-jar wide-mouth containers (125 mL) for extracting ignitable liquids from fire debris were obtained from Fisher Scientific (Fair Lawn, NJ). The same ignitable liquid samples and carpet samples as those used in Chapter 2 were also applied in this project.

3.2.2 Instruments

GC/MS analysis was achieved on an HP 5890A gas chromatograph coupled with an HP 5970 mass selective detector (Agilent Technologies, Palo Alto, CA, USA). GC was carried out on a 30 m × 0.25 mm × 0.25 µm Rtx®-5MS capillary column (Restek Corporation, Bellefonte, PA). Ultrapure helium as the carrier gas was maintained at a constant flow rate of 2.0 mL/min. The initial oven temperature was 40 °C which was held for 5 min. The temperature was increased from 40 °C to 120 °C at the rate of 3 °C/min, then from 120 °C to 250 °C at the rate of 30 °C/min and held at 250 °C for 1 min. Splitless SPME injection was performed at 270 °C. The mass spectrometer transfer line was maintained at 280 °C. Mass spectra were scanned from 45 to 200 m/z.
For GC–DMS analysis of ignitable liquids, the same chromatographic condition as GC/MS analysis was also utilized. The differential mobility spectrometer (Model SDP-1, Sionex Corporation, Bedford, MA) was coupled to the GC using a segment of stainless steel tubing and a Swagelok® (Solon, OH) tee union. The total transfer line length was 14 cm. The GC capillary column penetrated 2.5 cm into the DMS inlet. The DMS detector was equipped with a UV photoionization source (10.6 eV). The drift gas, which was dry air at a flow rate of 250 mL/min, passed through heated stainless steel tubing to warm the drift tube to 80 °C. The DMS was operated at a separation voltage of 1300 V. Compensation voltages were scanned from -15 V to +6 V. In this work, only positive ion mode data were reported because no peaks were observed in the negative ion mode.

### 3.2.3 Neat Sample Collection

A 20 µL aliquot of neat ignitable liquid sample was placed into a 4 mL amber vial followed by a 30 min equilibration. A SPME fiber was exposed to the headspace for 30 s adsorption and then transferred immediately to the GC injector, where desorption was kept for 3 min. The SPME fiber was conditioned between runs for 10 min at 270 °C. A random block design was applied for collecting data of five replicates of each sample. The same sampling method was performed on the same sample for both GC/MS and GC–DMS analyses.
3.2.4 Matrix Exposed Sample Preparation and Collection

To simulate fire debris samples, 5.5 cm × 5.5 cm square-shaped pieces of nylon and polyester carpets were burned and had a 50-µL aliquot of neat ignitable liquid samples randomly distributed across the surface carpet. The carpet squares then equilibrated for 5 min in the air. The burnt carpet square was placed into a 125 mL wide-mouth septa-jar. SPME sampling was accomplished by inserting the SPME needle through the septum of the jar where 3 min was given for equilibration at room temperature.

In another experiment, a 200 µL aliquot of one of the neat ignitable liquids was added randomly across the surface of the same size carpet square which was ignited and burned for three different periods of time (30, 60, and 90 s) on a ceramic evaporating dish. The burning carpet was extinguished by placing under an inverted 500 mL beaker. The burned carpets with residual ignitable liquids were transferred to the container after a 5 min cooling time. Three replicates were collected from both spiked and burned samples. The combusted samples characterize the weathering process that may occur in a fire or over time.

3.3 Data Processing

GC/MS data were read by MATLAB 7.3 (The MathWorks Inc., MA) and stored as a data matrix in which the rows correspond to the GC retention time and the columns correspond to the MS mass-to-charge ratio. Linear interpolation was used to convert the data matrix to the same retention time and mass-to-charge ratios. The two-way GC/MS data objects comprised
2221 retention times and 156 mass-to-charge ratios to furnish 346 476 points. Each object was unfolded into a vector representation and normalized to unit Euclidean length. The principal component transform (PCT) was used to reduce the dimensionality before implementing the FuRES algorithm, in which the object classes were presented as binary encoded matrices (i.e., the rows correspond to the object and the columns for the class). The performance of FuRES was validated using bootstrap Latin partitions (BLP). Processing method of GC–DMS data was the same as the one described Chapter 2, except for that the intensities of the two-way data matrices were not squared used when compared with the GC/MS data.

All evaluations were run with MATLAB 7.3 without any specialized toolboxes, except that the FuRES algorithm used the “fminunc” (function minimization unconstrained) function from the Optimization Toolbox 3.1. This function is basically a conjugate gradient optimization in MATLAB, the nonlinear simplex is used to find a good initial weight vector for the fminunc which helps avoid local minima.

3.4 Results and Discussion

3.4.1 Two-way GC/MS and GC–DMS of Ignitable Liquids

In one-way data, such as total ion chromatograms and integrated spectra, intensity is plotted with respect to retention time or mass-to-charge ratio. Two-way data objects can be visualized in images that combine information from each order or way, for which intense peaks are displayed with dark color, as given in Figure 3-1. The top image presents the two-way
GC/MS data of neat gasoline, which was extracted by headspace SPME. The horizontal axis corresponds to retention time and the vertical axis corresponds to mass-to-charge ratio. The bottom image in this figure is the two-way data profile of neat gasoline sample obtained from GC–DMS, where the axes correspond to retention time and compensation voltage, respectively. Gasoline contains many highly volatile components that elute during the first 15 min with high intensity. Under electron ionization, most of the fragment ions reaching the mass spectrometer have small mass-to-charge ratios between 50 and 110 Th. Within this range, chromatographic peaks at different retention times may have similar fragmentations, which actually do not contribute much useful information in differentiating the chromatographic peaks. The ion source of DMS applied in this work is photoionization. The energy is 10.6 eV, which is less than the energy of electron ionization of 70 eV. The two-way GC–DMS profile presents the compensation voltages of the most intense peaks that range from -5 to 0 V, most of which correspond to highly volatile components with shorter retention times. From these two-way profiles, it is apparent that the peaks at the same retention time may contain peaks at different mass-to-charge ratios or compensation voltages. Thus, the conventional classification approach of using total ion chromatograms omits selective information from the spectral orders.
Figure 3-1. Two-way data profiles of GC/MS (top) and GC–DMS (bottom) analysis of neat gasoline adsorbed on SPME fibers.
3.4.2 Projected Difference Resolutions

PDR provides a measure of interclass resolution between pairs of classes. PDR was used to compare the performances of GC/MS and GC–DMS in differentiating the seven classes of ignitable liquids. Each measurement generated a two-way data matrix $\mathbf{x}$, which was taken into the equations mentioned in the introduction. The minimum average PDR values with 95% confidence interval calculated from ten bootstraps for the seven neat ignitable liquid samples are given in Table 3-1. The minimum value is the smallest resolution among the 21 combinations of pairs of classes. In addition to two-way data, PDR values were also calculated from one-way total ion chromatograms (TIC) and integrated spectral data. The results show that one-way MS data have higher resolutions than DMS data, while one-way DMS chromatograms have higher resolutions than the MS chromatograms. This difference can be attributed to the fact that DMS photoionization yields more reproducible ion currents for hydrocarbons than MS electron ionization. The abundant fragment ions provide enough information for classifying different ignitable liquids by using integrated mass spectra. Although the DMS alone could not provide as much information as MS for classification, it helps to improve the resolution among the classes for two-way data. The resolutions calculated from two-way and one-way neat sample datasets are all above 1.5, indicating a classification with 100% accuracy should be obtainable. The higher resolution value between two classes, the easier it is to differentiate them from each other. Comparing the
minimum resolution values calculated from different datasets, smaller classification error can be expected from two-way GC/MS and GC–DMS data than one-way TIC or integrated spectra. For one-way datasets, mass spectra data could have higher classification efficiency than DMS data, while DMS chromatographic data could provide more information than MS chromatographic data in separating different ignitable liquids.

Table 3-1. Minimum projected difference resolution values with 95% confidence interval, calculated from neat sample datasets.

<table>
<thead>
<tr>
<th></th>
<th>GC/MS</th>
<th>GC–DMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated Spectra</td>
<td>5.6±0.2</td>
<td>2.9±0.1</td>
</tr>
<tr>
<td>Total Ion Chromatogram</td>
<td>3.9±0.2</td>
<td>5.1±0.2</td>
</tr>
<tr>
<td>Two-way</td>
<td>6.3±0.5</td>
<td>5.4±0.2</td>
</tr>
</tbody>
</table>

3.4.3 FuRES Analysis

To verify the prediction based on minimum PDR values, FuRES was applied to classify seven ignitable liquids from fire debris. Two FuRES classification tree models built from two-way GC/MS and GC–DMS neat sample datasets are given in Figure 3-2. There are six rules in each tree for differentiating the seven samples. Based on the separations and absolute fuzzy entropy values (H), it is found there is no difference between the first four rules of the two classification models. From the top of the tree to the bottom, the rules become more precise and the entropy values smaller.
Figure 3-2. FuRES classification trees for seven neat ignitable liquid samples on two-way GC/MS (top) and GC–DMS (bottom) data with a 95% classification rate. Nc is the number of replicates and H is the absolute fuzzy entropy of classification.
Confusion matrices with 95% confidence intervals obtained from the FuRES analysis of two-way GC/MS and GC–DMS data of ignitable liquids burned with nylon and polyester carpets are given in Table 3-2 and Table 3-3, respectively. The confidence intervals were calculated from 10 BLPs of the burned samples. The prediction class was assigned in each column of the matrix and the actual class listed in each row, thus numbers at confusion matrix diagonal indicate correct classification. The prediction error is caused by the evaporation of volatile components and production of pyrolyzates during the fire. Especially for ignitable liquids containing large amount of highly volatile hydrocarbon compounds, like gasoline, it is difficult to identify the liquids after combustion. The results listed in Table 3-4 showed that there is no error for classifying neat samples when using either two-way GC/MS or GC–DMS data. But for classifying spiked and burned debris samples, there is less error when applying two-way GC/MS data to build the FuRES model.
Table 3-2. Confusion matrix with a 95% confidence interval for FuRES analysis of two-way GC/MS data of burned samples, nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP FuRES models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>0.8±0.1</td>
<td>2.2±0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>2.98±0.05</td>
<td>0</td>
<td>0.02±0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2.2±0.3</td>
<td>0</td>
<td>0.8±0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>2.98±0.05</td>
<td>0</td>
<td>0.02±0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 3-3. Confusion matrix with a 95% confidence interval for FuRES analysis of two-way GC-DMS data of burned samples, nylon carpet (top) and polyester carpet (bottom). The predictions were obtained from 40 BLP FuRES models. A. gasoline, B. diesel, C. mineral spirits, D. paint thinner, E. paint remover, F. lighter fuel, G. turpentine.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>1.8±0.2</td>
<td>0</td>
<td>1.2±0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>0.05±0.08</td>
<td>0</td>
<td>0</td>
<td>2.95±0.08</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>2.1±0.3</td>
<td>0</td>
<td>0.9±0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
It was found that the PDR values for any neat sample data set are always higher than 1.5, which is considered as a limit for baseline separation in chromatographic analysis. The evaluation of two-way GC/MS data showed higher resolution values, but both MS and DMS are effective gas chromatographic detectors for neat ignitable liquid analysis and provided 100% classification accuracy.

Table 3-4. Two-way prediction errors by FuRES analysis using 10 bootstraps of four Latin partitions training and prediction pairs of neat samples, spiked samples, and burned samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Data</th>
<th>GC/MS</th>
<th>GC–DMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neat</td>
<td>Total Ion Chromatogram</td>
<td>(0.2±0.3)/35</td>
<td>0/35</td>
</tr>
<tr>
<td></td>
<td>Integrated Spectra</td>
<td>0/35</td>
<td>(1.5±0.6)/35</td>
</tr>
<tr>
<td></td>
<td>Two-way Data</td>
<td>0/35</td>
<td>0/35</td>
</tr>
<tr>
<td>Spiked</td>
<td>Total Ion Chromatogram</td>
<td>(4.4±1.0)/21</td>
<td>(2.6±0.3)/21</td>
</tr>
<tr>
<td></td>
<td>Integrated Spectra</td>
<td>(3.0±0.2)/21</td>
<td>(9.2±0.2)/21</td>
</tr>
<tr>
<td></td>
<td>Two-way Data</td>
<td>(0.32±0.04)/21</td>
<td>(0.33±0.06)/21</td>
</tr>
<tr>
<td>Burned</td>
<td>Total Ion Chromatogram</td>
<td>(5.9±0.5)/21</td>
<td>(3.5±0.3)/21</td>
</tr>
<tr>
<td></td>
<td>Integrated Spectra</td>
<td>(0.9±0.1)/21</td>
<td>(12.0±0.1)/21</td>
</tr>
<tr>
<td></td>
<td>Two-way Data</td>
<td>(0.8±0.2)/21</td>
<td>(3.9±0.1)/21</td>
</tr>
</tbody>
</table>
3.5 Conclusions

The performances of GC–DMS and GC/MS for classifying ignitable liquids were compared using the PDR measurement. The minimum resolution values calculated from GC/MS and GC–DMS datasets indicated that one-way MS data exhibited higher resolution than DMS data, while DMS chromatograms are more informative in differentiating ignitable liquids than MS chromatograms. The two-way datasets always have higher resolution than any one-way datasets for both detection methods. The prediction errors calculated from FuRES analysis of both neat sample and fire debris sample datasets are consistent the PDR results, which demonstrates that using two-way data is more efficient in classifying different ignitable liquids, and GC/MS could have smaller prediction errors than GC–DMS. Considering the high resolution values of these two methods, there is no statistic difference between two-way GC/MS and GC–DMS data in classifying neat ignitable liquid samples and spiked samples. However, a smaller prediction error was achieved with burned samples when using two-way GC/MS data, in that evaporation of volatile components and produce of pyrolyzates during the fire process complicated the analysis system, thus method with higher PDR values would be more effective.
Chapter 4  Classification of Bacteria by Simultaneous Methylation—Solid Phase Microextraction and Gas Chromatography/Mass Spectrometry Analysis of Fatty Acid Methyl Esters

4.1  Introduction

The U.S. Centers for Disease Control and Prevention (CDC) estimates that around 76 million people a year in the U.S. infect foodborne diseases. Although the large majority of such cases are mild and cause symptoms for only a day or two, some cases are more serious. CDC estimated that there are about 325,000 hospitalizations and 5,000 deaths related to foodborne diseases each year. Harmful pathogenic bacteria are the most common causes of foodborne diseases. Because most of the bacteria that cause such diseases have no odor and generate no change in the color or the texture of the food, they always go undetected. Therefore it would be quite easy to cause an epidemic in a large population when major food consumed in the U.S. is contaminated. From homeland security point of view, biological weapons that may result mass destruction have been rated as the greatest terrorism threat. Thus, researches on developing fast bacterial identification methods have attracted more and more attentions.

Lipid profiling has been considered as a viable approach for detecting bioorganisms. The types and the relative abundances of bacterial lipids provide plenty information for bacterial classification. The most common method in the classification of bacteria through lipid profiling is the analysis
of bacterial fatty acid methyl esters (FAMEs).\textsuperscript{93-96} Commericially available system from Microbial Identification Inc. (MIDI) has been successfully used for bacterial identification. This gas chromatography (GC)-based analytical method includes a 60-min bacterial lipid saponification/methylation/extraction step followed with a 15-20-min GC separation. Bacteria are identified by comparing the resulting chromatograms of FAME profiles with an existing library database. Labs using GC to identify bacteria are often faced with running large number of samples under strict time constraints. Buyer has demonstrated the reduction in chromatographic retention times by using fast GC analysis, which used a smaller column giving a threefold reduction in run time with no loss of resolution.\textsuperscript{97} An inserted oven was also applied for the completion of faster oven heating rate.\textsuperscript{98} In addition to using fast GC analysis to shorten the sample analysis time, the sample preparation process using conventional derivatization and liquid-liquid extraction (LLE) is the most time-consuming step and needs more efforts on developing fast and effective approaches.

An alternative technique to derivatize whole-cell bacteria is quaternary ammonium hydroxide compound-induced thermally assisted hydrolysis and methylation (THM), which derivatized the fatty acids contained in the bacterial lipids into their corresponding methyl esters. This method avoids applying multistep LLE to saponify and derivatize bacterial lipids, which greatly shortens the sample preparation time. THM has been used with GC to study the bacterial fatty acid patterns.\textsuperscript{99} TMH-mass spectrometry (MS)
technique enabled fast determination of acyl groups in bacterial lipids with the application of TMAH as an organic alkali.\textsuperscript{100, 101} Thermal desorption ion mobility spectrometry (IMS) has also been used to differentiate bacteria with TMAH as a derivatization reagent.\textsuperscript{102} Using supercritical fluid carbon dioxide with phenyltrimethylammonium hydroxide (PTMAH) to prepare FAMEs from whole-cell bacteria, Gharaibeh demonstrated a bacterial classification method based on statistical analysis of the GC/MS profiles of bacterial FAMEs.\textsuperscript{103}

Solid phase microextraction (SPME) is an innovative, solvent free sample preparation technology that is fast, economical, and versatile. It uses a fused silica fiber coated with a polymer to extract analytes from liquid, gas, and solid matrices. Coupled with GC, SPME has been applied to the analysis of environmental, food, and beverage samples, etc.\textsuperscript{10} Trace analysis of fatty acids in liquid and gas samples has been investigated by coupling SPME with derivatization of the target analytes to more volatile species prior to GC analysis.\textsuperscript{104, 105} Liu used SPME and on-line methylation GC to analyze aliphatic carboxylic acids.\textsuperscript{106} SPME was also utilized to directly preconcentrate volatile compounds from bacteria\textsuperscript{107} and bacteria-infected food\textsuperscript{108} for GC analysis.

Chemometric methods provide powerful tools in analyzing overwhelming and complicated data. The wealth of information provided by the GC/MS spectra of extracted bacterial FAMEs contains abundant signals in both total ion chromatogram and averaged mass spectra. To visualize the difference between spectra, find characteristic peaks for each sample, or
build classification model to differentiate samples of interest, a variety of chemometric methods have been used to analyze chromatographic and mass spectrometric data, such as principal component analysis (PCA), analysis of variance-principal analysis (ANOVA-PCA), partial least square (PLS), hierarchical analysis, and fuzzy rule-building expert system (FuRES), etc. The loadings of the principal components in PCA and PLS models provide invaluable information for detecting biomarker peaks which are significant in indicating the chemical differences between different subjects. 

In this work, headspace SPME was applied with TMAH for the simultaneous saponification/derivatization/extraction of whole-cell bacterial lipids in the form of FAMEs, which were directly analyzed by GC/MS. Classification and differentiation of bacteria were achieved by chemometric analysis of the GC/MS profiles of bacterial FAMEs. Based on the GC/MS data, an 87±3% correct classification rate was obtained using bootstrap Latin partition (BLP) cross validation with FuRES analysis.

4.2 Experimental Section

4.2.1 Bacterial Agents and Growth Conditions

Freeze-dried stock cultures of bacteria were purchased from the American Type Culture Collection (ATCC, Manassas, VA). Six bacterial agents were analyzed in this study: Enterococcus faecalis (ATCC 19433), Pseudomonas aeruginosa (ATCC 14207), Escherichia coli O157:H7 (ATCC 43895), Salmonella enterica (ATCC 13314), Vibro parahaemolyticus (ATCC 17802), and Listeria innocua (ATCC 51742). Growth condition of the six
bacteria is listed in Table 4-1. Brain-heart infusion broth (BHIB) and nutrient broth (NB) were purchased from Difco™ (Sparks, MD). Nutrient agar (NA), tryptic soy agar (TSA), tryptic soy broth (TSB), and brain-heart infusion agar (BHIA) were purchased from Sigma-Aldrich (St. Louis, MO). The received bacterial pellets were dissolved in 5 mL of TSB, NB, or BHIB, which were previously autoclaved at 121 °C for 15 min. All bacteria were rehydrated in the infusion at 36±1 °C for 24 h. Several drops of cloudy bacterial suspension were used to inoculate on previously autoclaved agar plates, on which the bacteria were then incubated at 36±1 °C for another 24 h before analysis.

Table 4-1. Bacteria investigated in this study.

<table>
<thead>
<tr>
<th>Microorganism</th>
<th>ATCC® No.</th>
<th>Re-hydrating Broth</th>
<th>Culture Agar</th>
<th>Gram Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Enterococcus faecalis</td>
<td>19433</td>
<td>BHIB</td>
<td>BHIA</td>
<td>+</td>
</tr>
<tr>
<td>B Pseudomonas aeruginosa</td>
<td>14207</td>
<td>BHIB</td>
<td>BHIA</td>
<td>+</td>
</tr>
<tr>
<td>C Escherichia coli O157:H7</td>
<td>43895</td>
<td>TSB</td>
<td>TSA</td>
<td>−</td>
</tr>
<tr>
<td>D Salmonella enterica</td>
<td>13314</td>
<td>NB</td>
<td>NA</td>
<td>−</td>
</tr>
<tr>
<td>E Vibrio parahaemolyticus</td>
<td>17802</td>
<td>BHIB</td>
<td>BHIA</td>
<td>−</td>
</tr>
<tr>
<td>F Listeria innocua</td>
<td>51742</td>
<td>BHIB</td>
<td>BHIA</td>
<td>+</td>
</tr>
</tbody>
</table>

BHIB/A = Brain Heart Infusion Broth/Agar  
NB/A = Nutrient Broth/Agar  
TSB/A = Tryptic Soy Broth/Agar
4.2.2 Instrumentation and Apparatus

GC/MS analysis was performed on an HP 5890A gas chromatograph coupled with an HP 5970 mass selective detector (Agilent Technologies, Palo Alto, CA, USA). A 30 m × 0.25 mm × 0.25 µm Rtx®-5MS capillary column (Restek Corporation, Bellefonte, PA) was used in this work. Ultrapure helium as the carrier gas was maintained at a constant flow rate of 1.0 mL/min. The initial oven temperature was 140 °C and then increased to 230 °C at the rate of 5 °C/min, then increased to 260 °C at the rate of 30 °C/min and held at 260 °C for 2 min. Splitless SPME injection was performed at 270 °C. The MS was run under electron ionization (EI) mode with electron energy of 70 eV. The mass spectrometer transfer line was maintained at 280 °C. Mass spectra were scanned from 45 to 360 m/z.

The SPME devices were purchased from Supelco Inc. (Bellefont, PA). A 100 µm polydimethylsiloxane (PDMS) fiber was used in this work. The fiber was conditioned as recommended by the manufacturer in the GC injector at 270 °C for 1 h before first use.

During extraction, samples were heated in a thermowell placed on Optitherm™ system from Chemglass (Vineland, NJ). An electronic contact thermometer (CG-1995-01 Opti-Chem®) was inserted into the thermowell for monitoring reaction temperature.

4.2.3 Sample Preparation

Using a sterilized inoculating loop, about 10 mg of bacterial cells were harvested from the agar plates and transferred to a 4 mL amber glass vial,
which has screw top and a hole-cap with PTFE/Silicone septa. Ten microliters of 0.1 M TMAH methanol solution was added onto the bacterial samples, and then capped vials were inserted into the thermowell for 10 minutes at 80 °C. Meanwhile, the SPME fiber was exposed to the sample headspace.

To compare the derivatization results, bacterial FAMEs were also prepared using traditional methylation method. The same amount of bacterial cells was transferred to an 8 mL amber glass vial. One milliliter of 3.75 N NaOH (1:1, methanol: distilled water) was added to each vial containing the bacterial cells to saponify the bacterial lipids. The vials were capped, briefly vortexed and heated in thermowell for 5 min, and then the vials were vigorously vortexed for 5-10 s, heated for an additional 25 min. The cooled vials were uncapped and 2 mL of 3.25 N HCl (1:1.18, methanol: 6 N HCl) were added for methylating the fatty acids. The vials were capped and briefly vortexed, heated for 10 min at 80 °C afterwards. After the vials were cooled down, the FAMEs were extracted by the addition of 1.25 mL of 1:1 hexane :methyl tert-butyl ether with 10 min of gentle tumbling. The lower phase (aqueous) was pipette off, and the upper phase (organic) was washed with 3 mL of 0.3 N NaOH, tumbling for 5 min. The organic phase was then ready for GC analysis.

**4.2.4 Data Analysis**

All calculations were performed using in-house scripts with MATLAB 2007b (The MathWorks Inc., Natick, MA). GC/MS data was read and stored as a data matrix, in which the values of each row correspond to retention
times and the values of each column corresponded to mass-to-charge ratio. Total ion chromatograms of all samples were interpolated to the same unit vector length and then the retention time was aligned using a third order polynomial. Normalized data matrices were applied for further multivariate data analysis.

4.3 Results and Discussion

4.3.1 Derivatization of Bacterial Fatty Acid

Bacterial fatty acids mainly present in the form of lipids in the bacterial cells. As a strong base, TAMH methanol solution in this experiment was used as a dual-function reagent to saponify and derivatize whole-cell bacterial fatty acids into FAMEs in one step. Headspace SPME was used for sample collection and then injecting the extracts into GC directly. Figure 4-1 shows the headspace SPME/GC/MS chromatogram of bacterial cells before and after the addition of TMAH methanol solution. Blank extractions of bacterial cells using SPME showed only few chromatographic peaks at very short retention times, which are from the nonpolar volatile compounds present in the bacterial cells. When TMAH methanol solution was added, after 10 min SPME extraction time, significantly large amounts of fatty acid methyl esters were observed.
Figure 4-1. Total ion chromatogram of Gram-negative bacterium *Ecoli* (ATCC 43895) FAMEs using SPME as the sampling method, without (top) and with (bottom) derivatizing bacterial cells using TMAH methanol solution.
Figure 4-2 shows the influence of temperature to derivatization and extraction of bacterial FAMEs. The methyl esters of two main fatty acids (C16:0 and Δ C17:0) present in Gram-negative bacteria were examined. It was found that the chromatographic peaks had highest intensity at about 80 °C. The increase of temperature helps the evaporation of derivatized FAMEs into the headspace of the sample vial, where the FAMEs can be adsorbed on the SPME fiber. However, TMAH has a boiling point at about 100 °C, if the temperature is close to this point, the reaction of saponifying and derivatizing the bacterial cells may not complete before the TMAH evaporated. In this experiment, although the peak intensity is not a critical parameter influencing the classification results, extraction temperature of 80 °C was applied in all the samplings.

Figure 4-3 compared two typical total ion chromatograms of *Ecoli*. O157:H7 FAMEs obtained through headspace SPME extraction in the presence of TMAH (upper figure) and conventional alkaline saponification and solvent extraction (bottom figure), as described in the experimental section. A series of peaks of fatty acid methyl esters derived from bacterial lipids were clearly observed from both the chromatograms. Based on the GC/MS profiles, the main fatty acid components present in *Ecoli*. O157:H7 were identified as C16:0 and Δ C17:0 fatty acids, which have been reported as the characteristic fatty acids in the Gram-negative bacteria. Comparing the two chromatograms, it was observed that the upper chromatogram contained more peaks with shorter retention time, which were not detected in the
chromatogram using conventional derivatization method. Also, the relative intensity of long-chain fatty acid (C18:1 and ΔC19:0) methyl esters achieved by TMAH derivatization/SPME is smaller than that by traditional LLE for both Gram-negative (Figure 4-3) and Gram-positive bacteria (Figure 4-4). The reason for that is probably because ten minutes extraction time is not enough for long-chain fatty acid methyl esters to reach optimum equilibrium between the SPME fiber and the headspace. To accomplish the goal of classifying bacteria with shortened sample preparation time, final equilibrium was sacrificed in this experiment as long as characteristic FAME GC/MS profiles could be achieved for all bacterial samples.

Figure 4-2. Influence of adsorption temperature on the performance of SPME.
Figure 4-3. Total ion chromatogram of Gram-negative bacterium *Ecoli* (ATCC 43895) FAMEs with different sample preparation methods. Top: TMAH derivatization and headspace SPME; bottom: Traditional alkaline saponification and LLE.
Figure 4-4. Total ion chromatogram of Gram-positive bacteria Enterococcus (ATCC 19433) FAMEs with different sample preparation methods. Top: TMAH derivatization and headspace SPME; bottom: Traditional alkaline saponification and LLE.
4.3.2 FAME GC/MS Profiles

It was observed from the GC/MS profiles that different bacterial Gram types have obviously various fatty acid distributions. More FAME chromatographic peaks can be achieved in Gram-negative bacteria than in Gram-positive bacteria, because the latter have more rigid structure and is difficult to be broken down even in alkaline solution. The total ion chromatograms of all the bacterial samples showed that saturated straight chain fatty acids contribute to a large amount in all bacterial lipids. It was found that the peak of C16:0 methyl ester was notably intensive in all bacterial samples, but C18:0 was found only in Gram-positive bacteria. The saturated branched-chain fatty acids (like i-C16:0) were found mainly in Gram-positive bacteria but absent in Gram-negative bacteria. The unsaturated straight-chain fatty acids (like C16:1) and cyclopropyl fatty acids (like Δ C17:0) were detected mainly in Gram-negative bacteria. The types of fatty acids may be similar among the same Gram type, but with some of the fatty acids, there are differences in the relative amounts. Thus, the differences in the FAME GC/MS profiles provide useful information for differentiating various bacterial species.

4.3.3 Multivariate Analysis

PCA was applied to the GC/MS total ion chromatograms of the extracted bacterial FAMEs. The profiles of five replicates of six bacteria were visualized in a score plot containing two principal components, which account for 69% of total variance, as seen in Figure 4-5. The same letter in this plot
corresponds to the replicates of each bacteria listed in Table 4-1. A clear separation between the Gram-positive (A, B, and F) and Gram-negative bacteria (C, D, and E) based on their fatty acid profiles was obviously noticeable, except that one replicate of bacterium B was much closer to the center of Gram-negative bacteria. This replicate can be explained as an outlier in the dataset. The separation among the bacteria in the same Gram type was not complete using two principal components because of their structural similarity. The Gram-negative bacteria have positive scores on the first principal component, while the Gram-positive bacteria in the left cluster have negative scores. The chromatographic peaks with positive intensities indicate the biomarker FAMEs of Gram-negative bacteria, while the peaks with negative intensities are characteristic of the Gram-positive bacteria. Because the scores of the spectra are projections onto the first principal component, the chromatogram patterns reveal the fatty acid biomarkers for the bacteria of different Gram types.
Figure 4-5. Top: PCA score plot of six bacteria. The first two PCs account for 69% variance in total. The same letter corresponds to the replicates of each bacterium listed in Table 1. Bottom: The 1st PCA loading plot.
To further classify the six bacterial species, FuRES was applied to the GC/MS total ion chromatograms of the extracted bacterial FAMEs as well. A house-write function of “nmvid3” was used for calculating the FuRES classification model. Figure 4-6 is a FuRES classification tree which clearly indicated the relationship between each bacterial sample. All samples were separated using this five-rule tree without splitting the samples in the leaf nodes. The first rule separated Gram-positive and Gram-negative bacteria to the left and the right side of the tree. The depth of the tree encodes the similarities between the bacterial samples. A cross-validation was applied using BLP and generated a confusion matrix of the prediction results (Table 4-2). In this matrix, each row is the actual class of the bacterial sample and each column is the predicted class, thus only the numbers on the diagonal indicate correct classification. The confidence intervals were calculated based on ten BLPs. Classification accuracy of 87±3% was achieved.
Figure 4-6. FuRES classification tree of six species of bacteria. Nc is the number of replicates and H is the absolute fuzzy entropy of classification.

Table 4-2. Confusion matrix with 95% confidence interval for FuRES analysis of GC/MS data from bacterial fatty acid methyl esters. 10 bootstraps and 5 Latin partitions.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2.7±0.7</td>
<td>0.4±0.5</td>
<td>0.1±0.2</td>
<td>0</td>
<td>0.4±0.4</td>
<td>1.4±0.7</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>4.9±0.2</td>
<td>0</td>
<td>0.1±0.2</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0.2±0.3</td>
<td>0</td>
<td>0</td>
<td>4.8±0.3</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>1.4±0.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.6±0.7</td>
</tr>
</tbody>
</table>
4.4 Conclusions

In this project, SPME was used for headspace extracting polar fatty acids from whole cells of six bacterial species by derivatizing bacterial fatty acids into their corresponding methyl esters. TMAH was functioned as the saponification and methylation reagent. The extracts were directly injected into GC/MS without extra sample treatment. Compared with traditional LLE, this method avoids using inorganic and organic solvents and simplifies the sample preparation as well. Characteristic GC/MS bacterial FAME profiles could be achieved with much shorter sample preparation time. This study also demonstrated the possibility of constructing a unique library for classifying different bacteria based on their FAME GC/MS profiles using simultaneous methylation and SPME sample preparation method.
Chapter 5  Detection of Cocaine and Its Metabolites in Urine Using Solid Phase Extraction–IMS with Alternating Least Square


Copyright 2009 Elsevier

5.1  Introduction

Increasingly, drug testing is being carried out under federally regulated programs. Biological fluids such as urine, semen, blood, sweat, or oral fluid samples are examined to determine the presence or absence of specified drugs or their metabolized traces. Cocaine is a powerfully addictive drug because of its instantaneous and overwhelming effects on the central nervous system. Cocaine is one of the most commonly abused illicit drugs in the United States. According to the National Survey on Drug Use and Health (NSDUH), in 2006, there were 2.4 million cocaine users aged 12 or older and the rate of cocaine use remained stable between 2002 and 2006.

Cocaine is metabolized too rapidly to be measured for routine screening, so benzoylecgonine is the most commonly monitored compound when urine is being screened for the presence of cocaine. When people consume cocaine with alcohol, a different metabolite, cocaethylene, is produced and should also be tested in screenings. The chemical structures of cocaine and its metabolites are given in Figure 5-1. Because of the highly hydrophilic nature of benzoylecgonine, other water-soluble metabolites in urine may interfere with its extraction, increasing the background noise and
generating interfering signals. Solid phase extraction (SPE) is an effective sample treatment technique in forensic analysis as it gives high recoveries and clean extracts. The presence of cocaine and its metabolites in different matrices, like urine, hair, oral fluid samples, human plasma, amniotic fluid, and meconium, has already been investigated using SPE coupled with chromatographic analysis.

Figure 5-1. Chemical structures for (top) cocaine and its metabolites, (bottom left) benzoylecgonine, (bottom right) cocaethylene.
Urine is the most commonly monitored biological fluid for detection of abused drugs in preliminary testing. The Substance Abuse and Mental Health Services Administration (SAMHSA) sets the initial and confirmatory test cut-off levels for cocaine metabolites in urine at 300 ng/mL and 150 ng/mL, respectively. Forensic testing of urine specimens generally consists of an immunoassay screening followed by GC/MS confirmatory analysis. To subvert a drug test, people may use commercial products to adulterate their urine, preventing screening immunoassays from giving the positive results. Some adulterants may cause interference with the detection scheme of the screening immunoassays or during the GC/MS analysis.\textsuperscript{113}

IMS measurement of a single sample produces a mixture of reagent and analyte product ions.\textsuperscript{114} The presence of small amounts of solvent and impurities in urine samples will make the interpretation of IMS spectra of drug metabolites complicated. Therefore, IMS spectral datasets were modeled using SIMPLISMA and ALS\textsuperscript{62-64,115} to separate product ions from reactant ions and background noises. As a result, the sensitivity and selectivity of the method can be further enhanced.

In this work, cocaine metabolites in urine samples were extracted by SPE reversed-phase C\textsubscript{18} cartridges and then analyzed using a bench top IMS, with a 20 s analysis time. Coupling SPE with IMS can decrease the detection limits of drug metabolites in urine while removing salts and other polar compounds that suppress ionization during the measurement. The chemometric method of ALS was applied to all IMS datasets for mathematical
extraction of the target compound spectra. The presence of adulterants did not influence detection of cocaine metabolites. The approach described in this paper provides forensic chemists a fast, simple drug screening technique for use as an alternative or supplement to existing methods.

5.2 Experimental

5.2.1 Chemicals and Reagents

Standard solutions of benzoylecgonine and cocaethylene as well as synthetic negative urine sample were purchased from Cerilliant Corporation (Round Rock, TX). Standard cocaine sample was acquired from Sigma-Aldrich (Milwaukee, WI). Methanol (HPLC grade), ammonium hydroxide (Certified ACS Plus grade) and isopropanol (Certified ACS grade) were from Fisher Scientific (Fair Lawn, NJ). Hydrochloric acid (ACS grade) and methylene chloride (ACS grade) were from EM Science (Gibbstown, NJ). Purified water used in all experiments was generated from a Milli-Q unit (Millipore, Bedford, MA). Adulterants tested in this work were Add-it-ive® (Health Tech, Champlain, NY) and Urine Luck® detoxifying agent (Spectrum Lab, Cincinnati, OH). Discovery® SPE reversed-phase C₁₈ cartridges (500 mg/3 mL) were obtained from Supelco (Bellefonte, PA).

5.2.2 SPE Procedures

The SPE cartridges were conditioned with 3 mL of methanol and 3 mL of distilled water, followed by 1 mL of 0.1 M phosphoric acid buffer solution (pH 6.0). The phosphoric buffer solution (2mL) was added to 2 mL of urine sample to optimize the absorption of the analyte molecules before loading
onto the SPE cartridges, which were then washed with purified, distilled water (2 mL) and 0.1 M hydrochloric acid (2 mL). All washes were discarded. Finally, the cartridges were eluted with 3 mL of methylene chloride–isopropanol–ammonium hydroxide solution (78: 20:2). The suction flow was maintained at 1 mL/min during cartridge conditioning, sample loading, washing, and elution. The final eluates were evaporated to 0.5 mL under pure nitrogen flow. The concentrated eluates were transferred to IMS glass fiber disks and dried for direct analysis.

5.2.3 IMS Parameters

The ion mobility spectrometer used in this work was a Barringer Ionscan® Model 350 (Barringer Instruments Inc., NJ) using nicotinamide as the internal calibrant in positive ion mode. The IMS was operated with the desorption temperature set at 260 °C, the inlet temperature at 260 °C, and the drift tube temperature at 233 °C. The drift flow and sample flow were set at 300 cm³/min and 228 cm³/min, respectively. Ion mobility spectra were collected after a 5 ms delay with a 200 µs shutter grid width. The scan period was set to 20 ms and desorption time to 20 s.

5.2.4 Data Acquisition and Processing

The ion mobility spectra were collected on a National Instruments™ AT-MIO-16XE-10 board using a lab constructed LabVIEW™ 7.0 virtual instrument (VI) as the data acquisition system. The data acquisition frequency was set at 100 kHz and each spectrum consisted of 2000 points. Removal of the gating pulses and baseline correction were performed before
data treatment. The IMS spectral datasets were then processed with ALS to obtain separated spectra and concentration profiles.

5.3 Results and Discussion

5.3.1 Ion Mobility Spectra

Due to the IMS parameters of this experiment, there are 1000 spectra collected during each 20 s IMS analysis. The result is illustrated in Figure 5-2, where the black color shows the most intense peak. In this figure, the peak at approximately 9.5 ms is the reactant ion peak (RIP), produced by the internal calibrant (nicotinamide). Once desorption process begins, the RIP intensity decreases substantially, which is caused by the charge transfer reactions between reactant ions and analyte. Another peak with drift time at approximately 15 ms is detected as well. However, its intensity is far less intensive than that of the reactant ion. Therefore, it is difficult to identify the peaks when only small amount of analytes present. The principle of ALS has been introduced above. The application of ALS can generate ion mobility spectra and concentration profiles, simultaneously, as shown in Figure 5-2. As a result, the normalized IMS spectra of reactant ion and analyte ions will be displayed separately, but in the same profile. Thus, any one of them can be extracted out independently. All IMS spectra in this paper are treated with ALS, unless noted otherwise.
Figure 5-2. Ion mobility spectra image (top), ALS spectra (bottom left), and concentration profile (bottom right) of urine sample containing 100 ng/mL benzoylecgonine with SPE treatment. Reactant ion (---); analyte ion (--).
5.3.2 Reduced Mobilities

Figure 5-3 gives the ion mobility ALS spectra acquired from 1 μg/mL methanol solutions of cocaine, benzoylecgonine, and cocaethylene. The drift times of the three compounds are very similar, but the differences allow for identification via the reduced mobilities for the compounds. As shown in Figure 5-4, a mixture of benzoylecgonine and cocaethylene was separated from each other although they have very similar drift times. Reduced mobility, calculated according to equation (1-4), is an important factor for indentifying different compounds in IMS. Table 5-1 lists the molecular masses, drift times, and reduced mobilities of the three analytes studied. When analyzing urine samples containing drug metabolites, these reduced mobilities are critical for analyte identification.

Table 5-1. The molecular masses, drift times, and measured reduced mobility values of cocaine and its metabolites investigated using SPE and IMS.

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Molecular Mass (g/mol)</th>
<th>Drift Time (ms)</th>
<th>Reduced Mobility (cm²/V·s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cocaine</td>
<td>303.4</td>
<td>15.1</td>
<td>1.16</td>
</tr>
<tr>
<td>Benzoylecgonine</td>
<td>289.4</td>
<td>14.9</td>
<td>1.19</td>
</tr>
<tr>
<td>Cocaethylene</td>
<td>317.4</td>
<td>15.6</td>
<td>1.13</td>
</tr>
</tbody>
</table>
Figure 5-3. ALS spectra of 1 μg/mL cocaine (top), benzoylecgonine (middle) and cocaethylene (bottom) added on IMS disk directly. Reactant ion peak (---); analyte ion peak (——).

Figure 5-4. ALS spectra of mixture of benzoylecgonine and cocaethylene. Reactant ion peak (---); analyte ion peaks (—).
5.3.3 Solid Phase Extraction

The application of SPE in urine analysis can increase the instrument’s sensitivity to target compounds by decreasing the amount of interfering compounds. As can be seen from Figure 5-5, the ion mobility spectra of the negative urine sample without SPE contains peaks that have drift times similar to that of the RIP and another peak that has a drift time of about 12.6 ms. The urine sample containing 300 ng/mL benzoylecgonine has a very similar profile. As listed in Table 5-1, under the same instrumental parameters, benzoylecgonine should have a drift time of approximately 14.9 ms. The corresponding peak does not appear in the data. This result indicates that IMS alone cannot detect benzoylecgonine in a urine sample. Compounds contained in urine may compete with the analyte in the charge transfer process and suppress the formation of product ions.
Figure 5-5. ALS spectra of (a) negative urine sample without SPE process (top); (b) urine sample containing 100 ng/mL benzoylecgonine without SPE process (middle); (c) urine sample containing 100 ng/mL benzoylecgonine with SPE process (bottom). Reactant ion peak (---); analyte ion peaks (—).
5.3.4 Alternating Least Squares

In the previous discussion, only two components were chosen to produce the ion mobility spectra. In addition to RIP, all the other peaks were all considered as from the analyte. Although SPE can remove most of the interference, it is inevitable that some molecules, other than the drug metabolites, remain in the sample. Choosing more components is an effective method to mathematically purify the target compound spectra. As shown in Figure 5-6, compared with the ion mobility spectra with two components, the bottom graph with three components separates the interferences from the RIP and the analyte ions.

Figure 5-7 and Figure 5-8 compare the average spectrum, single spectrum with highest intensity, and ALS treated spectra of a urine sample containing benzoylecgonine and cocaethylene, respectively. In the average spectrum, the intensity of the RIP is more intense than that of the analyte peak, while the single spectrum with the most intense analyte peak shows a very high signal-to-noise ratio. It is clear that the ALS spectra give both intense analyte peak and clean background, with larger signal-to-noise ratio as well, which mathematically increases the sensitivity of this method. Detection limits were defined as three times standard deviation of the baseline noise between 4 and 6 ms drift time interval in the ALS spectrum. By testing the urine samples with decreasing drug metabolites concentrations, the detection limits of benzoylecgonine and cocaethylene using this method were found to be 10 ng/mL and 4 ng/mL, respectively.
Without the process of ALS, as shown in the middle graphs of Figure 5-7 and Figure 5-8, the method sensitivity was much lower. Method detection limits also depend on the amount of urine sample tested. Better performance could be achieved if more urine samples are provided. However, IMS would be used as a screening method to test cocaine metabolites in urine. Confirmatory analysis would still be necessary for forensic purpose.

5.3.5 Adulterants Influence

The adulterants tested in this work were Add-it-ive® and Urine Luck® detoxifying agent. Neither of them influenced the results of the analysis using this method. Figure 5-9 shows the ion mobility spectra of urine sample containing 150 ng/mL benzoylecgonine with the manufacturer’s suggested amount of Add-it-ive® added to the sample. No extra peaks are observed, allowing for accurate detection of the metabolite.

5.4 Conclusions

Sample preparation using SPE combined with sample analysis using IMS is an effective method for drug analysis. Applying ALS to IMS data sets provides a system with enhanced sensitivity and selectivity for the detection of trace amounts of drug metabolites in urine samples. The detection limits of benzoylecgonine and cocaethylene using our method are 10 ng/mL and 4 ng/mL, respectively. The presence of adulterants in a urine sample does not influence the detection of cocaine metabolites after pretreatment with SPE. This SPE-IMS method provides forensic chemists an alternative approach for fast and simple drug screening.
Figure 5-6. ALS spectra of urine sample containing 100 ng/mL cocaethylene after SPE process. Number of components: two (top); three (bottom). Reactant ion peak (−−−−); analyte ion peak (—); interference (−−−−).
Figure 5-7. Comparison of average spectrum (top), maximum spectrum (middle) and ALS spectra (bottom) of urine sample containing 150 ng/mL of benzoylcegonine.
Figure 5-8. Comparison of average spectrum (top), maximum spectrum (middle) and ALS spectra (bottom) of urine sample containing 150 ng/mL of cocaethylene.
Figure 5-9. ALS spectra of urine sample containing 150 ng/mL cocaethylene and the manufacturer’s suggested amount of Add-it-ive® after the SPE process. Three components were chosen when applying ALS. Reactant ion peak (---); analyte ion peak (—); interference ( - - ).
Chapter 6  Summary and Future Work

Classification of ignitable liquids from fire debris using two-way GC–DMS profiles was investigated in Chapter 2. The performance of DMS as a gas chromatographic detector was also compared with MS using a chemometric tool PDR in Chapter 3. The combined information afforded by GC and DMS provided unique two-way patterns for each sample of ignitable liquid. This method could be implemented in arson analysis by forensic chemists.

Portable and easy-to-use analytical instruments are always desirable in forensic analysis. Although many instruments like GC/MS/MS, FTICR high resolution MS, and tandem GC have been successfully applied for ignitable liquids analysis, the cost and maintenance have to be considered. The microfabrication of DMS makes itself a very promising analyzer for field analysis. The applications of DMS alone have proved its capability in detecting various volatile and semi-volatile compounds. The use of a differential mobility spectrometer as a GC detector provides two-way information for pattern recognition. However, the GC used in this study was a bench top model, although the proposed method has the potential to be used for onsite analysis, future work will investigate an integrated micro GC–DMS system by coupling DMS to a portable GC.

Traditionally, total ion currents are mainly used for pattern recognition when analyzing data collected from hyphenated instruments. The proposed method used two-way GC–DMS data for classifying seven ignitable liquids
from fire debris using FuRES. Obviously, two-way data provided more
classification information than one-way total ion currents or integrated
spectra, and also help to correct the retention time shifting when aligning the
chromatographic peaks. In this project, pattern recognition was only tested
between seven types of ignitable liquids. To optimize this analysis method,
collecting more ignitable liquid sample profiles are needed to build a data
library, which would make this method more reliable in identifying unknown
samples.

The third project involved studies on classification of bacteria by
simultaneous methylation and solid phase microextraction of whole bacterial
cell fatty acid methyl esters, which were directly analyzed using GC/MS.
Compared with traditional method of saponifying and methylating bacterial
cells, this method could greatly decrease sample preparation time. Also, it
avoids using amount of acid, base, and organic solvents. Data preprocessing
is a key step for pattern recognition. In this case, chromatographic peak
alignment is necessary because manual injections always cause peak shifting
between different runs. Although small shifting can be corrected using
proper data processing skills, inconsistency in retention time would influence
the accuracy of classification results. Analyzing samples using a GC/MS
instrument with a SPME autosampler might reduce this one source of
variation.

The bacterial cells analyzed in this project were cultured following
ATCC instructions. Future research shall also investigate whether any
differences associated with culture media or ages are comparable to the differences in GC/MS profiles among various bacteria. To establish analytical protocols for rapid bacterial identification, not only the experimental factors that influence the sample preparation efficiency and GC/MS profiles need to be optimized: complex sample matrices or samples containing mixtures of bacteria should also be studied. For analyzing such complicated samples, multivariate curve resolution (MCRs) methods will be used to mathematically resolve mixtures by determining the number of components, the response profiles, and the concentration profiles. Future studies will also be extended to the analysis of specific bacteria strains in food products such as meat and milk. The procedure to extract bacterial fatty acids and minimize the influence from background matrices will be a key challenge.

The last project presented a comparatively rapid and facile method for detection of cocaine and its metabolites, benzoylecgonine and cocaethylene in urine sample using SPE coupled with IMS. Applying SPE could effectively extract cocaine metabolites from urine samples, and the addition of adulterants did not influence the detection of analytes as well. Compared with chromatographic analysis techniques, IMS provides the advantages of atmospheric pressure operation, simple sample preparation, portability, fast analysis, and high sensitivity. Another novel aspect of this project is the application of SIMPLISMA and ALS to mathematically separate ion mobility spectrum of calibrant and analyte, which greatly increase the method sensitivity. Although IMS analysis time is only 20 s for each measurement,
the sample preparation is still time-consuming. C18 SPE cartridges were used in sample preparation step, which includes solid-phase conditioning, sample loading, washing, analyte eluting, and elute concentrating. Future research will investigate the possibility of directly using SPE disks as IMS sample membranes, thus the analyte eluting and concentrating steps can be avoided during sample preparation step. In addition to the desired solid phase, the structure components of the SPE disk must be able to withstand the high IMS inlet temperature and durable disks would help. Considering the cost and specificity of the commercially available SPE disks, it would be ideal to synthesize thermally stable disks with selective functional groups for extracting drug metabolites from biological matrices. In addition to SPE, SPME can also be investigated to conduct the sample preparation step in this experiment. To desorb the analytes from the SPME fiber directly, the IMS sample inlet region has to be designed to fit the SPME fiber. This modification can be realized by transpiring a small hole (about 1 mm) through the wall of an IMS sample holder. Thus, the analytes on the SPME fiber can be desorbed directly. Future studies will also focus on the analysis of different drug metabolites in other biological matrices, like hair, blood, and serum. Choosing a suitable solid phase and optimizing the extraction conditions will be the key steps in these new projects. Note that, applying IMS could provide a viable approach for comparatively fast and simple drug screening; a confirmatory technique is still needed when evidence is required by court.
References


Appendix A

Publications


Appendix B

Presentations


