

2014

High-throughput analysis of contrived cocaine mixtures by direct analysis in real time/single quadrupole mass spectrometry and post acquisition chemometric analysis

<https://hdl.handle.net/2144/15351>

Downloaded from DSpace Repository, DSpace Institution's institutional repository

BOSTON UNIVERSITY
SCHOOL OF MEDICINE

Thesis

**HIGH-THROUGHPUT ANALYSIS OF CONTRIVED COCAINE MIXTURES BY
DIRECT ANALYSIS IN REAL TIME/SINGLE QUADRUPOLE MASS
SPECTROMETRY AND POST ACQUISITION CHEMOMETRIC ANALYSIS**

by

ANDREW BLAIR HORSLEY

B.S., University of Puget Sound, 2011

Submitted in partial fulfillment of the
requirements for the degree of

Master of Science

2014

© 2014 by
ANDREW BLAIR HORSLEY
All rights reserved

Approved by

First Reader _____

Adam B. Hall, Ph.D.
Instructor, Biomedical Forensic Sciences Program

Second Reader _____

Brian Musselman, Ph.D.
Chief Executive Officer
IonSense Inc.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my thesis advisor and mentor throughout this project Adam Hall. Without his patience, guidance, and placing this project into my hands I never would have had been able to tackle this project and reap all of the benefits that have come from it. From the classroom, to the laboratory, to my research and for giving me the chance to give a presentation at AAFS, thank you for teaching me for the last two years and giving me a firm foundation to grow from as a forensic scientist.

Secondly, to everyone at IonSense, to Rob, Taylor, Doug, Sam, Maureen, thank you so much for helping throughout this project where each and everyone of you have helped bring this thesis together, aid in my experiments, provide me insight into my problems, and help when I have hit roadblocks along the way. I would like to thank in particular Brian Musselman, Joseph Tice and Joseph LaPointe. To Brian for pushing me everyday to get results, work harder, and continually improve as a student, speaker, and scientist. To Joseph LaPointe for selfishly taking too much of your work time when it came to answering all of my questions, helping me fix the instrument, aid in running my experiments, and helping me through problems that arose along the way. And finally to Joseph Tice for mentoring me at IonSense, helping me with my research and the DHS program, pushing me to do better and waiting very patiently at Broadway station.

Thank you to Adam Hall, Brian Musselman, and Daniel Burgard for bringing my thesis to completion by giving me feedback, comments, and making edits draft after draft until it was perfected. I would also like to thank my friends and my faculty and peers in the Boston University Biomedical Forensic Sciences Program for helping me accomplish and succeed for the last two years. You have helped me through the class work, tests, projects, labs, and the day-to-day commitments for the last two years. Thank you to all my friends in Boston for making it a remarkable and amazing experience.

Lastly, I would like to thank my mom, step-dad, and sister for keeping me sane during this thesis project and program. You have always been just a phone call away. Thank you for continuing to support me 3,000 miles away and for all of the loving reminders along the way.

**HIGH-THROUGHPUT ANALYSIS OF CONTRIVED COCAINE MIXTURES BY
DIRECT ANALYSIS IN REAL TIME/SINGLE QUADRUPOLE MASS
SPECTROMETRY AND POST ACQUISITION CHEMOMETRIC ANALYSIS**

ANDREW BLAIR HORSLEY

ABSTRACT

Direct Analysis in Real Time (DART) ionization/mass spectrometry allows for the high throughput analysis of a wide range of materials including but not limited to: solids, liquids, powders, tablets, and plant materials. The ability to detect cocaine was established in a reproducible manner with the use of a DART ionization source (IonSense Inc., Saugus, MA) interfaced to a modified single quadrupole mass spectrometer. Development of a methodology for the detection of cocaine within contrived street quality drug mixtures involved the optimization of the ionization source, sample introduction mechanism, ion guide, and mass analysis parameters. An analytical method was created that utilized ionized helium carrier gas heated to 300°C and an automated sample introduction apparatus consisting of a Linear Rail Enclosure that holds consumable QuickStrip™ sample cards. Ionized molecules were then fragmented by manipulation of voltage levels within the ion guide to gain more structural information prior to detection by a single quadrupole mass spectrometer.

Cocaine was detected by the modified DART/MS analytical platform and gave two peaks within the mass spectrum at m/z 304 and 182. Optimization of in-source fragmentation by manual adjustment of the skimmer focus voltage allowed for the reproducible fragmentation of cocaine and the ability to increase or decrease the amount of fragmentation seen between the two peaks detected for cocaine. With the use of fragmentation, this analytical platform can be classified as a Category A technique as defined by the Scientific Working Group for the Analysis of Seized Drugs.

The robust detection of cocaine was demonstrated for reference samples at concentrations as low as 10 ng/ μ L (50 ng) with high signal abundance greater than ten times the signal to noise ratio. Furthermore, the detection of cocaine at 10 ng/ μ L was demonstrated for multi component mixtures of up to 14 additional components containing common adulterants and diluents found within street quality samples. In total, 25 common excipients were tested using the same method parameters as optimized for cocaine analysis. Of these 25 excipients tested, five were not detected in positive ion mode (one could be detected in negative ion mode). Of the twenty excipients that could be detected by mass spectrometry, two pairs of excipients (levamisole/tetramisole and creatine/creatinine) could not be differentiated from each other. There were no excipients tested that had equivalent m/z values as those of cocaine. Experimentation into the effects of various excipients at multiple concentrations on the abundance of the two cocaine peaks was performed. Regardless of

excipient amount (up to 10 times more concentrated than cocaine) and the number of components (up to 15 total components) the ratio of abundance between the m/z 304 to 182 peaks did not vary greater than 22% relative standard deviation.

A match criteria protocol was developed for the ability of an analyst to confirm the presence of cocaine within unknown forensic case samples that have previously tested positive for the presumptive identification of cocaine. The identification of cocaine was based on various factors such as the signal to noise ratio at m/z 304 and 182, the ratio of abundance between those two peaks as well as positive and negative controls. This match criteria protocol was utilized for 25 double blind mock forensic casework samples was performed. Determination for the presence of cocaine within these unknown samples gave an analyst error rate of 0%, with no false positives or false negatives predicted.

To further aid human interpretation and identification of compounds within mixtures, the advanced chemometric software, Analyze IQ, was utilized. Development of predictive classification models using a combination of pre-processing steps, principle component analysis and machine learning techniques was achieved. Models were built using 381 unique samples for the purposes of identifying the presence of cocaine within unknown samples. Of all methods available within the Analyze IQ software, the optimization of a model using principle component analysis with support vector machine regression with a radial basis function kernel yielded an initial error rate of 0% for 72 samples

tested. Furthermore, of the samples tested against the model, 20 samples were comprised of excipients that were not incorporated into the initial model development process. The inclusion of these samples (10 spiked with cocaine, 10 absent of cocaine), shows that predictive modeling based software can provide an accurate, robust, and evolving approach to the identification of cocaine within sample compositions that have not previously been tested and stored in a database of known reference samples. Predictive modeling has advantages over current mass spectral libraries, which are limited to the identification of pure compounds. To further test the abilities of predictive models, optimized machine learning models were applied to 25 double blind mock forensic casework samples. The predictive modeling error rate was identical to the human interpretation rate for the double blind mock casework samples with a 0% error rate.

Using the DART/MS analytical platform, 25 mock forensic casework samples along with positive and negative controls were analyzed and identified for the presence of cocaine within 30 minutes. On the order of 15 to 30 times faster than modern GC/MS and LC/MS methods, the ability to analyze and identify samples faster would allow for an increase in samples being processed on a daily basis and allow for the reduction of case backlogs that currently plague controlled substances sections of forensic science laboratories throughout the United States.

TABLE OF CONTENTS

	Page
Title	i
Copyright Page	ii
Reader's Approval Page	iii
Acknowledgements	iv
Abstract	vi
Table of Contents	x
List of Tables	xiv
List of Figures	xvi
List of Abbreviations	xxii
1 INTRODUCTION	1
1.1 Cocaine	1
1.1.1 Common Cocaine Adulterants and Diluents	3
1.1.2 Detection Methods for Controlled Substances	4
1.1.3 Increases in Case Backlogs	5
1.2 Direct Analysis in Real Time	8
1.2.1 Sample Introduction to DART Ionization Region	10
1.3 DART Interface to Mass Spectrometer	12
1.3.1 Interfaced Ion Guide Optics to Mass Spectrometer	13
1.4 Mass Spectrometry	15
1.4.1 Quadrupole Mass Analyzer	15

1.5 Analyze IQ Lab Software	17
1.5.1 Creating a Model within Analyze IQ	17
1.5.2 Pre-Processing of Data within Analyze IQ	19
1.5.3 Principle Component Analysis	20
1.5.4 Analyze IQ Machine Learning Techniques	21
1.6 Aim of This Thesis	26
2 EXPERIMENTAL	26
2.1 Materials	26
2.2 DART Cocaine Analysis Methodology Development	27
2.2.1 DART Source Parameter Settings	28
2.2.2 Sample Introduction Considerations	28
2.2.3 Optimization of the DART Carrier Gas Temperature	30
2.2.4 Optimization of the DART Skimmer Voltage	30
2.3 Ion Optics Parameters	30
2.4 Mass Spectrometer Method Parameters	31
2.5 Optimized Cocaine Parameters	31
2.6 Cocaine Mass Spectral Peak Ratio Testing	32
2.7 Cocaine Excipient Library Development	32
2.8 Analysis of Multi-Component Cocaine Mixtures	32
2.9 Analyze IQ Cocaine Identification Development	33
2.9.1 Creation and Analysis of Reference Samples	33
2.9.2 Pre-Processing Techniques	34

2.9.3 Analyze IQ Model Creation and Development	34
2.10 Cocaine Double Blind Study	34
3 RESULTS AND DISCUSSION	35
3.1 DART/Single Quadrupole Method Parameters for Cocaine	35
3.1.1 DART Carrier Gas Temperature Optimization	35
3.1.2 Analysis of Cocaine by DART/Single Quadrupole MS	36
3.1.3 DART Skimmer Voltage Optimization	38
3.2 Cocaine Excipient Library Analysis	41
3.3 Multi-Component Cocaine Mixture Analysis	45
3.4 Analysis of Peak Ratio Variability for Cocaine	50
3.5 Cocaine Double Blind Study Analysis	55
3.6 Cocaine Identification Model Development with Analyze IQ	65
3.6.1 Reference samples	65
3.6.2 Evaluation of Machine Learning Techniques	66
3.6.3 Early Model Errors	69
3.6.4 Optimization of Machine Learning Techniques	69
3.7 Evaluation of Cocaine Identification Model	71
4 CONCLUSIONS	72
5 FUTURE DIRECTIONS	74
LIST OF JOURNAL ABBREVIATIONS	76

REFERENCES	78
CURRICULUM VITAE	85

LIST OF TABLES

Table 1. Scientific Working Group for the Analysis of Seized Drugs recommendations. Techniques in bold are for the detection of cocaine ¹⁷ .	5
Table 2. Ion optic voltages between MS inlet and single quadrupole MSD.	31
Table 3. Multi-component cocaine mixtures ^a .	33
Table 4. Common cocaine excipients tested by DART/ single quadrupole mass spectrometry using in-source fragmentation	42
Table 5. Comparison of DART/MS peaks to the NIST EI peaks.	45
Table 6. Statistical measurements for ratio of peak area abundance between cocaine <i>m/z</i> 304 and 182 peaks.	54
Table 7. Statistical measurements for ratio of peak area abundance between <i>m/z</i> 304 and 182 peaks for unknown mock case samples and reference samples. Bold numbers are values less than 10x S/N (as defined by the average abundance of the negative control peaks). All statistics shown are based on triplicate measurement of sample.	62
Table 8. Reference samples for building cocaine identification models using Analyze IQ.	66
Table 9. Adulterant mixtures in comparison to multi-component cocaine mixtures as shown in Table 3.	66
Table 10. Samples for testing against Analyze IQ built models.	68

LIST OF TABLES, CONTINUED

Table 11. Model training and testing error rates for cocaine identification machine learning techniques at default values.	68
Table 12. Model training and testing error rates for optimization of SVM/RBF/Euclidean distance cocaine identification model.	70
Table 13. Evaluation of Analyze IQ optimized models against unknown samples from cocaine double blind study (including positive and negative controls).	71

LIST OF FIGURES

- Figure 1. Molecular structure and physical form of cocaine. Image (right) courtesy of the Drug Enforcement Agency. 1
- Figure 2. National trend of cocaine and cannabis/THC drug cases from 2001 to 2012⁵. Figure courtesy of National Forensic Laboratory Information System 2012 annual report. 2
- Figure 3. Regional trend of cocaine drug reports (per 100,000) people aged 15 or older from 2001 to 2012⁵. Figure courtesy of National Forensic Laboratory Information System 2012 annual report. 3
- Figure 4. Total Ion Chronogram of samples being analyzed by DART/MS analysis. Each peak corresponds to a unique sample and they are clustered in groups of twelve due to the nature of the twelve sample spots on a QuickStrip™ card (Figure 5). 7
- Figure 5. DART QuickStrip™ card in metal holder. 7
- Figure 6. DART schematic of metastable formation by glow discharge. Figure courtesy of IonSense Inc. 9
- Figure 7. Linear Rail Enclosure of QuickStrip™ for sample introduction into the ionization region. 11
- Figure 8. Linear Rail Enclosure attached between MS inlet (left of image) and DART ionization source (right of image) with QuickStrip™ card in holder. 12

LIST OF FIGURES, CONTINUED

- Figure 9. DART/single quadrupole MSD ion optics, where 1 is the heated capillary of the mass spec inlet, 2 is the skimmer focus lens, 3 is the skimmer focus, 4 is the quadrupole ion guide, and 5 is the ion guide exit lens. Figure courtesy of IonSense Inc. 14
- Figure 10. Analyze IQ Lab system architecture^{53,54,55}. Background operations of Analyze IQ are outlined in blue. Graphical user interface is outlined in red. 19
- Figure 11. A linear SVM showing the marginal hyperplane used to classify positive (black) versus negative (white) samples where the Support Vectors (shaded) are on the boundary of the margin⁶⁰. 22
- Figure 12. Hyper sphere containing target data with support vectors on boundary. The complexity factor C is represented by the radius of the sphere R versus the number of data accepted within. If C is high, R increases and the model might overfit the data and vice versa⁶⁰. 23
- Figure 13. DART/MS block diagram. 28
- Figure 14. Contec CONSTIX® SP-5 polyurethane swabs in holder. 29
- Figure 15. Extracted Ion Chronogram at 304 m/z of cocaine at 100 ng/uL analyzed with increasing DART heater temperatures. 36
- Figure 16. Mass spectrum of cocaine at 100 ng/uL analyzed by DART/single quadrupole MSD. Skimmer focus set to 12 volts. 37

LIST OF FIGURES, CONTINUED

- Figure 17. Mass spectrum of cocaine at 100 ng/ μ L analyzed by DART/single quadrupole MSD with in-source fragmentation. Skimmer focus set to 54 volts. Inset of cocaine in-source fragmentation mechanism by DART/single quadrupole mass spectrometry⁶⁴. 37
- Figure 18. Electron impact mass spectrum of cocaine acquired from NIST Library (Version 2.0, December 2005) with m/z on the x-axis. 38
- Figure 19. In-source fragmentation of cocaine at different skimmer focus voltages. 40
- Figure 20. Molecular structure of creatine (left) and creatinine imine (middle) and amine (right) tautomers. 43
- Figure 21. Induced fragmentation of a 3-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L. 47
- Figure 22. Induced fragmentation of a 6-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L. 47
- Figure 23. Induced fragmentation of a 9-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L. 48
- Figure 24. Induced fragmentation of a 12-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L. 48
- Figure 25. Induced fragmentation of a 15-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L. The asterisk refers to

LIST OF FIGURES, CONTINUED

- atmospheric/background contribution of decamethylcyclopentasiloxane at m/z 371 (personal communication, Joseph LaPointe). 49
- Figure 26. Mass spectrum of procaine fragmentation at a skimmer focus of 54V and at 100 ng/ μ L. 51
- Figure 27. Mass spectrum of levamisole fragmentation at a skimmer focus of 54V and at 100 ng/ μ L. 52
- Figure 28. Ratio of peak area abundance for cocaine peaks of reference samples and unknown samples. For the procaine dilutions series, the order from left to right: 1:1, 1:2, 1:5, and 1:10 dilution samples (cocaine to procaine). For the levamisole dilutions series, the order from left to right: 1:1, 1:2, 1:5, and 1:10 dilution samples (cocaine to levamisole). For the multi-component mixtures, the order from left to right is the 3, 6, 9, 12, and 15-component samples. 53
- Figure 29. Total Ion Chromogram of 25 unknown samples analyzed using the ICE method for the identification of cocaine. All samples were analyzed in triplicate with a substrate blank before and after sample analysis. The first and last samples are positive controls consisting of a cocaine standard at 100 ng/ μ L. The second and second to last samples are negative controls consisting of pure methanol. 56

LIST OF FIGURES, CONTINUED

- Figure 30. Extracted Ion Chromogram of m/z 304 (top) and 182 (bottom) for the detection of cocaine from 25 unknown samples. Numbers shown are for samples with abundances above the baseline noise for their respective ions. All samples analyzed in triplicate with a substrate blank before and after sample analysis. The first and last samples are positive controls consisting of a cocaine standard at 100 ng/ μ L. 57
- Figure 31. Sample 13 from cocaine double blind study. Example of unknown sample positive for the presence of cocaine from the identification of m/z 304 and 182 (shown with asterisks). 59
- Figure 32. Sample 22 from cocaine double blind study. Example of unknown sample negative for cocaine due to lack of observable peaks at m/z 304 and 182. 59
- Figure 33. Ratio of peak area abundance for cocaine peaks of reference samples and unknown samples. For the procaine dilutions series, the order from left to right is the 1:1, 1:2, 1:5, and 1:10 dilution samples (amount cocaine to procaine). For the levamisole dilutions series, the order from left to right is the 1:1, 1:2, 1:5, and 1:10 dilution samples (amount cocaine to levamisole). For the multi-component mixtures, the order from left to right is the 3, 6, 9, 12, and 15 component samples. For the unknown samples, the order from left to right is sample 1 through 25. For the positive control (cocaine standard

LIST OF FIGURES, CONTINUED

at 100 ng/ μ L the order from left to right is the pre-run and post-run sample.

For the negative control (methanol) the order from left to right is the pre-run
and post-run sample. 61

Figure 34. Match criteria flow chart for the determination of cocaine within
unknown samples. 64

ABBREVIATIONS

.csv	Comma Separated Value
°C	Degrees Celsius
DART	Direct Analysis in Real Time
EIC	Extracted Ion Chromogram
eV	Electron Volts
GC/MS	Gas Chromatography/Mass Spectrometry
He	Helium
He*	High Energy Helium Metastable
ICE	Identification of Cocaine in Excipients Method
LRE	Linear Rail Enclosure
MS	Mass Spectrometry
MSD	Mass Selective Detector
[M+H] ⁺	Protonated Molecular Ion
<i>m/z</i>	Mass to Charge Ratio
ng	Nanograms
NFLIS	National Forensic Laboratory Information System
PCA	Principle Component Analysis
PCR	Principle Component Regression
sec	Second
S/N	Signal to Noise Ratio

SAV	Spectral Attribute Voting
SVM	Support Vector Machine
TIC	Total Ion Chromogram
μL	Microliters

1 INTRODUCTION

1.1 Cocaine

Cocaine, (methyl (1R,2R,3S,5S)-3-(benzoyloxy)-8-methyl-8-azabicyclo[3.2.1] octane-2-carboxylate) in the United States is a Schedule II narcotic as defined by the Single Convention on Narcotic Drugs of 1961 and Comprehensive Drug Abuse Prevention and Control Act of 1970 (Figure 1)^{1,2}.

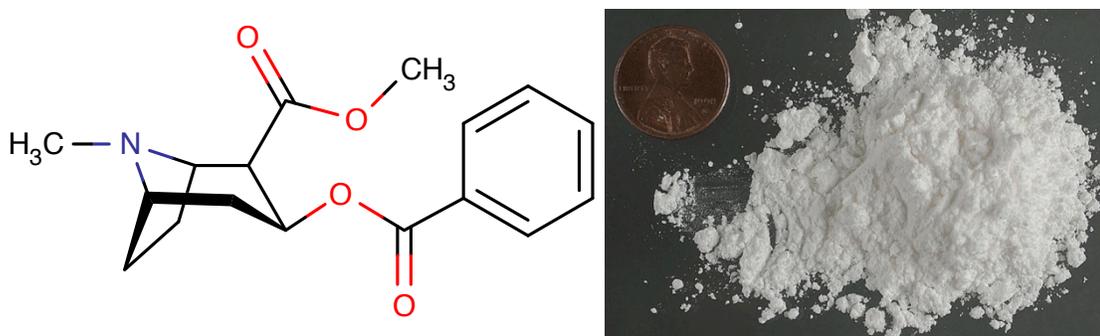


Figure 1. Molecular structure and physical form of cocaine. Image (right) courtesy of the Drug Enforcement Agency.

Cocaine is obtained from the leaves of the native South American coca plant (*Erythroxylon coca* and *Erythroxylon novogranatense*), and is an alkaloid stimulant that acts on the central nervous system creating feelings of euphoria, alertness, energy and competence^{2,3}. Due to its lipophilic and hydrophilic pockets, it can easily cross the blood brain barrier and become biologically active. Cocaine is considered to be one of the most addictive and dangerous drugs and according to the 2012 National Forensic Laboratory Information

System report (NFLIS), the number of cocaine case reports has been decreasing since 2006 (Figure 2)^{4,5}.

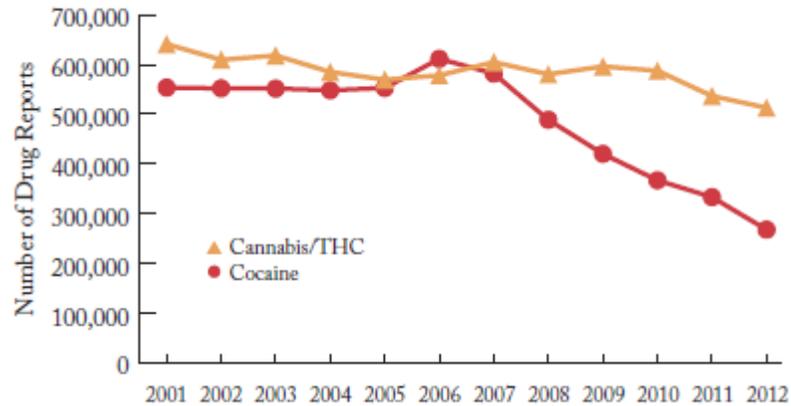


Figure 2. National trend of cocaine and cannabis/THC drug cases from 2001 to 2012⁵. Figure courtesy of National Forensic Laboratory Information System 2012 annual report.

Based on the NFLIS report, in 2012 an estimated 1.6 million drug reports were identified by state and local forensic labs within the United States⁵. With almost 17% of all cases, cocaine was the second most identified drug⁵. Within national laboratories from the DEA's System To Retrieve Information from Drug Evidence II (STRIDE) and the U.S Customs and Border Protection (CBP) laboratories, cocaine was the most identified drug⁵. As shown in Figure 3, cocaine has been on a steady decline regionally in the United States, where the majority of drug cases were seen in the South, Northeast, Midwest, and West (from highest to lowest)⁵.

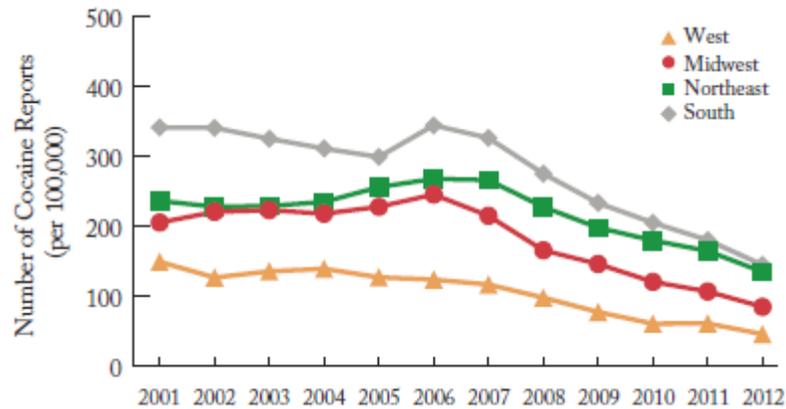


Figure 3. Regional trend of cocaine drug reports (per 100,000) people aged 15 or older from 2001 to 2012⁵. Figure courtesy of National Forensic Laboratory Information System 2012 annual report.

1.1.1 Common Cocaine Adulterants and Diluents

Within illicit substances such as cocaine or heroin, there are additional compounds known as excipients. Drug dealers frequently increase the bulk amount of a drug to be sold by cutting the drug with excipients that may or may not be biologically active. Adulterants are excipients that biologically mimic the effect of the drug of abuse. Diluents are excipients that do not have any associated biological effects. Common adulterants (i.e. xylazine, lidocaine, and caffeine) and diluents (i.e. mannitol, inositol, and boric acid) can have adverse effects on the cocaine user⁶⁻⁹. One example, xylazine, has led to increased injection rates, increased rates of skin lesions and an overdose rate above 20% within the cocaine/heroin drug use community of Puerto Rico¹⁰. Identification of excipients can have probative value for source determination (number of

manufacturers and routes of production) and public health⁷. There are many compounds that have been used as excipients in forensic cocaine samples^{3,7-15}.

1.1.2 Detection Methods for Controlled Substances

Under guidelines from the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) analytical techniques are based on three tiers of classification¹⁶. Category A, B, and C are ranked from highest to lowest maximum potential discriminating power for the analysis of controlled substances. The categorization of a technique can be reduced if the mode of operation diminishes its discriminating power¹⁶. Additionally, SWGDRUG recommends that labs follow certain minimum standards when determining the identity of unknown casework samples. If a Category A technique is used then at least one other technique from any category must be performed and corroborate the results of the previous testing. If a Category A technique is not used at least three other uncorrelated techniques, two of which must be from Category B, need to be performed.

Table 1. Scientific Working Group for the Analysis of Seized Drugs recommendations. Techniques in bold are for the detection of cocaine¹⁷.

Category A	Category B	Category C
Infrared Spectroscopy	Capillary Electrophoresis	Color Tests
Mass Spectrometry	Gas Chromatography	Immunoassay
Nuclear Magnetic Resonance Spectroscopy	Ion Mobility Spectrometry	Fluorescence Spectroscopy
Raman Spectroscopy	Liquid Chromatography	Melting Point
X-ray Diffractometry	Microcrystalline Tests	Ultraviolet Spectroscopy
	Pharmaceutical Identifiers	
	Thin-Layer Chromatography	

The standard analytical technique for the confirmatory identification of illicit substances is gas chromatography/mass spectrometry (GC/MS)¹⁸. Detection of cocaine/excipient mixtures is possible using chromatographic separation of each component prior to analysis by mass spectrometry. With GC/MS methods, the separation of drug mixtures is on the scale of 15 to 45 minutes per sample¹⁹. Furthermore standard operating procedures state a solvent blank is required before each sample to be tested, to eliminate any doubt introduced by column carryover effects.

1.1.3 Increases in Case Backlogs

Using the most recent Census of Publicly Funded Forensic Crime Laboratories, in 2009 there was an estimated 4.1 million requests for forensic services from publicly funded crime labs²⁰. Of these requests, one third of them were for controlled substances analysis. Furthermore, in 2009, crime labs

performed controlled substance analysis 82% of the time²⁰. Of the more than 1.19 million cases that were not completed at the end of 2009, approximately 137,000 (12%) cases were controlled substances requests²⁰. Although controlled substances examinations are not as time consuming as other forensic examinations, they represent the major types of requests in forensic labs (requested only 1% less than biological evidence requests)²⁰. By one research estimate, using data from 2007, crime labs would need to hire at least 275 forensic examiners to eliminate the case backlog by processing only controlled substance cases²¹. Ultimately the need for more analysts, more instrumentation, and faster techniques are continually being sought after within the forensic crime laboratory. Even factors outside of the laboratory, such as *Melendez-Diaz v. Massachusetts*, now require forensic analysts to be present for testimony, leading to slower evidence processing and increases in the number of case backlogs^{6,21}.

Direct Analysis in Real Time is one option that allows for the needs of the continuously evolving forensics laboratory. As shown in Figure 4, the ability to analyze a large amount of samples is limited more by the ability of the analyst to change sampling cards than by the speed of the DART instrumental platform. In this example, 60 samples are analyzed by DART/MS in sixteen minutes by analyzing five QuickStrip™ (Figure 5) cards in succession. According to SWGDRUG guidelines, DART/MS can be considered a Category A technique, if the mode of operation allows for fragmentation of the sample to produce mass

data in addition to the molecular ion. If the mode of operation produces only the molecular ion DART/MS would then be considered a Category B technique.

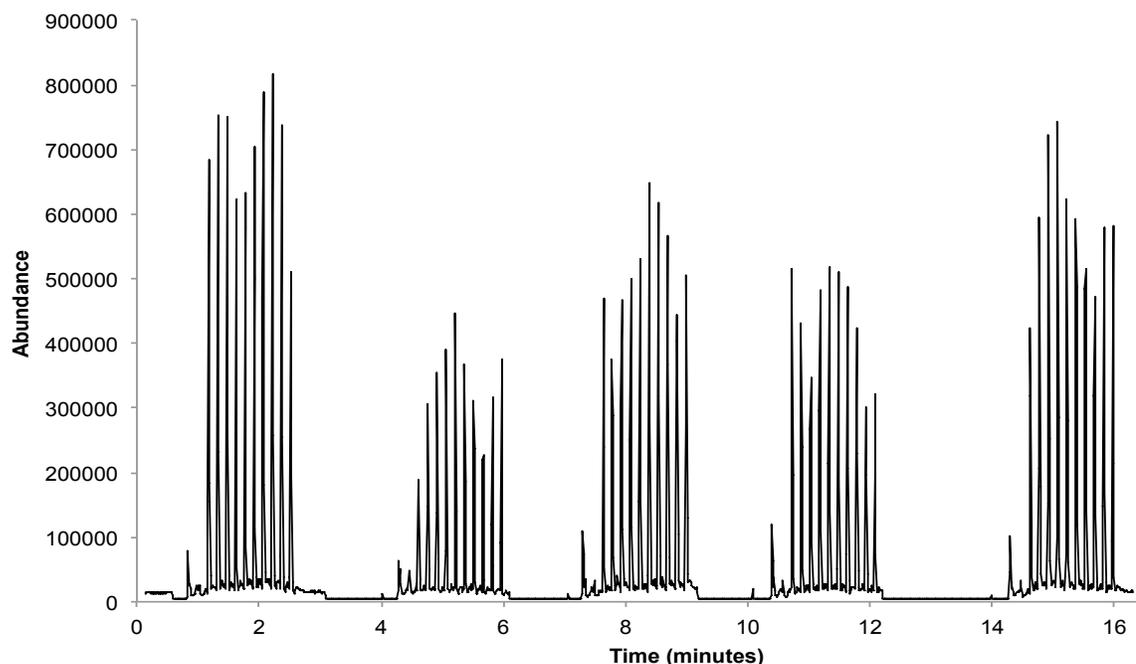


Figure 4. Total Ion Chromatogram of samples being analyzed by DART/MS analysis. Each peak corresponds to a unique sample and they are clustered in groups of twelve due to the nature of the twelve sample spots on a QuickStrip™ card (Figure 5).

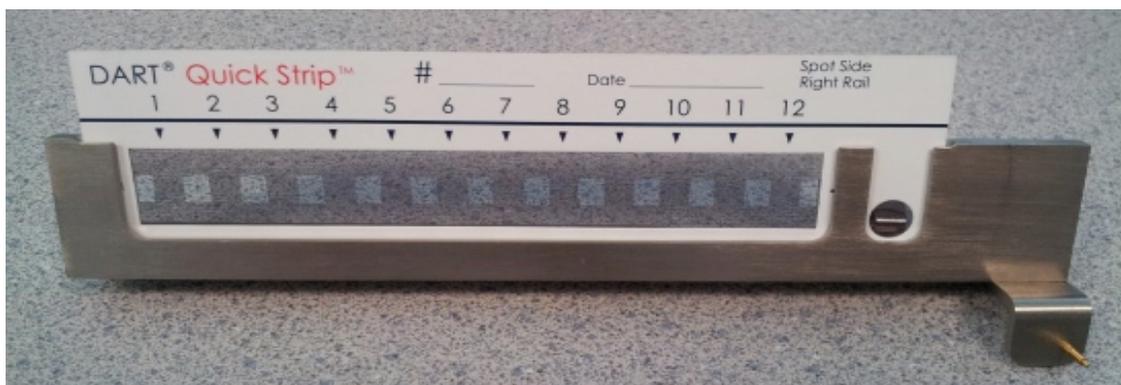


Figure 5. DART QuickStrip™ card in metal holder.

1.2 Direct Analysis in Real Time

Direct analysis in real time (DART) is a versatile ambient ionization technique. Invented in 2003, Cody et al. showed that an inherent advantage of metastable ambient ionization of analytes no longer requires the use of radioactive ionization sources²². DART is a very robust and versatile technique that has been shown to detect a wide variety of compounds including the detection of drugs^{23–26}, cocaine in urine²⁷, drugs from thin layer chromatography²⁸, chemical warfare agents²⁹, nerve agents^{30,31}, explosive residues³², bank dyes³³, pepper spray components³³, condom lubricants³⁴, synthetic cannabinoids^{35,36}, pesticides on produce³⁷, and ignitable liquids³⁸.

The DART ionization source used in this study consists of a carrier gas inlet, an atmospheric glow discharge region, a perforated grounded electrode, a gas heater region, a perforated grid electrode and a ceramic end cap on the atmospheric exit aperture (Figure 6). Through the formation of metastable species by a glow discharge, several advantages over other corona/glow discharge interfaces are realized. By preventing oxygen and other contaminants in the discharge region, degradation of the electrode is minimized and more easily maintained^{22,39,40}. Additionally, a grounded counter electrode can be utilized to remove ions from entering the gas heater region^{22,39,40}. The biasing of the perforated grid electrodes reduces the formation of unwanted recombination reactions from entering the ambient atmosphere^{40,41}. With positive or negative

values set for the grid electrode, the utilization of positive and negative ion modes for detection by mass spectrometry can be accomplished^{40,41}.

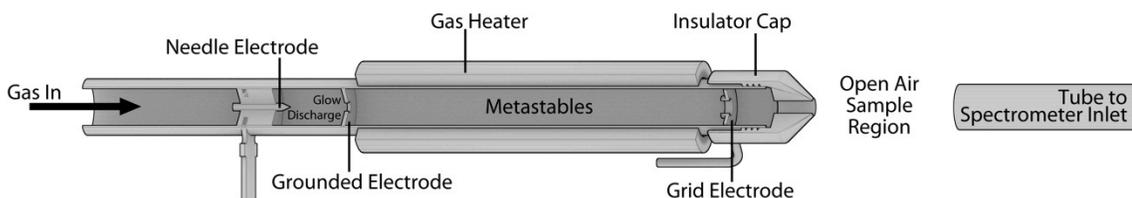


Figure 6. DART schematic of metastable formation by glow discharge.
Figure courtesy of IonSense Inc.

DART ionization occurs through a series of competing ionization techniques. The first step occurs through Penning ionization in which a long lived metastable (He^*) interacts with water vapor in the air to form a protonated water cluster and a hydroxyl radical⁴¹.



It should be stated that this reaction will occur with any analyte molecule in the atmosphere that has an ionization energy lower than the internal energy of the excited metastable⁴¹. The long-lived helium metastable has an internal energy of 19.8 eV, which is higher than the ionization energies of common atmospheric gases and organic molecules⁴¹. Proton transfer from the protonated water cluster to the sample (S in equation 2) will occur if the molecule has a higher proton affinity than the ionized water clusters⁴¹. Soft ion formation through proton transfer results in abundant intact molecular ions $[\text{M}+\text{H}]^+$, depending on the carrier gas temperature.



Equation 2

It should be stated that negative ion formation of electrophilic species through hydride abstraction is possible through this ionization technique, although testing under these conditions was not performed during this thesis^{22,39}. The heated sample region allows for the heating of the metastable species to permit thermal desorption of molecules from the sample surface.

1.2.1 Sample Introduction to DART Ionization Region

One of the significant advantages of any ambient ionization technique is the wider variety of ways to analyze samples than common chromatographic systems. With an ambient ionization source, the ability to introduce samples is not limited to liquid or volatile samples through injection or headspace sampling systems. Solid, liquid, and gas samples can all be analyzed with DART simply by introducing the sample into the ionization region. Most sample introduction systems interfaced to DART involve the sampling of molecules off of a substrate such as TLC plates²⁸, cardstock paper³³, glass capillaries^{27,30,35,42}, sorbent fabric^{25,33,37}, metal wire mesh⁴³, or the physical sample itself such as, drywall⁴⁴, tablets/pills^{23,24,26,36,45}, writing ink⁴⁶, condoms³⁴, and plant materials^{36,47}. All of these techniques allow for the direct analysis of a wide variety of samples without the need for any separation, concentration, or chromatographic techniques. One such set-up, a linear rail enclosure (LRE) developed by IonSense Inc. (Saugus, MA), allows for high throughput analysis. This is accomplished by the use of a computer-controlled holder that introduces the sample into the ionization region.



Figure 7. Linear Rail Enclosure of QuickStrip™ for sample introduction into the ionization region.

Specifically within the LRE is a removable sampling substrate that allows for the rapid analysis of multiple samples. The QuickStrip™ (Saugus, MA) consumable sampling device is a cardboard rectangle that sandwiches a sheet of wire mesh. The wire mesh is composed of consecutive sampling regions separated by empty gaps where the wire mesh has been removed. This allows for the easy application of up to twelve samples within the designated sampling regions on the card (Figure 5).

The linear rail slides horizontally into the sample ionization region directly between the DART ionization source and MS inlet (Figure 8). Computer control of the speed and position of the cardholder on the linear rail is possible through premade automated programs based on desired method development

parameters or in a free run mode involving the manipulation of speed and location via human interaction with the computer interface.

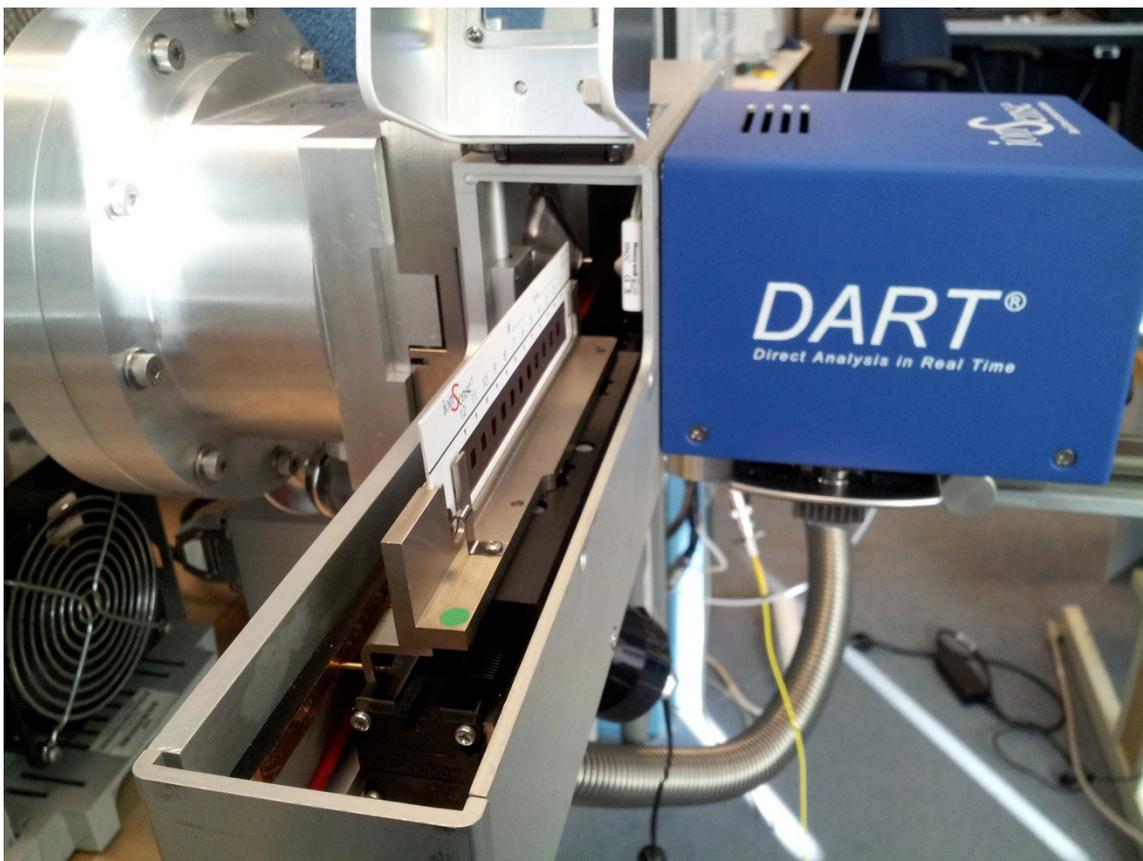


Figure 8. Linear Rail Enclosure attached between MS inlet (left of image) and DART ionization source (right of image) with QuickStrip™ card in holder.

1.3 DART Interface to Mass Spectrometer

Depending on the type of sample introduction and experimental considerations, the ionization source can theoretically be positioned at most angles (0 to 45 degrees) to the mass spectrometer inlet^{34,48}. Ionization source placements for angled injection of ions have been employed for analysis of liquid

samples in a 96-well plate⁴⁸. These DART set-ups attempt to increase the sensitivity and reproducibility of the analysis, while decreasing the effects and contribution of the sampling surface or matrix. The most common DART ionization source placement to the mass spectrometer is directly in-line with the mass spectrometer inlet (0 degrees off axis). By simple diffusion and pressure from the carrier gas, ions are carried into the mass spectrometer for detection.

1.3.1 Interfaced Ion Guide Optics to Mass Spectrometer

Prior to entrance into the mass spectrometer, there is a three stage vacuum system with a capillary inlet and a stage zero ion guide for optimum transfer of analytes into the mass selective detector (MSD)⁴⁹.

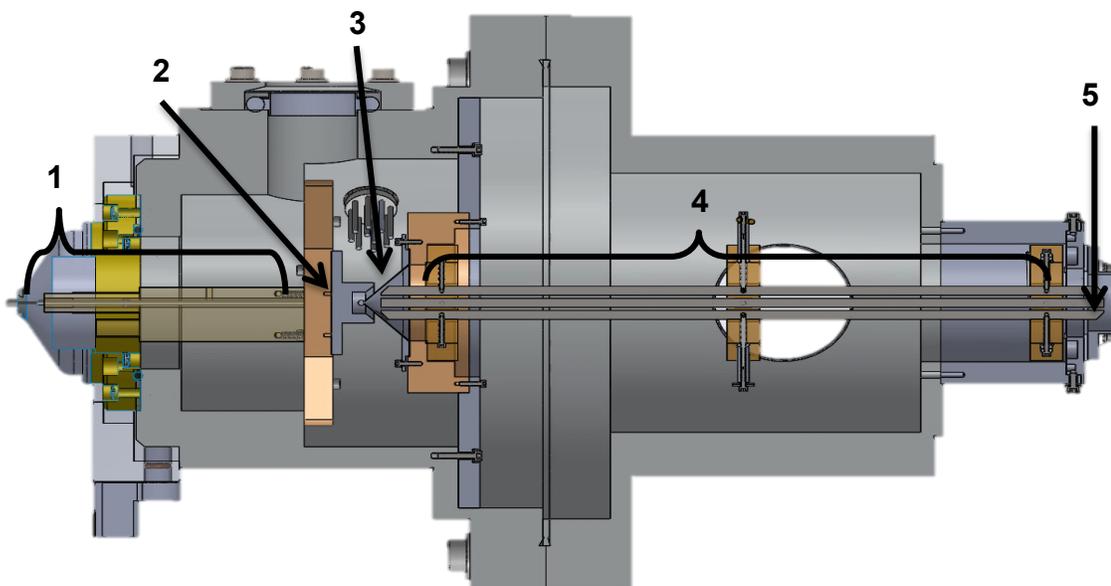


Figure 9. DART/single quadrupole MSD ion optics, where 1 is the heated capillary of the mass spec inlet, 2 is the skimmer focus lens, 3 is the skimmer focus, 4 is the quadrupole ion guide, and 5 is the ion guide exit lens. Figure courtesy of IonSense Inc.

DART allows for ambient ionization of molecules and formation of molecular ions similar to chemical ionization (CI), atmospheric pressure chemical ionization (APCI) and electrospray ionization (ESI) techniques. The ability to induce fragmentation without use of an additional collision gas is possible with the DART/MS platform utilized for this thesis. Fragmentation occurs by collisional induced dissociation (CID) in a voltage dependent fashion. This occurs by increasing the orifice and skimmer voltages of the ion optics in order to produce more energetic collisions with gaseous molecules inducing greater fragmentation (Figure 9, part 3)^{26,50,51}.

1.4 Mass Spectrometry

Integral to every mass spectrometer is the mass analyzer. Molecules are ionized prior to entrance into the mass analyzer in order to allow separation and detection of analytes. Using either a combination of electrical and magnetic fields or field free drift tubes, ions are separated based on their mass to charge ratio (m/z). The ability to differentiate between two analytes with similar mass to charge ratios is directly tied to the resolving power of the mass analyzer. There are numerous types and configurations of mass spectrometers that have been developed for the analysis of a wide range of substances. There is a wide range in price for a new mass spectrometer based on the type and specifications of the mass analyzer(s) within the mass spectrometer. Many mass spectrometers have been developed and are currently used that have inherent limitations and advantages. One of the simplest and most common mass analyzers in use today is the quadrupole. For the purposes of drug detection and identification in many routine analyses, a high-resolution mass spectrometer with exact mass identification is not required. Nominal mass data provides sufficient mass accuracy to confidently identify analytes routinely encountered in the controlled substance section of a forensic laboratory.

1.4.1 Quadrupole Mass Analyzer

A quadrupole consists of four hyperbolic or cylindrical shaped rods. They are mounted parallel in the z-axis and equally spaced apart creating a square in the x- and y-axis, which allows for the flight of ions to pass through the center of

the four rods along the z-axis. Opposing rods of the quadrupole are held at the same DC and AC potentials. Based on Mathieu stability region, an ion's trajectory will be either stable or unstable through the quadrupole based on its interaction with the AC/DC field. If an ion is stable (resonant) it will be guided through the poles and if the ion is unstable (non-resonant) it will collide into one of the poles and become neutral or be ejected between the poles and not reach the detector. By scanning through the ratio of DC to AC component, the ability to select specific m/z values allows for the separation of ions in a quadrupole. This scanning speed is on the millisecond scale per m/z value and allows for scanning hundreds of peaks in a few seconds. One way to increase the sensitivity of detection is to limit the scan range from hundreds of peaks scanned to less than ten. The decrease in peaks scanned increases the dwell time (time spent on a specific AC/DC value relating to a certain m/z) and is commonly referred to as Selective Ion Monitoring (SIM).

Although the quadrupole mass analyzer doesn't possess the highest resolution possible it has several key advantages. As the least expensive mass analyzer, most labs can afford to purchase and maintain instruments that contain this mass analyzer. Quadrupole mass analyzers have high transmission/scan speeds and are light weight and compact compared to time of flight and Fourier-transform ion cyclotron resonance mass spectrometers⁵². However, several disadvantages of quadrupole mass analyzers exist and should be mentioned. In addition to low mass accuracy, the high mass cut-off for a quadrupole is limited

to analytes of about m/z 2000. With these concessions, the quadrupole mass selector is an inexpensive mass selector capable of fast ion separations required of high-throughput platforms⁵². DART ionization interfaced to a single quadrupole mass spectrometer allows for faster analysis in comparison to chromatographic based mass spectrometry techniques (GC/MS and LC/MS) currently employed.

1.5 Analyze IQ Lab Software

Analyze IQ is an advanced chemometric software platform that allows for the identification of unknown samples. Through a model building approach unknown compounds can be qualitatively and quantitatively identified within complex mixtures. Statistical interpretation of data is an area of forensic science that is absent in many types of analysis. Error rates, correlation of variation, and measures of uncertainty are necessary to aid in the contextual evaluation of experimental results. The need for an expansive role of statistical interpretation within forensic science was heavily discussed in the 2009 National Research Council report on the state of forensic science in the United States¹⁸. Discussed as being currently lacking in scope and execution, statistical interpretation of data aids investigators by reducing bias and remaining objective in their conclusion¹⁸. The Analyze IQ platform allows for faster and easier interpretation of mass spectral data than human interpretation when analyzing mixtures.

1.5.1 Creating a Model within Analyze IQ

Analyze IQ Lab software allows for the interpretation of complex spectroscopy and spectrometry data using established chemometric techniques

and machine learning techniques^{53,54,55}. As shown in Figure 10, Analyze IQ Lab uses spectra from references samples for predictive model development. The pool of reference samples used to make a model is the most important factor in a predictive model's ability to correctly identify unknown spectra. A model should be created using the most appropriate and widest scope of reference samples possible. The inclusion of mixtures at a variety of concentrations that both contain and omit the analyte(s) of interest is crucial to the successful creation of a predictive model.

After a model is built, the user can apply different models to an unknown spectrum. Once a model is selected and applied to an unknown sample, the results are shown with a notification if the model identifies the unknown compound or not.

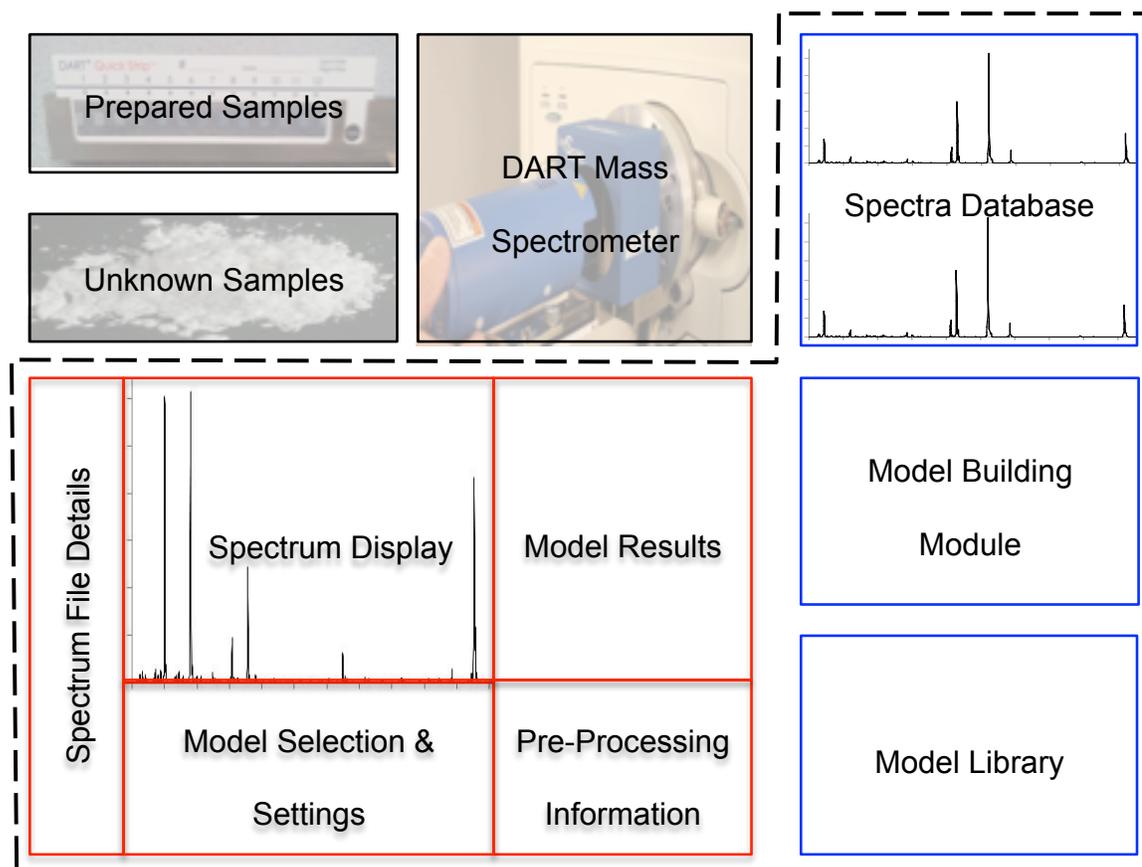


Figure 10. Analyze IQ Lab system architecture^{53,54,55}. Background operations of Analyze IQ are outlined in blue. Graphical user interface is outlined in red.

1.5.2 Pre-Processing of Data within Analyze IQ

Prior to selection of a chemometric or machine learning technique, the user has the ability to pre-process the data. There are several available options that can be performed such as local smoothing, applying an n th degree derivative, normalization, standard normal variate, equal area scale, internal standard normalization, and truncation/notching⁵⁴. Only two pre-processing techniques were utilized in this thesis: first order derivative and normalization.

Selection of a derivative for pre-processing computes a derivative of degree k over the curve with a sliding window. A zero order derivative would result in the local smoothing of the curve to remove small, high frequency fluctuations, while a first order derivative, allows for the removal of baseline effects^{54,56}. Normalization re-scales the data so that each spectrum's lowest and highest value is 0 and 1, respectively⁵⁴. This allows for reduced variances in mass spectral abundances, which can occur commonly with DART ionization platforms based on changes in gas flow rates and the ionization source to MS inlet distance⁴¹.

1.5.3 Principle Component Analysis

One of the biggest challenges in creating a predictive model is the effective use of large data sets to create models. Large data sets require longer computation times and can hinder the model's ability to differentiate between different sample sets. Through the use of a common chemometric technique, principle component analysis (PCA) can transform a set of possibly correlated variables into a set of uncorrelated (orthogonal) variables called principle components (PCs)^{57,58}. In PCA, reduction of data occurs with minimal information loss if enough principle components are utilized⁵⁵.

The first PC accounts for the greatest variation in the data set while the second accounts for the second largest amount of variation, etc.⁵⁶. The user can set the number of PCs, while the default value of Analyze IQ is set at 10 principle components. An increase in the number of PC should retain more of the data and

complexity thereby reducing the classification error rate of the model, but there isn't a linear correlation between the number of PCs and the classification error rate^{57,58}. Prior to selection and evaluation of a machine learning technique, a user can choose to apply principle component analysis on the reference data or proceed without PCA.

1.5.4 Analyze IQ Machine Learning Techniques

Within Analyze IQ Lab, there are 4 types of machine learning techniques: linear regression, support vector machine (SVM), k-nearest neighbor (k-NN), and spectral attribute voting (SAV). In linear regression a multiple least squares linear regression between the variables of the spectral dataset and the predictor variable occurs⁵⁴. When PCA is applied prior to linear regression, a multiple linear regression between the scores identified by PCA and the predictor variables is performed⁵⁸. This is more commonly referred to as principle component regression (PCR). The predictor variable is a value chosen by Analyze IQ to represent if the target is present (1) or absent (-1)⁵⁴.

Support vector machine is capable of representing non-linear relationships by projecting data into high dimensional space that will separate data based on a binary classification⁵⁷⁻⁵⁹. One advantage of SVM's and the most important method design choice is the selection of a kernel parameter. There are four different types of kernels that can be chosen within Analyze IQ. This kernel will define the structure of the high dimensional feature space where a maximal margin will be found. The greater the distance of the margin the greater the

ability for the model to separate between the binary classification (Figure 11)⁵⁸.

Two possibilities occur when a structure is created in the high dimensional feature space; too rich a feature space would cause overfitting of the model and allow for false negatives. Conversely, too poor a feature space would cause underfitting and allow for false positives^{58,59}. To regulate over/underfitting SVMs have a complexity factor (C), which determines the trade-off between choosing a classifier that might overfit the data and the amount by which misclassified samples are tolerated^{56-58,60}. This factor C can be represented as the tradeoff between the radius of the high dimensional feature space versus the number of data points accepted within the feature space (Figure 12).

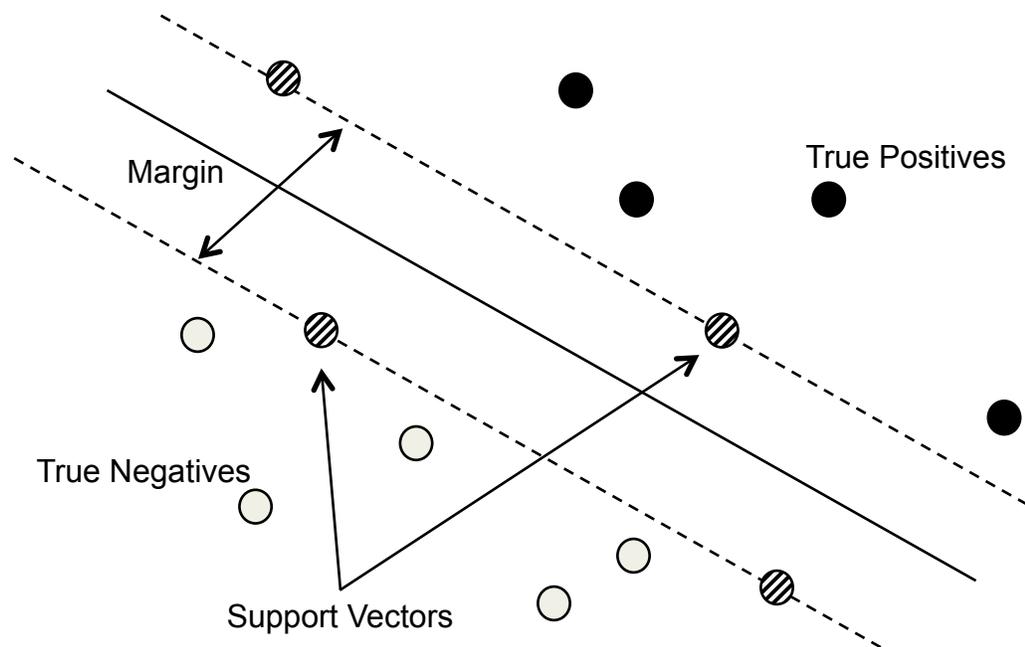


Figure 11. A linear SVM showing the marginal hyperplane used to classify positive (black) versus negative (white) samples where the Support Vectors (shaded) are on the boundary of the margin⁶⁰.

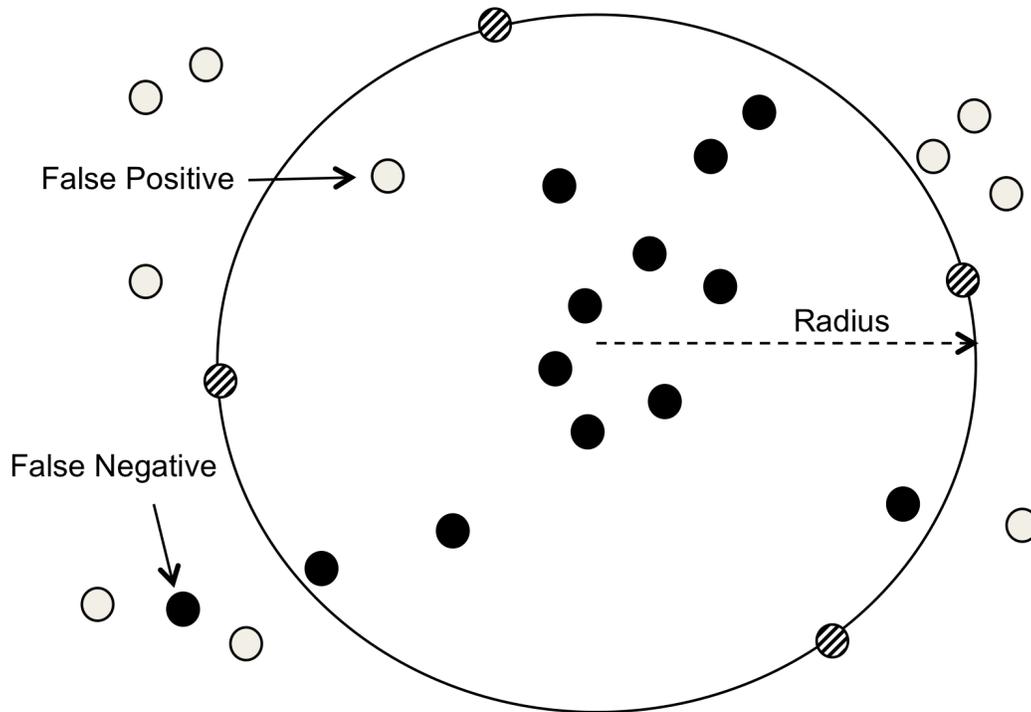


Figure 12. Hyper sphere containing target data with support vectors on boundary. The complexity factor C is represented by the radius of the sphere R versus the number of data accepted within. If C is high, R increases and the model might overfit the data and vice versa⁶⁰.

Within Analyze IQ there are 4 kernels: Linear, Polynomial, Radial Basis Function (RBF), and Sigmoid, where a polynomial kernel of degree 0 is a linear kernel⁵⁴. Specific to each kernel there are additional parameters that can be changed to further adjust the fit of the model to the data in the hopes of reducing classification error rates. SVM's allow for the analysis of high-dimensional data if that data is based on a binary classification system.

k-Nearest Neighbor (*k*-NN) is a learning algorithm that classifies an unknown spectrum by referencing the *k* closest reference samples (nearest neighbors) and returning a prediction. A value is assigned for each reference sample based on its similarity to the unknown sample. The difference in intensity at each spectral data point between the unknown and reference sample is squared and summed to give this value. *k*-NN works best for cases when *k* is odd, where the values of the reference samples (positive or negative for the target analyte) are averaged and assigned to the unknown. For example, if *k*=7, then the seven nearest neighbors would be compared to the unknown sample. If five of the nearest neighbors were positive for the target analyte and two were negative then the unknown sample would be classified as being positive for the target. In Analyze IQ, the user can choose the value of *k*, where the default value is three⁵⁴.

Spectral Attribute Voting is a proprietary machine learning technique developed by Analyze IQ that takes into account both the spectral profile (peaks and troughs) and locally correlated nature of spectra when generating a model⁵⁴. As an ensemble method, theoretically it should perform better than individual parameters in that errors arising from individual parameters should tend to cancel out when used together^{54,61}. SAV has been shown to yield lower error rates than established chemometric techniques and SVM's⁵³. SAV cannot be used for quantitative modeling, but works well for high dimensional data and requires no optimization⁵⁴.

There are three ways within Analyze IQ Lab that models can be evaluated (Testing, Import Test File, and Cross-Validation). Testing uses the entire data set to build the model and then tests the completed model against each reference sample in turn⁵⁴. This evaluation process is the fastest but is the least robust and prone to higher error rates than the other evaluation methods. Import test file is similar to Testing in that, the model is built with the entire reference data set, but the completed model is then tested against each element of a user imported test file. This test file can contain positive or negative reference samples, which can be similar or distinct from the reference pool.

Cross validation is an evaluation method that works by dividing the entire reference pool into x equal folds and then creating a model based on folds 2 through x (omitting fold 1) which is subsequently tested against fold 1. This is repeated again (fold 1 and 3 through x create a model, and then test against fold 2) and again until all folds have been tested against^{54,61,62}. Completion of this process is considered a “run” and the user can dictate how many runs to perform. Multiple runs would divide the reference pool into different initial folds, where the process of testing using all but one fold is then tested against the withheld fold. Errors are assigned for each incorrect assignment of the reference sample against the model, this is then averaged for each run and the average of each run error is computed to yield the final error rate. Although computationally intensive, this evaluation technique is the most robust technique and the best when compared to other evaluation techniques⁶¹.

1.6 Aim of This Thesis

Based on the number of case backlogs, greater public awareness of forensic science, and a large percentage of forensic casework, there is a need for higher throughput analytical techniques for controlled substance analysis. DART/MS provides an analytical platform for the identification of controlled substances faster than current GC/MS techniques by analyzing samples in 30 seconds rather than 30 minutes. Additionally, the use of statistical interpretation of data is advantageous within forensic science to reduce bias and aid in an analyst's interpretation of the results. Use of statistical software such as Analyze IQ can provide an unambiguous interpretation of experimental data for future forensic applications. The evaluation and efficacy of the DART/single quadrupole mass spectrometer was performed for the analysis of contrived cocaine samples. Identification of cocaine within unknown samples was performed by both human interpretation means, simulating that of a forensic analyst and by the use of predictive computer modeling software.

2 EXPERIMENTAL

2.1 Materials

Chemicals used in this thesis were purchased from the following manufactures: cocaine, phenacetin, acetaminophen, diltiazem HCl, and phenobarbital were obtained from Cerilliant (Round Rock, Texas, USA). Tetramisole HCL, xylazine, hydroxyzine, and atropine were obtained from MP Biomedicals (Santa Ana, California, USA). Diphenhydramine HCl, acetylsalicylic

acid, thiamine HCl, creatine, and inositol were purchased from Acros Organics (part of Thermo Fisher, Waltham, Massachusetts, USA). Mannitol, boric acid, caffeine, prilocaine and piracetam were purchased from Fischer Scientific, (part of Thermo Fisher, Waltham, Massachusetts, USA). Acetaminophen and creatinine were purchased from Sigma-Aldrich (St. Louis, Missouri, USA). For use in the cocaine double blind study chemicals were purchased from previously mentioned manufactures unless specified below. Phenacetin was purchased from ChemCruz™ (Santa Cruz, California, USA) and benzocaine was purchased from Chem Service Inc. (West Chester, Pennsylvania, USA).

The DART source was interfaced to a modified Agilent 5973 series single quadrupole mass spectrometer. MSD Chemstation® software, version E.02.02.1431, was used for mass spectrometer data collection and analysis. Contec CONSTIX® SP-5 polyurethane swabs (Figure 14) along with the QuickStrip™ cards (Figure 5) were utilized for sample introduction into the ionization region. The DART™ ionization source, Linear Rail Enclosure, and QuickStrip™ cards utilized during this thesis were provided by IonSense Incorporated (Saugus, MA).

2.2 DART Cocaine Analysis Methodology Development

The creation of an analytical method for the analysis of cocaine involved a stepwise optimization of parameters that could be separated into four different groups. These parameters as discussed below were optimized from left to right as shown in Figure 13. This approach was taken so that the first parameter

optimized had the greatest effect on the detection of cocaine and the second parameter optimized had the second greatest effect, etc.

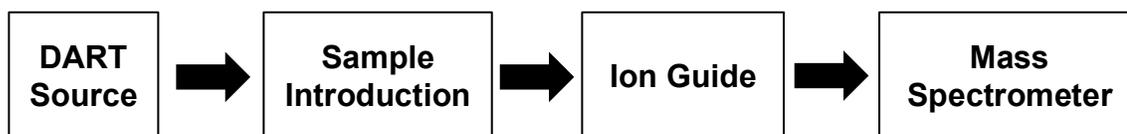


Figure 13. DART/MS block diagram.

2.2.1 DART Source Parameter Settings

All samples were tested utilizing the following DART source parameters. An atmospheric glow discharge is created with a voltage set initially at -6000 volts. Once the discharge is formed the current increases to maintain the discharge which subsequently decreases the voltage to approximately -1100 volts (personal communication, Doug Simmons). The grid voltage was set at 350 volts. Helium carrier gas was used during the analysis of samples, while nitrogen gas was used as the carrier gas during standby mode operations.

2.2.2 Sample Introduction Considerations

Several sample substrates were tested to determine the most efficient sample substrate carrier with the lowest background mass spectrum contribution. Contec CONSTIX® swabs and QuickStrip™ cards were tested for this purpose. For all samples, a computer guided linear rail (set between 0.8-1.0 mm/sec) introduced the samples into the DART ionization region.

For all testing using Contec swabs, a swab holder shown in Figure 14 was utilized to maintain the same horizontal and vertical spacing between each

sample. The distance from the DART source to the MS inlet was approximately 15 mm or less.

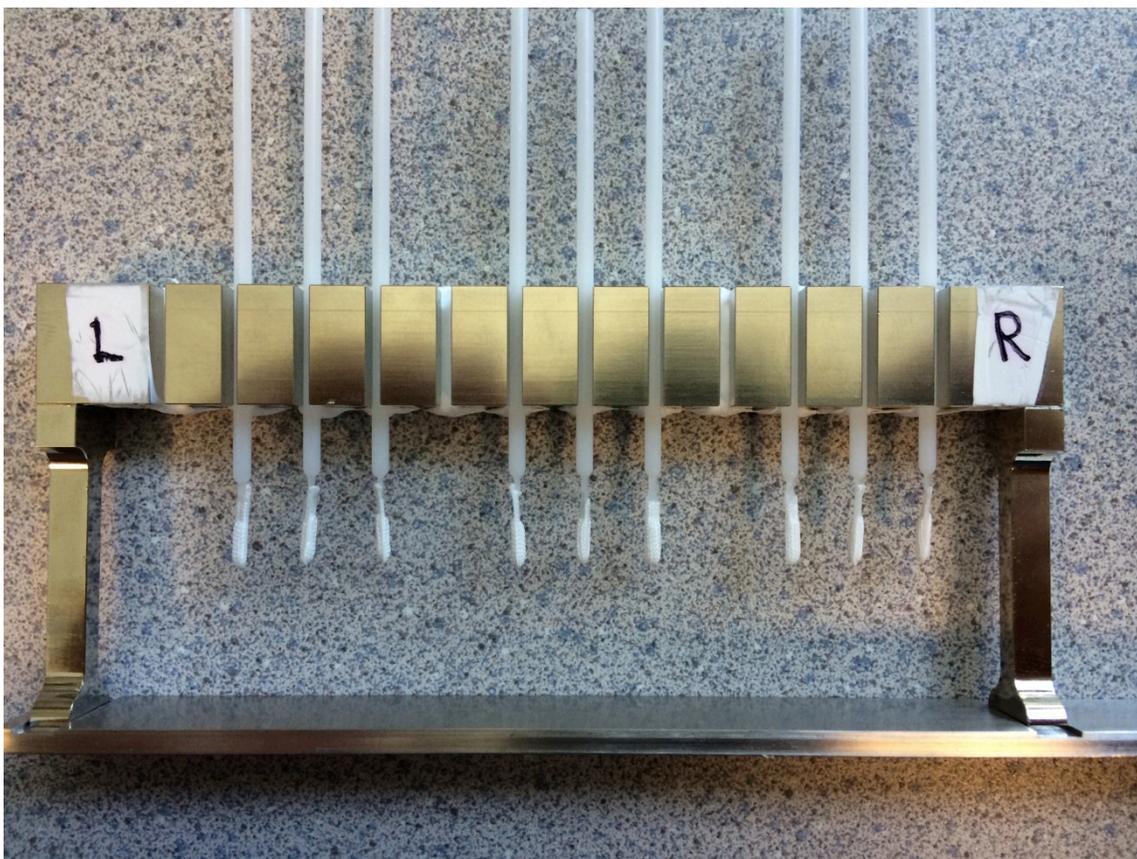


Figure 14. Contec CONSTIX® SP-5 polyurethane swabs in holder.

For all testing using the QuickStrip™, the strip was placed in the open metal holder of the Linear Rail Enclosure (LRE) as shown in Figure 7. A voltage of 250 V was applied to the QuickStrip™. The distance between the DART source and MS inlet was approximately 10 millimeters or less.

For either sampling substrate, 5 μ L of the mixture to be analyzed was pipetted onto the surface of the substrate and allowed to air dry before analysis.

2.2.3 Optimization of the DART Carrier Gas Temperature

The optimal DART carrier gas heater temperature was determined by analysis of 5 μL of cocaine at 100 ng/ μL spotted onto the QuickStrip™ at 100, 200 and 300°C, performed in triplicate. The linear rail speed was set at 1.0 mm/s.

2.2.4 Optimization of the DART Skimmer Voltage

The optimal skimmer focus voltage was determined by analysis of 5 μL of cocaine at 100 ng/ μL spotted onto QuickStrip™ cards, performed in triplicate. The skimmer focus voltage was set at 12 V and increased to 54 V in steps of 6 volts.

2.3 Ion Optics Parameters

The ion optics (Table 2) were held constant, while the skimmer focus was adjusted to induce fragmentation (12 V for minimal fragmentation, 54 V for greatest fragmentation). The capillary temperature was set at 150 °C, while the R_f of the quadrupole ion guide prior to the mass analyzer was held between 490 and 500 volts.

Table 2. Ion optic voltages between MS inlet and single quadrupole MSD.

Name	Voltage (volts)
Capillary Inlet	79
Skimmer Focus Lens	60
Skimmer Focus ^a	12 or 54
Quadrupole Ion Guide Offset	8
Quadrupole Ion Guide Exit Lens	5
Quadrupole Ion Guide Focus Lens	-299
Mass Analyzer Entrance Lens	-27

^aSkimmer focus adjusted to induce more or less fragmentation of analytes.

2.4 Mass Spectrometer Method Parameters

With no solvent delay, the data was collected in raw scan mode from m/z 40 to 430, (1.24 scans/second) at a threshold of 0 counts and a sampling rate of 2^x where x equaled 2. An absolute EM voltage was set at 1235 V with an MS source and MS quadrupole temperature set at 150°C and 200°C, respectively.

2.5 Optimized Cocaine Parameters

The optimized parameters determined for the identification of cocaine in excipients (ICE) method were as follows: DART source parameters as stated in section 2.2.1 *DART Source Parameter Settings*, DART heater temperature set at 300°C using a QuickStrip™ placed in the LRE with an applied voltage of 250 volts to the strip. Ion optic parameters were set at values given in Table 2, and mass spectrometer parameters were set at values given in section 2.4 Mass Spectrometer Method Parameters

2.6 Cocaine Mass Spectral Peak Ratio Testing

Variability in the ratio of abundance between the molecular ion and a fragment ion for cocaine was determined. A dilution series of 1:1, 1:2, 1:5, and 1:10 (cocaine : excipient) was created to determine the effects of increasing concentration of excipient on the ion abundances at m/z 304 and 182.

Levamisole and procaine were used to model the extreme differences in the amount of fragmentation seen in the excipients. Further testing was performed with the addition of cocaine to increasingly complex excipient mixtures. Data collected from section 2.8 Analysis of Multi-Component Cocaine Mixtures testing was utilized for this purpose.

2.7 Cocaine Excipient Library Development

All of the excipients tested in this thesis were analyzed with the ICE method and the skimmer focus set at 12 volts and 54 volts for minimal and greatest fragmentation, respectively. Five microliters of an excipient at 100 ng/ μ L was pipetted onto either a Contec CONSTIX® swab or QuickStrip™.

2.8 Analysis of Multi-Component Cocaine Mixtures

Multi-component mixtures spiked with cocaine were made in methanol so that the final concentration of each component (including cocaine) was 10 ng/ μ L. Three, six, nine, twelve and fifteen component mixtures were made in increasing complexity. The twelve component mixture contained the same analytes as the fifteen component mixture, but with the removal of three excipients, as shown in Table 3. The nine component mixture contained the same analytes as that of the

twelve component mixture, but with the removal of three more excipients. This process was continued for the six and three component mixtures. Each multi-component mixture was analyzed using the ICE method.

Table 3. Multi-component cocaine mixtures^a.

Number of Components	Cocaine	Tetramisole	Prilocaine	Benzocaine	Procaine	Lidocaine	Caffeine	Diphenhydramine	Thiamine	Piracetam	Phenacetin	Acetaminophen	Hydroxyzine	Acetyl salicylic acid	Diltiazem
3	X	X	X												
6	X	X	X	X	X	X									
9	X	X	X	X	X	X	X	X	X						
12	X	X	X	X	X	X	X	X	X	X	X	X			
15	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

^aEach component was at a final concentration of 10 ng/μL.

2.9 Analyze IQ Cocaine Identification Development

2.9.1 Creation and Analysis of Reference Samples

Models were built for the identification of cocaine using a DART/ single quadrupole mass spectrometer collected in raw scan mode from m/z 40 to 430. Analyze IQ requires that each reference spectrum contain the same amount of data points, therefore the mass spectrometer was set in raw scan mode during the data collection process⁵⁴. Three hundred eighty-one reference samples (120 positive, 261 negative) were collected using the ICE method. These mass spectra were imported into a single comma separated value (.csv) file.

2.9.2 Pre-Processing Techniques

Prior to the application of each machine learning technique, pre-processing and PCA was carried out on the reference sample data. Two pre-processing steps, a first derivative using the default values within Analyze IQ followed by normalization, were used in conjunction. A default value of 10 PCs was used for all models.

2.9.3 Analyze IQ Model Creation and Development

A predictive model was made using each machine learning technique at default values for each parameter set within Analyze IQ. Each model was evaluated using a 5 run 10 fold cross-validation. The models were then tested against seventy-two samples (49 positive, 23 negative) collected using the ICE method. The model with the lowest testing error rate was selected for further optimization of parameters within that machine learning technique. For the SVM/RBF/Euclidean Distance method the complexity factor and kernel width were adjusted in factors of 10 from the default values^{57,58,63}.

2.10 Cocaine Double Blind Study

A cocaine double blind study was performed for the evaluation of the DART/single quadrupole instrumental platform and subsequent identification by human interpretation and Analyze IQ optimized predictive models. A double blind methodology was used to eliminate confirmation bias on part of the analyst running the samples and ensure the closest approximation to real world cocaine samples that would be submitted to a forensic laboratory. Each sample (excipient

and cocaine) was purchased or prepared at 1 mg/mL concentration prior to use in the double blind study. The 24 excipients were divided randomly in half by two students in the Boston University Biomedical Forensic Science Program. Five mixtures each of a three, five, nine, fifteen, and twenty-one component mixtures were created. Cocaine was spiked at random in sixteen of the twenty-five total samples. The final concentration of cocaine within each spiked solution was greater than 10 ng/uL. The identities of the twenty-five unknown mixtures were revealed to the analyst only after instrumental detection, data analysis, and both human and computer identification methods were completed.

3 RESULTS AND DISCUSSION

3.1 DART/Single Quadrupole Method Parameters for Cocaine

3.1.1 DART Carrier Gas Temperature Optimization

Cocaine was analyzed at a DART source heater temperature of 100, 200, and 300°C, with the skimmer focus set at 12 volts. All other parameters were set using the ICE method. The results from the EIC of m/z 304 show maximal peak abundance at 300°C with best peak shape, separation and reproducibility compared to the other heater temperatures (Figure 15).

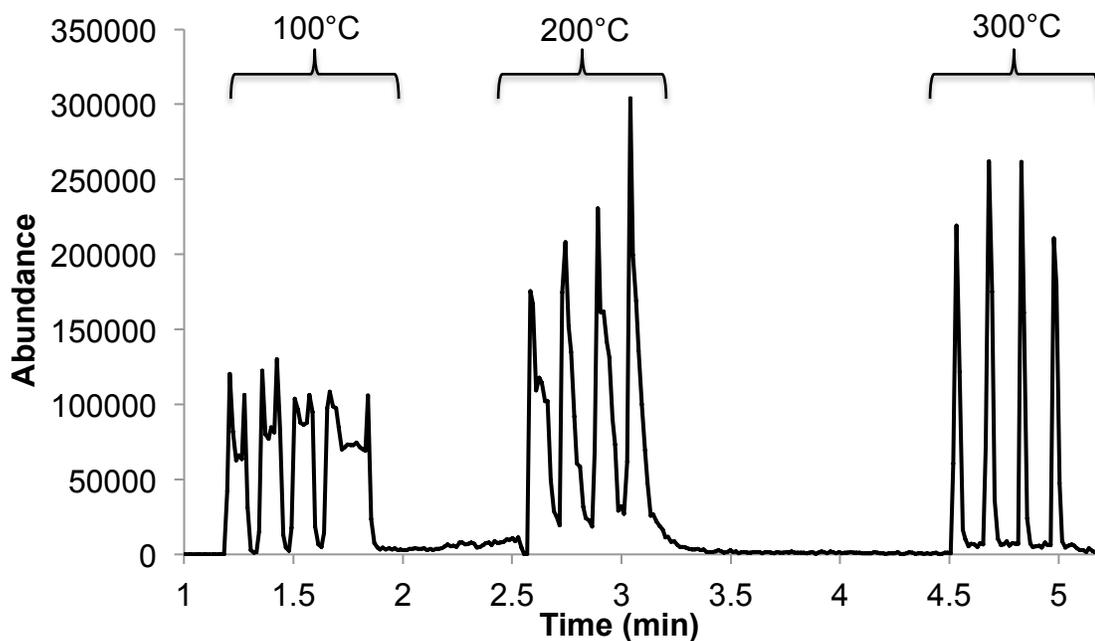


Figure 15. Extracted Ion Chromatogram at 304 m/z of cocaine at 100 ng/uL analyzed with increasing DART heater temperatures.

3.1.2 Analysis of Cocaine by DART/Single Quadrupole MS

When measured using the ICE method, the mass spectrum of cocaine was determined to have a difference in the number of peaks and ratio of peak abundance depending on the amount of induced fragmentation. For a skimmer focus of 12 volts, cocaine gave minimal fragmentation with only one abundant peak from the molecular ion of cocaine at m/z 304. For a skimmer focus set to 54 volts, additional fragmentation occurred with two abundant peaks at m/z 304 and 182.

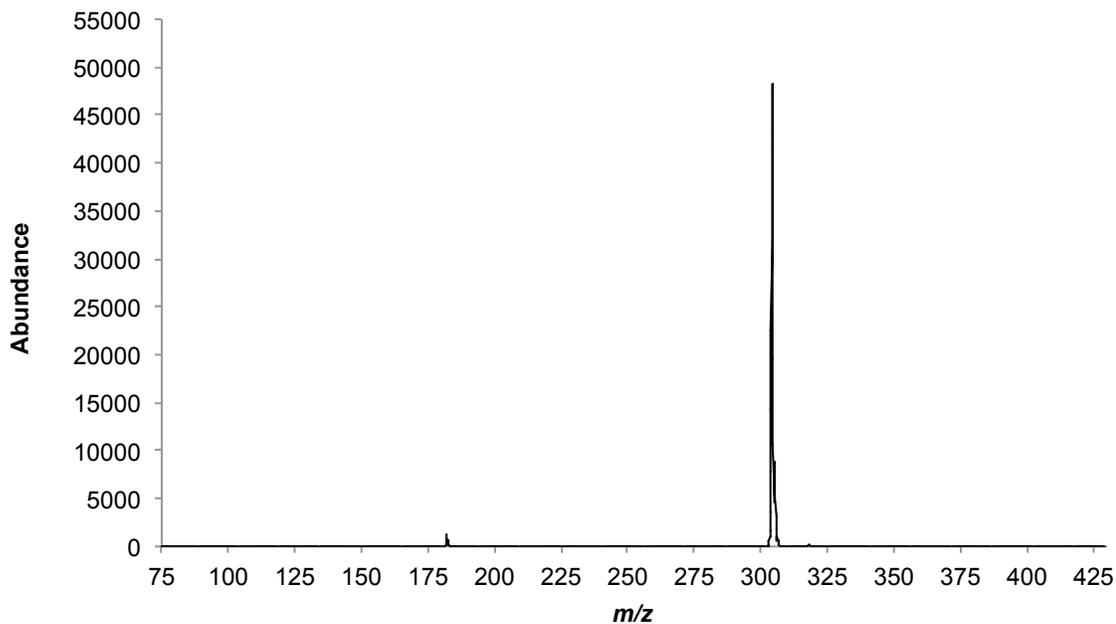


Figure 16. Mass spectrum of cocaine at 100 ng/uL analyzed by DART/single quadrupole MSD. Skimmer focus set to 12 volts.

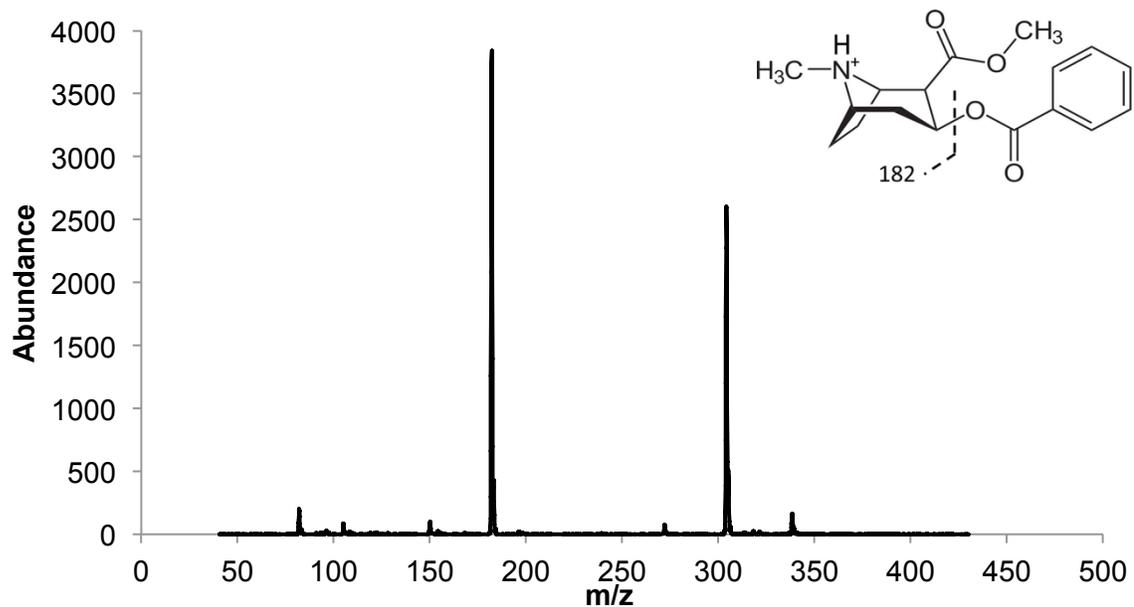


Figure 17. Mass spectrum of cocaine at 100 ng/uL analyzed by DART/single quadrupole MSD with in-source fragmentation. Skimmer focus set to 54 volts. Inset of cocaine in-source fragmentation mechanism by DART/single quadrupole mass spectrometry⁶⁴.

Mass spectral detection of cocaine is possible through the molecular ion and fragment peak of cocaine (at m/z 304 and 182) with the use of in-source fragmentation (Figure 17). This fragment peak at m/z 182 is due to the loss of benzoic acid (inset of Figure 17)^{6,7,64}. Minor peaks are also seen and could be from additional fragmentation of cocaine (m/z 82) as these are common fragments seen in a CID mass spectrum (analyzed by a DART/ion trap, data not shown) and in an EI mass spectrum of cocaine (Figure 18). Other minor peaks are likely from DART atmospheric contribution and other background contaminants such as erucamide at m/z 338 (personal communication, Joseph LaPointe).

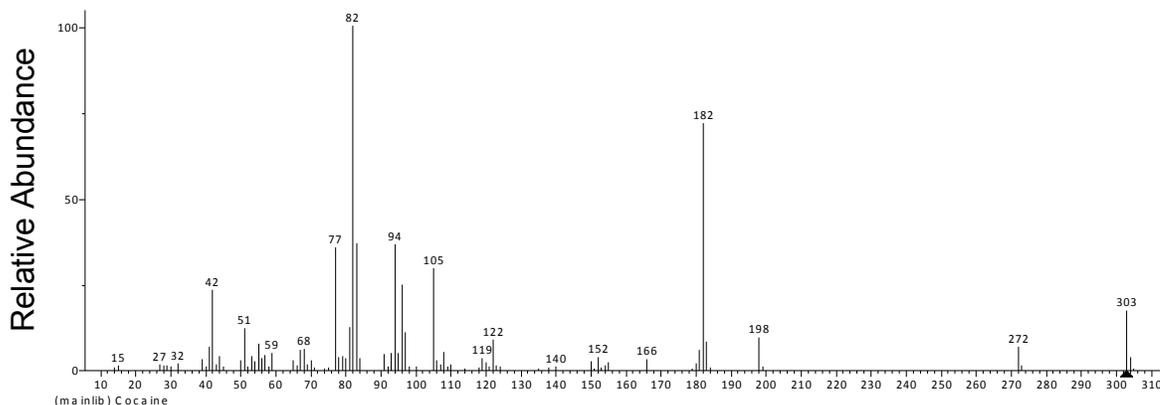


Figure 18. Electron impact mass spectrum of cocaine acquired from NIST Library (Version 2.0, December 2005) with m/z on the x-axis.

3.1.3 DART Skimmer Voltage Optimization

Optimization of the skimmer voltage was determined by finding the voltage that induced the maximum abundance of the cocaine fragment peak at m/z 182

(Figure 17). A 100 ng/ μ L cocaine standard was analyzed with a DART heater temperature of 300°C and a variable skimmer focus voltage in duplicate. The skimmer focus voltage was increased in ten volt increments from 20 volts to 50 volts. The ion abundances from duplicate samples were averaged and normalized to the average $[M+H]^+$ abundance of the eight samples tested. An overlay of the average mass spectrum for each voltage is shown in Figure 19. The data is presented ± 2 amu of the center mass (m/z 182 and 304) to enhance the differences seen.

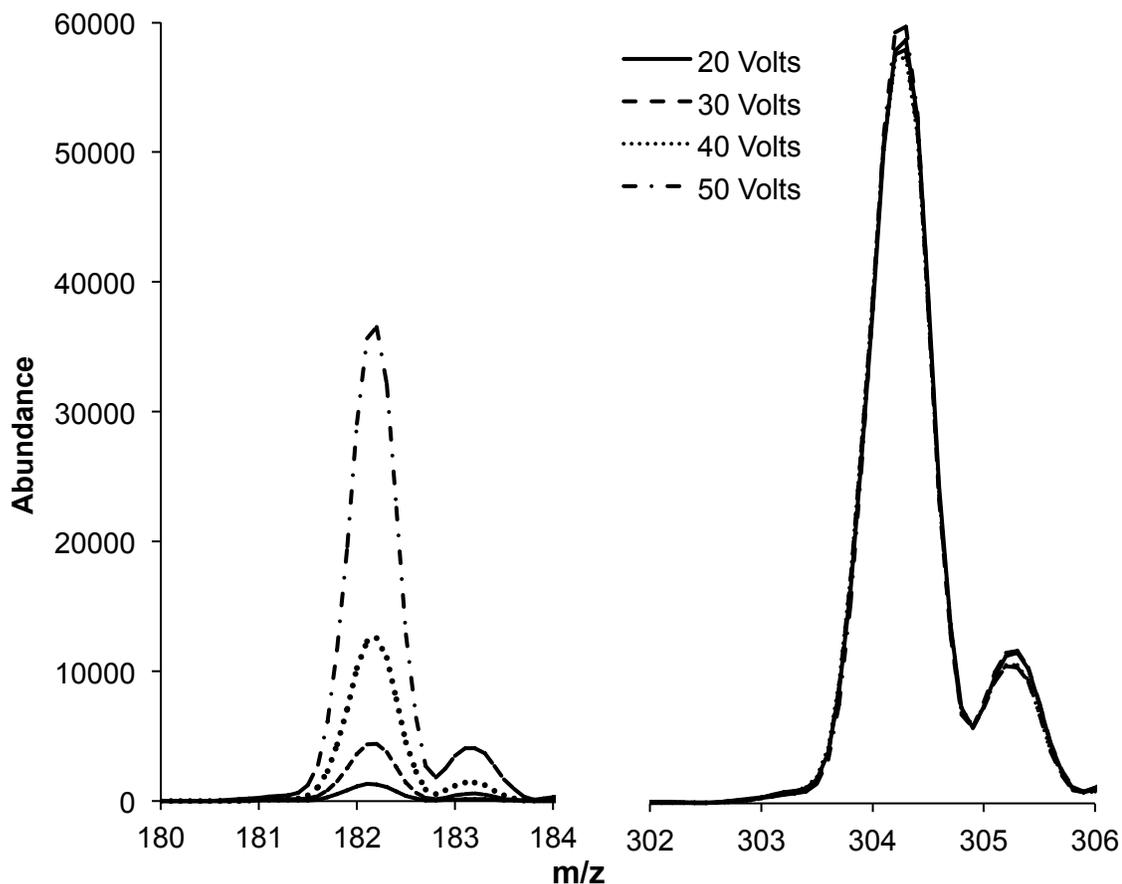


Figure 19. In-source fragmentation of cocaine at different skimmer focus voltages.

Further optimization of the skimmer focus voltage increased the abundance of a fragment ion of cocaine (m/z 182) at a setting of 54 volts. At this setting the ratio of abundance between the molecular ion and the fragment ion of cocaine was close to a 1:1 ratio as previously shown in Figure 17.

Through the manipulation of ion optic voltages, specifically the skimmer focus of the quadrupole ion guide, the ability to fragment in a reproducible manner was established for cocaine. Similar to the electron voltage setting of an

El mass spectrometer (usually 70 eV), the amount of fragmentation can be changed in a voltage dependent fashion with this DART/MS platform.

3.2 Cocaine Excipient Library Analysis

As shown in Table 4, twenty-four common cocaine excipients were analyzed using the ICE method on the DART/MS platform. Minor contributors were defined as any peak that was quantitatively less than half that of the largest peak detected for any given excipient. Of the 24 adulterants tested, no peaks were detected for boric acid, inositol, mannitol, phenobarbital, and xylazine when tested as individual components.

Table 4. Common cocaine excipients tested by DART/ single quadrupole mass spectrometry using in-source fragmentation

Compound	Molecular Weight ^a	Peaks Detected ^b (minor contributors)
Acetaminophen	151.06	152 , 110
Acetylsalicylic acid	180.04	121
Atropine	289.17	290 , (124, 96)
Benzocaine	165.08	(166) , 138, 120, 94
Boric Acid	62.02	<i>no peaks detected</i>
Caffeine	194.08	195 , 138
Creatine	131.07	(132) , 114
Creatinine	113.06	114
Diltiazem	414.16	415 , 178, (370, 312, 310, 223, 137, 72)
Diphenhydramine	255.16	(256) , 167
Hydroxyzine	374.18	(375) , 201, (203)
Inositol	180.06	<i>no peaks detected</i>
Levamisole	204.07	205 , (178), (146)
Lidocaine	234.17	235 , 86
Mannitol	182.08	<i>no peaks detected</i>
Phenacetin	179.09	180 , (138, 110)
Phenobarbital	232.08	<i>no peaks detected</i> (+) ion mode 231 , 188 (-) ion mode
Piracetam	142.07	(143) , 126, 98
Prilocaine	220.16	221 , 86
Procaine	236.15	237 , 164, 120, 100, (118, 152, 72)
Quinine	324.18	325, (136)
Tetramisole	204.07	205 , (178), (146)
Thiamine	265.11	(265) , 144, (154, 122, 81, 113)
Xylazine	220.10	<i>no peaks detected</i>

^aMonoisotopic mass

^b**Bold** fragments are molecular ion peaks

Phenobarbital did not produce any positive ions but could be detected by DART/MS in the negative ion mode. This was performed on a DART/LCQ Deca using an ion trap in negative ion mode. Phenobarbital has also been shown to be

detectable by GC/MS methods⁶⁵⁻⁶⁹. Xylazine has been previously detected using EI/GC/MS and ESI/LC/MS (positive ion mode) methods. The inability to detect xylazine could be due to the inability to ionize xylazine by DART. The inorganic compounds tested and carbohydrates, such as inositol and mannitol, did not produce any detectable ions by DART/MS and could be due to an inability to readily ionize these analytes as well.

Only two pairs of compounds, creatine/creatinine and levamisole/tetramisole could not be differentiated by mass spectrometry when measured using the ICE method. A peak in the mass spectrum at m/z 114 was observed for both creatine and creatinine. Creatine, at 18 mass units greater than creatinine, can undergo an intramolecular condensation reaction with the loss of H_2O and form creatinine (Figure 20)^{51,70}. Creatine is thermally labile and can undergo this reaction within the ionization region or mass spectrometer inlet at a DART gas heater temperatures of $300^\circ C$ ⁷¹. Two additional adulterants, levamisole and tetramisole, are enantiomers and cannot be separated solely by mass spectrometry means. Separation would have to occur on a chiral GC column or by other analytical means such as polarimetry.

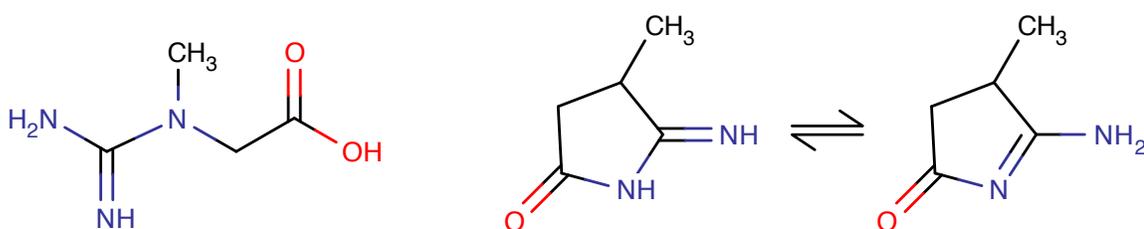


Figure 20. Molecular structure of creatine (left) and creatinine imine (middle) and amine (right) tautomers.

Comparison of DART/MS fragmentation data to that of the National Institute of Standards and Technology (NIST) mass spectral library (Version 2.0, December 2005) showed high correlation between DART and EI fragmentation data even though the mechanism of fragmentation differs (Table 5). This data suggests that when energy is introduced into the system, molecules will fragment in similar methodologies (such as breaking of the weakest bond and other thermodynamic properties) regardless of how the energy was introduced (CID versus EI methodologies).

Table 5. Comparison of DART/MS peaks to the NIST EI peaks.

Compound	Peaks Detected (minor contributors) ^a	NIST Library Peaks ^a
Acetaminophen	152, 110	109 151 43 80 108
Acetylsalicylic acid	121	120 138 43 92 121
Atropine	290, (124, 96)	124 82 83 94 42
Benzocaine	(166), 138, 120, 94	120 165 92 65 137
Boric Acid	<i>no peaks detected</i>	44 62 44 61 43
Caffeine	195, 138	194 109 55 67 82
Cocaine	304, 182	82 182 83 94 77
Creatine	(132), 114	42 43 113 84 112
Creatinine	114	113 42 43 112 84
Diltiazem	415,178, (370,312,310,223,137,72)	58 71 72 121 150
Diphenhydramine	(256), 167	58 73 165 42 45
Hydroxyzine	(375), 203, 201	201 165 203 299 202
Inositol	<i>no peaks detected</i>	73 60 71 102 43
Levamisole	205, 178, 146	148 204 73 101 203
Lidocaine	235, 86	86 58 30 87 42
Mannitol	<i>no peaks detected</i>	73 61 103 74 43
Phenacetin	180, (138, 110)	108 109 179 137 43
Piracetam	(143), 126, 98	98 70 84 41 42
Phenobarbital	<i>no peaks detected</i>	204 117 115 161 205
Prilocaine	221, 86	86 44 87 43 56
Procaine	237, 164, 120, 100, (118, 152, 72)	86 99 120 164 92
Tetramisole	205, 178, 146	204 148 73 101 203
Quinine	325, (136)	136 41 137 42 81
Thiamine	144, (265, 158, 154, 122, 81)	112 122 264 143 113
Xylazine	<i>no peaks detected</i>	205 220 130 145 177

^aPeaks listed in order of decreasing abundance. NIST Version 2.0 December, 2005

3.3 Multi-Component Cocaine Mixture Analysis

To determine the effects on the peak abundance of m/z 304 and 182 within mixtures, a series of multi-component excipient mixtures spiked with cocaine were tested and analyzed by the ICE method with a skimmer focus of 54

V to induce fragmentation as shown in Figure 21 through Figure 25. The final concentration of each component (including cocaine) was 10 ng/ μ L. The three, six, nine, twelve, and fifteen component mixtures were made in increasing complexity, as shown in Table 3. The twelve-component mixture included the same analytes as the fifteen component mixture, but with the removal of three excipients. The nine component mixture contained the same analytes as that of the twelve component mixture, but with the removal of three more excipients. This process was continued for the six and three-component mixtures. Using a stepwise addition of excipients allows for the ability to monitor ion suppression, and appearance of additional peaks to be attributed to the added excipients. Shown in Figure 21 through Figure 25 are the mass spectra from the multi-component mixtures in increasing complexity (three, six, nine, twelve and fifteen components) with identification of cocaine and excipients in the mass spectrum. The ability to consistently detect both peaks attributed to cocaine (m/z 304 and 182) at abundances greater than 10x the signal to noise ratio was demonstrated in every multi-component mixture tested.

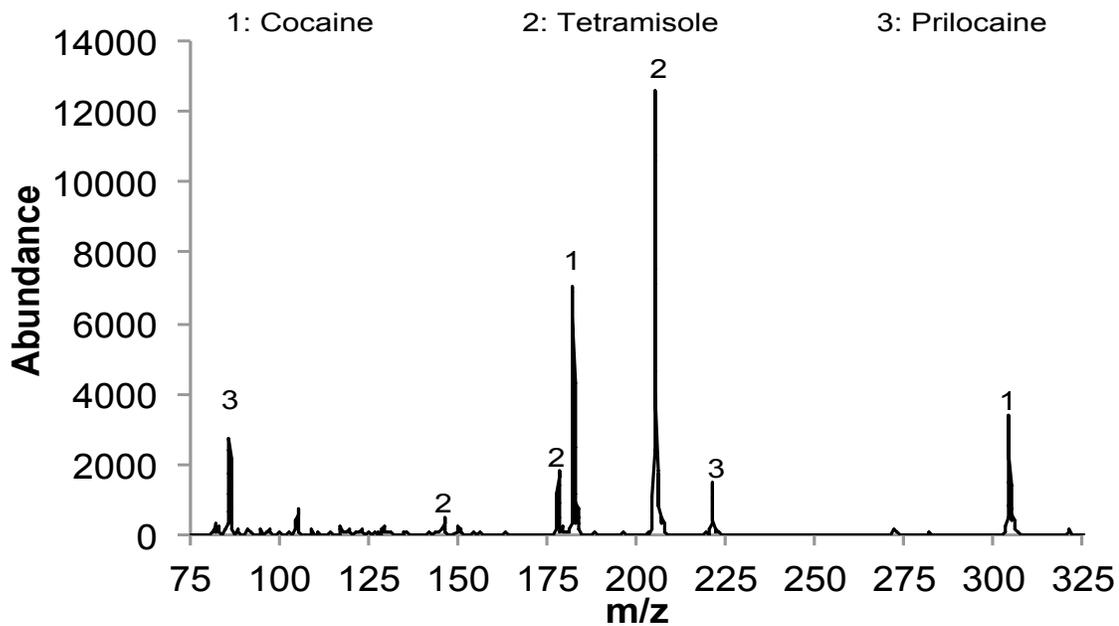


Figure 21. Induced fragmentation of a 3-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L.

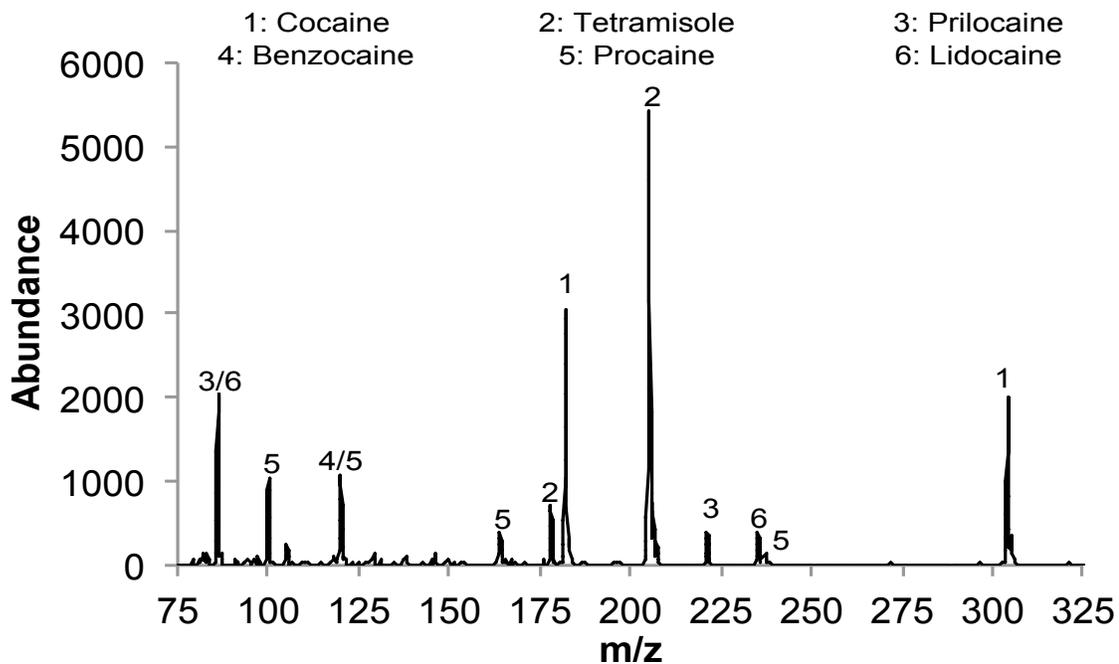


Figure 22. Induced fragmentation of a 6-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L.

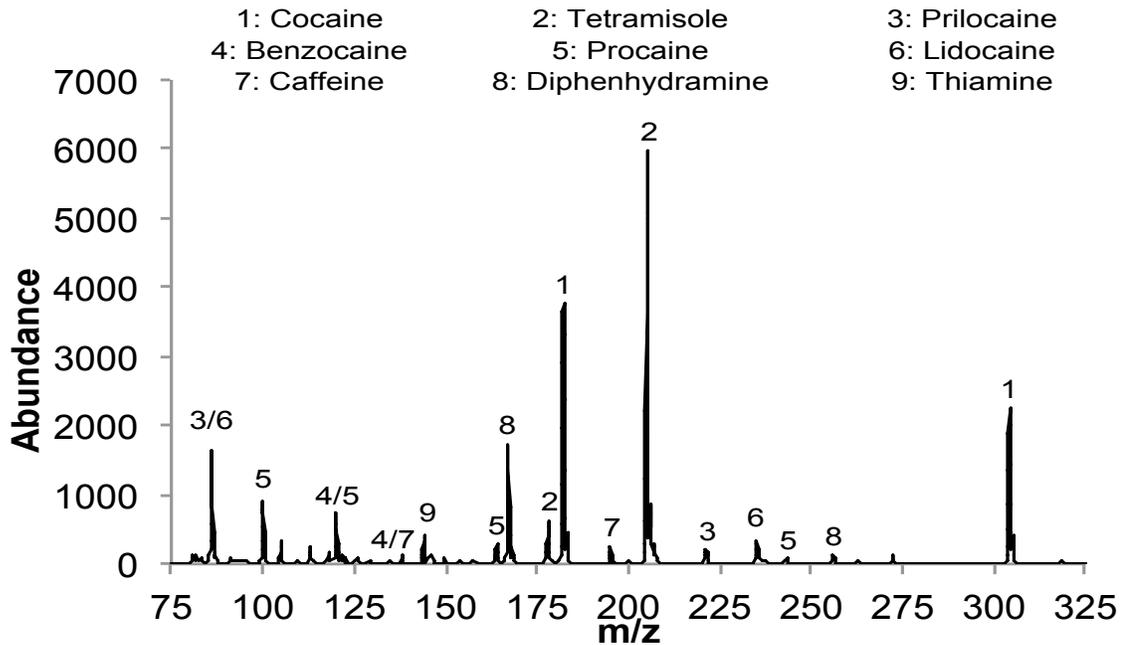


Figure 23. Induced fragmentation of a 9-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L.

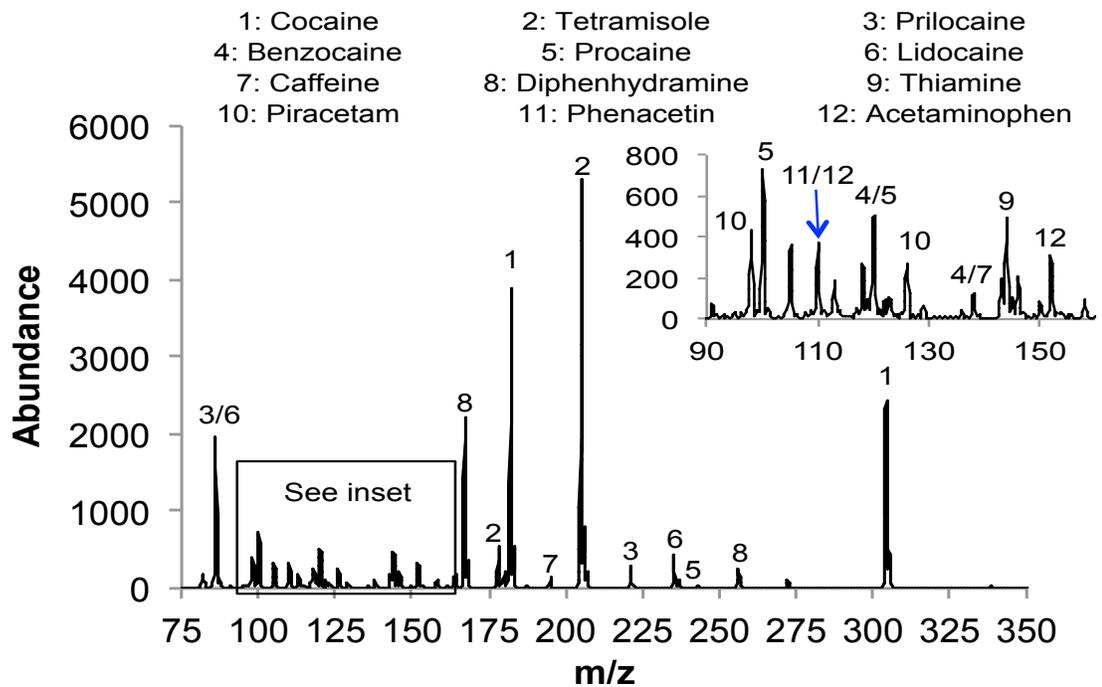


Figure 24. Induced fragmentation of a 12-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L.

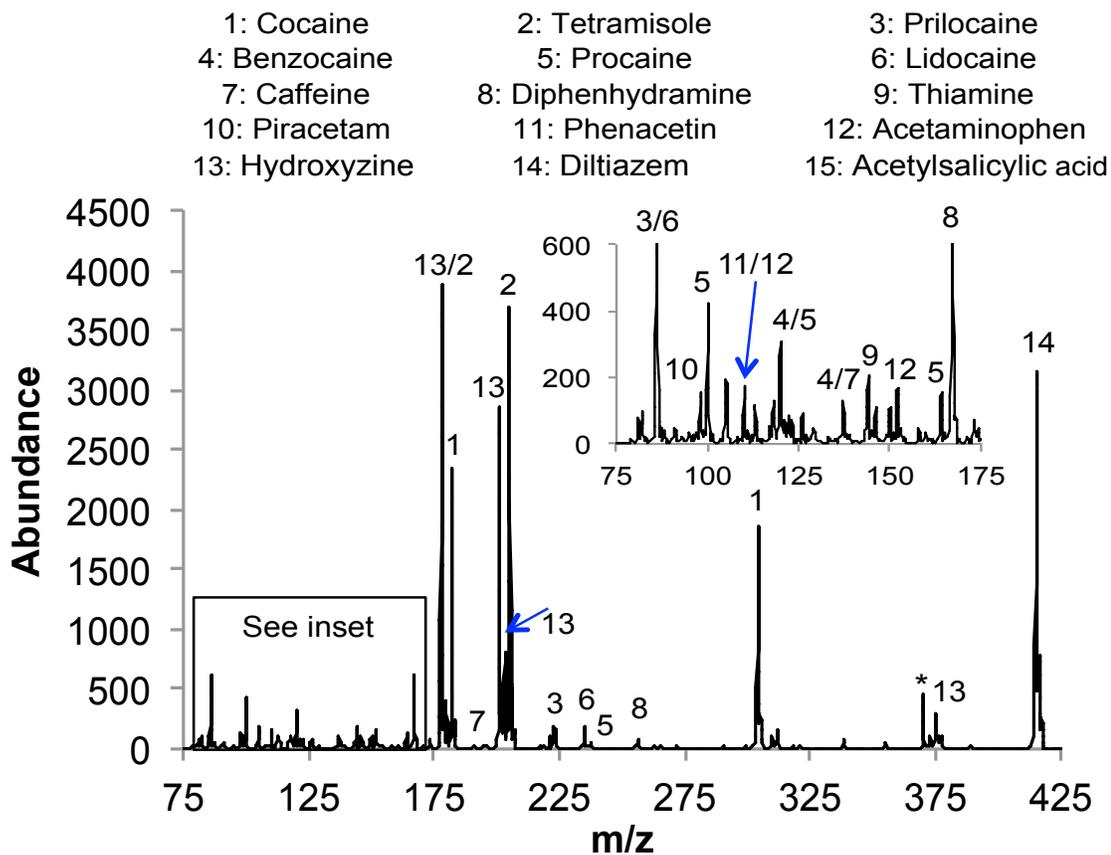


Figure 25. Induced fragmentation of a 15-component cocaine mixture. Each component is at a final concentration of 10 ng/ μ L. The asterisk refers to atmospheric/background contribution of decamethylcyclopentasiloxane at m/z 371 (personal communication, Joseph LaPointe).

Acetylsalicylic acid was not detected in the 15-component mixture.

However, a peak at m/z 121 was observed for pure acetylsalicylic acid by DART/single quadrupole analysis using the ICE method (Table 4). The loss of ion abundance below the analytical threshold for acetylsalicylic acid could be due to ionization suppression by other excipients or the preferential ionization of other

analytes in mixtures preventing acetylsalicylic acid from fragmenting into the ion with a m/z of 121.

Experimentation to determine the effects of matrix composition and concentration on the ratio of peak abundance between the molecular ion and fragment of cocaine was performed. The changes measured in the ratio of m/z 304 to 182 for cocaine varied the greatest for dilute cocaine mixtures. Similar to current toxicology methods for the identification of unknown compounds in GC/MS applications, the use of parent to spectator ion ratios can be used for identification purposes using unknown peaks. With the use of fragmentation the DART/MS platform can be considered as a Category A technique based on SWGDRUG guidelines, as opposed to the lesser Category B technique if only molecular weight information was produced¹⁶. One possible area of concern for the detection of cocaine by fragmentation could occur if the presence of cocaine is within a mixture including an excipient that exhibits preferential fragmentation by DART/MS. This preferential fragmentation could lead to the reduced abundance of either the molecular ion or fragment ion of cocaine. This phenomenon was not observed with any of the mixtures tested.

3.4 Analysis of Peak Ratio Variability for Cocaine

Further analysis into the variability in the ratio of abundance between the molecular ion and single fragment ion for cocaine was performed for two scenarios. Scenario one involved the analysis of cocaine in the presence of a single excipient mixture at increasingly greater concentrations compared to the

same concentration of cocaine at 10 ng/ μ L. Scenario two involved the analysis of cocaine in the presence of additional excipients where the concentration of each component was at 10 ng/ μ L. For the single excipient mixtures, levamisole and procaine were chosen to approximate the extreme differences in the amount of fragmentation seen within all excipients tested. Levamisole shows minimal fragmentation with a minor contributor compared to the molecular ion (Figure 26), while procaine gave extensive fragmentation with several peaks in high abundance and almost the complete loss of procaine's molecular ion at m/z 237 (Figure 27). These two analytes were chosen to determine if fragmentation from excipients could alter the ratio of m/z 304 to 182 through either ion suppression or preferential fragmentation.

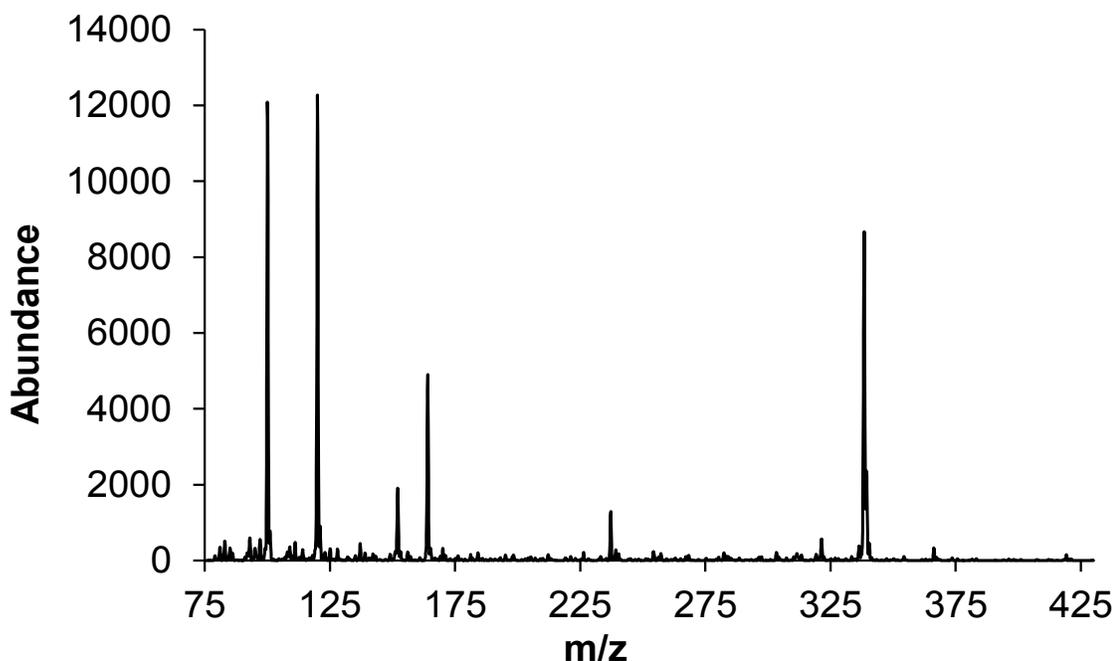


Figure 26. Mass spectrum of procaine fragmentation at a skimmer focus of 54V and at 100 ng/ μ L.

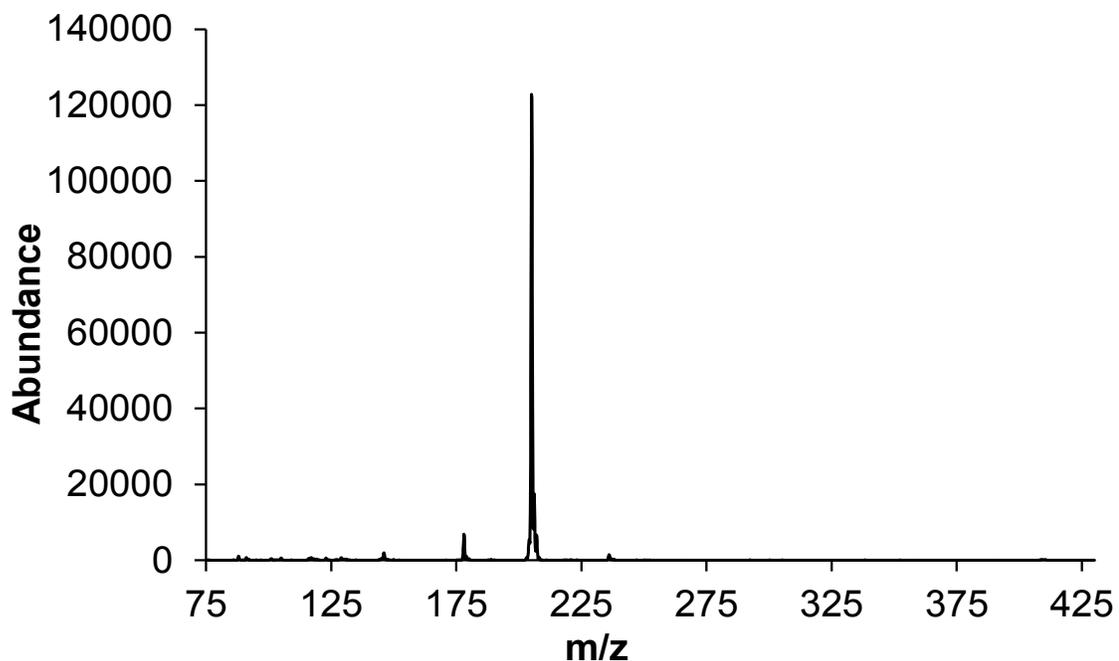


Figure 27. Mass spectrum of levamisole fragmentation at a skimmer focus of 54V and at 100 ng/ μ L.

Single excipient mixture solutions of 1:1, 1:2, 1:5, and 1:10 concentrations (cocaine: excipient) as well as three, six, nine, twelve, and fifteen component mixtures were utilized to mimic the wide range of sample composition and concentration in order to challenge the DART/MS platform. The ratio between the molecular ion abundance and main fragment for cocaine was calculated as follows: the abundance ± 0.5 m/z from the nominal mass peak was summed for each replicate. The average was calculated across the triplicate samples and then the ratio of abundance between m/z 304 to 182 was calculated (Figure 28).

Values from this calculation are shown in Table 6 for all single excipient mixtures multi-component mixtures and pure cocaine.

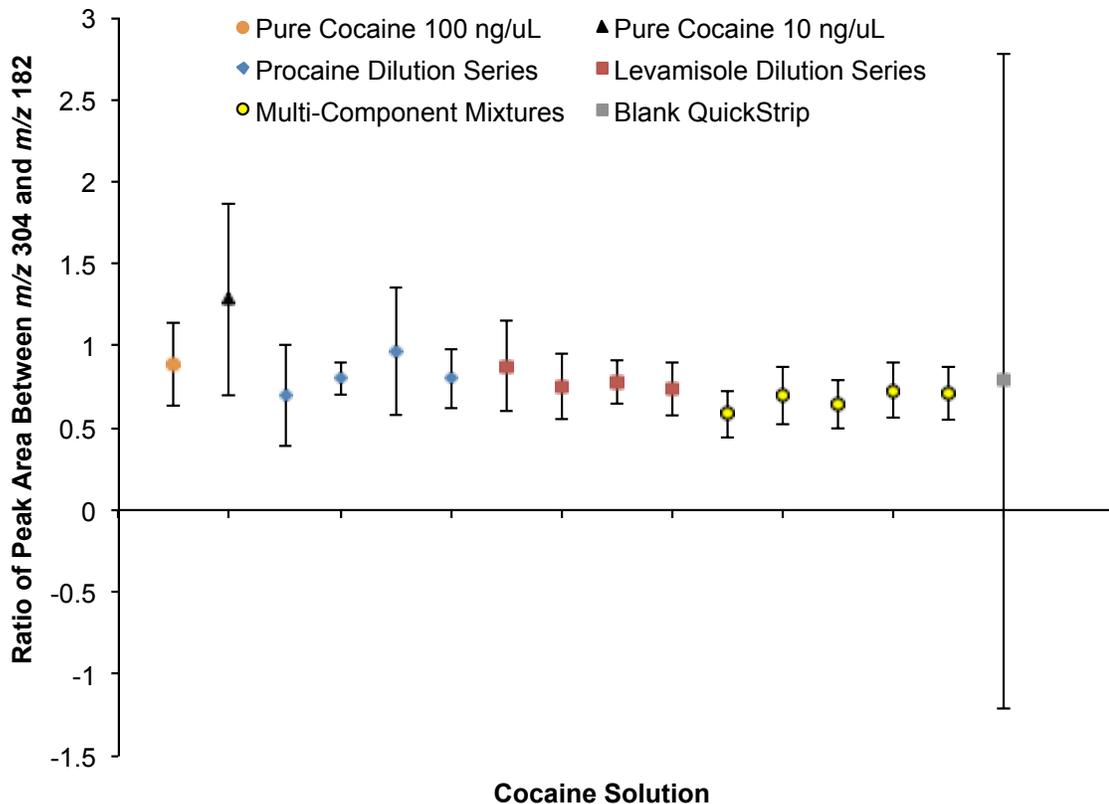


Figure 28. Ratio of peak area abundance for cocaine peaks of reference samples and unknown samples. For the procaine dilutions series, the order from left to right: 1:1, 1:2, 1:5, and 1:10 dilution samples (cocaine to procaine). For the levamisole dilutions series, the order from left to right: 1:1, 1:2, 1:5, and 1:10 dilution samples (cocaine to levamisole). For the multi-component mixtures, the order from left to right is the 3, 6, 9, 12, and 15-component samples.

Table 6. Statistical measurements for ratio of peak area abundance between cocaine *m/z* 304 and 182 peaks.

Sample	Type	Replicates	Average	Standard Deviation	Relative Standard Deviation
[Cocaine] to [Procaine] Mixtures	1 to 1	4	0.70	0.15	22
	1 to 2	4	0.80	0.05	6
	1 to 5	4	0.97	0.20	20
	1 to 10	4	0.80	0.09	11
[Cocaine] to [Levamisole] Mixtures	1 to 1	4	0.88	0.14	16
	1 to 2	4	0.75	0.10	13
	1 to 5	4	0.78	0.07	8
	1 to 10	4	0.73	0.08	11
Cocaine Multi-Component Mixtures (Number of Components)	3	10	0.58	0.07	12
	6	10	0.70	0.09	13
	9	10	0.64	0.07	11
	12	10	0.73	0.08	12
	15	10	0.71	0.08	11
Pure Cocaine	100 ng/ μ L	35	0.89	0.13	14
	10 ng/ μ L	27	1.28	0.29	23
Blank QuickStrip™		49	0.79	1.00	127

For all samples with cocaine present, the ratio of peak area between *m/z* 304 and 182 was consistently within the range of 0.4 to 1.9. The greatest relative standard deviation was for the 10 ng/ μ L pure cocaine data set at 23%. Stepwise addition of more components or increasing concentration of excipient added to cocaine mixtures did not drastically alter the abundance ratio between the molecular ion and the fragment. Furthermore, the stepwise addition of additional components did not exhibit any changes to the relative abundance of cocaine's peaks within the overall mass spectrum. Even within the most complex multi component mixture, cocaine did not exhibit ion suppression or reduced

fragmentation efficiency. The abundance at m/z 304 and 182 were greater than the 10x S/N for all samples tested. As predicted, high variability was seen in blank QuickStrip™ cards, which had inherently low or negligible mass abundances at m/z 304 and 182. An average ratio of 0.79 was calculated between m/z 304 to 182, with a RSD of 127%.

3.5 Cocaine Double Blind Study Analysis

A double blind study consisting of cocaine and all excipients tested, was utilized for the evaluation of three different facets: the DART/MSD platform, the ability to identify cocaine from DART mass spectrum data by human means and the ability to identify cocaine from DART mass spectrum data by computer modeling using Analyze IQ. Twenty-five contrived street quality cocaine samples containing a combination of excipients either spiked with or without cocaine were created in ambiguity to the chemist testing the samples. The chemist analyzed the samples via a batch analysis approach using ICE methodology on the DART/single quadrupole platform. Positive and negative controls were run prior to and after all unknown samples. Through the use of multiple QuickStrip(s)™ a substrate blank sample was run in between each triplicate unknown sample as to prevent carryover contamination from occurring.

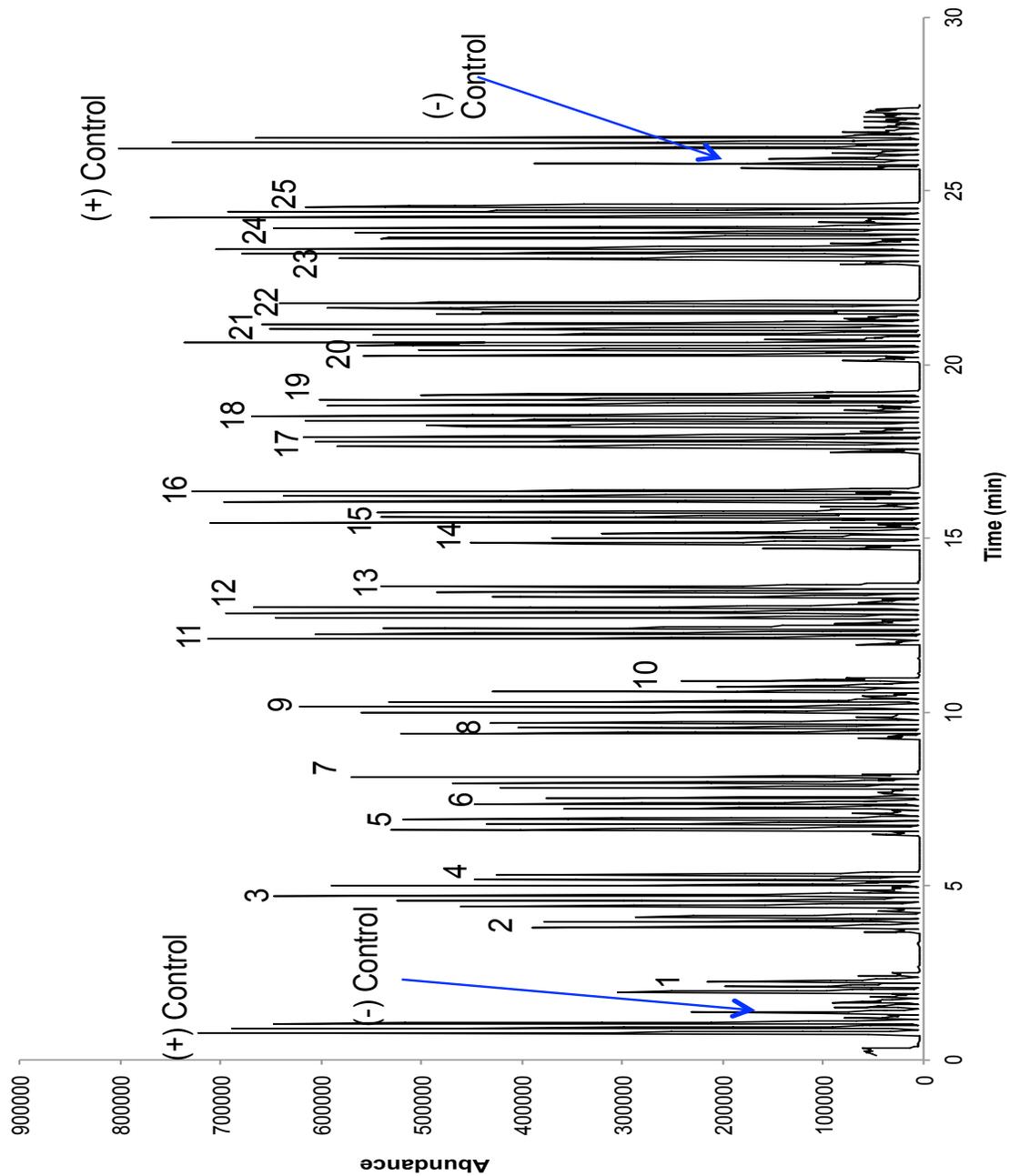


Figure 29. Total Ion Chromatogram of 25 unknown samples analyzed using the ICE method for the identification of cocaine. All samples were analyzed in triplicate with a substrate blank before and after sample analysis. The first and last samples are positive controls consisting of a cocaine standard at 100 ng/ μ L. The second and second to last samples are negative controls consisting of pure methanol.

Determination for the presence of cocaine by human means involved the identification of the molecular ion at m/z 304 and the main fragment at m/z 182. Using the EIC feature in Chemstation, the ability to select for only these two ions reduced the time required to manually interpret every peak. Samples 3, 4, 5, 7, 9, 11, 12, 13, 15, 16, 18, 20, 21, 22, 23, 24, 25, and both positive controls (cocaine standard at 100 ng/ μ L) all contained these two ions (Figure 30).

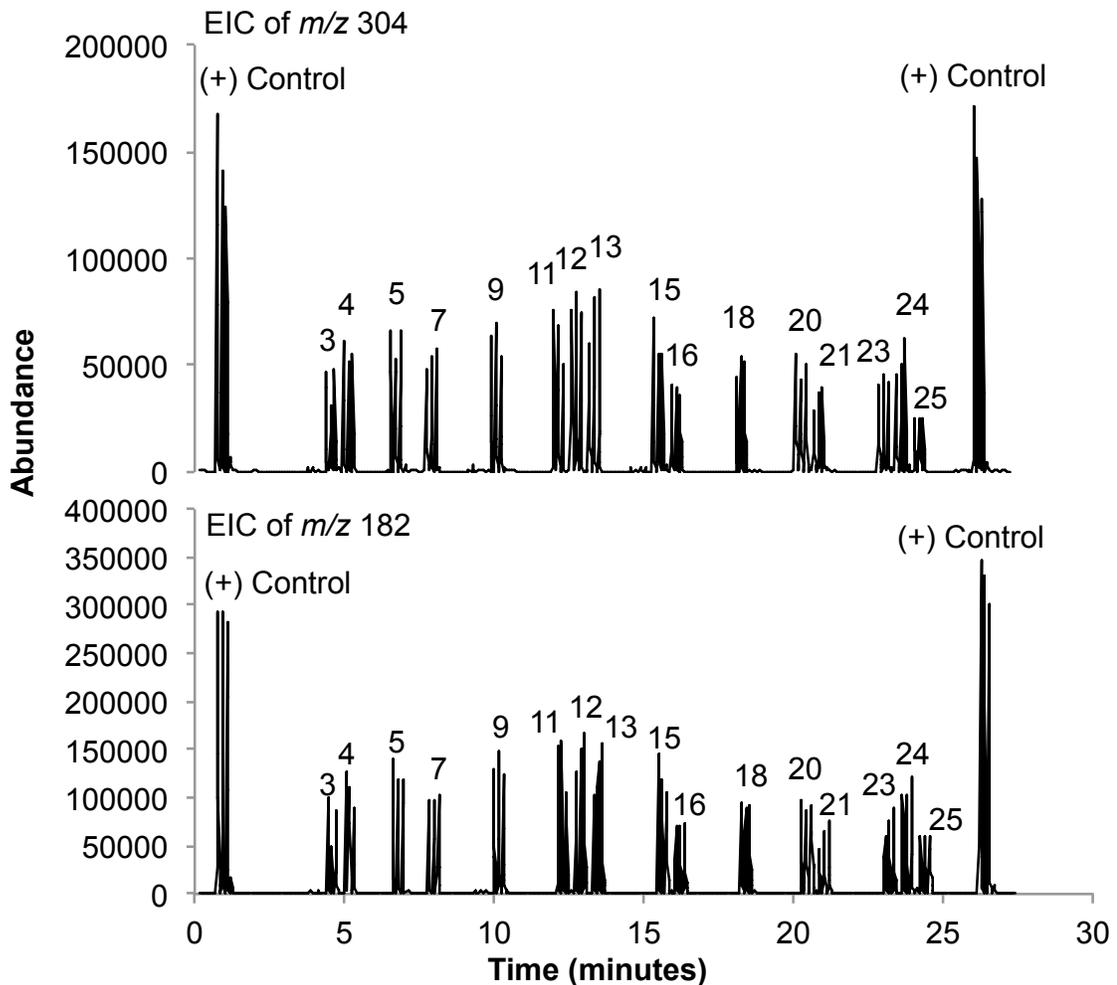


Figure 30. Extracted Ion Chromonogram of m/z 304 (top) and 182 (bottom) for the detection of cocaine from 25 unknown samples. Numbers shown are for samples with abundances above the baseline noise for their respective

ions. All samples analyzed in triplicate with a substrate blank before and after sample analysis. The first and last samples are positive controls consisting of a cocaine standard at 100 ng/ μ L.

Two mass spectra (Figure 31 and Figure 32) have been included to show the ability of an EIC to simplify the identification of selected ions from a complex spectrum or from a spectrum for which the ion of interest is of relatively low abundance. Utilization of the full mass spectrum allows for the presumptive identification of possible excipients as well, if further identification of excipients is needed for additional probative information. Figure 31 shows the presence of m/z 304 and 182 in high abundance compared to the presumptive identification of caffeine, piracetam, hydroxyzine, and either/both creatine and/or creatinine (cannot distinguish between the two using the ICE method), in the unknown sample #13. Whereas, Figure 32 has abundance below 3x S/N at m/z 304 and 182 so can be deemed negative for the presence of cocaine. It should be stated that the presumptive identification of diltiazem, hydroxyzine, levamisole/tetramisole, lidocaine, procaine, atropine, and quinine can be determined by the presence of their respective fragments within unknown sample #22 (see Table 4 for excipient DART fragment peaks).

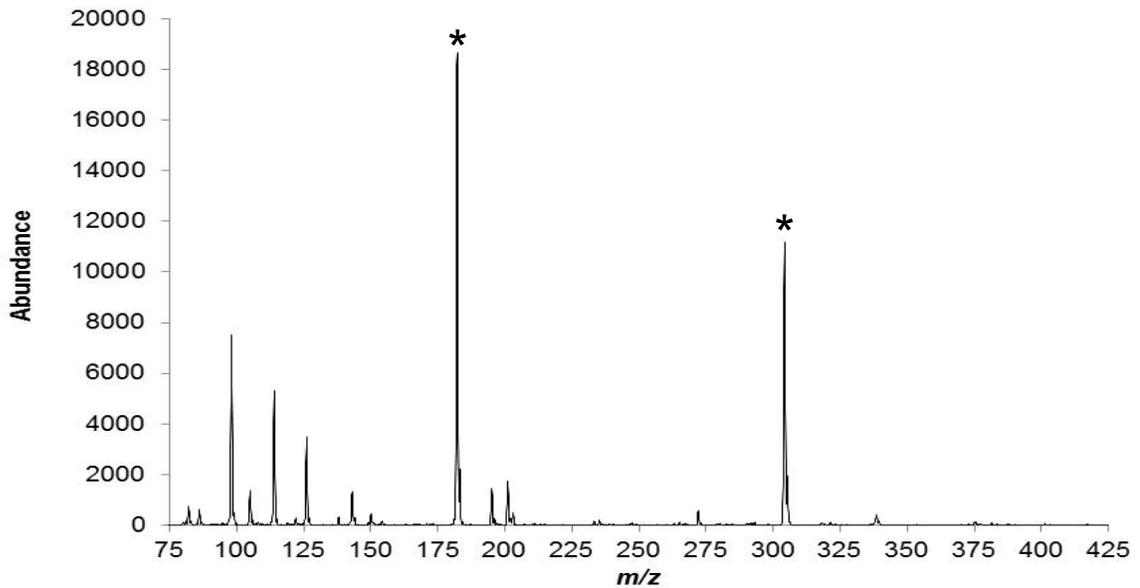


Figure 31. Sample 13 from cocaine double blind study. Example of unknown sample positive for the presence of cocaine from the identification of m/z 304 and 182 (shown with asterisks).

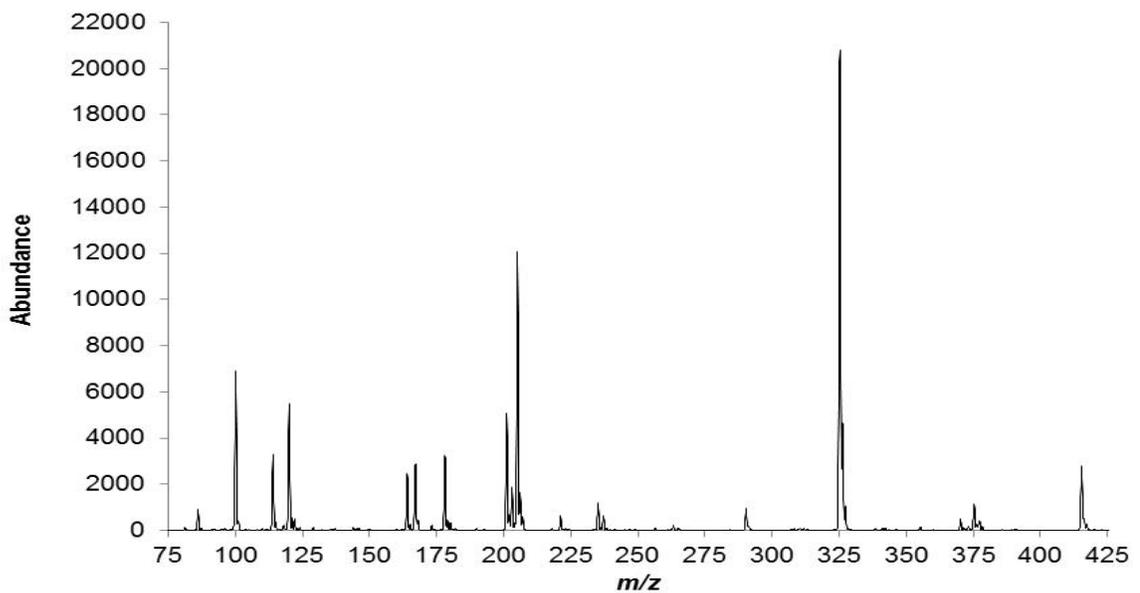


Figure 32. Sample 22 from cocaine double blind study. Example of unknown sample negative for cocaine due to lack of observable peaks at m/z 304 and 182.

Further testing for the presence of cocaine involved the peak area abundance ratio testing for all unknown samples as outlined in 2.6 Cocaine Mass Spectral Peak Ratio Testing. Using the same procedure with the raw data, the peak area abundance was averaged across all replicates of the sample for both m/z 304 and 182. The ratio was then calculated for each unknown sample and plotted in Figure 33. All statistics performed in Table 7 are based on triplicate measurements of the respective samples. Numbers shown as bold are less than ten times the average S/N abundance of the negative control samples (pre- and post-run) at m/z 304 and 182, respectively.

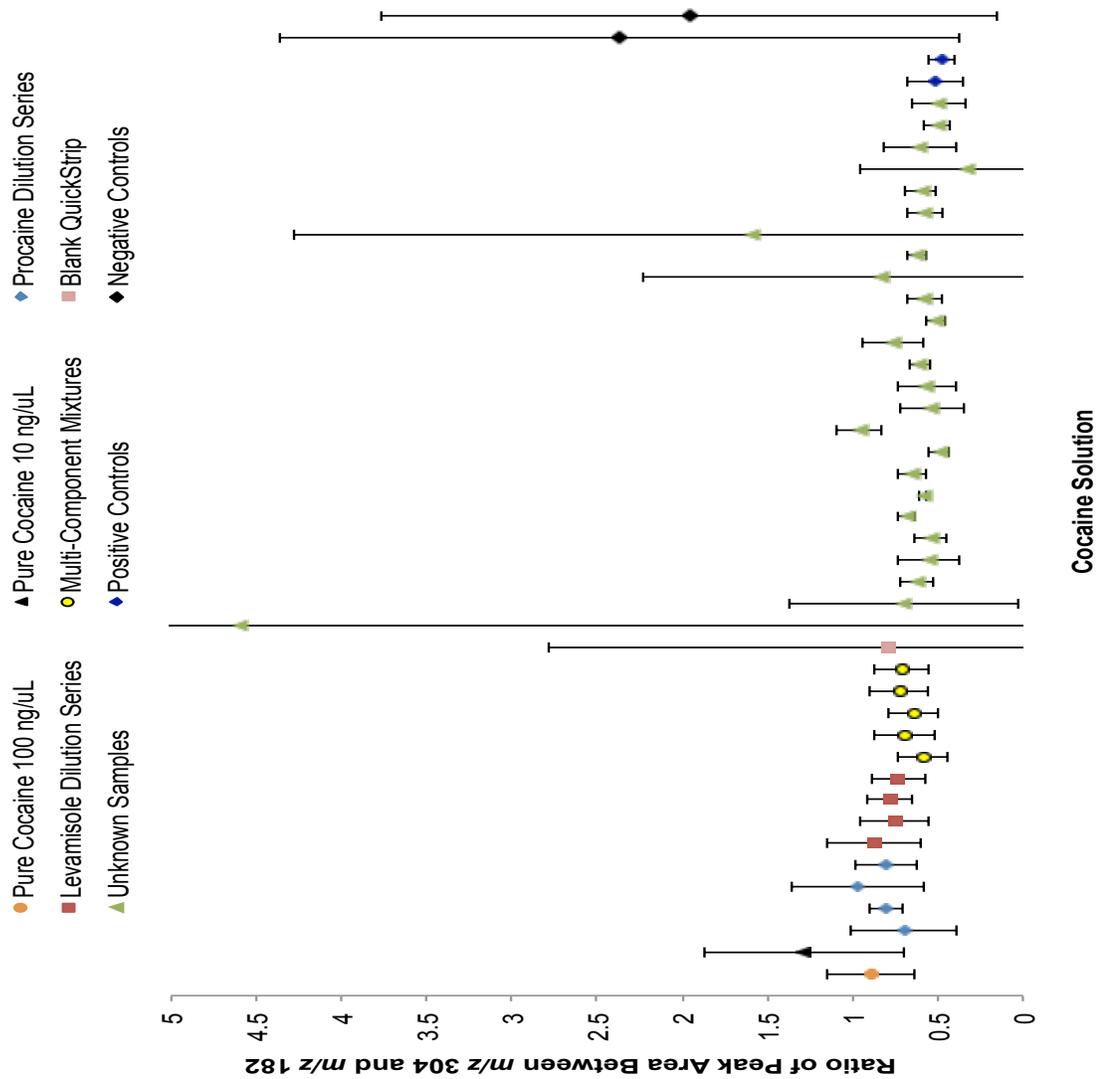


Figure 33. Ratio of peak area abundance for cocaine peaks of reference samples and unknown samples. For the procaine dilutions series, the order from left to right is the 1:1, 1:2, 1:5, and 1:10 dilution samples (amount cocaine to procaine). For the levamisole dilutions series, the order from left to right is the 1:1, 1:2, 1:5, and 1:10 dilution samples (amount cocaine to levamisole). For the multi-component mixtures, the order from left to right is the 3, 6, 9, 12, and 15 component samples. For the unknown samples, the order from left to right is sample 1 through 25. For the positive control (cocaine standard at 100 ng/ μ L the order from left to right is the pre-run and post-run sample. For the negative control (methanol) the order from left to right is the pre-run and post-run sample.

Table 7. Statistical measurements for ratio of peak area abundance between *m/z* 304 and 182 peaks for unknown mock case samples and reference samples. Bold numbers are values less than 10x S/N (as defined by the average abundance of the negative control peaks). All statistics shown are based on triplicate measurement of sample.

Sample	Average	Standard Deviation	Relative Standard Deviation (%)	Average Peak Area Abundance (304)	Average Peak Area Abundance (182)
1	4.60	3.05	66	712	247
2	0.70	0.34	48	2235	3539
3	0.62	0.05	8	46834	76275
4	0.55	0.09	16	62996	116063
5	0.54	0.05	9	68810	127601
6	0.69	0.03	4	946	1379
7	0.58	0.01	2	59141	101320
8	0.65	0.04	6	2135	3310
9	0.49	0.03	6	70358	142717
10	0.96	0.06	7	1160	1214
11	0.54	0.10	18	72687	140657
12	0.56	0.08	15	87344	157520
13	0.61	0.03	5	85111	141338
14	0.76	0.09	12	1978	2543
15	0.52	0.03	6	67768	131710
16	0.58	0.05	9	43505	75539
17	0.83	0.70	85	268	542
18	0.62	0.03	5	55778	90103
19	1.60	1.34	84	389	233
20	0.58	0.05	9	50918	88843
21	0.60	0.05	8	39682	67166
22	0.34	0.31	91	106	444
23	0.60	0.11	18	47529	80409
24	0.51	0.04	8	59194	116480
25	0.50	0.08	16	28761	58878
(+) Control ^a	0.51	0.08	16	159342	310224
(+) Control ^b	0.48	0.04	8	166079	348407
(-) Control ^c	2.37	1.00	42	1152	564
(-) Control ^d	1.96	0.90	46	1409	852

^aCocaine standard at 100 ng/μL analyzed before analysis of unknown samples

^bCocaine standard at 100 ng/μL analyzed after analysis of unknown samples

^cMethanol analyzed before analysis of unknown samples

^dMethanol analyzed after analysis of unknown samples

A set of match criteria was developed for the confirmatory identification of cocaine within contrived mixtures. As outlined in Figure 34, the use of a multi-step approach based on peak m/z value, peak abundances, peak ratio, and use of positive and negative controls gives the ability to identify cocaine in highly complex and dilute mixtures. By using Chemstation software and exporting the raw data into data processing software such as Excel, the ability to determine which samples had the presence of cocaine greater than 10 ng/ μ L was shown to be effective for twenty-five unknown samples. Identification for the presence of cocaine was determined for 16 of the 25 samples (Sample 3, 4, 5, 7, 9, 11, 12, 13, 15, 16, 18, 20, 21, 23, 24, and 25). All other samples were determined to be negative for the presence of cocaine. Based on the analyst's predictions a success rate of 100% with no false positives or false negatives was determined.

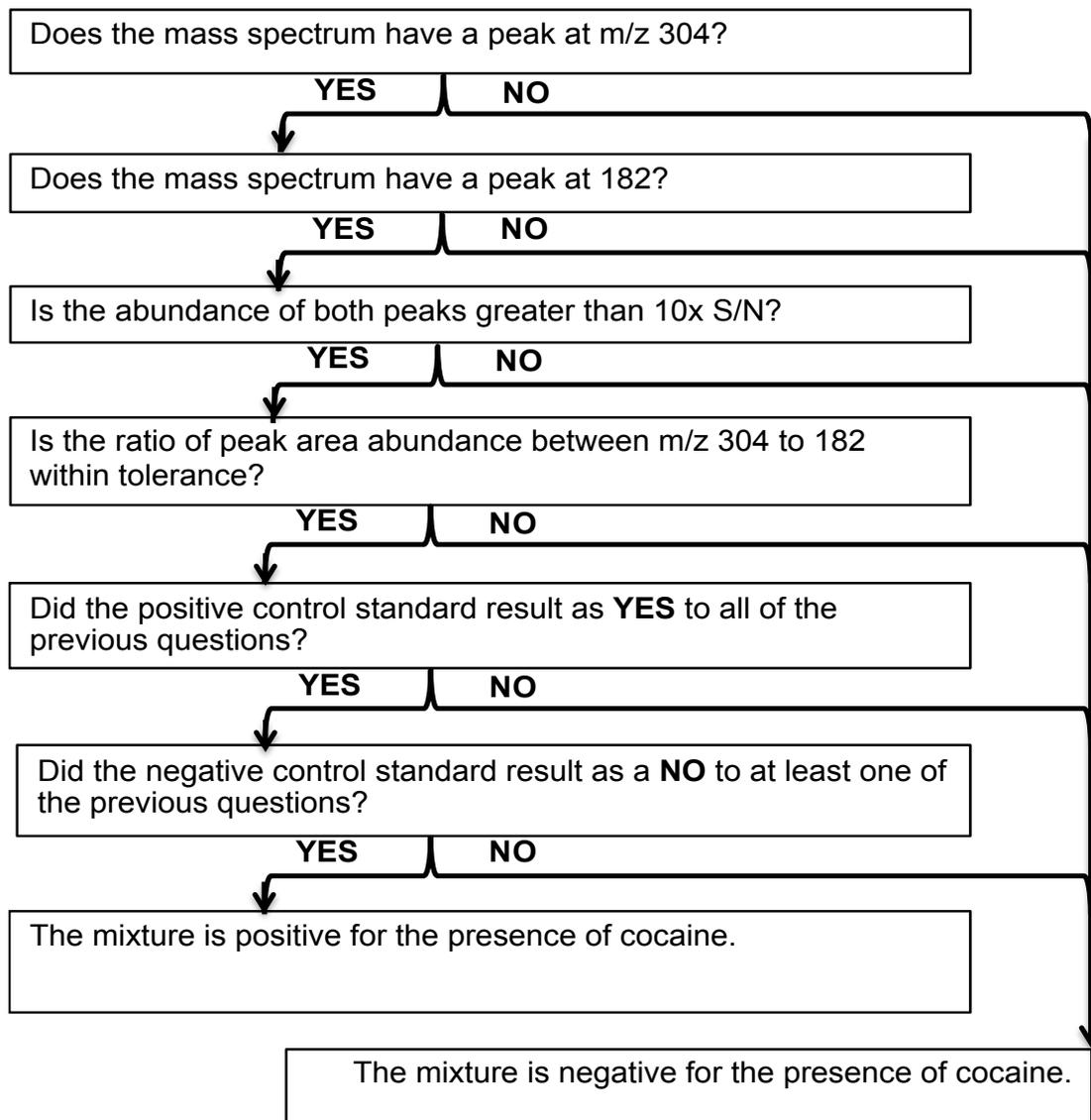


Figure 34. Match criteria flow chart for the determination of cocaine within unknown samples.

The greatest discriminating factor for the identification of cocaine was the determination that both peak abundances (m/z 304 and 182) were greater than ten times the defined signal-to-noise ratio. This measurement was found to be more discriminating between positive and negative cocaine samples than the

ratio of peak abundance between m/z 304 and 182 and the associated relative standard deviation of the triplicate samples. Several samples (6, 8, 10, and 14) were within the expected cocaine peak abundance ratio and had associated relative standard deviations below 22.7% (the highest RSD seen of the known reference samples under the 10 ng/ μ L cocaine conditions). But with the inclusion of the 10x S/N threshold, all of these samples were found to have peak abundances below the calculated noise threshold and deemed to be negative for the presence of cocaine. The use of a 10x S/N threshold and calculation of a cocaine peak abundance ratio within the expected range, allows for a high power of discrimination for the identification of cocaine within an unknown sample.

3.6 Cocaine Identification Model Development with Analyze IQ

3.6.1 Reference samples

Models were built for the identification of cocaine using DART/ single quadrupole mass spectral data collected in raw scan from m/z 40 to 430. Analyze IQ requires that each reference spectrum contain the same amount of data points⁵⁴, therefore raw scan was used during the data collection process. 381 reference samples were used to build the model. These 381 samples contained a broad scope of positive and negative cocaine samples in order to create a robust model that would not overfit the data (Table 8).

Table 8. Reference samples for building cocaine identification models using Analyze IQ.

Positive for Cocaine		Negative for Cocaine	
Sample Name ^{a,b}	Replicates	Sample Name ^{a,b}	Replicates
Cocaine	35	Tetramisole	37
Cocaine (10 ng/ μ L)	27	Prilocaine	40
3-Component Mixture ^c	14	Caffeine	39
6-Component Mixture ^c	14	Procaine	37
9-Component Mixture ^c	10	Adulterant Mixture A ^d	30
12-Component Mixture ^c	10	Adulterant Mixture B ^d	29
15-Component Mixture ^c	10	Blank QuickStrip™	49
Total	120	Total	261

^aFor all samples the final concentration was 100 ng/ μ L unless otherwise stated.

^bFor all mixtures, each component was at a final concentration of 10 ng/ μ L.

^cAs shown in Table 3.

^dAs shown in Table 9.

Table 9. Adulterant mixtures in comparison to multi-component cocaine mixtures as shown in Table 3.

Name of Adulterant Mixture ^a	Cocaine	Tetramisole	Prilocaine	Benzocaine	Procaine	Lidocaine	Caffeine	Diphenhydramine	Thiamine	Piracetam	Phenacetin	Acetaminophen	Hydroxyzine	Acetyl salicylic acid	Diltiazem
A		X	X	X	X	X	X								
B								X	X	X	X	X	X		

^aEach component was at a final concentration of 10 ng/ μ L.

3.6.2 Evaluation of Machine Learning Techniques

In order to determine the best machine learning technique for predictive classification modeling of cocaine in unknown samples, each machine learning

technique was evaluated at its default values set by Analyze IQ⁵⁴. Prior to the application of each machine learning technique, pre-processing and PCA were applied to the reference pool to correct for variances and reduce the size of the data.

Pre-processing of the data allows for the minimization of background noise and fluctuations in intensity⁶². Two pre-processing techniques were used in conjunction: a first order derivative followed by normalization. With these two pre-processing techniques, model prediction rates are the lowest compared to either pre-processing technique applied by itself or without any pre-processing techniques applied^{56–59,62}. Furthermore, implementation of PCA is another useful tool for the reduction of data while retaining as much complexity as possible. The use of PCA has been shown to lower model error rates prior to selection of a machine learning technique^{57,58}. The default value of 10 PCs was used for all models. A 5 run 10 fold cross-validation was utilized to evaluate all models.

After each model was built, they were tested against 72 samples. This testing pool was comprised of samples included in the model building process along with two sets of samples that contained excipients that were not incorporated in the reference pool to build the model. One set was a mixture of cocaine and creatine, creatinine, atropine, and xylazine, while another set was a mixture of only the new excipients. Both mixtures had a final concentration of each component at 10 ng/ μ L (Table 10). As shown in Table 11 each machine learning technique was evaluated once using the cross-validation (Training Error

Rate) technique and once with a set of reference samples, some of which were outside the scope of the built model (Testing Error Rate).

Table 10. Samples for testing against Analyze IQ built models.

Positive for Cocaine		Negative for Cocaine	
Sample Name ^{a,b}	Replicates	Sample Name ^{a,b}	Replicates
Cocaine	4	Prilocaine	1
Cocaine at 10 ng/ μ L	10	Adulterant Mixture A ^d	6
3-Component Mixture ^c	5	Adulterant Mixture B ^d	4
5-Component Mixture ^c	10	Adulterant Mixture C ^e	10
6-Component Mixture ^c	5	Blank QuickStrip TM	2
9-Component Mixture ^c	5		
12-Component Mixture ^c	5		
15-Component Mixture ^c	5		
Total	49	Total	23

^aFor all samples the final concentration was 100 ng/ μ L unless otherwise stated.

^bFor all mixtures, each component was at a final concentration of 10 ng/ μ L.

^cAs shown in Table 3.

^dAs shown in Table 9.

^eFour part mixture of creatine, creatinine, atropine, and xylazine.

Table 11. Model training and testing error rates for cocaine identification machine learning techniques at default values.

Model Type ^a	Training Error Rate ^b	Testing Error Rate ^c
k-Nearest Neighbors	2.20%	6.94%
SVM/RBF/Euclidean Distance	2.46%	0%
SVM/RBF/Weighted Spectral Distance	19.01%	27.78%
SVM/Sigmoid	36.44%	47.22%
SVM/Weighted Spectral Linear	0.05%	6.94%
SVM/Linear	0%	2.78%
Spectral Attribute Voting	0.52%	5.55%

^aPCA was applied on each model prior to the application of a machine learning technique except for Spectral Attribute Voting.

^bModel built with 120 (+) cocaine and 261 (-) cocaine samples.

^cModel tested against 59 (+) cocaine samples and 23 (-) cocaine samples.

3.6.3 Early Model Errors

Many of the originally developed models stemmed from an inadequate reference pool for model development. Early models utilized only pure cocaine and pure excipient samples at 100 ng/ μ L. When tested against simulated mixtures, all of the models had error rates too high to be deemed acceptable (above 50%). Two problems were identified with the model development protocol. One problem with the predictive model was due to the lack of mixtures in the reference pool. The model could not identify cocaine in multi-component mixtures when it was built using only pure concentrations of cocaine. Another problem was the concentration of the reference and test samples. The reference samples were 10x more concentrated than the cocaine test samples, which decreased the ability of the model to correctly identify when cocaine was present. Arbitrarily setting the cutoff threshold too high was detrimental to finding dilute samples containing cocaine. Only after introduction of mixtures and diluted samples (equal in concentration to the simulated mixtures) to the reference pool, did the models begin to yield error rates below 50%.

3.6.4 Optimization of Machine Learning Techniques

Upon evaluation, the Support Vector Machine (SVM) with a Radial Basis Function (RBF) based on Euclidean distance yielded the lowest testing error rate at default values. Further optimization of the SVM with RBF kernel was performed. Within the RBF kernel is the complexity factor, C , and sigma value (σ) that represents the kernel width (1.5.4 Analyze IQ Machine Learning

Techniques). These parameters were adjusted until a minimum training error rate was found (Table 12). Models 1, 4, 5, 7, and 9 were determined to have the lowest training and testing error rates associated with them. All of these models were able to successfully differentiate when cocaine was within mixtures of excipients that were not incorporated within the reference pool for model development.

Table 12. Model training and testing error rates for optimization of SVM/RBF/Euclidean distance cocaine identification model.

Model Type ^a	C	Sigma	Training Error Rate ^b	Testing Error Rate ^c
Model 1 (Default)	1	0.01	2.46%	0%
Model 2	1	0.1	13.93%	23.61%
Model 3	1	1	28.90%	40.28%
Model 4	1	0.001	0.58%	2.78%
Model 5	5	0.001	0.52%	5.56%
Model 6	0.1	0.01	19.74%	48.61%
Model 7	10	0.01	1.99%	0%
Model 8	10	0.001	0.52%	6.94%
Model 9	100	0.001	1.99%	0%
Model 10 ^d	10	0.001	0.52%	2.78%
Model 11	0.1	0.001	12.77%	48.61%
Model 12	0.5	0.001	1.52%	11.11%
Model 13	0.75	0.001	0.63%	2.78%
Model 14	1	0.0001	0.89%	13.89%

^a PCA was applied on each model prior to the application of SVM with RBF kernel using Euclidean distance machine learning technique.

^b Model built with 120 (+) cocaine and 261 (-) cocaine samples.

^c Model tested against 59 (+) cocaine samples and 23 (-) cocaine samples.

^d Model built with 20 PC's as opposed to the default value of 10 PC's.

3.7 Evaluation of Cocaine Identification Model

Evaluation of the optimized Analyze IQ models for the purposes of cocaine identification of unknown samples was performed with the 25 unknown samples from the cocaine double blind study. Testing of the optimized PCA/SVM regression with RBF kernel models using Euclidean distance, were applied to the 25 unknown double blind cocaine study samples.

Table 13. Evaluation of Analyze IQ optimized models against unknown samples from cocaine double blind study (including positive and negative controls).

Model Type ^a	C	Sigma	Error Rate	False Positives	False Negatives
Model 1 (Default)	1	0.01	24.14%	2	6
Model 2	1	0.1	48.28%	1	12
Model 3	1	1	62.07%	1	17
Model 4	1	0.001	13.79%	4	0
Model 5	5	0.001	14.94%	3	0
Model 6	0.1	0.01	41.38%	0	10
Model 7	10	0.01	22.99%	2	5
Model 8	10	0.001	16.09%	3	0
Model 9	100	0.001	22.99%	1	5
Model 10 ^b	10	0.001	16.09%	4	0
Model 11	0.1	0.001	41.38%	0	10
Model 12	0.5	0.001	25.29%	1	7
Model 13	0.75	0.001	13.79%	4	0
Model 14	1	0.0001	9.20%	2	0
Spectral Attribute Voting			0%	0	0

^a PCA was applied on each model prior to the application of SVM with RBF kernel using Euclidean distance machine learning technique.

^b Model built with 20 PC's as opposed to the default value of 10 PC's.

Evaluation of optimized models resulted in the lowest error rate of 9.20% error from an SVM with an RBF kernel of complexity value of 1 and a sigma value of 0.0001 (Model 14). Four additional models (Models 11 through 14) were

created to address the problem of underfitting in the original optimized models (Models 1 through 10) as shown in Table 13. Although the issue of underfitting was corrected in Models 11 and 12, the unwanted inclusion of false negatives occurred, suggesting that with reduced C values, and overfitting of the model occurred. Finally the decrease in sigma proved to decrease both types of errors better than adjusting the value of C as shown with the lowest error rate in Model 14. Model 14 had all false positives suggesting underfitting was occurring within this model as well.

Although the optimized models developed had a relatively low percent error, the problem with underfitting within a forensic context is concerning. To parallel the criminal law system with the United States, predictive models should be developed that purposefully overfit rather than underfit the data to be in concordance with the idea that a false positive is more detrimental than a false negative. Through the application of the Spectral Attribute Voting ensemble machine learning technique, a predictive model was created that correctly identified the presence of cocaine within 25 unknown samples 100% of the time with no false positives or false negatives.

4 CONCLUSIONS

The ability to detect cocaine in an accurate and robust manner for the purposes of forensic controlled substances by Direct Analysis in Real Time/ single quadrupole mass spectrometry was demonstrated. Cocaine was identified at concentrations as low as 10 ng/ μ L and in complex samples with many added

excipients. The ability to determine the presence of cocaine can be accomplished by the identification and interpretation of abundances for two peaks (m/z 304 and 182) relating to the structure of cocaine. Abundance values for the two peaks relating to cocaine were shown to have a small variance for different concentrations of pure cocaine (Figure 28 and Table 6). The development of a match criteria flow chart was enabled by the statistical interpretation of mass spectrometer data for the presence of cocaine Figure 34. This framework was applied to 25 unknown mock casework samples for the presence of cocaine within 24 possible excipients.

The development and optimization of a predicative classification model for the detection of cocaine within unknown samples was evaluated. Several models based on chemometric techniques were utilized to determine which model was most effective at identifying cocaine within complex mixtures. A model based on principle component analysis with support vector machine regression using a radial basis function kernel with Euclidean distance was determined to be the most effective modeling technique. Further optimization of this model resulted in error rate of less than 10% against 25 contrived street quality cocaine samples that were unknown to the analyst at the time of analysis. Further analysis showed that all samples that were incorrectly predicted by this optimized model were false positive identifications.

A model based on the proprietary machine learning method known as Spectral Attribute Voting gave an initial error rate of 0.52% after a 5 run 10 fold

cross validation. When the model was tested against 25 unknown contrived street quality cocaine samples the model correctly predicted the presence or absence of cocaine 100% of the time.

5 FUTURE DIRECTIONS

Increased efforts in the ability to fragment cocaine would aid in further probative value similar to the amount and variance of fragmentation seen in electron impact mass spectra of cocaine. Further development of the ion optics could increase the voltage settings of the skimmer focus, which in turn would allow for greater induced voltage imparted onto cocaine allowing for increased fragmentation. Adjustments to the ion optics could have an upper limit, if increasing the ion voltages hinders ions from being efficiently guided into the mass spectrometer. This could result in decreased abundance of ions and inherently lower the sensitivity of the DART/MS platform.

This analytical platform could easily be expanded to include other drugs of abuse and forensic interest that are normally cut with excipients as well. Common controlled substances such as amphetamines, ecstasy, heroin, and cathinones (bath salts) are all instances of illicit substances that may be part of complicated multi component mixtures. The ability to identify common drugs of abuse within mixtures without the use of chromatography would further aid forensic laboratories in faster screening and confirmatory techniques compared to GC/MS and LC/MS analytical platforms.

Additional software and modeling development using the Analyze IQ software could enable the rapid detection and identification of the excipients present within cocaine samples as well. Although not routinely examined, the ability to detect these compounds could have probative value for route determination and public health awareness. The use of this software is not limited to the scope of cocaine mixtures either. Application to other illicit substances (heroin, ecstasy, and cathinones) would aid in faster identification of the illicit components as well.

It should be stated that quantitative modeling was never developed or tested during this thesis. Quantification would enable the identification of percent purity for active compounds and excipients within the sample. Although this does not commonly occur within forensic laboratories today, the inclusion of these quantitative measurements could also be of further probative value. This type of software in both of its forms would greatly aid forensic laboratories as another means to corroborate the interpretive findings of forensic analysts that can be scrutinized for their subjective nature and lends itself to a more statistical interpretive approach.

LIST OF JOURNAL ABBREVIATIONS

Anal Chem	Analytical Chemistry
Anal Bioanal Chem	Analytical and Bioanalytical Chemistry
Artif Intell Cogn Sci Conf	Artificial Intelligence and Cognitive Science Conference
Artif Intell Rev	Artificial Intelligence Review
Forensic Sci Int	Forensic Science International
Instrum Sci Technol	Instrumental Science & Technology
Int Forensic Res Inst	International Forensic Research Institute
Int J Drug Policy	International Journal of Drug Policy
Int J Mass Spectrom	International Journal of Mass Spectrometry
J Am Soc Mass Spectrom	Journal of the American Society for Mass Spectrometry
J Anal Toxicol	Journal of Analytical Toxicology
J Chromatogr B Biomed Sci App	Journal of Chromatography B: Biomedical Sciences and Applications
J Forensic Identif	Journal of Forensic Identification
J Forensic Sci	Journal of Forensic Science
J Mass Spectrom	Journal of Mass Spectrometry
J Urban Health	Journal of Urban Health
Pharm Chem J	Pharmaceutical Chemistry Journal

Rapid Commun Mass Spectrom Rapid Communications in Mass Spectrometry

REFERENCES

1. Comprehensive Drug Abuse Prevention and Control Act of 1970. 1970
2. U.S. Drug Enforcement Administration. DEA Drugs of Abuse 2011 [Internet]. 2011; Available from: http://www.justice.gov/dea/pr/multimedia-library/publications/drug_of_abuse.pdf
3. Penido CAF de O, Silveira L, Pacheco MTT. Quantification of binary mixtures of cocaine and adulterants using dispersive Raman and FT-IR spectroscopy and principal component regression. *Instrum Sci Technol* 2012;40(5):441–56.
4. Nutt D, King LA, Saulsbury W, Blakemore C. Development of a rational scale to assess the harm of drugs of potential misuse. *The Lancet* 2007;369(9566):1047–53.
5. U.S. Drug Enforcement Administration, Office of Diversion Control. National Forensic Laboratory Information System: Year 2012 Annual Report. 2013
6. Hall AB, Coy SL, Nazarov EG, Vouros P. Rapid separation and characterization of cocaine and cocaine cutting agents by differential mobility spectrometry-mass spectrometry. *J Forensic Sci* 2012;57(3):750–6.
7. Maietti S, Castagna F, Molin L, Ferrara SD, Traldi P. Cocaine adulterants used as marker compounds. *J Mass Spectrom* 2009;44(7):1124–6.
8. Evrard I, Legleye S, Cadet-Taïrou A. Composition, purity and perceived quality of street cocaine in France. *Int J Drug Policy* 2010;21(5):399–406.
9. Dr. Sc. Miguel Lopez-Arts, Dr. Pharm. Ana Camean, Dr.Sc. Dr.M. Manuel Repetto. Unequivocal identification of several common adulterants and diluents in street samples of cocaine by infrared spectroscopy. *J Forensic Sci* 1995;40(4):602–10.
10. Reyes JC, Negrón JL, Colón HM, Padilla AM, Millán MY, Matos TD, et al. The emerging of xylazine as a new drug of abuse and its health consequences among drug users in Puerto Rico. *J Urban Health* 2012;89(3):519–26.

11. Bernardo NP, Siqueira MEPB, de Paiva MJN, Maia PP. Caffeine and other adulterants in seizures of street cocaine in Brazil. *Int J Drug Policy* 2003;14(4):331–4.
12. Fucci N, De Giovanni N. Adulterants encountered in the illicit cocaine market. *Forensic Sci Int* 1998;95(3):247–52.
13. Fucci N. Unusual adulterants in cocaine seized on Italian clandestine market. *Forensic Sci Int* 2007;172(2).
14. Brunt TM, Rigter S, Hoek J, Vogels N, van Dijk P, Niesink RJM. An analysis of cocaine powder in the Netherlands: content and health hazards due to adulterants. *Addiction* 2009;104(5):798–805.
15. Schneider S, Meys F. Analysis of illicit cocaine and heroin samples seized in Luxembourg from 2005–2010. *Forensic Sci Int* 2011;212(1-3):242–6.
16. Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG). Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) recommendations [Internet]. United States Department of Justice Drug Enforcement Agency, 2013; Available from: <http://www.swgdrug.org/Documents/SWGDRUG%20Recommendations%20Version%206-1.pdf>
17. Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG). Cocaine Monographs 2005 [Internet]. Available from: <http://www.swgdrug.org/Monographs/COCAINE.pdf>
18. National Research Council (U.S.), Committee on Identifying the Needs of the Forensic Science Community, National Research Council (U.S.), Committee on Science T, and Law, National Research Council (U.S.), Committee on Applied and Theoretical Statistics. Strengthening forensic science in the United States: a path forward. Washington, D.C.: National Academies Press, 2009
19. Hall AB. High throughput differential mobility spectrometry - mass spectrometry: Fundamental considerations and application development for forensic science [Internet]. 2012 [cited 2014 Mar 15]; Available from: <http://search.proquest.com/docview/1011002123>
20. Matthew R. Durose, Kelly A. Walsh, Andrea M. Burch. Census of Publicly Funded Forensic Crime Laboratories 2009 [Internet]. 2012 [cited 2014 Mar 14]; Available from: <http://www.bjs.gov/content/pub/pdf/cpffcl09.pdf>

21. Browning K, Strom KJ, Ropero-Miller J, Jones S, Sikes N, Pope M, et al. The 2007 survey of law enforcement forensic evidence processing. 2009 [cited 2014 Mar 15]; Available from: <http://www.ccasa.org/wp-content/uploads/2014/01/2007-Survey-of-LE-Forensic-Exam-Processing.pdf>
22. Cody RB, Laramée JA. Atmospheric pressure ion source [Internet]. 2005 [cited 2013 May 11]; Available from: <http://www.google.com/patents?id=YIAVAAAAEBAJ>
23. Steiner RR, Larson RL. Validation of the direct analysis in real time source for use in forensic drug screening. *J Forensic Sci* 2009;54(3):617–22.
24. Samms WC, Jiang YJ, Dixon MD, Houck SS, Mozayani A. Analysis of alprazolam by DART-TOF mass spectrometry in counterfeit and routine drug identification cases. *J Forensic Sci* 2011;56(4):993–8.
25. Grange AH, Sovocool GW. Detection of illicit drugs on surfaces using direct analysis in real time (DART) time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 2011;25(9):1271–81.
26. Easter JL, Steiner RR. Pharmaceutical identifier confirmation via DART-TOF. *Forensic Sci Int* 2014;240:9–20.
27. Jagerdeo E, Abdel-Rehim M. Screening of cocaine and its metabolites in human urine samples by direct analysis in real-time source coupled to time-of-flight mass spectrometry after online preconcentration utilizing microextraction by packed sorbent. *J Am Soc Mass Spectrom* 2009;20(5):891–9.
28. Howlett SE, Steiner RR. Validation of thin layer chromatography with AccuTOF-DARTTM detection for forensic drug analysis. *J Forensic Sci* 2011;56(5):1261–7.
29. Nilles JM, Connell TR, Durst HD. Quantitation of chemical warfare agents using the direct analysis in real time (DART) technique. *Anal Chem* 2009;81(16):6744–9.
30. Harris GA, Falcone CE, Fernández FM. Sensitivity “hot spots” in the direct analysis in real time mass spectrometry of nerve agent simulants. *J Am Soc Mass Spectrom* 2011;23(1):153–61.
31. Harris GA, Kwasnik M, Fernández FM. Direct analysis in real time coupled to multiplexed drift tube ion mobility spectrometry for detecting toxic chemicals. *Anal Chem* 2011;83(6):1908–15.

32. Rowell F, Seviour J, Lim AY, Elumbaring-Salazar CG, Loke J, Ma J. Detection of nitro-organic and peroxide explosives in latent fingermarks by DART- and SALDI-TOF-mass spectrometry. *Forensic Sci Int* 2012;221(1-3):84-91.
33. Pfaff AM, Steiner RR. Development and validation of AccuTOF-DART™ as a screening method for analysis of bank security device and pepper spray components. *Forensic Sci Int* 2011;206(1-3):62-70.
34. Musah RA, Cody RB, Dane AJ, Vuong AL, Shepard JRE. Direct analysis in real time mass spectrometry for analysis of sexual assault evidence. *Rapid Commun Mass Spectrom* 2012;26(9):1039-46.
35. Musah RA, Domin MA, Walling MA, Shepard JRE. Rapid identification of synthetic cannabinoids in herbal samples via direct analysis in real time mass spectrometry. *Rapid Commun Mass Spectrom* 2012;26(9):1109-14.
36. Musah RA, Domin MA, Cody RB, Lesiak AD, John Dane A, Shepard JRE. Direct analysis in real time mass spectrometry with collision-induced dissociation for structural analysis of synthetic cannabinoids. *Rapid Commun Mass Spectrom* 2012;26(19):2335-42.
37. Crawford E, Musselman B. Evaluating a direct swabbing method for screening pesticides on fruit and vegetable surfaces using direct analysis in real time (DART) coupled to an Exactive benchtop orbitrap mass spectrometer. *Anal Bioanal Chem* 2012;403(10):2807-12.
38. Coates CM, Coticone S, Barreto PD, Cobb AE, Cody RB, Barreto JC. Flammable solvent detection directly from common household materials yields differential results: an application of direct analysis in real-time mass spectrometry. *J Forensic Identif* 2008;58(6):624-31.
39. Laramée JA, Cody RB. Method for atmospheric pressure analyte ionization [Internet]. 2006 [cited 2013 May 11]; Available from: <http://www.google.com/patents?id=ENZ6AAAAEBAJ>
40. Cody RB, Laramée JA, Durst HD. Versatile new ion source for the analysis of materials in open air under ambient conditions. *Anal Chem* 2005;77(8):2297-302.
41. Cody RB. Observation of molecular ions and analysis of nonpolar compounds with the direct analysis in real time ion source. *Anal Chem* 2009;81(3):1101-7.

42. Wang Y, Liu L, Ma L, Liu S. Identification of saccharides by using direct analysis in real time (DART) mass spectrometry. *Int J Mass Spectrom* 2014;357:51–7.
43. IonSense Inc. IonSense | Products [Internet]. IonSense. 2014 [cited 2014 Mar 8]; Available from: <http://www.ionsense.com/products>
44. Curtis ME, Jones PR, Sparkman OD, Cody RB. Determination of the presence or absence of sulfur materials in drywall using direct analysis in real time in conjunction with an accurate-mass time-of-flight mass spectrometer. *J Am Soc Mass Spectrom* 2009;20(11):2082–6.
45. Chernetsova ES, Bochkov PO, Zatonskii GV, Abramovich RA. New approach to detecting counterfeit drugs in tablets by DART mass spectrometry. *Pharm Chem J* 2011;45(5):306–8.
46. Adams J. Analysis of printing and writing papers by using direct analysis in real time mass spectrometry. *Int J Mass Spectrom* 2011;301(1-3):109–26.
47. Lancaster C, Espinoza E. Analysis of select Dalbergia and trade timber using direct analysis in real time and time-of-flight mass spectrometry for CITES enforcement. *Rapid Commun Mass Spectrom* 2012;26(9):1147–56.
48. Bai Y, Zhang J, Bai Y, Liu H. Direct analysis in real time mass spectrometry combined with single-drop liquid–liquid–liquid microextraction for the rapid analysis of multiple phytohormones in fruit juice. *Anal Bioanal Chem* 2012;403(8):2307–14.
49. Musselman B, LaPointe J, Goguen R. Direct analysis in real time (DART) analysis with a modified GC/MSD system for rapid drug screening. *Int Forensic Res Inst* 2013
50. Harris GA, Hostetler DM, Hampton CY, Fernández FM. Comparison of the internal energy deposition of direct analysis in real time and electrospray ionization time-of-flight mass spectrometry. *J Am Soc Mass Spectrom* 2010;21(5):855–63.
51. Vail TM, Jones PR, Sparkman OD. Rapid and unambiguous identification of melamine in contaminated pet food based on mass spectrometry with four degrees of confirmation. *J Anal Toxicol* 2007;31(6):304–12.
52. Jurgen H. Gross. *Mass Spectrometry: A Textbook*. Second. Springer, 2011
53. Madden M, Howley T. *Analyze IQ: an innovative model-driven approach to analysis of chemical spectra*. 2008

54. Analyze IQ Limited. Analyze IQ Lab version 2.0 user manual. 2011
55. Madden MG, Howley T. A machine learning application for classification of chemical spectra [Internet]. In: Applications and Innovations in Intelligent Systems XVI. Springer, 2009 [cited 2013 Nov 30]; 77–90. Available from: http://link.springer.com/chapter/10.1007/978-1-84882-215-3_6
56. O’Connell M-L, Howley T, Ryder AG, Leger MN, Madden MG. Classification of a target analyte in solid mixtures using principal component analysis, support vector machines, and Raman spectroscopy [Internet]. In: Opto-Ireland. 2005 [cited 2013 Nov 30]; 340–50. Available from: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=865137>
57. Ryder AG, O’Connell M-L, Madden MG, Howley T. The effect of principal component analysis on machine learning accuracy with high dimensional spectral data. 2005 [cited 2013 Nov 30]; Available from: <http://vmserver14.nuigalway.ie/xmlui/handle/10379/194>
58. Howley T, Madden MG, O’Connell M-L, Ryder AG. Investigation into the use of PCA with machine learning for the identification of narcotics based on Raman spectroscopy. *Artif Intell Cogn Sci Conf* 2005;:277.
59. Howley T, Madden MG. The genetic kernel support vector machine: description and evaluation. *Artif Intell Rev* 2005;24(3-4):379–95.
60. Khan SS. Kernels for one-class nearest neighbour classification and comparison of chemical spectral data [Internet]. 2010 [cited 2013 Dec 2]; Available from: <http://aran.library.nuigalway.ie/xmlui/handle/10379/1358>
61. Dietterich TG. Ensemble methods in machine learning [Internet]. In: Multiple Classifier Systems. Springer Berlin Heidelberg, 2000; 1–15. Available from: http://dx.doi.org/10.1007/3-540-45014-9_1
62. Conroy J, Ryder AG, Leger MN, Hennessey K, Madden MG. Qualitative and quantitative analysis of chlorinated solvents using Raman spectroscopy and machine learning [Internet]. In: Opto-Ireland. 2005 [cited 2013 Nov 30]; 131–42. Available from: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=865115>
63. Hsu C-W, Chang C-C, Lin C-J. A practical guide to support vector classification [Internet]. Department of Computer Science and Information Engineering, National Taiwan University, 2003 [cited 2014 Mar 11]; Available from: <https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf>

64. Castiglioni S, Zuccato E, Chiabrando C, Fanelli R, Bagnati R. Mass spectrometric analysis of illicit drugs in wastewater and surface water. *Mass Spectrom Rev* 2008;27(4):378–94.
65. Meyer GMJ, Maurer HH. Qualitative metabolism assessment and toxicological detection of xylazine, a veterinary tranquilizer and drug of abuse, in rat and human urine using GC–MS, LC–MSn , and LC–HR-MSn. *Anal Bioanal Chem* 2013;405(30):9779–89.
66. Ruiz-Colon K, Martinez MA, Silva-Torres LA, Chavez-Arias C, Melendez-Negron M, Conte-Miller MS, et al. Simultaneous determination of xylazine, free morphine, codeine, 6-acetylmorphine, cocaine and benzoylecgonine in postmortem blood by UPLC-MS-MS. *J Anal Toxicol* 2012;36(5):319–26.
67. Stillwell ME. A reported case involving impaired driving following self-administration of xylazine. *Forensic Sci Int* 2003;134(1):25–8.
68. Wong SC, Curtis JA, Wingert WE. Concurrent detection of heroin, fentanyl, and xylazine in seven drug-related deaths reported from the philadelphia medical examiner’s office. *J Forensic Sci* 2008;53(2):495–8.
69. Staerk U, Külpmann WR. High-temperature solid-phase microextraction procedure for the detection of drugs by gas chromatography–mass spectrometry. *J Chromatogr B Biomed Sci App* 2000;745(2):399–411.
70. Cajka T, Danhelova H, Vavrecka A, Riddellova K, Kocourek V, Vacha F, et al. Evaluation of direct analysis in real time ionization–mass spectrometry (DART–MS) in fish metabolomics aimed to assess the response to dietary supplementation. *Talanta* 2013;115:263–70.
71. Hunt DF, Shabanowitz J, Botz FK. Chemical ionization mass spectrometry of salts and thermally labile organics with field desorption emitters as solids probes. *Anal Chem* 1977;49(8):1160–3.

CURRICULUM VITAE

ANDREW BLAIR HORSLEY

89 Northampton St. Apt 3A Boston MA 02118
Phone: 541-908-4854 E-mail: dhorsley@bu.edu
Year of Birth: 1989

HIGHER EDUCATION

Master of Science in Biomedical Forensic Sciences Anticipated: Sept. 2014

Boston University School of Medicine, Boston, MA

Thesis Advisor: Adam Hall, Ph.D.

High-Throughput Analysis Of Contrived Cocaine Mixtures By Direct Analysis In Real Time/Single Quadrupole Mass Spectrometry And Post Acquisition Chemometric Analysis

Bachelor of Science in Biochemistry

Graduation: Aug. 2011

University of Puget Sound, Tacoma, WA

Thesis Advisor: Daniel Burgard, Ph.D.

On-Road Evaluation of New, Heavy-Duty Diesel Truck Emission Standards

PRACTICAL SKILLS

Instrumentation

- Gas Chromatography/Mass Spectrometry (GC/MS)
- Fourier Transform Infrared Spectroscopy (FTIR)
- Liquid Chromatography/Mass Spectrometry (LC/MS)
- Direct Analysis in Real Time Ionization/Mass Spectrometry (DART/MS)

RELEVANT PROFESSIONAL EXPERIENCE

Laboratory Internship

Sept. 2013 – Jan. 2014

IonSense™, Saugus, MA

- Determined efficacy of various sorbent coated wire mesh substrates for the rapid extraction of toxic industrial chemicals from ignitable liquids for Direct Analysis in Real Time Ionization/Mass Spectrometry
- Developed and presented both written and oral reports of experimental findings for supervisors to aid in future project developments
- Awarded through the Massachusetts Life Sciences Center's Internship Challenge

Laboratory Research for Undergraduate Thesis

June 2009 – May 2011

University of Puget Sound, Tacoma, WA

- Measured on-site exhaust emission data of Heavy-Duty trucks to determine the efficacy of new governmental regulations

Laboratory Course Assistant for Chemistry Labs Sept. 2008 – Dec. 2009
University of Puget Sound, Tacoma, WA

- Aided the professor in teaching and monitoring a weekly lab of students conducting general chemistry experiments

Biology Laboratory Storeroom Assistant Sept. 2007 – May 2011
University of Puget Sound, Tacoma, WA

- Prepared solutions, biological media, and took care of animals used in biology lab for research

ACHIEVEMENTS & AWARDS

Massachusetts Life Sciences Center's Internship Challenge Sept. 2013
IonSense™, Saugus, MA

- Workforce development program that facilitates the placement of students and recent graduates into paid internships with Massachusetts companies engaged in life sciences.

University of Puget Sound Summer Research Grant Summer 2009 & 2010
University of Puget Sound, Tacoma, WA

- Received funding for two research projects totaling in \$6,859

ASSOCIATIONS & CONFERENCES

66th American Academy of Forensic Science Annual Meeting Feb. 2014
Seattle, WA

- Oral presentation on high throughput analysis of street-quality drug mixtures by DART® analysis and Analyze IQ™ post acquisition characterization

Vice President, Boston University Forensic Students Society

Boston University School of Medicine, Boston, MA Jan. – Dec. 2013

- Contact and coordinate with members in the forensic science community to give lectures at Boston University and to the Biomedical Forensic Science program
- Create and lead events that spurred community engagement among peers
- Helped fundraise through charitable events for the American Cancer Society and Boston 1 Fund

President, Student Affiliates of the American Chemical Society

University of Puget Sound, Tacoma, WA Sept. 2010 – May 2011

- Facilitated community outreach and helped coordinate elementary school visits around the chemistry department
- Created and lead events that spurred chemistry education among peers

- Helped fundraise for the American Cancer Society through Relay for Life

241st American Chemical Society National Meeting Mar. 2011
Anaheim, CA

- Poster presentation on the efficacy of new governmental regulations for Heavy-Duty diesel truck exhaust emissions

Member, American Chemical Society Sept. 2010 – Sept. 2012

References Available Upon Request