

Detection and Profiling of Synthetic Opioids

by

Joshua C. Klingberg

A thesis submitted for the

Degree of Doctor of Philosophy (Science)

University of Technology Sydney

Certificate of authorship and originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all the information sources and literature used are indicated in the thesis.

This research is supported by an Australian Government Research Training Program Scholarship.

Production Note:

Signature removed prior to publication.

NAME: Joshua C. Klingberg

DATE: 23 February 2021

Acknowledgements

This PhD would not have been possible without the help I received along the way from numerous people, and I would like to take the opportunity now to thank them for their support.

Firstly, I would like to thank my primary supervisor Shanlin Fu. It has been a pleasure working with you throughout my Honours and PhD work. Your constant support and encouragement were invaluable in helping me push through, especially when things weren't going to plan.

To my co-supervisor Ron Shimmon, my time working in the labs would not have been the same (and would have been a lot more difficult) without you. No matter how busy you were, you always made time to help, and I will be forever grateful. Your chemical knowledge and encouragement helped get me through when things weren't going my way in the laboratory.

Finally, to my external supervisor Adam Cawley, thank you for everything that you did for me throughout my PhD. From providing equipment and materials, to being on the receiving end of many, many emails, your constant guidance and advice were always welcome and helped me lift my work to a higher standard.

I would also like to thank the Drug and Toxicology Research Group at UTS who have always been supportive, and in particular Laura and Rhiannon, who I've had the pleasure to work alongside since Honours. Always having someone to bounce ideas off was amazing, and really helped keep me sane throughout this whole process. I would also like to thank Beth for all of her help while we were out at the racing lab. It would have been a very different experience without someone else to annoy (and teach me how to do stuff). I would also like to acknowledge Dan Pasin, who laid the groundwork for much of the work that I did throughout my PhD and was always happy to answer any questions that I had and help me get everything working properly.

I would like to thank all the staff at the Australian Racing Forensic Laboratory for allowing me to work alongside them in the lab and taking time out of their busy day to help me. In particular, I would like to thank Lauren McClure for all of her help on the instrument and for answering my many, and probably sometimes stupid, questions.

I would like to thank Chris Fouracre from Agilent Technologies for his help and support throughout my project. Having an industry contact that was willing and able to assist me with developing my

work was invaluable and made the whole process a lot smoother. Similarly, I would like to thank Emmanuel Blanchard from MathWorks for his help developing the statistical models used in this thesis. Your advice and explanations of how things worked really helped me improve my understanding of statistics and how it could be applied to my project.

Finally, I would like to thank my family for putting up with me and supporting me throughout my PhD journey. The constant encouragement I received was amazing and especially to my Mum who would just sit and listen while I talked about whatever it was I happened to be working on at the time.

While I am glad to be finishing, I will always be grateful for all the support I received along the way and the many memories that came with this PhD.

Table of contents

CERTIFICATE OF AUTHORSHIP AND ORIGINALITY	II
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	V
LIST OF FIGURES	XI
LIST OF TABLES.....	XV
ABBREVIATIONS	XVII
PUBLICATIONS AND CONFERENCE PROCEEDINGS	XXI
ABSTRACT.....	XXII
CHAPTER 1: INTRODUCTION	2
1.1 BACKGROUND	2
1.1.1 <i>Impact of Drug Use on Society</i>	2
1.1.2 <i>New Psychoactive Substances</i>	4
1.1.2.1 International Drug Trends	4
1.1.2.2 Australian Drug Trends.....	6
1.1.3 <i>Legal and Law Enforcement Issues</i>	6
1.2 SYNTHETIC OPIOIDS.....	10
1.2.1 <i>Structure and Chemistry</i>	12
1.2.1.1 Non-pharmaceutical Fentanyl's	12
1.2.1.2 Novel Synthetic Opioids	14
1.2.1.3 Pseudo- <i>Novel Synthetic Opioids</i>	17
1.2.2 <i>Pharmacology and Toxicology</i>	18
1.2.2.1 Non-pharmaceutical Fentanyl's	18
1.2.2.2 Novel Synthetic Opioids	20
1.2.2.3 Pseudo- <i>Novel Synthetic Opioids</i>	21

1.2.2.4	Opioid Effects in Horses	21
1.2.3	<i>Synthesis</i>	22
1.3	RECOMMENDATIONS FOR THE ANALYSIS OF SEIZED MATERIAL	24
1.4	NON-TARGETED ANALYSIS	25
1.4.1	<i>Collision-induced Dissociation</i>	26
1.4.2	<i>Data Acquisition Methods</i>	27
1.4.2.1	Data-dependant Acquisition	28
1.4.2.2	Data-independent Acquisition	29
1.4.3	<i>Data Processing</i>	30
1.4.4	<i>Statistical Analysis and Machine Learning</i>	34
1.5	SYNTHETIC ROUTE PROFILING.....	35
1.5.1	<i>Value of the Chemical Profile</i>	37
1.5.2	<i>Profiling of Synthetic Opioids</i>	39
1.6	PROJECT OBJECTIVES	40
1.7	REFERENCES.....	41
CHAPTER 2:	COLLISION-INDUCED DISSOCIATION STUDIES	56
2.1	RATIONALE	56
	COLLISION-INDUCED DISSOCIATION STUDIES OF SYNTHETIC OPIOIDS FOR NON-TARGETED ANALYSIS	57
2.2	INTRODUCTION.....	58
2.3	MATERIALS AND METHODS.....	60
2.3.1	<i>Solvents and Reagents</i>	60
2.3.2	<i>Opioid Standards</i>	61
2.3.3	<i>Sample Preparation</i>	61
2.3.4	<i>Instrumental Analysis</i>	62
2.4	RESULTS AND DISCUSSION	63
2.4.1	<i>Collision-induced Dissociation Pathways</i>	63
2.4.1.1	Non-pharmaceutical Fentanyls	63

2.4.1.2	AH Series Opioids	66
2.4.1.3	U Series Opioids	67
2.4.1.4	W Series Pseudo-Opioids	69
2.4.1.5	MT-45	70
2.4.2	<i>Non-targeted Screening</i>	71
2.5	CONCLUSION.....	76
2.6	REFERENCES.....	77
CHAPTER 3: NON-TARGETED ANALYSIS – COMPOUND DETECTION.....		80
3.1	RATIONALE	80
FINDING THE PROVERBIAL NEEDLE: NON-TARGETED SCREENING OF SYNTHETIC OPIOIDS IN EQUINE PLASMA.....		81
3.2	INTRODUCTION.....	82
3.3	EXPERIMENTAL	84
3.3.1	<i>Solvents and Reagents</i>	84
3.3.2	<i>Sample Preparation</i>	84
3.3.3	<i>Instrumental Analysis</i>	86
3.3.4	<i>Data Analysis</i>	87
3.3.4.1	Targeted Compound Extraction Limit of Detection.....	87
3.3.4.2	Product Ion Searching (PIS)	87
3.3.4.3	Kendrick Mass Defect.....	88
3.3.4.4	Non-targeted Workflow	89
3.3.5	<i>Blind Trial Setup</i>	89
3.4	RESULTS AND DISCUSSION	89
3.4.1	<i>Implementation of Non-targeted Workflow</i>	89
3.4.2	<i>Application of Screening Techniques</i>	90
3.4.2.1	Product Ion Searching	90
3.4.2.2	Kendrick Mass Defect Analysis	93
3.4.3	<i>Blind Trial</i>	95

3.4.4	Comparison of Techniques.....	98
3.5	MOLECULAR FEATURE EXTRACTION	102
3.5.1	Background.....	102
3.5.2	Experimental.....	102
3.5.3	Results and Discussion.....	104
3.6	CONCLUSION.....	107
3.7	REFERENCES.....	107
CHAPTER 4:	NON-TARGETED ANALYSIS – COMPOUND IDENTIFICATION.....	112
4.1	RATIONALE	112
	TOWARDS COMPOUND IDENTIFICATION IN NON-TARGETED SCREENING USING MACHINE LEARNING TECHNIQUES.....	113
4.2	INTRODUCTION.....	114
4.3	EXPERIMENTAL	115
4.3.1	Solvents and Reagents.....	115
4.3.2	Drug Standards	115
4.3.3	Sample Preparation	116
4.3.3.1	Class Prediction Samples.....	116
4.3.3.2	Retention Time Repeatability Studies	116
4.3.3.3	Retention Time Prediction Samples	117
4.3.3.4	Plasma Extraction Method	117
4.3.4	Instrumental Analysis	118
4.3.5	Data Analysis	118
4.3.5.1	Class Prediction Samples.....	119
4.3.5.2	Retention Time Prediction Samples	119
4.3.6	Molecular Features	119
4.3.7	Statistical Analysis	120
4.3.7.1	Class Prediction Modelling	121
4.3.7.2	Retention Time Prediction	122

4.4	RESULTS AND DISCUSSION	122
4.4.1	<i>Class Prediction Modelling</i>	122
4.4.2	<i>Retention Time Repeatability Studies</i>	128
4.4.3	<i>Retention Time Prediction</i>	131
4.5	CONCLUSION.....	136
4.6	REFERENCES.....	136
CHAPTER 5: NON-TARGETED SCREENING WORKFLOW		141
5.1	RATIONALE	141
5.2	IMPLEMENTATION OF NON-TARGETED SCREENING PROCESSES IN FORENSIC TOXICOLOGY	142
5.2.1	<i>Compound Detection</i>	142
5.2.2	<i>Compound Identification</i>	144
5.3	CONCLUSION.....	147
5.4	REFERENCES.....	147
CHAPTER 6: SYNTHETIC ROUTE PROFILING.....		150
6.1	RATIONALE	150
6.2	INTRODUCTION.....	151
6.3	EXPERIMENTAL	152
6.3.1	<i>Synthetic Strategy</i>	152
6.3.1.1	Materials	153
6.3.2	<i>Instrumentation</i>	153
6.3.2.1	Gas Chromatography – Mass Spectrometry.....	153
6.3.2.2	Liquid Chromatography – Mass Spectrometry.....	154
6.3.3	<i>Recursive Feature Extraction</i>	155
6.3.3.1	Gas Chromatography – Mass Spectrometry.....	155
6.3.3.2	Liquid Chromatography – Mass Spectrometry.....	155
6.3.4	<i>Statistical Modelling</i>	156
6.4	RESULTS AND DISCUSSION	157

6.4.1	Identification of Major Impurities.....	157
6.4.1.1	Common Synthetic Impurities.....	157
6.4.1.2	Route-specific Impurities	160
6.4.2	Statistical Approaches to Route Classification.....	163
6.5	CONCLUSION.....	169
6.6	REFERENCES.....	169
CHAPTER 7:	OVERALL CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK	172
7.1	OVERALL CONCLUSIONS	172
7.2	RECOMMENDATIONS FOR FUTURE WORK.....	173
APPENDIX 1	176
APPENDIX 2	178
APPENDIX 3	205
APPENDIX 4	212
APPENDIX 5	225

List of figures

Figure 1-1: Trends in number of drug users by drug class (Adapted from UNODC) ^[1]	2
Figure 1-2: Changes to the NPS market reported in 2017 (Adapted from UNODC) ^[1]	5
Figure 1-3: Chemical structure of fentanyl	12
Figure 1-4: Chemical structure of MT-45	16
Figure 1-5: Chemical structures of benzimidazole isomers, including etonitazene (A), isotonitazene (B), and clonitazene (C)	17
Figure 1-6: Synthesis of fentanyl as patented by Janssen ^[110]	22
Figure 1-7: Generic synthesis procedure for fentanyl derivatives	23
Figure 1-8: Comparison of data-dependent and data-independent acquisition techniques in high-resolution mass spectrometry (Adapted from Zhu et al.) ^[132]	28
Figure 1-9: Comparison of the different workflows for biased non-targeted screening as proposed by Pasin et al. ^[122] . The * denotes that specialised software is required	31
Figure 1-10: Chemical profiling of MDMA, as carried out by NMI (Adapted from Morelato et al.) ^[183]	36
Figure 1-11: Framework for processing of case information to support the three different levels of forensic intelligence (Adapted from Marclay et al.) ^[196]	37
Figure 2-1: General chemical structure of fentanyls (A), AH series opioids (B), W series opioids (C), U series opioids with (D), and without (E) a methylene spacer and MT-45 (F).....	59
Figure 2-2: Proposed structures of fragments observed for NPFs containing differing tails and C9 side chains, showing molecular ion (1) and common product ions (1a – 1c)	64
Figure 2-3: Proposed structures of fragments observed for NPFs containing a phenylethyl tail and lacking a C9 side chain, showing molecular ion (2) and common product ions (2a – 2g)	66

Figure 2-4: Proposed structures of fragments observed for AH series opioids, showing molecular ion (3) and common products ions (3a – 3e)	67
Figure 2-5: Proposed structures of fragments observed for U series opioids containing a methylene spacer, showing molecular ion (4) and common product ions (4a – 4c)	68
Figure 2-6: Proposed structures of fragments observed for W series opioids, showing molecular ion (5) and common product ions (5a – 5c)	70
Figure 2-7: Proposed structures of fragments observed for MT-45, showing molecular ion (6) and product ions (6a – 6d)	71
Figure 2-8: Extracted ion chromatograms for diagnostic product ions 105.0704 (top) and 188.1439 (bottom) showing the identification of acetyl fentanyl and fentanyl in 0.1 ng/mL mixed spike	74
Figure 2-9: Extracted ion chromatograms for diagnostic product ions identified for the AH series opioids: 144.9612, 172.9561, 189.9827, 201.9827, and 284.0609 (A), and for the U series opioids with a methylene spacer: 158.9768, 218.0140 and 298.0766 (B), showing the identification of AH-7921 and U-50488 in an equine plasma sample spiked with a mixture of NPFs and NSOs at 0.1 ng/mL	75
Figure 3-1: Chemical structures of opioids included in the study including acetyl fentanyl (A), fentanyl (B), carfentanil (C), AH-7921 (D) and U-50488 (E)	85
Figure 3-2: EICs showing the targeted extraction of known precursor masses showing the detection of the spiked compounds in a 0.05 ng/mL spike. The common product ions used for the PIS screening also showed detection of the samples peaks and could be easily distinguished from the matrix blank, showing the specificity of the technique.	91
Figure 3-3: Example overlaid EICs for common U series product ions m/z 158.9768 (red), 218.0140 (blue) and 298.0766 (green) from a spiked plasma sample with 0.05 ng/mL U-50488 (A) and a blank sample (B)	92
Figure 3-4: EICs for some of the identified masses from DefectDetect showing selectivity to the spiked compounds in comparison to the blank (A) and the results table given by DefectDetect for the spiked sample (B). The detected class for each mass in the results table is given according to the filters listed in Table 3-2. AC: Amide Chain, Cy: Cyclic, FA: F analogues	94

Figure 3-5: PIS results obtained from the blind trail showing the detection of both monitored ions in the 0.5 ng/mL spike and only detection of the m/z 188.1439 ion in the 0.05 ng/mL spike	96
Figure 3-6: PIS results for the spiked acetyl fentanyl standards (A) compared with the associated KMD results for each sample (B).....	97
Figure 3-7: EICs showing common fentanyl product ions at m/z 105.0704 and 188.1439 (A) showing a peak for carfentanil at a lower abundance at m/z 105.0704 (B) but not m/z 188.1439	99
Figure 3-8: Layout of the Profinder outputs showing: (A) list of compound groups; (B) compound information within each sample; (C) EICs for selected compound group; (D) Find by Formula mass spectra for selected compound group.....	105
Figure 3-9: Overview of the effects of different filtering processes on the number of compounds extracted by the recursive feature extraction approach	107
Figure 4-1: Generic structures of the opioid subclasses used in the class prediction model, including fentanyl analogues (A), AH series (B), and U series with (C), and without (E) a methylene spacer .	123
Figure 4-2: Confusion matrix showing the prediction accuracy of the developed Naïve Bayes model for each opioid subclass	126
Figure 4-3: Model accuracy following removal of individual predictors in comparison to the accuracy of a model trained with all features (red). A lower RMSE indicates a higher model accuracy.....	132
Figure 4-4: Predicted (RRT_p) vs. experimental (RRT_e) for the developed Gaussian Process Regression model	133
Figure 4-5: Residuals produced from the retention time prediction model.....	134
Figure 5-1: Proposed non-targeted screening workflow for detection of unknown compounds	143
Figure 5-2: Non-targeted screening workflow for proposing putative identifications of detected unknown compounds.....	146
Figure 6-1: Overall synthetic route for the production of acetyl fentanyl. Compounds in black are those specific to the majority of synthetic methods. Compounds specific to each method are shown in blue, Valdez ^[43] (Method 1); brown, One Pot ^[87] (Method 2); and green, Siegfried ^[111] (Method 3).	

TBAB: tetra-N-butylammonium bromide; ACN: acetonitrile; DCM: dichloromethane; DIPEA: diisopropylethylamine.....	153
Figure 6-2: Chemical structures of some of the common impurities identified. (A) 1-phenethyl-piperidine-4-ol; (B) 4-anilino-N-phenethylpiperidine; (C) acetanilide	158
Figure 6-3: Chemical structures of N-ethylacetanilide	159
Figure 6-4: Chemical structures of (2-chloroethyl)benzene (left) and (2-bromoethyl)benzene (right)	160
Figure 6-5: Chemical structures of some impurities specific to the Valdez method. (A) diisopropylethylamine (DIPEA); (B) N,N-diisopropylacetamide; (C) styrene	160
Figure 6-6: Chemical structures of some impurities specific to the One-pot method. (A) phenacetaldehyde; (B) phenacetaldehyde dimethyl acetal; (C) phenethyl alcohol	161
Figure 6-7: Principal component analysis showing the separation between different synthetic pathways based on GC-MS analysis for 2 components (top) and 3 components (bottom). Red: Valdez (Method 1); Blue: One-pot (Method 2); Yellow: Siegfried (Method 3). Each data point represents samples prepared in separate synthesis batches. The first 3 components accounted for 76.81% of the variance present.	164
Figure 6-8: PCA loading plot showing the separation of the entities identified after GC-MS analysis	166
Figure 6-9: Principal component analysis showing the separation between different synthetic pathways based on LC-MS analysis for 2 components (top) and 3 components (bottom). Red: Valdez (Method 1); Blue: One-pot (Method 2); Yellow: Siegfried (Method 3). Each data point represents samples prepared in separate synthesis batches. The first 3 components accounted for 80.99% of the variance present.	167

List of tables

Table 1-1: Summary of main regulatory options available (Adapted from Reuter and Pardo) ^[17]	9
Table 1-2: Chemical information of the fentanyl derivatives	13
Table 1-3: Chemical information for the AH series of novel synthetic opioids.....	14
Table 1-4: Chemical information for some of the U series of novel synthetic opioids.....	15
Table 1-5: Chemical information for the U series opioids containing a methylene spacer	16
Table 1-6: Chemical information for the W series of novel synthetic opioids.....	18
Table 1-7: Potency of various fentanyl derivatives compared with morphine and fentanyl.....	19
Table 1-8: Categories of Analytical Techniques as define by SWGDRUG ^[121]	25
Table 2-1: Summary of diagnostic product ions for each subclass of synthetic opioids	73
Table 3-1: Diagnostic product ions from each compound class used for screening ^[224]	88
Table 3-2: Kendrick mass defect (KMD) filters used for compound detection	89
Table 3-3: Kendrick mass defect (KMD) analysis of one set of 0.1 ng/mL samples, showing consistent detection of analytes of interest and total number of filtered results for each sample	95
Table 3-4: Overview of the estimated limits of detection (LODs) and screening cut-offs for the different detection techniques applied	101
Table 4-1: Molecular features used for retention time prediction	120
Table 4-2: Accuracy of each class prediction model trained.....	125
Table 4-3: F1 scores and Matthew's Correlation Coefficients calculated from the optimised classification model.....	128
Table 4-4: Overview of the average standard deviation (SD) and relative standard deviation (%RSD) values for the repeatability studies conducted on neat standards	129

Table 4-5: Overview of the average standard deviation (SD) and relative standard deviation (%RSD) values for both the absolute retention time (RT) and relative retention time (RRT) of the spiked plasma samples (n = 49)	130
Table 4-6: Accuracy of each retention time prediction model trained, measured by the root mean square error (RMSE)	131
Table 6-1: Prediction accuracy for each trained model	168

Abbreviations

%RSD	Relative Standard Deviation
2-BEB	(2-bromoethyl)benzene
2-CEB	(2-chloroethyl)benzene
4-ANPP	4-anilino-N-phenethylpiperidine
4-FBF	4-fluorobutyl fentanyl
ABS	Australian Bureau of Statistics
ACIC	Australian Criminal Intelligence Commission
ACMD	Advisory Council on the Misuse of Drugs
ACN	Acetonitrile
AFP	Australian Federal Police
AIDIP	Australian Illicit Drug Intelligence Program
AIHW	Australian Institute of Health and Welfare
ANN	Artificial Neural Network
AORC	Association of Official Racing Chemists
ARFL	Australian Racing Forensic Laboratory
ATS	Amphetamine-type Stimulants
CAS	Chemical Attribution Signature
CE	Collision Energy
CE-DAD	Capillary Electrophoresis – Diode Array Detector
CID	Collision-induced Dissociation
CRM	Certified Reference Material
CSV	Comma-separated Value
DALY	Disability-adjusted Life Years
DBE	Double Bond Equivalents
DCM	Dichloromethane
DDA	Data-dependent Acquisition
DEA	Drug Enforcement Administration
DIA	Data-independent Acquisition

DIPEA	Diisopropylethylamine
EIC	Extracted Ion Chromatogram
EMCDDA	European Monitoring Centre for Drugs and Drug Addiction
ESI	Electrospray Ionisation
ESI+	Positive Electrospray Ionisation Mode
EWA	Early Warning Advisory
FbF	Find by Formula
FbI	Find by Ion
FTIR	Fourier Transform Infra-red Spectroscopy
FWHM	Full Width at Half Maximum
GC-MS	Gas Chromatography – Mass Spectrometry
GPR	Gaussian Process Regression
HR	High Resolution
HRMS	High-resolution Mass Spectrometry
Hy	Hydrophilic Factor
ICP-MS	Inductively Coupled Plasma – Mass Spectrometry
IS	Internal Standard
KMD	Kendrick Mass Defect
LC-ELSD	Liquid Chromatography – Evaporative Light Scattering Detector
LC-MS	Liquid Chromatography – Mass Spectrometry
LC-QTOF-MS	Liquid Chromatography – Quadrupole Time of Flight – Mass Spectrometry
LD₅₀	Median Lethal Dose
LR	Low Resolution
<i>m/z</i>	Mass-to-charge Ratio
MALDI	Matrix-assisted Laser Desorption/Ionisation
MCC	Matthew's Correlation Coefficient
MCR	Multicomponent reaction
MDA	3,4-methylenedioxyamphetamine
MDF	Mass Defect Filtering
MDMA	3,4-methylenedioxymethamphetamine

MFE	Molecular Feature Extraction
MPP	Mass Profiler Professional
MRM	Multiple Reaction Monitoring
MS/MS	Tandem Mass Spectrometry
MSC	Molecular Structure Correlator
MS^E	Elevated Mass Spectrometry
MSE	Mean Square Error
NDSHS	National Drug Safety Household Survey
NIST	National Institute of Standards and Technology
NLF	Neutral Loss Filtering
NMI	National Measurement Institute
NMR	Nuclear Magnetic Resonance Spectroscopy
NPF	Non-pharmaceutical Fentanyl
NPP	N-phenethyl-4-piperidone
NPS	New Psychoactive Substances
NSO	Novel Synthetic Opioid
ONDCP	Office of National Drug Control Policy
PCA	Principal Component Analysis
PCDL	Personal Compound Database and Library
PIS	Product Ion Searching
PLS-DA	Partial Least Squares – Discriminant Analysis
PMMA	Para-methoxymethylamphetamine
QqQ	Triple Quadrupole
RFE	Recursive Feature Extraction
RMSE	Root Mean Square Error
RRT	Relative Retention Time
RT	Retention Time
S/N	Signal-to-noise Ratio
SD	Standard Deviation
SPE	Solid Phase Extraction

STRL	Special Testing and Research Laboratory
SVM	Support Vector Machine
SWATH	Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra
SWGDRUG	Scientific Working Group for the Analysis of Seized Drugs
TBAB	Tetra-N-butylammonium bromide
TIC	Total Ion Chromatogram
UHPLC-MS	Ultra High-performance Liquid Chromatography – Mass Spectrometry
UNODC	United Nations Office on Drugs and Crime
VBA	Visual Basic for Applications

Publications and Conference Proceedings

Refereed Journal Publications

1. **Klingberg, J.**, et al., *Collision-Induced Dissociation Studies of Synthetic Opioids for Non-targeted Analysis*. *Front. Chem.*, 2019. **7**(331).
2. **Klingberg, J.**, et al., *Finding the Proverbial Needle: Non-targeted Screening of Synthetic Opioids in Equine Plasma*. *Drug Test. Anal.* 2020. **DOI**: 10.1002/dta.2893

Refereed Conference Proceedings (presenting author underlined)

1. **Klingberg J.**, Shimmon R, Cawley A, Fu S. *Organic Impurity Profiling of Fentanyl Derivatives*. Oral presentation at the 8th European Academy of Forensic Science Conference (August 2018), Lyon, France.
2. **Klingberg J.**, Shimmon R, Cawley A, Fu S. *Collision-induced Dissociation Studies of Synthetic Opioids for Non-targeted Screening*. Poster presentation at the 24th Australian and New Zealand Forensic Science Society International Symposium (September 2018), Perth, Australia.
3. **Klingberg J.**, Cawley A., Pasin D, Fouracre C, Fu S. *Development of Non-targeted LC-HRMS Screening for Synthetic Opioids in Equine Plasma*. Oral presentation at the 72nd Meeting of the Association of Official Racing Chemists (May 2019), Paris, France.
4. **Klingberg J.**, Cawley A, Shimmon R, Pasin D, Fouracre C, Fu S. *Development of Non-targeted Screening Strategies for Synthetic Opioids*. Oral presentation at the Forensic and Clinical Toxicology Association Conference (June 2019), Adelaide, Australia.
5. **Klingberg J.**, Cawley A, Shimmon R, Pasin D, Fouracre C, Fu S. *Development of a Non-targeted Screening Workflow for the Detection of Synthetic Opioids in Equine Plasma*. Poster presentation at the 57th Annual Meeting of the International Association of Forensic Toxicologists (September 2019), Birmingham, United Kingdom.

Abstract

Synthetic opioids are a drug class of particular concern due to their incredibly high potency and the large public health threat that they pose. These compounds have also seen significant modification, highlighting the importance of developing techniques that can detect them without relying on databases or certified reference materials. This work provides a comprehensive investigation into the detection and profiling of synthetic opioids from the perspective of both drug screening in biological matrices and analysis of seized drug samples.

Collision-induced dissociation studies were conducted on a range of different synthetic opioid standards and common product ions belonging to each opioid subclass were identified for use in non-targeted screening strategies. Product ion searching, Kendrick Mass Defect analysis and recursive feature extraction approaches were evaluated for data analysis. Product ion searching and Kendrick mass defect analysis proved effective, with estimated screening cut-offs proposed of 0.05 ng/mL and 0.1 ng/mL, respectively. Recursive feature extraction was found to have a high sensitivity for the detection of spiked compounds, however unbiased extraction of all compounds within a sample presented issues with relevance for screening.

Machine learning approaches were investigated for the identification of unknown compounds. A Naïve Bayes classification model was trained, exploiting the common fragmentation pathways identified, to predict the opioid subclass of a sample with 89.5% accuracy. Additionally, a Gaussian Process Regression model was optimised to predict the experimental relative retention time of a compound based on its molecular features. Relative retention times were predicted for 79.7% of the samples within ± 0.1 of their experimental value. By using these models as complementary approaches putative identities of unknown compounds can be proposed with greater confidence before confirmation using certified reference materials.

A preliminary study was also conducted into the synthetic route profiling of acetyl fentanyl. Several common impurities were identified, as well as a number of impurities that were unique to a specific method. These impurities can provide an analyst with an indication of the method used in the synthesis of a seized sample. Furthermore, a statistical approach was taken, with the creation of principal component analysis plots and classification models. The PCA plots showed distinct separation between samples made with different methods and the trained classification models

displayed high accuracy. These results should be reviewed in context, however, as small sample sizes were used in this preliminary study.

Chapter 1:

Introduction

Chapter 1: Introduction

1.1 Background

1.1.1 Impact of Drug Use on Society

It is estimated that approximately 271 million people worldwide between the ages of 15 and 64 used drugs at least once in 2017 ^[1]. While the number of drug users has been increasing in recent years, this increase is in proportion with the rise in global population, meaning that the prevalence of drug use has remained relatively stable, at about 5.5% of global population between 15 and 64 ^[1]. Figure 1-1 provides an overview of the prevalence of different drugs in the current market. The Australian Bureau of Statistics (ABS) reported that, for the period of 2018-19, illicit drug offences were one of the most common principal offences, making up 20% of all offenders in Australia ^[2]. Illicit drug offences have consistently been one of the top 3 principal offences in Australia since 2013 ^[2]. The Illicit Drug Data Report, published by the Australian Criminal Intelligence Commission (ACIC), states that the number of illicit drug seizures reached 112,827 in 2017-18 (about 1 seizure every 5 minutes), which accounted for a record 30.6 tonnes of drugs seized ^[3].

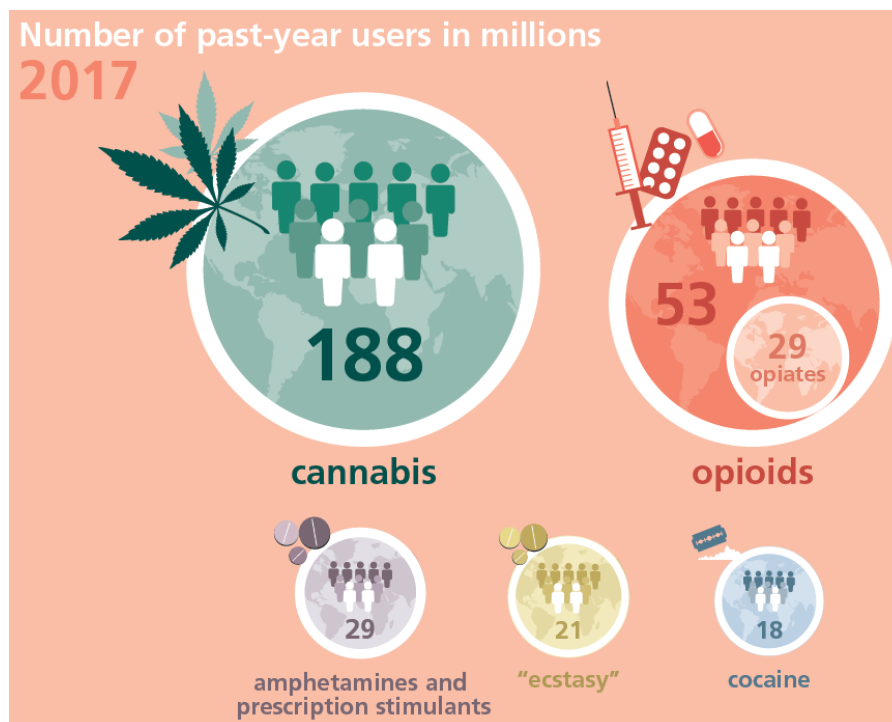


Figure 1-1: Trends in number of drug users by drug class (Adapted from UNODC) ^[1]

The widespread use of drugs in our society is, at its heart, a public health issue. The emergence of an extensive, lucrative and often violent black market as an unintended consequence of international drug control has caused a 'policy displacement' meaning that resources have been focused on law enforcement responses and drawn away from health care initiatives ^[4]. In addition to this, the use of the criminal justice system against those who use illicit drugs can often lead to increased marginalisation of these people, who often come from marginal groups of society originally. This can cause difficulties and can diminish the capacity for healthcare providers to provide help and treatment to those in most need of it ^[4]. The use of illicit drugs can lead to further health issues, especially if it progresses into a drug use disorder. It is estimated that more than half of all people who inject drugs are living with hepatitis C, and one in eight are living with HIV ^[1].

Problem drug use is defined by the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) as 'the use of drugs in a way that significantly increases the risk of serious, adverse physical, psychological or social consequences for the user' ^[5]. It is estimated that approximately 585,000 people lost their lives as a result of drug use around the world in 2017 ^[1]. The burden of disease is a measure of the health impact of diseases, and related risk factors, on a population ^[6]. This phenomenon is measured through disability-adjusted life years (DALYs), which combine the impact of dying early and living with illness ^[6]. As a part of these studies, the burden that can be attributed to various risk factors is studied. The attributable burden is defined as the reduction in burden that would have occurred if exposure to the risk factor had been avoided or reduced to the lowest possible level of exposure ^[6].

The Global Burden of Disease Study 2017 found that approximately 44.7 million years of life were lost worldwide because of drug use ^[7]. Of this amount, approximately 21.5 million years lost were linked to opioid use disorders, which accounts for about 48% of the burden of disease attributable to drug use disorders ^[7]. The Australian Institute of Health and Welfare (AIHW) conducted a study of the Australian society to determine common causes of illness and death. This study showed that, in 2015, the burden of disease and death attributable to drug use was 2.7% of the total burden of disease, with opioid use alone responsible for 1.0% of the total burden ^[8]. This accounts for an increase of 0.9% from the previous study conducted in 2011 ^[6]. The extent of drug use in society and its impact on community health demonstrates the need for measures to reduce the consumption and harm of illicit drugs.

1.1.2 New Psychoactive Substances

The United Nations Office on Drugs and Crime (UNODC) defines new psychoactive substances (NPS) as ‘substances of abuse, either in a pure form or a preparation, that are not controlled by the 1961 Single Convention on Narcotic Drugs or the 1971 Convention on Psychotropic Substances, but which may pose a public health threat’ ^[9]. In this context, the word ‘new’ does not necessarily mean new inventions, but merely substances that have appeared on the drug market relatively recently ^[9]. These NPS imitate the effects of ‘traditional’ illicit drugs such as cocaine, 3,4-methylenedioxymethamphetamine (MDMA), heroin and cannabis ^[10]. The producers of these compounds alter the chemical structures of these substances slightly in order to bypass current legislation while keeping the effects the same ^[10]. Marketing phrases such as ‘legal highs’, ‘bath salts’ and ‘research chemicals’ are often used to promote these substances and they often bear the disclaimers of ‘not for human consumption’ or ‘for research purposes only’ in order to circumvent drug abuse legislation ^[10-15]. One of the greatest concerns to have come out of the rise of NPS is that their effects on the human body are not yet fully understood ^[10]. Safety data relating to the toxicity of many of these compounds is limited and very little is known about the long-term side effects ^[1]. The potential health implications of these substances are a cause for concern, especially if their use continues to increase.

1.1.2.1 International Drug Trends

Recent years have seen the emergence of a huge array of different NPS onto the global drug market. As of December 2019, more than 950 different NPS have been reported to the UNODC Early Warning Advisory (EWA) ^[9]. The EMCDDA currently monitor more than 450 different NPS, with 101 of these being reported for the first time in 2014 ^[14]. The World Drug Report 2019 stated that 78 new substances have been reported to the UNODC in 2017, with a total of 492 different drugs being reported ^[1]. The success of many of these drugs can, in part, be attributed to the fact that they can be easily found through various internet sources or specialised ‘head shops’ where they were sold quite openly ^[14]. In Australia, the introduction of the Crimes Legislation Amendment (Psychoactive Substances and Other Measures) Act 2015 has led to further restrictions on the import and sale of these compounds where unknown psychoactive substances are prohibited ‘unless the seller can prove that they are in fact a substance which is permitted under a law or is otherwise subject to an exception’ ^[16]. The EMCDDA’s monitoring of various internet shops found 651 different sources which were selling products marketed as ‘legal highs’ or ‘research chemicals’ to consumers in Europe ^[14].

The constantly changing landscape of the NPS market provides a large challenge for law enforcement, with new drugs being introduced and previously reported drugs disappearing. Figure 1-2 provides an overview of the changes to the NPS market as reported by the UNODC in 2017.



Figure 1-2: Changes to the NPS market reported in 2017 (Adapted from UNODC) ^[1]

Along with the fluctuation of the types of compounds reported, Reuter and Pardo found that NPS generally appeal to four distinct niches of the drug market rather than its entirety ^[17, 18]. These four niches can be defined as ^[17, 18]:

- (1) Drug users wishing to avoid criminal or legal sanctions but desiring substances that mimic the effects of prohibited drugs.
- (2) Individuals who are under supervision that attempt to circumvent drug testing by using undetectable substances.
- (3) Drug users that are attracted to new altered states of mind produced by entirely novel drugs.
- (4) Drug suppliers who adulterate or potentiate currently controlled substances.

It should be noted that for the second niche defined, the legality of the substance in use is not always an issue. An NPS that is not detected by the tests being performed is helpful to the user regardless of whether or not it has been banned ^[18].

In addition to the large number of substances that are marketed as alternatives to controlled drugs, NPS are frequently sold in various compositions with a variety of additional compounds, including ‘traditional’ illicit drugs, pharmaceuticals and adulterants ^[1]. Some drug users can unwittingly end up using NPS due to these substances being marketed as something else. Various countries in Europe and South-East Asia have reported seizures of ‘ecstasy’ tablets which contained primarily a mixture of non-controlled substances, including NPS, with little or no MDMA present ^[1]. There have been reported cases of people purchasing pills with the street name of ‘Norco’, a medication containing hydrocodone and acetaminophen, which instead contained fentanyl and other synthetic opioids ^[19, 20]. This can complicate the situation with NPS, as users are often unaware of what they are actually taking, making health issues that may arise from poly drug use a real concern ^[1].

1.1.2.2 Australian Drug Trends

The availability of NPS, along with the diversity of the range of substances available, has been expanding rapidly in East Asia, South-East Asia and Oceania. The number of seizures of NPS at the Australian border has decreased by 29%, from 968 in 2016-17 to 687 in 2017-18 ^[3], however this is still higher than previous reporting periods. The total weight of NPS seized is also decreasing, with only 33.1 kg being seized in 2017-18 ^[3]. The National Drug Safety Household Survey (NDSHS) contained questions on NPS use for the second time in 2016 ^[3]. This survey found that 1% of Australians aged 14 or over reported using an NPS other than synthetic cannabinoids in their lifetime, showing an increase from 0.4% from the previous survey conducted in 2013 ^[3, 10]. A 2019 national study of regular ecstasy users indicated that 30% reported recent NPS use, similar to the 31% reported in 2018 ^[21].

1.1.3 Legal and Law Enforcement Issues

It has been stated that, ‘drug markets and drug use patterns change rapidly, so measures to stop them must also be quick to adapt. Thus the more comprehensive the drug data we collect and the stronger our capacity to analyse the problem, the better prepared the international community will be to respond to new challenges’ ^[22]. The large number of new substances being reported makes it difficult for policy makers and law enforcement to keep up with the changing market. A study carried

out in Germany by Maas *et al.* showed that the types of NPS found in serum samples changed in accordance with which substances were subjected to governmental control ^[23]. These findings demonstrate the rapid reactions of the recreational drug market to the newly introduced drug laws. As such, detection and analysis methods that rely on certified reference materials (CRMs) and database searching may no longer be adequate. Therefore, new techniques must be investigated and implemented to keep up with the ever-changing drug market.

The fact that resources and training opportunities are limited for law enforcement and healthcare professionals means that it is important to direct additional attention to the substances that are more likely to become widespread and pose identified harms ^[15]. As identified previously (Figure 1-2), many of the new NPS that emerge onto the drug market are relatively short-lived, meaning that they never really develop beyond a minor problem ^[1, 15]. While Reuter and Pardo identified the main niches responsible for the use of NPS ^[17, 18], it is necessary to try and predict the likely prevalence of newly emerging substances. This information can assist with effective and efficient decision-making and resource allocation concerning the response to the changing drug market ^[15]. To this end, Stogner developed a five-step method for forecasting the magnitude of the use of different NPS ^[15]. The framework that was suggested relies on a series of criteria, namely existence of a potential user base, cost analysis, relative subjective experience, ease of acquisition and dependence potential ^[15]. Stogner suggests that the potential widespread use of a particular NPS appears to be directly tied to the existence of individuals who use a pharmacologically similar substance as a drug of choice ^[15]. Additionally, the potential users' cost analysis must be considered relative to existing drugs and take into account financial and health costs, as well as perceived costs, which can be affected by factors such as the legal status of the NPS in question ^[15]. Substances which have positive effects and minimal immediate psychological drawbacks are more likely to see continual use, as well as those which are easy to access initially, such as in head shops ^[15]. Finally, compounds that cause physical and/or psychological dependence are more likely to be associated with repeated use and greater personal and societal harms ^[15].

While the above factors assist with the prediction of the prevalence of a given substance, the appropriate reaction to the emergence of said substance should also be governed by the potential for harm that it presents ^[15]. For example, a particular substance may become highly popular and turn out to have a large negative impact on public health, such as long-term dependence or behavioural or health consequences, yet it may displace a less harmful existing drug ^[17]. Alternatively, a new substance may fail in the long term due to harmful effects; however a large number of people may have tried the substance before those effects were publicised, leading to a substantial number of

victims ^[17]. Consideration of the various factors, and the implementation of the forecasting framework set out by Stogner, can allow law enforcement agencies, policy makers and healthcare industries to more strategically allocate the resources they have available to them to combat the rise in NPS available on the drug market.

A number of different approaches to the regulation of NPS have been considered and, in some cases, trialled with varying degrees of success. A paper published by Reuter and Pardo looked at the main options available to legislators and looked at the different pros and cons of each approach ^[17]. The findings from their paper can be found in Table 1-1. The best known and most common approach to the restriction of illicit drugs is to schedule them under specific legislative controls ^[17]. Some countries have begun to use a generic system for scheduling, where a group of similar substances are banned, rather than having to consider each substance separately ^[17, 24]. Alternatively, an analogue system may be used which ‘addresses more general aspects of similarity in chemical structure to a ‘parent’ compound; these aspects might be supplemented by a requirement for similarity in pharmacological activity, attempting a more specific delineation of the analogue system’s sphere of control’ ^[17, 24].

A more drastic measure to the control of NPS is imposing a complete ban on all psychoactive substances other than those that have been listed as exemptions, including things such as medicines, alcohol and tobacco ^[17]. This style of legislation was introduced into Australia under the Crimes Legislation Amendment (Psychoactive Substances and Other Measures) Act 2015 ^[25] and, more recently, to the UK under the British Psychoactive Substances Bill, which was passed in 2016 ^[18]. This approach to the regulation of NPS has some clear advantages. Firstly, if this approach were successful, it would reduce the number of new NPS introduced in a given period of time ^[18]. New variations of NPS are often introduced into the market as a response to the scheduling of current substances under government restriction ^[18, 26]. Under a total ban, however, the government is no longer simply reactive. It would no longer be necessary for law enforcement agencies to confirm the molecular composition of a seized substance against the list of scheduled substances, rather, any non-exempt psychoactive substance would be considered illegal ^[18].

Table 1-1: Summary of main regulatory options available (Adapted from Reuter and Pardo) ^[17]

Type of regulatory framework	Examples of use	Problems/Considerations
Foodstuff regulations (presumption of safety)	Weight loss products Performance enhancement products	<ul style="list-style-type: none"> • Rapid product turnover • Difficulty/cost of evaluating individual products • Public presumes that government has approved • Allows too much exposure to dangerous products
Regulation of specific commodities	Alcohol regulation Tobacco regulation Solvents	<ul style="list-style-type: none"> • No relevant examples of applicability to 'legal highs' due to low product turnover
Medicines regulations (safety/efficacy must be proven)	Mephedrone and BZP in several EU states	<ul style="list-style-type: none"> • Deceptive since NPS not medicines • Manufacturers/potential users advocate for approval • Production and ingredients tightly controlled • Only substances with proven medical use can be approved
Scheduling and regulation of illicit substances	U.N. Conventions National laws	<ul style="list-style-type: none"> • Slow to address new substances, though new temporary restraining orders have proven effective • Furthers development of substances outside of controls • Biased towards prohibition
Total ban of any new psychoactive substance, with exemption for medical research	Irish and U.K. Psychoactive Substances Act	<ul style="list-style-type: none"> • No possibility of developing less harmful recreational drugs • Difficulty of operationalising psychoactivity

There have been many criticisms for the use of a total ban, however. Referring to the NPS niches defined by Reuter and Pardo ^[17, 18], the implementation of a total ban would only reduce the first of these four niches. The other 3 do not rely on the legal status of the substance and therefore would still serve as sources of demand for the NPS market ^[18]. The first major issue that has been identified, however, is the concept of defining psychoactivity ^[18]. Different legislation defines a 'psychoactive effect' differently. The Crimes Legislation Amendment (Psychoactive Substances and Other Measures) Act 2015 in place in Australia defines a psychoactive effect as '(a) stimulation or depression of the person's central nervous system, resulting in hallucinations or in a significant disturbance in, or significant change to, motor function, thinking, behaviour, perception, awareness or mood; or (b) causing a state of dependence, including physical or psychological addiction' ^[16]. The issue with definitions such as this is that it implies that any substance, other than those exempted, has the potential to cause harm and to the same extent ^[18].

The other issue comes when trying to operationalise psychoactivity. The Advisory Council on the Misuse of Drugs (ACMD) in the UK has stated that ‘Psychoactivity in humans cannot be definitively established in many cases in a way that would definitely stand up in a court of law where a high threshold of evidence is required. There is currently no way to define psychoactivity through a biochemical test, therefore there is no guarantee of proving psychoactivity in a court of law. The only definitive way of determining psychoactivity is via human experience, which is usually not documented’ ^[27]. While there have been many critics of the way that total bans have been implemented, few disagree with the principle of having a total ban in place ^[18].

International control will also have an effect on the availability of illicit substances on a drug market. The United Nations General Assembly produced an outcome document from their Special Session on the World Drug Problem in 2016, which contained over 100 resolutions intended to address the drug problem at large, including the threat posed by synthetic opioids ^[28]. One of the major themes of this outcome document was the reduction of demand and supply for illicit substances ^[28]. With international cooperation, the production of illicit substances, such as synthetic opioids, can be reduced through legislation and regulation in the producer countries, which can have a drastic effect on the scope of a drug market at the endpoint.

An issue that becomes apparent under any of the regulatory options available is that the decision maker will always be biased towards prohibition ^[17]. When considering the case of NPS, the consequences for refraining from prohibiting a substance that turns out to be dangerous far outweigh the benefits gained from allowing a new substance to enter the market with appropriate regulatory controls ^[17]. Regardless of the regulatory system employed, the motivation of the policy makers must always be the wellbeing of society. If a blanket ban were workable, it would help to establish public confidence in the government’s ability to stay ahead of the problem ^[17]. It is likely that the best approach to the regulation of illicit drugs, and NPS in particular, will be highly context specific and therefore should be based on the individual country ^[17].

1.2 Synthetic Opioids

The term ‘opioid’ literally means ‘opium-like substance’ ^[29]. Opium is the dried juice of the opium poppy, *Papaver Somniferum*, which has been used for the relief of pain for thousands of years ^[30]. Synthetic opioids are man-made compounds which mimic the effects of natural opiates such as morphine and codeine, which occur naturally in opium ^[29]. Pharmacologically, opioids produce an analgesic effect on the user through the binding to specific opioid receptors ^[29, 31]. These compounds

mimic the effects of endogenous opioid peptides, such as endorphins, enkephalins and dynorphins, which are present in the central and peripheral nervous systems ^[32]. There are three main subtypes of opioid receptors in the body, mu (μ), kappa (κ) and delta (δ) ^[33]. While opioids may act on any of these, most clinically available opioid analgesics act upon the μ -opioid receptor ^[33]. Fentanyl is one of the most well-known of the synthetic opioids and has been used medicinally since the 1960s for pain management and anaesthesia ^[34]. Fentanyl and its related analogues, such as sufentanil and carfentanil, are among the most potent μ -opioid receptor agonists known ^[33].

In addition to the illicit use of non-pharmaceutical fentanyls (NPFs), there is a vast array of novel synthetic opioids (NSOs) that continue to emerge onto the drug market ^[19]. While there can be some overlap between these classifications, NSOs usually refer to the newer classes of opioids, such as the AH and U series', which were originally developed to act as new opioid agonists, but never brought to the market for human use ^[19]. In addition, the W series of compounds have recently been proven to show no action at the opioid receptors ^[35]. Due to their similar effects and structural similarities, however, these compounds are often considered to be 'pseudo-NSOs'. The UNODC have identified opioids as the class of compounds which cause the most harm, accounting for nearly 68% of all deaths attributed to drug use worldwide in 2017 ^[1]. Fentanyl analogues alone accounted for over 58% of all overdose deaths in the United States in 2017, representing an increase of 45% from 2016 ^[36]. While the majority of these reported overdose deaths may come from fentanyl itself, these figures still demonstrate the large issues posed by these compounds. There is also an increasing trend of fentanyl derivatives, and other synthetic opioids, being mixed into heroin, along with other illicit drugs, or being sold as prescription opioids ^[1].

Traditional approaches to illicit drug analysis in biological samples may fail to detect these compounds and routine analyses rarely screen for these compounds ^[37]. This may be due to there being little to no cross-reactivity of these compounds with traditional immunoassay tests ^[38]. Additionally, data for these compounds may not be present in the mass spectral libraries that are used for drug screening ^[38]. Very little is known about the effects of these drugs on humans, which makes them even more dangerous, especially in cases of poly drug use and where the users do not know what they are taking.

1.2.1 Structure and Chemistry

1.2.1.1 Non-pharmaceutical Fentanyl

Fentanyl and its related analogues belong to a class of compounds called 4-anilidopiperidines, with fentanyl being the representative compound of this class ^[39]. Fentanyl (Figure 1-3), also known by its IUPAC name N-phenyl-N-[1-(2-phenylethyl)piperidin-4-yl]propanamide ^[40], was first discovered by Dr Paul Janssen in 1960 ^[41, 42]. The compound has the molecular formula $C_{22}H_{28}N_2O$ and an exact mass of 336.22016 u in its free base form ^[40]. It appears as a light yellow oil in its free base form and has a melting point of 83-84 °C ^[40]. Often fentanyl can be found as hydrochloride or citrate salts which have the appearance of a white powder ^[43].

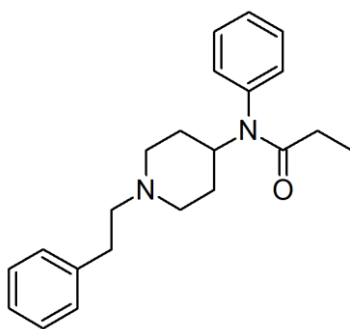
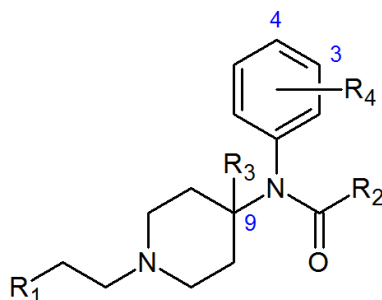


Figure 1-3: Chemical structure of fentanyl

Table 1-2 provides chemical information for some of the major compounds belonging to the 4-anilidopiperidine class. An important advance in the chemistry of this class of analgesic compounds was the incorporation of small polar moieties at the 9-carbon of the piperidine ring (Table 1-2) ^[41]. The compounds carfentanil (also known by the IUPAC name methyl 1-(2-phenylethyl)-4-(N-propanoylanilino)piperidine-4-carboxylate) ^[44], remifentanyl (methyl 1-(3-methoxy-3-oxopropyl)-4-(N-propanoylanilino)piperidine-4-carboxylate) ^[45], and sufentanyl (N-[4-(methoxymethyl)-1-(2-thiophen-2-ylethyl)piperidin-4-yl]-N-phenylpropanamide) ^[46] were found to be much more potent and exhibited longer lasting effects than fentanyl ^[41, 47]. On the other hand, alfentanil is a very short-acting opioid anaesthetic ^[41]. The reason for this difference is that the tetrazolinone ring of alfentanil has a large concentration of electronegative atoms due to electron withdrawal through the connected ethylene chain ^[41]. This results in the piperidino nitrogen of this compound being far less basic ($pK_a = 6.5$) than fentanyl ($pK_a = 8.5$) ^[41]. Therefore, at a physiological pH of 7.4, alfentanil is mostly unionised, while fentanyl is mostly protonated, allowing for a faster absorption rate ^[41].

The compounds presented in Table 1-2 do not represent the entirety of the range of fentanyl derivatives; there are many other compounds with a similar core structure synthesised for their analgesic and euphoric properties [41, 43, 48-50]. The compounds presented provide a snapshot of the range of different substances, which have made their way onto the illicit drug market.

Table 1-2: Chemical information of the fentanyl derivatives



Functional Groups	Common Name	Molecular Formula	Exact Mass (u)	Melting Point (°C)
R ₁ – C ₆ H ₅ R ₂ – CH ₃	Acetyl fentanyl	C ₂₁ H ₂₆ N ₂ O	322.20451	256.6 (HCl) [51-53]
R ₁ – C ₆ H ₅ R ₂ – CH ₂ CH ₃ R ₃ – CO ₂ CH ₂ CH ₃	Carfentanil	C ₂₄ H ₃₀ N ₂ O ₃	394.22564	152.2 (Citrate salt) [54]
R ₁ – C ₄ H ₉ S R ₂ – CH ₂ CH ₃ R ₃ – CH ₂ OCH ₂ CH ₃	Sufentanil	C ₂₂ H ₃₀ N ₂ O ₂ S	386.20280	96.6 (Base) [46, 55]
R ₁ – CO ₂ CH ₃ R ₂ – CH ₂ CH ₃ R ₃ – CO ₂ CH ₂ CH ₃	Remifentanil	C ₂₀ H ₂₈ N ₂ O ₅	376.19982	192 – 197 (HCl) [56]
R ₁ – CN ₄ OCH ₂ CH ₃ R ₂ – CH ₂ CH ₃ R ₃ – CH ₂ OCH ₂ CH ₃	Alfentanil	C ₂₁ H ₃₂ N ₆ O ₃	416.25359	140.8 (Base) [57]
R ₁ – C ₆ H ₅ R ₂ – CHCH ₂	Acryl fentanyl	C ₂₂ H ₂₆ N ₂ O	334.20451	101 – 103 (Base) [58]
R ₁ – C ₆ H ₅ R ₂ – CH ₂ CH ₂ CH ₃	Butyr fentanyl	C ₂₃ H ₃₀ N ₂ O	350.23581	206.5 (HCl) [59]
R ₁ – C ₆ H ₅ R ₂ – CH ₂ CH ₂ CH ₃ R ₄ – 4-F	4-fluorobutyr fentanyl	C ₂₃ H ₂₉ FN ₂ O	368.22639	114 – 115 [60]
R ₁ – C ₆ H ₅ R ₂ – C ₄ H ₃ O	Furanyl fentanyl	C ₂₄ H ₂₆ N ₂ O ₂	374.19943	232.7 (HCl) [61]
R ₁ – C ₆ H ₅ R ₂ – CH ₂ CH ₂ CH ₂ CH ₃	Valeryl fentanyl	C ₂₄ H ₃₂ N ₂ O	364.25146	—

*All R groups are assumed to be H unless otherwise specified

1.2.1.2 Novel Synthetic Opioids

While many synthetic opioids that have been developed are derived from pharmaceutically used compounds, such as fentanyl and meperidine, NSOs with distinctly different structures have recently been emerging ^[62]. There are many different types of these NSOs, and they can be organised into different families or series based on similar chemical structures. The AH series, named for Allen and Handburys Limited, the company which patented the first compound AH-7921 ^[63, 64], are NSOs which have recently emerged onto the illicit drug market ^[63]. This family of NSOs are *N*-substituted cyclohexylmethylbenzamides, where the benzamide group is either mono or dichlorinated ^[63, 65]. Table 1-3 provides an overview of some of the compounds that form part of this series.

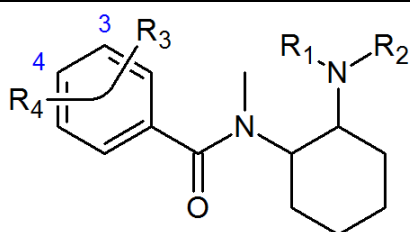
Table 1-3: Chemical information for the AH series of novel synthetic opioids

Functional Groups	Common Name	Molecular Formula	Exact Mass (u)	Melting Point (°C)
R ₁ – CH ₃ R ₂ – CH ₃ R ₃ – 3-Cl R ₄ – 4-Cl	AH-7921	C ₁₆ H ₂₂ Cl ₂ N ₂ O	328.11093 (³⁵ Cl/ ³⁵ Cl) 330.10798 (³⁷ Cl/ ³⁵ Cl) 332.10503 (³⁷ Cl/ ³⁷ Cl)	219.6 (HCl) ^[66]
R ₁ & R ₂ – C ₅ H ₁₀ R ₃ – 3-Cl R ₄ – 4-Cl	AH-7959	C ₁₉ H ₂₆ Cl ₂ N ₂ O	368.14223 (³⁵ Cl/ ³⁵ Cl) 370.13928 (³⁷ Cl/ ³⁵ Cl) 372.13633 (³⁷ Cl/ ³⁷ Cl)	–
R ₁ & R ₂ – C ₄ H ₈ NCH ₃ R ₃ – 3-Cl R ₄ – 4-Cl	AH-8507	C ₁₉ H ₂₇ Cl ₂ N ₃ O	383.15313 (³⁵ Cl/ ³⁵ Cl) 385.15018 (³⁷ Cl/ ³⁵ Cl) 387.14723 (³⁷ Cl/ ³⁷ Cl)	–
R ₁ – CH ₃ R ₂ – CH ₃ R ₃ – 4-Cl	AH-8529	C ₁₆ H ₂₃ ClN ₂ O	294.14990 (³⁵ Cl) 296.14695 (³⁷ Cl)	–
R ₁ – CH ₃ R ₂ – CH ₃ R ₃ – 3-Cl	AH-8532	C ₁₆ H ₂₃ ClN ₂ O	294.14990 (³⁵ Cl) 296.14695 (³⁷ Cl)	–
R ₁ – CH ₃ R ₂ – CH ₃ R ₃ – 2-Cl	AH-8533	C ₁₆ H ₂₃ ClN ₂ O	294.14990 (³⁵ Cl) 296.14695 (³⁷ Cl)	–

*All R groups are assumed to be H unless otherwise specified

Another family of NSOs that has made an appearance on the illicit drug market recently is the U series. The most common compound in this series, U-47700, also known as ‘pink’, is a structural isomer of AH-7921 ^[67]. This compound was originally developed and patented, as part of a series of *N*-(2-aminocycloaliphatic)benzamides, in 1978 by Jacob Szmuszkovicz who worked for the Upjohn Company, hence the name of the series ^[62, 67, 68]. Compounds in this family of NSOs feature two chiral centres and therefore can occur as two geometrical diastereoisomers (*cis* and *trans*) which may have different pharmacological properties ^[67, 69]. The *trans* isomer in particular has been targeted for pharmacological studies ^[70]. Table 1-4 and Table 1-5 provide an overview of different compounds, which have been found to be a part of the U series of opioid agonists.

Table 1-4: Chemical information for some of the U series of novel synthetic opioids



Functional Groups	Common Name	Molecular Formula	Exact Mass (u)	Melting Point (°C)
$R_1 - \text{CH}_3$ $R_2 - \text{CH}_3$ $R_3 - 3\text{-Cl}$ $R_4 - 4\text{-Cl}$	U-47700	$\text{C}_{16}\text{H}_{22}\text{Cl}_2\text{N}_2\text{O}$	$328.11093 (^{35}\text{Cl}/^{35}\text{Cl})$ $330.10798 (^{37}\text{Cl}/^{35}\text{Cl})$ $332.10503 (^{37}\text{Cl}/^{37}\text{Cl})$	97 – 98.5 ^[67, 71]
$R_1 - \text{CH}_2\text{CH}_3$ $R_2 - \text{CH}_2\text{CH}_3$ $R_3 - 3\text{-Cl}$ $R_4 - 4\text{-Cl}$	U-49900	$\text{C}_{18}\text{H}_{26}\text{Cl}_2\text{N}_2\text{O}$	$356.14223 (^{35}\text{Cl}/^{35}\text{Cl})$ $358.13928 (^{37}\text{Cl}/^{35}\text{Cl})$ $360.13633 (^{37}\text{Cl}/^{37}\text{Cl})$	169 – 171 ^[72]

Contrary to the other families of NSOs, the U-series contains two different generic structures for its compounds. Table 1-5 shows the compounds of this series that contain an additional methylene spacer between the aromatic ring and the amide carbonyl group. It has been suggested that the addition of this spacer affects κ -opioid receptor specificity, since U-47700 is known to be a strong μ -opioid agonist, with very limited affinity for the κ -opioid receptor ^[67], while U-50488 has been found to preferentially bind to the κ -opioid receptor ^[33, 38]. A study published in 1992 by Takemori and Portoghesi suggested that the dimensions of the spacer would play an important role in receptor selectivity, as bivalent binding of a compound to the receptor should be favoured over univalent binding ^[73]. To confirm the role of the methylene spacer in the generic structure of the U series

opioids, a number of different U series analogues were synthesised to determine where they displayed receptor activity ^[74].

Table 1-5: Chemical information for the U series opioids containing a methylene spacer

Functional Groups	Common Name	Molecular Formula	Exact Mass (u)	Melting Point (°C)
R ₁ – CH ₃ R ₂ – CH ₃ R ₃ – 2-Cl R ₄ – 4-Cl	U-48800	C ₁₇ H ₂₄ Cl ₂ N ₂ O	342.12658 (³⁵ Cl/ ³⁵ Cl) 344.12363 (³⁷ Cl/ ³⁵ Cl) 346.12068 (³⁷ Cl/ ³⁷ Cl)	–
R ₁ & R ₂ – C ₄ H ₈ R ₃ – 3-Cl R ₄ – 4-Cl	U-50488	C ₁₉ H ₂₆ Cl ₂ N ₂ O	368.14223 (³⁵ Cl/ ³⁵ Cl) 370.13928 (³⁷ Cl/ ³⁵ Cl) 372.13633 (³⁷ Cl/ ³⁷ Cl)	115 – 188 (HCl) ^[75]
R ₁ – CH ₃ R ₂ – CH ₃ R ₃ – 3-Cl R ₄ – 4-Cl	U-51754	C ₁₇ H ₂₄ Cl ₂ N ₂ O	342.12658 (³⁵ Cl/ ³⁵ Cl) 344.12363 (³⁷ Cl/ ³⁵ Cl) 346.12068 (³⁷ Cl/ ³⁷ Cl)	– 98 ^[76]

In addition to the major subclasses of NSOs, there are a number of other NSOs that have been encountered. MT-45 is another opioid analgesic that was developed in the 1970s, however has recently been introduced into the recreational drug market ^[77]. While this compound is known to toxicologists, it doesn't fit with any of the previously described subclasses. The structure of MT-45, shown in Figure 1-4, displays potential for the development of further analogues, however, to date, only a fluorinated analogue has been reported ^[78].

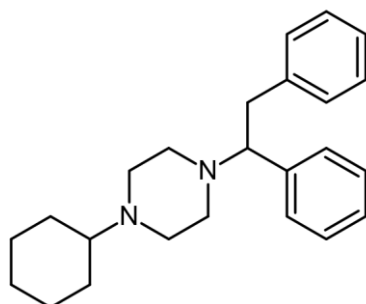


Figure 1-4: Chemical structure of MT-45

More recently, Blanckaert *et al.* reported on a newly emerging subclass of synthetic opioids, namely the benzimidazole class ^[79]. Similar to the other NSOs, these compounds are not new, with the first benzimidazole derivative with analgesic activity being synthesised in 1957, however they have only recently been detected on the illicit market. Figure 1-5 shows the chemical structures of some analogues that have been found as part of this opioid subclass ^[79].

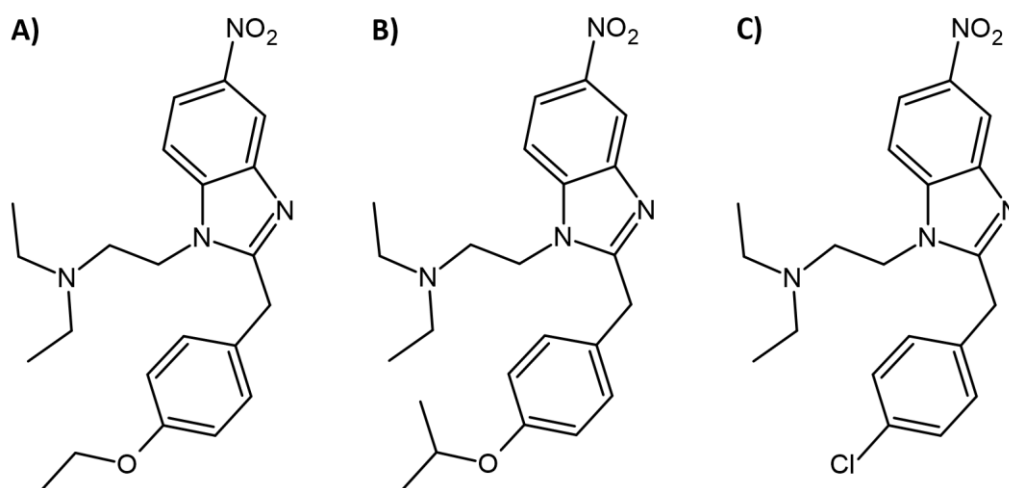
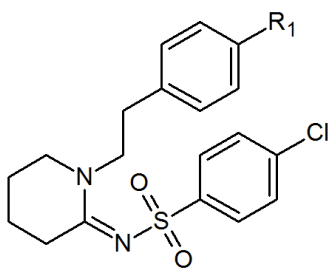


Figure 1-5: Chemical structures of benzimidazole isomers, including etonitazene (A), isotonitazene (B), and clonitazene (C)

1.2.1.3 Pseudo-Novel Synthetic Opioids

The W-series of NSOs was first developed in 1981 at a Canadian university and contains 32 different compounds ^[19, 80, 81]. The compounds found in this series of NSOs are substituted piperidylidene-2-sulfon(cyan)amide derivatives ^[81]. Very little is known about this series and, while there are many compounds which belong to this series, only a small number have been found in the illicit drug market. Table 1-6 provides an overview of the different compounds from this series that have been found.

Table 1-6: Chemical information for the W series of novel synthetic opioids

Functional Groups	Common Name	Molecular Formula	Exact Mass (u)	Melting Point (°C)
R ₁ – H	W-15	C ₁₉ H ₂₁ ClN ₂ O ₂ S	376.10124 (³⁵ Cl) 378.09829 (³⁷ Cl)	123.5 ^[82]
R ₁ – NO ₂	W-18	C ₁₉ H ₂₀ ClN ₃ O ₄ S	421.08632 (³⁵ Cl) 423.08337 (³⁷ Cl)	162.6 ^[83]
R ₁ – NH ₂	W-19	C ₁₉ H ₂₂ ClN ₃ O ₂ S	391.11213 (³⁵ Cl) 393.10918 (³⁷ Cl)	—

The information presented above demonstrates the variability of NSOs in terms of their chemical structure. There are a number of gaps in the knowledge, however, indicating that little is known about some of these compounds. This is especially apparent in the compounds that have not been detected very often or have been detected for the first time recently. More recently, a number of different studies have been conducted in an attempt to map the metabolic pathways of various synthetic opioids ^[84-86]. This information is critical for the detection of these compounds in a clinical or forensic toxicological context. On the other hand, techniques for the detection of the parent drug itself are still required for many policing and border control operations. It is important to keep in mind the structural diversity of this class of drugs as it may influence techniques that are implemented to detect and analyse these compounds. While there is a large amount of diversity between the different classes of synthetic opioids, the structural similarity of different compounds within a particular class presents an interesting possibility for the development of targeted techniques that may be specific to a particular class of compounds.

1.2.2 Pharmacology and Toxicology

1.2.2.1 Non-pharmaceutical Fentanyl

The 4-anilidopiperidine class of μ -receptor opioid agonists are characterised by high analgesic potency and a relatively short duration of action ^[87]. Fentanyl has been reported to have 50 – 100x the potency of morphine and its lethal dose (LD₅₀) is estimated to be 2 mg ^[43, 47, 87-89]. This, along with

its fast onset and short duration of action, has made it the primary choice for a fast acting anaesthetic [43]. The high lipophilicity of fentanyl and its related analogues allows for rapid diffusion through membranes, including the blood brain barrier, which accounts for its rapid onset [47]. While the potency and pharmacology of many fentanyl derivatives has been well documented [47, 88, 90], some of the newer derivatives not approved for medicinal use have very little information available. Estimations of potency ratios can be made for some of the compounds, such as acetyl fentanyl, based on animal studies [88], however other analogues have no data reported at all. Table 1-7 shows the known potency ratios of some common fentanyl analogues.

Table 1-7: Potency of various fentanyl derivatives compared with morphine and fentanyl

Name	Potency ratio to morphine	Potency ratio to fentanyl
Acetyl fentanyl	15.7 [47, 88]	0.29 [47, 88]
Carfentanil	10,000 [47, 91, 92]	100 [47, 91]
Sufentanil	500 – 1,000 [43, 47]	5 – 10 [47, 90]
Remifentanil	100 – 200 [43]	1 – 2 [47]
Alfentanil	10 – 20*	0.1 – 0.2 [47]
Acryl fentanyl	169.5 [58, 93]	1.7*
Butyr fentanyl	1.5 – 7 [47, 88, 94]	0.03 – 0.13 [47, 88]
Furanyl fentanyl	100*	1 [^] [38]

*Values estimated from the potency ratio of fentanyl to morphine ($\approx 1:100$)

[^]Value estimated from similar ED₅₀ value in studies conducted on mice

While many of these compounds have less than, or about the same, potency as fentanyl, there are some with a much higher potency, such as carfentanil. When compared to heroin, which has a potency approximately 4 – 5x that of morphine [95], it shows that these compounds are still quite potent. Carfentanil has no approved human uses, however it is used in the veterinary industry for the rapid immobilisation of large animals, such as elephants [47]. It is estimated that the LD₅₀ of this drug is only 20 µg [91], which highlights the incredible public health threat it poses. Even though a specific potency for some analogues, such as 4-fluorobutyryl fentanyl (4-FBF) has not been extensively studied, it has been found that fluorinated fentanyl analogues generally had a lower potency than their un-fluorinated counterparts [41, 96].

The high potency of these compounds demonstrates the public health threat that they can pose. This is especially pertinent when the end user might not know what substances they are taking. If traditional opiates, such as heroin, are cut or substituted with these highly potent synthetic opioids, it can drastically increase the chance of an overdose due to the increased potency. Additionally, the

small window between a safe and fatal dose is of even more concern, especially since the manufacture of these compounds in clandestine laboratories is not likely to be quantitatively accurate ^[97].

1.2.2.2 *Novel Synthetic Opioids*

Many of the NSOs have very little pharmacological information available. Human studies have not been conducted on many of these compounds due to them not being brought to the market for medicinal use and only limited animal studies have been conducted. The only information available about the effects and/or toxicities of many of these compounds comes from online drug user forums and are unsubstantiated by analytical studies ^[84]. From the series of NSOs synthesised by the Allen & Hanbury Company, AH-7921 was found to be the most potent ^[63, 65, 98]. The first reports of AH-7921 on the illicit drug market came from Europe in 2012, quickly followed by its emergence in Japan in 2013 ^[86]. This compound has been shown to exhibit potency equal to, or slightly higher than, that of morphine from limited animal studies ^[63, 65, 98, 99]. In just 10 months following its first detection, this compound was implicated as a contributing cause of death in a number of fatal overdose cases in four different countries ^[86]. Although AH-7921 shows preferential selectivity for the μ -opioid receptor, its analgesic activity may also involve the κ -opioid receptors ^[63, 65, 98-100]. While there have been no clinical studies conducted on humans, it has been suggested that the LD₅₀ for this compound is higher than 10 mg/kg based on results obtained from studies in rats ^[65].

U-47700, the most common substance in the U series of NSOs, has been found to be approximately 7.5 times more potent than morphine based on limited animal models ^[20, 38, 67, 80]. This compound has been found to be a μ -selective agonist, with some evidence of limited κ -receptor agonism ^[20, 38, 62, 67, 69]. Reports from users of this compound have suggested that its effects last longer than its isomer AH-7921 ^[101]. There have been some recent studies looking at the pharmacological activity of some previously reported U-series compounds ^[64], however the continual detection of new compounds makes it difficult to stay ahead of the curve. U-50488, on the other hand, has been found to exhibit κ -selectivity, contrary to U-47700 ^[38]. This distinction between the receptor selectivity of the two compounds is important as κ -receptor agonists may not exhibit side effects like respiratory depression and constipation, which can be found in μ -agonists such as morphine ^[74]. On the other hand, κ -opioid receptor agonists cause strong dysphoria, which is an unpleasant effect and, therefore, are often avoided by drug users.

The potencies of other NSOs not belonging to either of these main classes have been studied more recently. MT-45 has been found to have a potency that is relatively low when compared to fentanyl^[78, 102]. A recent study by Baptista-Hon *et al.*^[78], however, has shown that the introduction of a fluorine substituent at various positions on the first benzene ring can increase the potency of the compound by over 20-fold. In addition, isotonitazene, a member of the recently reported benzimidazole class of opioids^[79], has been found to have a potency higher than fentanyl^[79], with one study suggesting that the potency is twice that of fentanyl^[102]. The presence of a nitro group in the structures of these compounds (Figure 1-5) has been suggested to be important for high analgesic activity, with removal of this group resulting in a reduction of the antinociceptive activity^[79, 103].

As further NSOs are detected on the illicit drug market, more research will be conducted in order to determine the pharmacological potency of these compounds.

1.2.2.3 *Pseudo-Novel Synthetic Opioids*

The W series has had very little work done to confirm the toxicology and pharmacology of its compounds. The compound W-18 has been suggested to be the most potent, with a potency approximately equal to that of carfentanyl (i.e. 10,000 the potency of morphine)^[19, 81]. There has been recent evidence, however, that suggests that these compounds do not exhibit any effects on the opioid receptors at all, and therefore are not actually opioids^[35, 104, 105]. Due to their high potency, however, it has been suggested that they are still considered dangerous analgesics that could be sold as opioids or mixed with drugs that are claimed to be opioids^[80]. The fact that there appears to be no action on the opioid receptors may, in fact, increase the danger of these compounds. People who abuse this class of drugs may be provided a false sense of security by the fact that its effects can be reversed by naloxone, when this is not actually the case^[104].

1.2.2.4 *Opioid Effects in Horses*

In addition to the large human health risk posed by these compounds, opioids have a propensity for causing stimulant-like effects when administered to horses, in addition to their pain depressing effects^[106, 107]. This stimulant effect was found to be common to all narcotic analgesics^[107], however it has been theorized that κ -opioid receptor agonists should provide analgesia without stimulation^[108]. The effect of the κ -opioid receptor agonist U-50488 on horses was evaluated and it was found that, while there was slight stimulation, it was markedly less than potent μ -agonists, such as fentanyl^[108]. The stimulant effects have been confirmed to be the result of action at the opioid receptors, as

administration of opioid receptor antagonists, such as naloxone, suppresses the stimulation ^[107]. The combination of the analgesic and stimulant properties of these substances make them of importance from an anti-doping perspective. The first detection of a synthetic opioid in equine anti-doping occurred in 2015, when AH-7921 was detected in a post-race sample in the New York Equine Drug Testing Program ^[109].

1.2.3 Synthesis

The synthesis of fentanyl first described by Dr Janssen can be found in Figure 1-6 ^[110]. This synthetic method involves a substitution reaction of (2-chloroethyl)benzene (2-CEB) with N-(4-piperidyl)-propionanilide ^[110]. This results in the ethylbenzene group being attached to the piperidino nitrogen to give the fentanyl product. The N-(4-piperidyl)-propionanilide starting material can be synthesised from simpler starting materials, such as 4-piperidone, however a protecting group for the piperidino nitrogen would be required to make sure that it remained available for the substitution reaction.

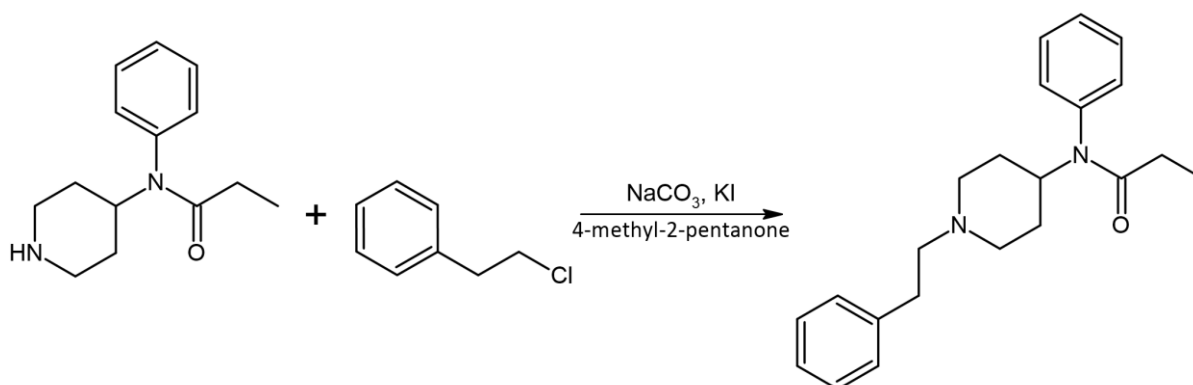


Figure 1-6: Synthesis of fentanyl as patented by Janssen ^[110]

A few different synthetic methods have been proposed for the synthesis of fentanyl from 4-piperidone ^[43, 87, 111]. While these methods differ in the starting material used and the reaction conditions applied, they share a similar generic pathway, shown in Figure 1-7. In general, the synthetic route first involves a substitution reaction at the piperidino nitrogen, similar to the Janssen method, to give the N-phenethyl-4-piperidone (NPP, **1**) intermediate. This is followed by a reaction with aniline at the 9-carbon position of the piperidine ring (see Table 1-2) to give an imine, which can then be reduced to provide the secondary amine, 4-anilino-N-phenethylpiperidine (4-ANPP, **5**). Finally, a reaction with propionyl chloride is carried out to provide the fentanyl product (**7**) via the

formation of the amide bond. The Janssen route ^[110], seen in Figure 1-6, and the Siegfried route ^[111], which utilises the general reaction scheme seen in Figure 1-7 using (2-bromoethyl)benzene (2-BEB) as a starting material, are the most common routes used for the illicit manufacture of fentanyl ^[112].

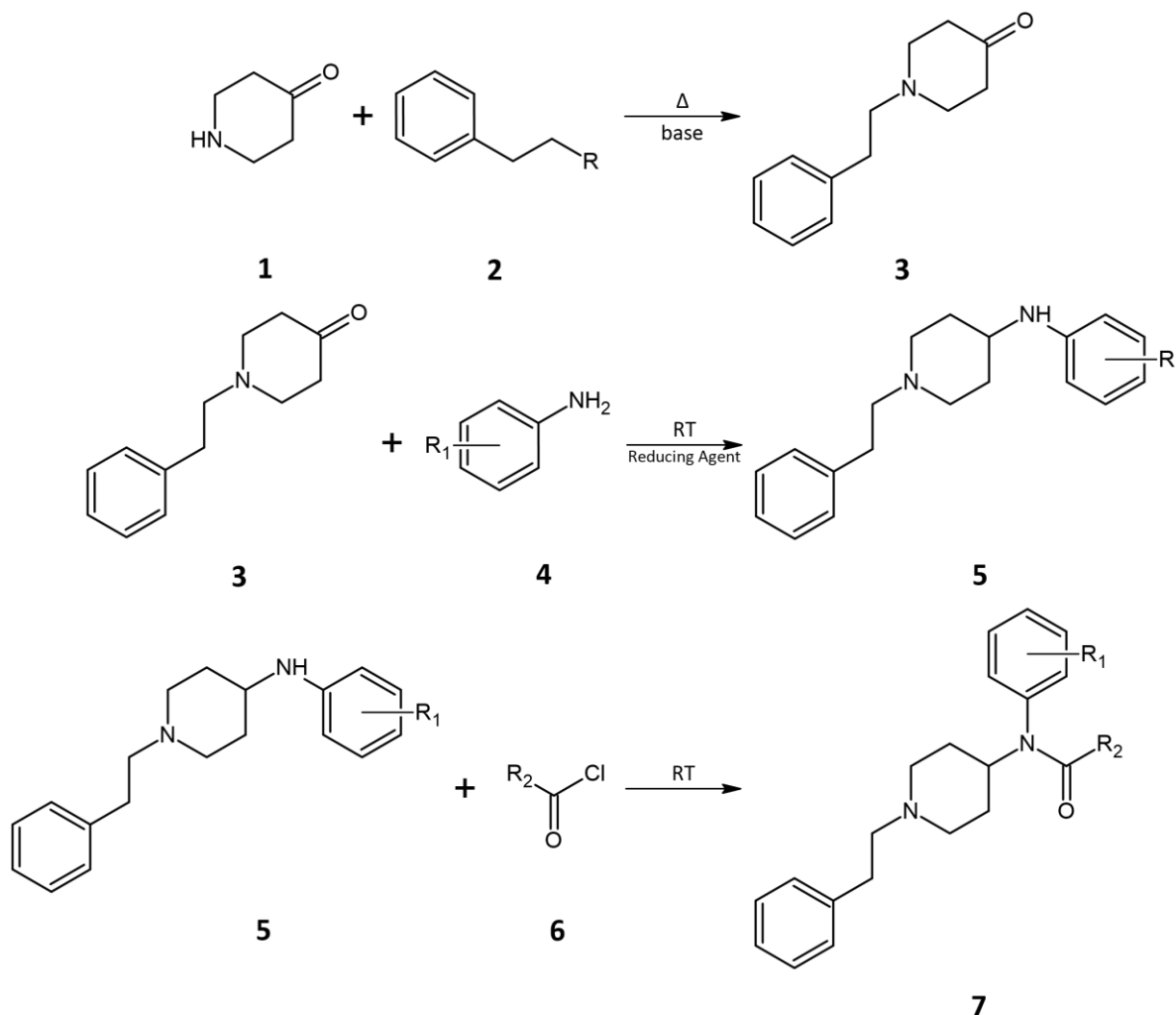


Figure 1-7: Generic synthesis procedure for fentanyl derivatives

Since the different fentanyl analogues share a similar core structure, their synthetic method can be adapted from the generic procedure presented in Figure 1-7. For example, many fentanyl analogues, such as acetyl fentanyl are created by altering the functional group attached to the amide (see Table 1-2). These different analogues can be prepared by altering the reagent used in the final step of the fentanyl synthesis (**6**). Valdez *et al.* reported the synthesis of acetyl fentanyl by replacing the propionyl chloride in the final step with acetic anhydride ^[43]. Similarly, the moiety attached to the piperidino nitrogen can be changed by altering the starting material used during the first synthetic step (**2**). Valdez *et al.* also reported the synthesis of thiofentanyl where the thiophene group is

introduced by the reaction of 2-(thiophen-2-yl)ethyl methanesulfonate with 4-piperidone monohydrate hydrochloride ^[43]. Halogenated fentanyl analogues, such as 4-fluorofentanyl, can also be formed through the substitution of the aniline used with a halogenated reagent (**4**) ^[113].

The introduction of additional side chains, usually either an ester or ether group, at the 9-carbon position of the piperidine ring (see Table 1-2) makes the synthetic procedure a little more complicated. There are a number of different methods that can be used to introduce these additional functional groups. The addition of ester groups, such as those found in carfentanil and remifentanil, includes methods such as the formation of methyl esters ^[114] and multicomponent reactions (MCRs) ^[115]. The development of MCRs has provided a powerful tool for organic synthesis ^[115, 116]. While there are many possible uses for these types of reaction, the most common use is for drug discovery ^[115, 117]. The major benefit of MCRs is that they allow for the straightforward development of large libraries of compounds while also accepting a variety of chemical functionality ^[115]. On the other hand, the introduction of ether groups to the fentanyl core structure involves quite a lengthy and complicated procedure ^[118, 119]. This method often led to very low yields of the final products (~2%) and utilised potassium cyanide which, especially in large-scale synthesis, is not ideal from a safety perspective ^[119]. More recently, Mathew and Killgore patented an improved method for the synthesis of alfentanil and sufentanil which has fewer steps and a greater yield (approximately 16%) than what had previously been reported ^[119].

While the procedures discussed are not the only methods that can be used for the synthesis of fentanyl analogues, they provide an indication of the breadth of different synthetic procedures available. The combination of the different techniques that have been described allow for the synthesis of a vast range of different fentanyl analogues, and the possibility of new compounds which haven't been developed yet. There are numerous synthetic procedures that can be used for the production of NSOs, and these are often taken and adapted from the patents published during the discovery of these compounds. These procedures have not been presented as this study focuses on the synthesis and profiling of fentanyl derivatives. Understanding the basic processes behind the synthesis of some common synthetic opioids can help with the analysis and profiling of seized samples and the investigation of suspected clandestine laboratory environments.

1.3 Recommendations for the Analysis of Seized Material

The Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) was established in 1997 and was jointly sponsored by the United States Drug Enforcement Administration (DEA) and the

Office of National Drug Control Policy (ONDCP) ^[120]. This group was created to improve the quality of the examination of seized drugs and respond to the needs of the forensic community by supporting the development of internationally accepted minimum standards, identifying best practices, and providing resources to help laboratories meet these standards ^[121]. In service of this objective, SWGDRUG has arranged different analytical techniques into categories based on their maximum discriminating power ^[121]. Table 1-8 provides an outline of the different techniques recommended.

Table 1-8: Categories of Analytical Techniques as define by SWGDRUG ^[121]

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

The categorisation of these techniques is accompanied by recommendations for the minimum standards that should be applied by forensic laboratories in relation to the type and number of techniques used. Where an analytical sequence includes the use of a Category A technique, such as Nuclear Magnetic Resonance spectroscopy (NMR), at least one other technique from any category must be used ^[121]. If a Category A technique is not used, at least three techniques must be employed, two of which must be based on uncorrelated techniques from Category B ^[121]. All Category A techniques, or Category B when no Category A techniques are used, must have data which is considered reviewable such as printed spectra ^[121]. When hyphenated techniques are used, such as Gas Chromatography-Mass Spectrometry (GC-MS), they can be treated as two separate techniques provided that data from each individual technique is used ^[121]. Finally, the analytical scheme to be used must avoid false positive identifications and minimise any possible false negative results ^[121].

1.4 Non-targeted Analysis

The continuous introduction of a large number of new NPS, along with the lack of available CRMs for many of these compounds, highlights the inadequacy of traditional targeted screening methods to

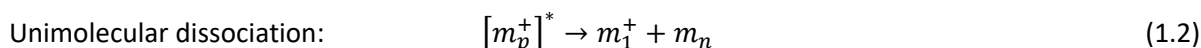
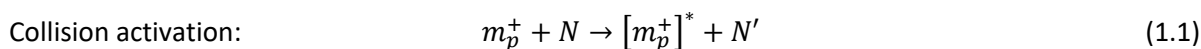
detect these compounds ^[122]. To combat this, the aim of non-targeted analysis is to identify compounds from samples where the molecular content is unknown ^[123]. Since illicit drug seizures are, by their very nature, unknown entities, this approach to their analysis provides some obvious advantages. While GC-MS has, historically, been the workhorse for forensic toxicological analyses, lately liquid chromatography – mass spectrometry (LC-MS) methods using triple quadrupole (QqQ) mass analyzers are considered valuable complementary, or even replacement, techniques ^[122, 124]. The use of high-resolution mass spectrometry (HRMS) has also been investigated for non-targeted analysis due to its ability to measure accurate masses and operate in data-independent acquisition modes (DIA) ^[122, 125]. The main advantage of HRMS over other mass spectrometry techniques is its ability to differentiate between ions and losses that have the same nominal mass ^[126]. For example, the loss of an –NH₃ group or an –OH group would equate to the same nominal mass change of 17 Da. This change may be indistinguishable using lower resolution MS techniques. If an MS technique with a high enough mass resolution was used, however, the accurate masses for those fragments of 17.0265 and 17.0027 Da, respectively, could be determined and therefore the losses differentiated ^[126]. Both data-dependent acquisition (DDA) and DIA modes can be used for acquisition of tandem mass spectrometry (MS/MS) data for non-targeted analysis ^[122, 124].

1.4.1 Collision-induced Dissociation

Ion activation and dissociation techniques are very important to assist with structure elucidation of unknown compounds ^[127]. There are a number of reasons that the extra activation/dissociation step is required. Firstly, the amount of ions that will fragment spontaneously from the initial ionisation source is usually quite low ^[127]. Additionally, the number of reaction pathways that could be detected from these spontaneously fragmenting ions would be very limited ^[127]. Finally, ions that are generated by common ionisation modes, such as electrospray ionisation (ESI) or matrix-assisted laser desorption/ionisation (MALDI), are generally either stable or exhibit weak fragmentation patterns that may not provide any meaningful structural information ^[127].

The most common method used for activation and dissociation of ions is known as collision-induced dissociation (CID) ^[127]. This technique involves a two-step process: collision activation followed by unimolecular dissociation ^[127]. The first involves the excitation of precursor ions to a higher energy state via high speed collisions with an inert gas such as helium or argon ^[127, 128]. This step usually occurs very quickly, occurring in around 10⁻¹⁴ to 10⁻¹⁶ s ^[128]. Immediately following this activation step, unimolecular dissociation begins to occur causing the precursor ions to be converted into fragment ions (or product ions) as the excitation energy causes the ion to fragment ^[127, 128]. Equations 1.1 and

1.2 provide an example of mechanism involved in CID processes, where m_p^+ is the precursor ion, N and N' are the pre- and post-collision forms of the collision gas and m_1^+ and m_n are fragments ^[127].



The fragmentation reactions undergone by the precursor ion contain a myriad of information which must be understood for structure characterisation ^[128]. The fragmentation spectrum of an ion reflects both the relative energies of the bonds present and the structure of the precursor and product ions ^[129]. The proliferation of NPS emphasises the need for class-based detection strategies to be developed ^[126]. Related compounds can often be found by identifying characteristic fragments or neutral losses ^[129]. The structural similarity of the NPFs and different families of NSOs presented previously alludes to the possibility of these characteristic fragments or neutral losses being present. The identification of these features would facilitate the development of this kind of class-based detection strategy. Therefore, CID studies of a wide variety of different synthetic opioids are vital in order to understand the fragmentation reactions they undergo.

1.4.2 Data Acquisition Methods

The efficacy of a developed non-targeted screening method relies, in part, on the data acquisition techniques applied ^[122]. The identification of NPS in forensic samples can often be complicated by the lack of CRMs for comparison, especially with novel analogues new to the market ^[130]. If appropriate data acquisition techniques are employed, however, then old samples can be reinvestigated when such materials become available ^[130]. The most basic data acquisition method that can be employed is full scan MS. This technique involves the measurement of all intact masses that reach the detector ^[122]. This information can help with the determination of chemical formulae and double bond equivalents (or degree of unsaturation, DBE); however it cannot assist with the identification of the chemical structures of unknown components ^[122]. This highlights the importance of MS/MS techniques and CID studies for putative structural elucidation ^[122].

Traditional targeted screening methods are often accomplished by multiple reaction monitoring (MRM) ^[124]. This process involves the detection of a number of selected precursor-product ion pairs for the analyte of interest ^[124, 127]. MRM techniques are able to be used for the detection and conformation of low amounts of analyte and are often able to provide quantitative information ^[124]. There are significant weaknesses of techniques such as this, however. It is necessary to maintain a

balance between duty cycles and sensitivity limits, therefore, the number of targets that can be included in a single experiment is usually much less than 1000 compounds ^[124, 131]. This means that any information that falls outside of these parameters is lost. This problem lead to the development of non-targeted data acquisition techniques to enable the collection of fragmentation information for nearly every featured detected ^[124]. Two different techniques were developed to accomplish this goal: DDA and DIA ^[124]. A simple comparison of these data acquisition techniques is provided in Figure 1-8.

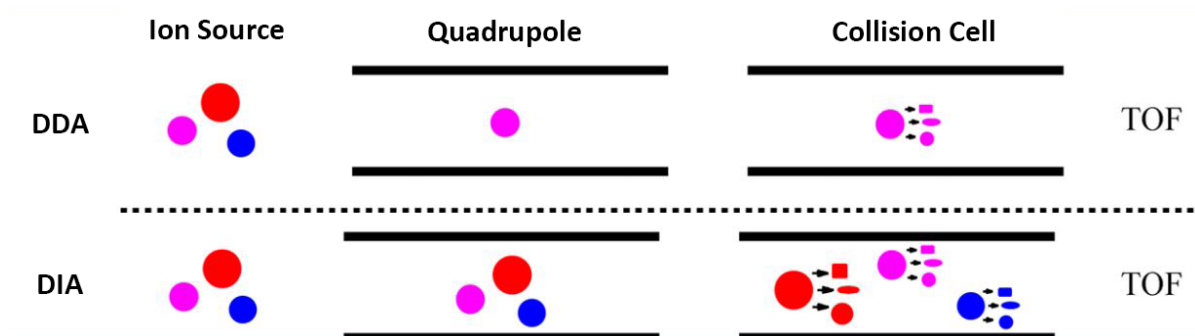


Figure 1-8: Comparison of data-dependent and data-independent acquisition techniques in high-resolution mass spectrometry (Adapted from Zhu et al.) ^[132]

1.4.2.1 Data-dependent Acquisition

Of the two different acquisition methods, DDA is more established and has successfully been applied to both low resolution (LR) ^[133] and high resolution (HR) ^[134] instruments ^[124]. This acquisition method operates by taking an initial MS scan and identifying precursor ions for MS/MS experiments ^[122, 124]. These MS/MS events are triggered when the precursor ions meet a set of pre-defined criteria, such as an intensity threshold ^[122]. When this method is used for non-targeted analysis, MS/MS experiments are conducted based on the abundance of precursor ions in the previous MS scan ^[122]. The main issue with this technique, however, is that abundant, yet forensically irrelevant, compounds such as the matrix background may be subjected to MS/MS events ^[122]. Since this acquisition method has a finite number of precursor ions that can be selected, this may cause the analytes of interest to go undetected, requiring additional analysis to be conducted. To combat this drawback, 'dynamic exclusion' and 'background subtraction' routines were developed which could be applied to DDA methods ^[124, 135]. These routines allowed a larger number of unique precursor ions to be subjected to MS/MS experiment by disallowing the reselection of already fragmented ions ^[124].

DDA can also be used for targeted analysis by using a preordained inclusion list of mass-to-charge (m/z) ratios for analytes of interest ^[122]. Additionally, the inclusion list can contain masses which should be excluded from MS/MS experiments, which can be useful for avoiding compounds such as common background ions ^[122]. While this application would prevent the collection of irrelevant data, making data analysis/interpretation procedures simpler, it also severely limits the range of analytes detected and prevents any retrospective data interrogation of new compounds ^[122]. Recently, complementary targeted and non-targeted DDA methods have been applied to a single analysis. This allows for the simultaneous use of an inclusion list and non-targeted interrogation of a set number of abundant precursor ions ^[122].

1.4.2.2 Data-independent Acquisition

On the other hand, DIA techniques have been more recently developed and have been used exclusively with HR instruments ^[124, 136-141]. These methods subject all precursor ions to CID which allows for the collection of full scan MS/MS ^[122]. Predetermined m/z ranges are analysed by either (1) fragmenting all ions entering the mass spectrometer (such as broadband DIA) or (2) separating the full range into smaller m/z ranges that are then sequentially analysed (such as Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra, SWATH) ^[124]. The most important advantage of this acquisition technique is that it allows for retrospective analysis of the collected data at a later date ^[122]. It has recently been determined that DIA is capable of identifying more compounds at lower concentrations than DDA through the analysis of forensic casework samples ^[136]. While retrospective data analysis may also be possible using DDA techniques, as demonstrated by Gundersen et al. ^[142], this relies on the fragmentation of a set number of abundant precursor ions. If compounds of interest have a lower abundance, or if there are numerous abundant background ions, they may be missed by DDA experiments. Using DIA techniques, however, would still allow for detection under these circumstances, as all precursor ions are being subjected to CID ^[122]. The main drawback for this technique is so called 'chimeric spectra', which are brought about by the simultaneous CID of multiple precursor ions ^[122, 124]. These spectra can be very difficult to interpret if the product ions produced cannot be correctly correlated with their precursor ions ^[122].

The applicability of this method of data acquisition to the detection of synthetic opioids was recently demonstrated by Noble *et al.* ^[143]. In this study, the authors analysed 50 different fentanyl analogues using liquid chromatography-quadrupole time of flight-mass spectrometry (LC-QTOF-MS) and a DIA method utilising an elevated mass spectrometry mode (MS^E). The results indicated the presence of a class-specific cleavage of the C-N bond between the amide moiety and the piperidine ring, which the

authors proposed could be used as a target for screening methods to detect this class of compounds [143]. The methods used in this study could be further expanded to include the other classes of NSOs and data gathered from these types of studies can be used to develop more comprehensive non-targeted screening methods.

1.4.3 Data Processing

The data acquisition method is only part of the non-targeted screening workflow [122]. Once the data is collected, appropriate processing techniques need to be applied in order to utilise the collected information to its full potential. Often software tools are applied for fast and automated data mining; however an expert is still required to effect a final decision on the identity of the unknown substance by reviewing the provided information [124]. Oberacher and Arnhard reported extensively on the use of MS/MS strategies for non-targeted screening by comparison with spectral databases [124]. With the vast array of new NPS being introduced to the illicit drug market, however, these databases can no longer be relied upon for the detection of novel analogues [122]. Non-targeted screening strategies can be broadly sorted into unbiased and biased techniques. Unbiased screening strategies involve the identification of all components detected in a particular sample [144, 145]. The workflow used for this type of non-targeted analysis involves the automated detection of peaks by exact mass filtering, assignment of an elemental formula to each exact mass and finally searching for plausible structures using chemical databases [145]. The main drawback with this process is the relevance of the compounds which are identified [122]. Ibáñez *et al.* provided an example whereby a peak in the chromatogram of a purchased NPS powder was assigned an elemental formula of $C_{11}H_{13}NO_3$, representing a m/z of 208.0976 Da [145]. When searching for possible chemical structures, 1549 possible identifications were found in the ChemSpider database [145, 146]. This then requires considerable effort to filter through the possible results to determine a definitive identify for the compound/s in question.

On the other hand, a compromise between targeted and non-targeted techniques is a biased non-targeted approach [145]. This approach focuses on the detection and identification of compounds in a sample that are related to known NPS [122, 145]. In this context, detection refers to the discovery of the components through data processing techniques, rather than the actual detection of ions with the MS [122]. This type of workflow may allow for the identification of novel analogues that are unreported or unknown at the time of analysis [145]. These techniques, however, rely on the assumption that the novel analogues will be extracted using the routinely employed sample preparation methods and are ionisable using the current instrumental analysis techniques [122]. Biased

non-targeted screening can be separated into two different steps: component discovery and putative identification ^[122]. Recently, Pasin *et al.* described two different workflows that could be used for discovery of any compounds of interest within a sample (Figure 1-9) ^[122].

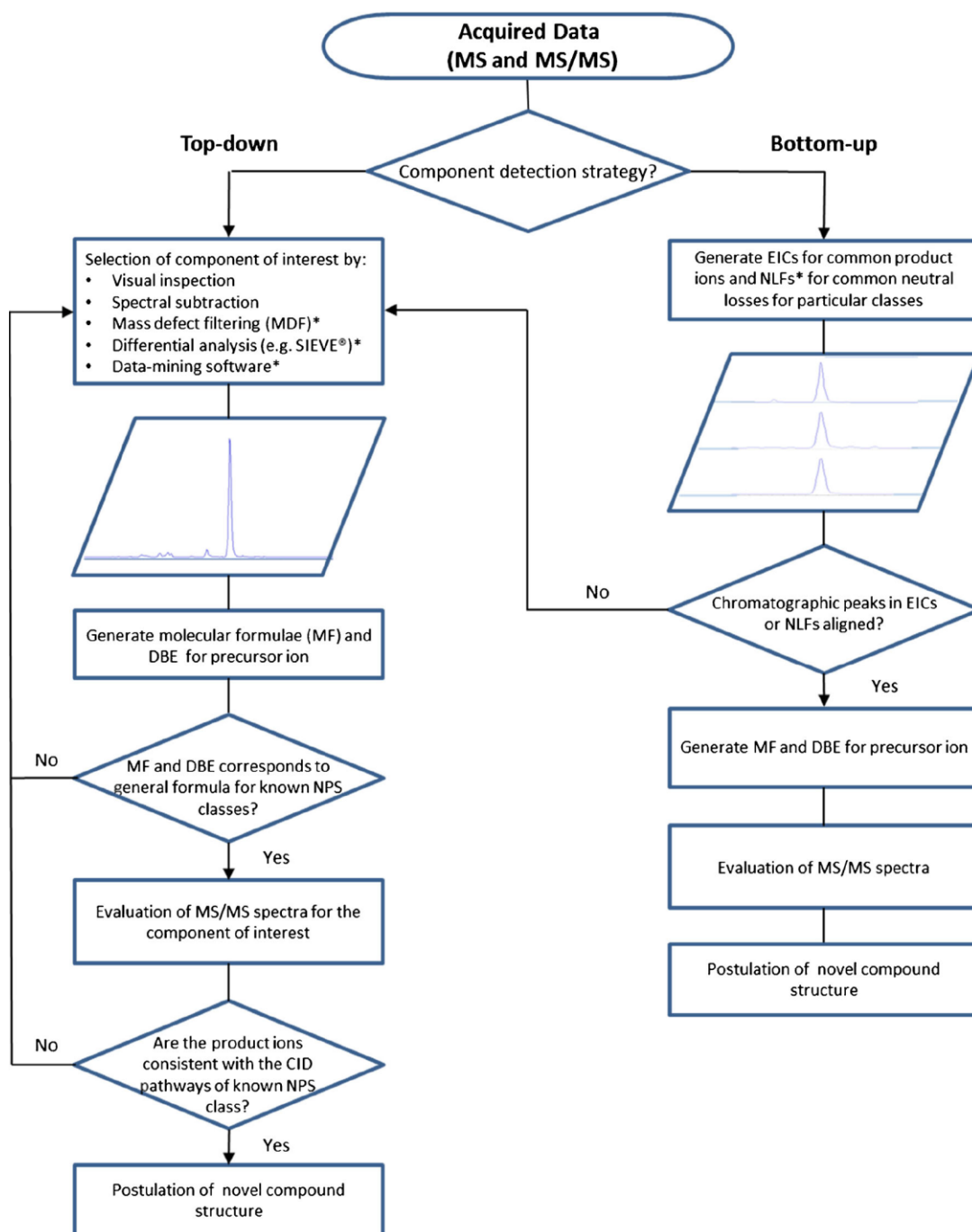


Figure 1-9: Comparison of the different workflows for biased non-targeted screening as proposed by Pasin *et al.* ^[122]. The * denotes that specialised software is required

The first, and most common, approach to biased non-targeted screening is the top-down approach^[122]. This approach involves the identification of abundant peaks from the visual inspection of a total ion chromatogram (TIC)^[122]. After the isolation of the peak, a molecular formula can be determined from the accurate mass of the precursor ion and other structural information such as DBE^[122]. The approach also has some limitations, which can lower its efficacy. Firstly, the use of visual inspections of the TICs is only appropriate for samples that produce very simple TICs with peaks which can be quickly located^[122]. This affects the usefulness of this type of technique for toxicological samples, where the matrix is quite complex and compounds of interest may not be immediately obvious and may overlap with endogenous compounds or background ions^[122].

Additionally, the main limitation of this approach is the requirement to assess the relevance of each peak identified to the compound class of interest. The molecular formula, DBE and MS/MS data must be critically appraised to see whether it correlates to the targeted class^[122]. This process can be labour-intensive and time consuming, especially when there are a number of peaks that need to be interrogated^[122].

In an attempt to improve the efficacy of this approach to non-targeted screening, innovative methods to filter the data and help highlight peaks of interest were developed^[122]. The use of mass defect filtering (MDF) was first reported by Grabenauer *et al.* who applied it to the analysis of synthetic cannabinoids in herbal products and were able to detect a compound that was not visible in the TIC^[147]. The mass defect of a particular compound is defined as the difference between its exact mass and its nominal integer mass^[148, 149]. Compound classes often have very similar mass defects, or specific trends with increasing mass, and therefore MDF can be used to selectively filter out related compounds^[148]. As previously mentioned, toxicological samples often have quite a complex matrix and, therefore, in addition to MDF, it is important to be able to quickly differentiate between endogenous and exogenous compounds to avoid spending time interrogating false positive components^[122]. Background peaks can be removed through subtraction of TIC data of a representative matrix or solvent blank from an authentic sample^[122, 150, 151]. In some circumstances, however, it may be difficult to obtain a representative sample that adequately accounts for the individual variations in a given population, which may limit the possible applications of this technique^[122]. Due to this limitation, Cawley *et al.* reported a more innovative approach that utilised differential analysis software^[152]. This technique relies on the comparison of m/z data from matrix blanks and authentic samples^[122, 152]. Finally, there are numerous data-mining software packages available from different instrument vendors that can be used, however these approaches are usually limited to suspect rather than non-targeted screening^[122, 153-156].

A sub-type of MDF, known as Kendrick mass defect (KMD) can also be useful in the development of non-targeted screening methods for drugs of abuse. The Kendrick mass scale allows for the recognition of a group of compounds that differ by a specific repeating mass unit ^[148, 157]. When calculating KMD, a normalisation factor is used to 'correct' the exact mass of a particular compound and give the Kendrick mass. Most commonly, the mass of the CH₂ group is used as the correction factor. Essentially, the mass of a CH₂ group is set to be exactly 14.00000, allowing the Kendrick mass of a compound to be calculated using equation 1.3 ^[157].

$$\text{Kendrick Mass} = \text{IUPAC mass} \times \frac{14}{14.01565} \quad (1.3)$$

$$\text{KMD} = (\text{nominal Kendrick mass}) - (\text{exact Kendrick mass}) \quad (1.4)$$

This correction is useful as series of compounds that have similar core structures, but vary in the number of methyl groups, will have exactly the same KMD. Kendrick masses are often used for complex samples, such as petroleum samples, as they can help to visualise different compound classes ^[148]. Different normalisation factors can also be used to identify compounds that differ by a different functional group. The use of KMD filtering has been investigated for drug screening. Anstett *et al.* ^[158] explored the application of KMD filtering to the screening of different phenethylamine classes, including 2C-, aminopropylbenzofuran- and 2,5-dimethoxy-N-(2-methoxybenzyl)-phenethylamines. The authors successfully implemented KMD filters that were able to distinguish between compounds from the different classes of phenethylamines ^[158]. If similar filters could be developed for the different classes of synthetic opioids and related compounds, this technique could prove a valuable asset in the development of non-targeted screening methods for the detection of these compounds.

Alternatively, a bottom-up approach can be used for compound detection that focuses on the interpretation of data using known characteristics of particular chemical classes, such as common product ions or neutral losses ^[122]. Since this approach focuses on known information about a class, it reduces the possibility of false positives ^[122]. This technique is based on the knowledge that compounds which are structurally related can produce common product ions. Novel analogues of existing chemical classes, therefore, could be identified by searching for the precursor ions of these common product ions ^[122, 147]. More recently, methods for monitoring common product ions and neutral losses were reported by Pasin *et al.* ^[126]. These methods utilised common data processing techniques such as product ion searching (PIS) and neutral loss filtering (NLF) ^[122, 126]. While the analysis of various different NPS has been thoroughly investigated over the years, the reporting of CID pathways is, sadly, much less common ^[122, 126, 159-166]. The usefulness of this approach to non-

targeted screening is dependent on the adequate collection of MS/MS data. For this reason, it has been suggested that DIA acquisition methods are used in preference over DDA methods ^[122, 126]. The main problem with this approach is that it requires specialised data processing software for complete interrogation of the MS/MS data ^[122]. Additionally, it is possible that future analogues of the different NPS classes may no longer share similar product ions, reducing the applicability of this approach ^[122]. NLF may still be appropriate for the detection of structurally related analogues, however the software required for this processing technique is not commonly included in standard data processing software ^[122].

Once a peak of interest has been isolated, using one of the techniques outlined above, the final step is the identification of the component present. A combination of different information can be used for the putative identification of unknown components, namely the molecular formula, DBE and CID pathways ^[122, 152]. Often the assessment of the collected MS/MS data requires experience and familiarity with the CID pathways of different NPS classes, which can increase the complexity of data processing workflows ^[122]. Finally, *in silico* fragmentation, using software such as Mass Frontier (Thermo Fisher Scientific) ^[152, 167] or Molecular Structure Correlator (MSC; Agilent Technologies) ^[156], can be used to confirm whether there is correlation between the experimental and theoretical fragmentation patterns achieved ^[122].

While the general workflow guidelines suggested by Pasin *et al.* (Figure 1-9) ^[122] provide a framework from which non-targeted screening methods should be built, work still needs to be conducted to demonstrate the applicability of these workflows to the different classes of NPS, such as the synthetic opioids. CID information for different chemical classes needs to be established and reported to assist with the development of these screening techniques. Biased non-targeted screening techniques are very powerful tools in the arsenal of forensic toxicologists and are instrumental in the attempt to stay ahead of the ever-expanding NPS market.

1.4.4 Statistical Analysis and Machine Learning

The recent shift towards using HRMS techniques means that there is a huge volume of data available from the analysis of unknown samples. When this is considered in addition to the availability of cheaper and more powerful computational processing ^[168], the use of artificial intelligence approaches, such as machine learning, for toxicological applications is much more viable. Machine learning allows a computer to 'learn' information directly from a given data set without needing predetermined equations to use as a model ^[169]. This means that the algorithm can adaptively

improve its performance over time with experience. Mitchell provided a definition for machine learning, stating ‘a computer is said to learn from experience with respect to some task and some performance measure, if its performance on said task improves with experience ^[170].’ Machine learning can be further divided into both supervised and unsupervised learning approaches. Supervised learning involves using a known set of input data (called the training set) and known responses to that data (the output) and using it to train a model to predict the response for new input data ^[168, 171]. Supervised learning models are categorised as either ‘classification’ or ‘regression’ models. Regression models try to predict an output within a continuous space, such as a temperature or time range, whereas classification models aims to predict a discrete output ^[168]. Unsupervised learning, on the other hand, is useful when attempting to explore data without specific goals or previous knowledge of the information contained within the data ^[168, 172].

Machine learning approaches have been used in the field of drug discovery previously, most notably within the context of determining structure activity relationships ^[173, 174]. The application of machine learning to case-based reasoning has even been investigated for forensic autopsies ^[175]. From a broader forensic perspective, machine learning is seeing increasing use in areas where its advanced pattern recognition can be exploited. For example, classification algorithms have been used to assist with the analysis of paint composition ^[168, 176, 177] and firearm identification ^[178]. Retention time prediction under a specific chromatographic system has also been explored using regression models from both an environmental ^[179] and drug analysis perspective ^[180, 181].

1.5 Synthetic Route Profiling

The goal of any analysis of bulk drug seizures is to provide comprehensive information on the chemical composition of an unknown sample or piece of evidence ^[124]. From a law enforcement perspective, the chemical, or impurity, profile of illicit substances can be of particular interest. In order to create these profiles, a broad range of information about both the physical and chemical properties of the sample in question needs to be collected. To this end, numerous different analytical techniques can be applied, including GC-MS ^[182, 183], LC-MS ^[112, 184], Capillary Electrophoresis ^[183], Inductively Coupled Plasma – Mass Spectrometry (ICP-MS) ^[183], and occasionally Fourier Transform Infra-red Spectroscopy (FTIR) ^[185] and Raman Spectroscopy ^[186-189].

In Australia, the use of impurity profiling has been demonstrated through the Australian Illicit Drug Intelligence Program (AIDIP) ^[190]. In this program, illicit substances such as amphetamine-type stimulants (ATS), heroin and cocaine are profiled ^[191, 192]. The Australian Federal Police (AFP) provides

the physical information for these seizures and the National Measurement Institute (NMI) performs the chemical profiling^[183, 190]. A variety of different analytical techniques are performed to construct a comprehensive profile of the substance in question, and the selection of these techniques often depends on the impurities expected to be present.

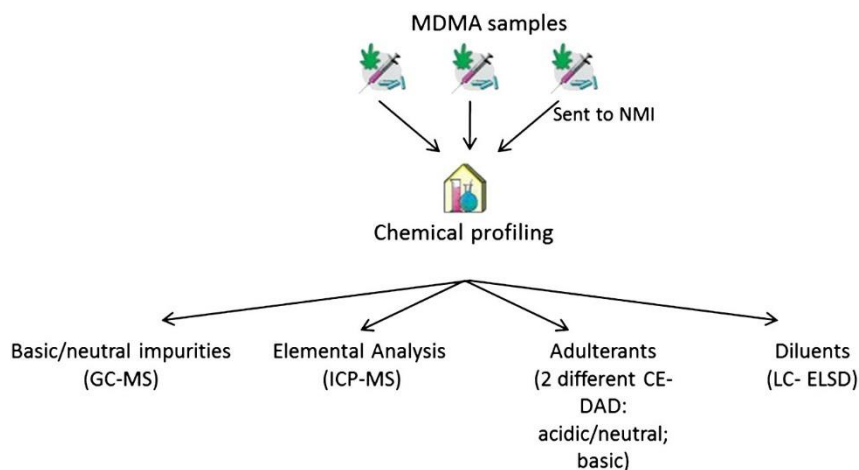


Figure 1-10: Chemical profiling of MDMA, as carried out by NMI (Adapted from Morelato et al.)^[183]

Figure 1-10 gives an example of how these concepts can be operationalised for the profiling of illicit material. This example highlights the profiling sequence that is employed by NMI for the profiling of MDMA^[183]. Initially, any basic and/or neutral impurities, such as para-methoxymethylamphetamine (PMMA) and 3,4-methylenedioxyamphetamine (MDA), are identified using a GC-MS method^[183]. This is followed by the identification of any inorganic impurities using ICP-MS^[183]. Acidic/neutral adulterants which may be present, such as caffeine or paracetamol, as well as basic adulterants, such as ketamine, are analysed using different Capillary Electrophoresis – Diode Array Detector (CE-DAD) techniques^[183]. Finally, any diluents present are analysed using Liquid Chromatography – Evaporative Light Scattering Detector (LC-ELSD)^[183]. All of this obtained data is then collated to produce a comprehensive profile of the analysed substance.

Of particular interest for the synthetic route profiling of an illicit substance is the organic impurity profile. The impurity profile can provide indicators of the synthetic route that was used to create the compound and possible precursors that might have been used through the identification of route-specific by-products^[193-195]. In addition to the organic impurity profile, chiral analysis can be conducted to provide an insight into the various enantiomers present in the sample^[193]. Finally, elemental analysis may be conducted, which can provide a better understanding of any catalysts that have been used as part of the synthesis^[193].

While much research is being conducted into the identification of NPS, such as the synthetic opioids, both through on-site presumptive tests and laboratory analysis, there has been less work done on the profiling of these compounds. Therefore, the collection of this information and construction of comprehensive profiles is crucial to ensuring that the data being obtained from drug seizures can be utilised to its full potential.

1.5.1 Value of the Chemical Profile

While more traditional roles of law enforcement are driven solely by the judicial system, more recently models focusing on intelligence-led policing have come into the spotlight ^[190]. These models recognise that the thorough and timely processing of case-related data, and its subsequent circulation to relevant parties, play an integral role in the framework of policing ^[190]. In this respect, the concept of forensic intelligence goes beyond the current paradigm of reactive and case-specific approaches to law enforcement and offers a more proactive and holistic response to criminal activities ^[196]. Marclay *et al.* discuss this intelligence-based framework in terms of three different levels that serve different areas of the policing process, namely tactical intelligence, operational intelligence and strategic intelligence (Figure 1-11) ^[196].

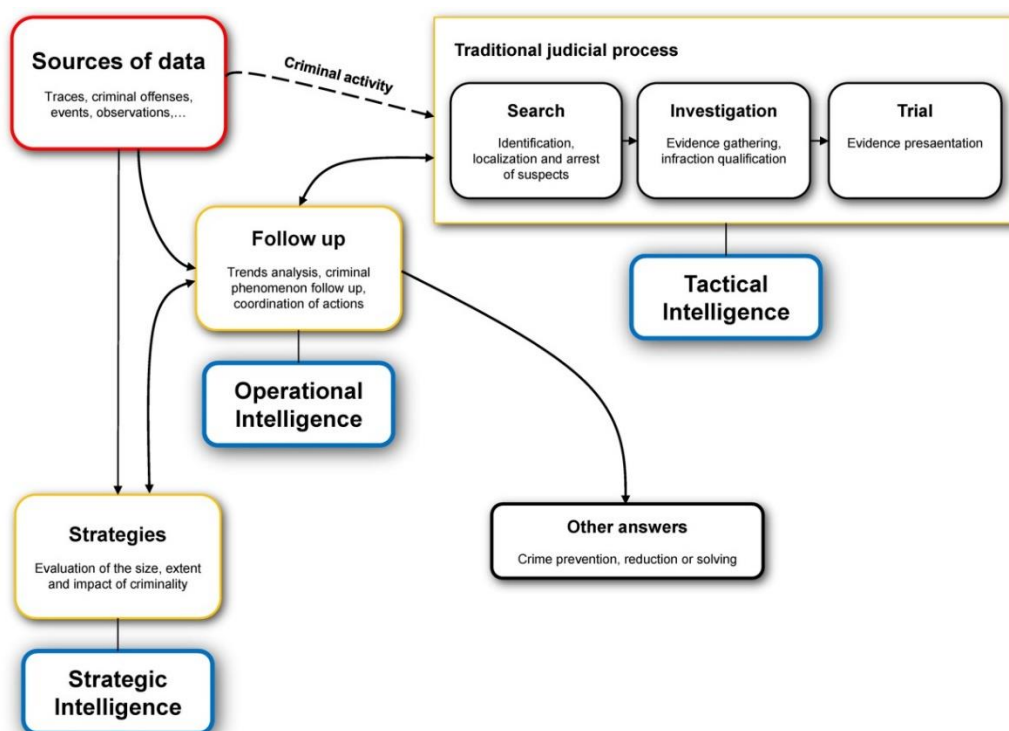


Figure 1-11: Framework for processing of case information to support the three different levels of forensic intelligence
(Adapted from Marclay et al.) ^[196]

In order to support the more traditional judicial process, tactical intelligence supports the real-time decision making of front-line law enforcement personnel and helps with the identification of leads for investigations at a case-level ^[196]. For this purpose, tactical intelligence often does not attempt to highlight criminal activities on a scale larger than the individual offender; the criminal environment under examination is a micro-level one ^[196, 197]. Stepping up to a broader perspective, operational intelligence focuses on an understanding of criminal trends to allow for effective follow up of incidents and assist with the coordination of police action ^[196]. Finally, strategic intelligence operates at a more global level. It is understood that criminality is a complex phenomenon that is constantly changing, therefore strategic intelligence attempts to provide a multivariate approach to study this criminality in the context of an environment that is in a constant state of change ^[196]. It is the goal of this aspect of forensic intelligence to foresee the rise of criminal threats in an attempt to plan crime reduction methods or, where possible, identify and rectify the vulnerabilities in the system that allow for the formation of these issues in the first place ^[196]. Overall, this approach to policing attempts to impact on the phenomenon as a whole, rather than specific incidents.

Chemical profiles of illicit substances can be of immense value to law enforcement agencies as they can provide a large amount of data to assist with investigations. The information provided by these profiles can include information about the manufacturing processes, the precursor chemicals that have been used or even confirm or refute links between different seizures ^[198]. While these profiles can provide limited information for tactical purposes, the true value of this information is the impact that it can have on operational and strategic intelligence, allowing for its use in a more proactive role, rather than simply assisting in the judicial process ^[196]. In this way, chemical profiling can contribute to intelligence-led policing, rather than merely on a case-by-case, reactionary basis ^[190]. An additional use for these developed chemical profiles from an intelligence perspective is comparative analysis. Comparative analysis makes use of in-depth examinations of the chemical and/or physical characteristics of various drug exhibits ^[199]. The purpose of this technique is to provide strategic and tactical intelligence in a timely manner to law enforcement agencies with the goal of assisting with the development and/or enhancement of enforcement initiatives ^[200].

To ensure that the chemical profile is used to its full potential, techniques are needed to analyse the profile once it has been collected and generate information that is useful for intelligence purposes. In order to ensure that information collected in the chemical profiles is utilised effectively, databases are compiled which can assist police in the identification of unsuspected links between seizures ^[201]. Principle component analysis (PCA) is a common unsupervised pattern recognition technique that can be used to sort the raw data contained within the profiles into more useable information. A study

conducted by Esseiva *et al.* at the 'Institut de Police Scientifique' of the University of Lausanne used PCA to process data from street samples of heroin and compiled their results into a database ^[202]. Any samples showing significant correlation were grouped together into different 'chemical classes.' New seizures could then be assigned to a particular class based on their PCA scores ^[202]. Once these kinds of models have been developed, class prediction models, such as partial least squares – discriminant analysis (PLS-DA), artificial neural networks (ANNs), naïve Bayes, decision trees and support vector machines (SVM), can be built to help identify what compounds may present within an unknown sample ^[202]. These models can then be used to make qualified inferences as to the synthetic routes used to make them and potential relationships between samples seized at different times and/or locations ^[202].

There are problems that can arise from these kinds of statistical analyses, however. Threshold values are often desired to assist experts to devise their opinion and to limit the number of false positives, while maximising the number of true positives ^[203]. These thresholds can be very difficult to determine and can be influenced by many factors such as data pre-treatment and the analytical methods employed for the creation of the profile ^[203]. This variability highlights the need for rigorous method validation and the need for the threshold values to be able to be adjusted based on the circumstances of the individual case.

1.5.2 Profiling of Synthetic Opioids

Lurie *et al.* ^[112] first investigated the chemical profiling of fentanyl in 2012 using ultra high performance liquid chromatography – tandem mass spectrometry (UHPLC-MS/MS) by comparing samples synthesised by two of the common methods at the time, the original Janssen method ^[110] and the Siegfried method ^[111]. This study identified 5 different markers specific to the Siegfried method, 10 different markers for the Janssen method, along with a number of different impurities common to both methods ^[112]. More recently, Mayer *et al.* ^[204] studied the chemical attribution signatures (CAS) of fentanyl synthesised by a number of different methods. This broader study included the use of the Siegfried method, along with two newer methods published by Gupta *et al.* ^[87] and Valdez *et al.* ^[43] and three different hybrid methods ^[204]. The authors opted to exclude the Janssen method from this study as it was decided that the method was too complex to realistically be used for clandestine synthesis. This study employed the use of multiple, complementary mass spectrometric techniques, including GC-MS, LC-MS and ICP-MS. The gathered data was then analysed using multivariate statistical analysis, namely PLS-DA ^[204]. The authors reported 87 different CAS and developed a statistical model that was capable of predicting the method of fentanyl synthesis. It was also found

that a combination of mass spectral techniques provided a better degree of discrimination than a single technique, however ICP-MS was not found to be useful ^[204]. Similar studies can be conducted for other synthetic opioids to expand the applicability of these techniques to samples containing synthetic opioids.

1.6 Project Objectives

In order to detect compounds which are not known to be present, non-targeted screening workflows will be developed for the different classes of synthetic opioids. Opioid standards will be analysed by HRMS in order to determine CID pathways and possible common fragments and neutral losses. The use of various “back-end” (i.e., post-acquisition) data analysis techniques to process acquired data will be evaluated and an optimised workflow will be described. In addition to differential analysis, the applicability of machine learning approaches to the detection and classification of synthetic opioids will be investigated. The developed method will be applied to the forensic toxicological setting, specifically non-targeted analysis of plasma samples for equine anti-doping.

Forensic intelligence relies on the collection of data which can be analysed and evaluated to provide relevant information for intelligence purposes. To this end, this study also aims to provide information which can be used to assist in the profiling of synthetic opioids. Fentanyl analogues will be synthesised utilising a variety of different synthetic methods. These samples will be analysed, and their impurity profile obtained in order to determine if there any route specific by-products present. This information may assist law enforcement agencies to determine how an illicit substance was made and the possible precursors and reagents that may have been used.

This thesis comprises of 7 chapters and presents the development of new data analysis methods for the non-targeted screening of synthetic opioids using HRMS techniques. Additionally, it contains a preliminary investigation into the synthetic route profiling of acetyl fentanyl. Chapter 2 investigates the CID pathways for a broad range of synthetic opioid derivatives and identifies diagnostic product ions belonging to the different opioid subclasses. Chapter 3 details the development of non-targeted screening methods that can be implemented to detect opioid derivatives using DIA methods and without relying on reference materials or library databases. Chapter 4 continues on from Chapter 3 and investigates the use of statistical modelling and machine learning approaches to assist in the identification of unknown opioid derivatives that are detected by the methods detailed in Chapter 3. Chapter 5 discusses the implementation of the methods developed in Chapters 3 and 4 and proposes a non-targeted screening workflow that may be used by routine laboratories. Chapter 6 details the

preliminary investigation that was conducted into the route profiling of acetyl fentanyl and the possible applications to forensic analysis. Finally, Chapter 7 provides a critical summary of the work and offers recommendations and potential future directions for the analysis of synthetic opioids and NPS in general.

1.7 References

1. United Nations Office on Drugs and Crime. *World Drug Report 2019*. Vienna: United Nations Office on Drugs and Crime; June 2019.
2. Australian Bureau of Statistics, *4519.0 - Recorded Crime - Offenders, 2018-19*. 2020, Australian Bureau of Statistics: Online.
3. Australian Criminal Intelligence Commission. *Illicit Drug Data Report 2017-18*. Australia: Australian Criminal Intelligence Commission;2019. ISSN 2205-1929.
4. Pietschmann, T., M. Tullis, and T. Leggett. *A Century of International Drug Control*. Austria: United Nations Office on Drugs and Crime;2008.
5. European Monitoring Centre for Drugs and Drug Addiction. *Annual report on the state of the drugs problem in the European Union*. Luxembourg: European Monitoring Centre for Drugs and Drug Addiction;1999.
6. Australian Institute of Health and Welfare. *Australian Burden of Disease Study: Impact and causes of illness and death in Australia 2011*. Canberra: Australian Institute of Health and Welfare;2016.
7. Kyu, H.H., et al., *Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017*. The Lancet, 2018. **392**(10159): p. 1859-1922.
8. Australian Institute of Health and Welfare. *Australian Burden of Disease Study: Impact and causes of illness and death in Australia 2015*. Canberra: Australian Institute of Health and Welfare;2019.
9. United Nations Office on Drugs and Crime. *UNODC Early Warning Advisory on New Psychoactive Substances*. 2020 [cited 2020 3rd June]; Available from: <https://www.unodc.org/LSS/Page/NPS>.
10. Australian Criminal Intelligence Commission. *Illicit Drug Data Report 2015-16*. Australia: Australian Criminal Intelligence Commission;2017.
11. Zawilska, J.B. and D. Andrzejczak, *Next generation of novel psychoactive substances on the horizon – A complex problem to face*. Drug Alcohol Depend., 2015. **157**: p. 1-17.

12. Elliott, S. and J. Evans, *A 3-year review of new psychoactive substances in casework*. Forensic Sci. Int., 2014. **243**: p. 55-60.
13. Favretto, D., J.P. Pascali, and F. Tagliaro, *New challenges and innovation in forensic toxicology: Focus on the "New Psychoactive Substances"*. J. Chromatogr. A, 2013. **1287**: p. 84-95.
14. European Monitoring Centre for Drugs and Drug Addiction. *New psychoactive substances in Europe. An update from the EU Early Warning System*. Luxembourg: Publication Office of the European Union;2015.
15. Stogner, J.M., *Predictions instead of panics: the framework and utility of systematic forecasting of novel psychoactive drug trends*. Am. J. Drug Alcohol Ab., 2015. **41**(6): p. 519-526.
16. *Crimes Legislation Amendment (Psychoactive Substances and Other Measures) Act 2015*, in C2015A00012. 2015, Federal Register of Legislation: Australia.
17. Reuter, P. and B. Pardo, *New psychoactive substances: Are there any good options for regulating new psychoactive substances?* Int. J. Drug Policy, 2017. **40**: p. 117-122.
18. Reuter, P. and B. Pardo, *Can new psychoactive substances be regulated effectively? An assessment of the British Psychoactive Substances Bill*. Addiction, 2016. **112**(1): p. 25-31.
19. Lucyk, S.N. and L.S. Nelson, *Novel Synthetic Opioids: An Opioid Epidemic Within an Opioid Epidemic*. Ann. Emerg. Med., 2017. **69**(1): p. 91-93.
20. Armenian, P., et al., *Fentanyl and a Novel Synthetic Opioid U-47700 Masquerading as Street "Norco" in Central California: A Case Report*. Ann. Emerg. Med., 2017. **69**(1): p. 87-90.
21. Peacock, A., et al. *Australian Drug Trends 2019: Key Findings from the National Ecstasy and Related Drugs Reporting System (EDRS) Interviews*. Sydney: National Drug and Alcohol Research Centre;2019. 978-0-7334-3890-5.
22. United Nations Office on Drugs and Crime. *World Drug Report 2011*. Vienna: United Nations Office on Drugs and Crime;2011.
23. Maas, A., et al., *Separation of ortho, meta and para isomers of methylmethcathinone (MMC) and methylethcathinone (MEC) using LC-ESI-MS/MS: Application to forensic serum samples*. J. Chromatogr. B, 2017. **1051**: p. 118-125.
24. Hughes, B. and T. Bildaru. *Legal Responses to New Psychoactive Substances in Europe*. Lisbon, Portugal: European Monitoring Centre for Drugs and Drug Addiction;2009.
25. Barratt, M.J., K. Seear, and K. Lancaster, *A critical examination of the definition of 'psychoactive effect' in Australian drug legislation*. Int. J. Drug Policy, 2017. **40**: p. 16-25.
26. *Oral Evidence: Psychoactive Substances*, in *Home Affairs Committee*. 2015, House of Commons: London, UK.
27. Iversen, L., *Re: Psychoactive Substances Bill*. 2015, Advisory Council on the Misuse of Drugs: Online.

28. Fornili, K.S., *International Control Efforts to Curb the Global Production and Trafficking of Fentanyl and Other Synthetic Opioids*. Journal of Addictions Nursing, 2019. **30**(1): p. 71.
29. World Health Organisation. *Ensuring balance in national policies on controlled substances*. World Health Organisation;2011.
30. Ballantyne, J.C. and K.S. LaForge, *Opioid dependence and addiction during opioid treatment of chronic pain*. Pain, 2007. **129**(3): p. 235-255.
31. Sandilands, E.A. and D.N. Bateman, *Opioids*. Medicine, 2016. **44**(3): p. 187-189.
32. Holden, J.E., Y. Jeong, and J.M. Forrest, *The endogenous opioid system and clinical pain management*. AACN clinical issues, 2005. **16**(3): p. 291-301.
33. Corbett, A.D., et al., *75 years of opioid research: the exciting but vain quest for the Holy Grail*. Br. J. Pharmacol., 2006. **147**(S1): p. S153-S162.
34. Busardò, F.P., et al., *Ultra-High-Performance Liquid Chromatography-Tandem Mass Spectrometry Assay for Quantifying Fentanyl and 22 Analogs and Metabolites in Whole Blood, Urine, and Hair*. Front. Chem., 2019. **7**: p. 184-184.
35. Huang, X.-P., et al., *Fentanyl-related designer drugs W-18 and W-15 lack appreciable opioid activity in vitro and in vivo*. JCI Insight, 2017. **2**(22): p. e97222.
36. Scholl, L., et al., *Drug and Opioid-Involved Overdose Deaths - United States, 2013-2017*. Morb. Mortal. Wkly. Rep., 2019. **67**(51-52): p. 1419-1427.
37. Breindahl, T., et al., *Identification of a new psychoactive substance in seized material: the synthetic opioid N-phenyl-N- 1-(2-phenethyl)piperidin-4-yl prop-2-enamide (Acrylfentanyl)*. Drug Test. Anal., 2017. **9**(3): p. 415-422.
38. Mohr, A.L.A., et al., *Analysis of Novel Synthetic Opioids U-47700, U-50488 and Furanyl Fentanyl by LC-MS/MS in Postmortem Casework*. J. Anal. Toxicol., 2016. **40**(9): p. 709-717.
39. Gupta, P.K., et al., *Vapor pressure and enthalpy of vaporization of fentanyl*. J. Chem. Eng. Data, 2008. **53**(3): p. 841-845.
40. Scientific Working Group for the Analysis of Seized Drugs. *Fentanyl*. Online: Drug Enforcement Administration;2005.
41. Bagley, J.R., et al., *Evolution of the 4-Anilidopiperidine Class of Opioid Analgesics*. Med. Res. Rev., 1991. **11**(4): p. 403-436.
42. Stanley, T.H., *The Fentanyl Story*. J. Pain, 2014. **15**(12): p. 1215-1226.
43. Valdez, C.A., R.N. Leif, and B.P. Mayer, *An Efficient, Optimized Synthesis of Fentanyl and Related Analogs*. PLoS ONE, 2014. **9**(9): p. 1-8.
44. PubChem. *Carfentanil*. 2017 [cited 2017 28th April]; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/62156#section=Top>.

45. PubChem. *Remifentanyl*. 2017 [cited 2017 4th May]; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/remifentanyl>.
46. PubChem. *Sufentanyl*. 2017 [cited 2017 28th April]; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/41693#section=Top>.
47. Suzuki, J. and S. El-Haddad, *A review: Fentanyl and non-pharmaceutical fentanyls*. Drug Alcohol Depend., 2017. **171**: p. 107-116.
48. Petrov, R.R., et al., *Synthesis and evaluation of 3-aminopropionyl substituted fentanyl analogues for opioid activity*. Bioorg. Med. Chem. Lett., 2006. **16**(18): p. 4946-4950.
49. Girón, R.o., et al., *Synthesis and opioid activity of new fentanyl analogs*. Life Sci., 2002. **71**(9): p. 1023-1034.
50. Varadi, A., et al., *Synthesis of Carfentanil Amide Opioids Using the Ugi Multicomponent Reaction*. ACS Chem. Neurosci., 2015. **6**(9): p. 1570-1577.
51. Katselou, M., et al., *Old opioids, new concerns: the case of acetylfentanyl*. Forensic Toxicol., 2016. **34**(2): p. 201-212.
52. Scientific Working Group for the Analysis of Seized Drugs. *Acetyl Fentanyl*. Online: Drug Enforcement Administration;2013.
53. World Health Organisation. *Acetylfentanyl - Critical Review Report*. Geneva: World Health Organisation;2015.
54. World Health Organisation. *Carfentanil - Critical Review Report*. Geneva: World Health Organisation;2017.
55. Babazade, R. and A. Turan, *Pharmacokinetic and pharmacodynamic evaluation of sublingual sufentanyl in the treatment of post-operative pain*. Expert Opinion on Drug Metabolism & Toxicology, 2016. **12**(2): p. 217-224.
56. ChemSpider. *Remifentanyl Hydrochloride*. 2017 [cited 2017 5th May]; Available from: <http://www.chemspider.com/Chemical-Structure.54802.html>.
57. PubChem. *Alfentanyl*. 2017 [cited 2017 5th May]; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/alfentanyl#section=Top>.
58. Ujváry, I., et al., *Acryloylfentanyl, a recently emerged new psychoactive substance: a comprehensive review*. Forensic Toxicol., 2017. **35**(2): p. 232-243.
59. Scientific Working Group for the Analysis of Seized Drugs. *Butyryl Fentanyl*. Online: Drug Enforcement Administration;2016.
60. Rojkiewicz, M., et al., *Identification and physicochemical characterization of 4-fluorobutyrylfentanyl (1-((4-fluorophenyl)(1-phenethylpiperidin-4-yl)amino)butan-1-one, 4-FBF) in seized materials and post-mortem biological samples*. Drug Test. Anal., 2017. **9**(3): p. 405-414.

61. Scientific Working Group for the Analysis of Seized Drugs. *Furanyl Fentanyl*. Online: Drug Enforcement Administration;2016.
62. Schneir, A., et al., *Near death from a novel synthetic opioid labeled U-47700: emergence of a new opioid class*. Clin. Toxicol., 2017. **55**(1): p. 51-54.
63. Katselou, M., et al., *AH-7921: the list of new psychoactive opioids is expanded*. Forensic Toxicol., 2015. **33**(2): p. 195-201.
64. Sharma, K.K., et al., *The search for the "next" euphoric non-fentanyl novel synthetic opioids on the illicit drugs market: current status and horizon scanning*. Forensic Toxicol., 2019. **37**(1): p. 1-16.
65. European Monitoring Centre for Drugs and Drug Addiction and Europol. *EMCDDA–Europol Joint Report on a new psychoactive substance: AH-7921 3,4-dichloro-N-[[1-(dimethylamino)cyclohexyl]methyl]benzamide*. Luxembourg: Publication Office of the European Union;2014.
66. Scientific Working Group for the Analysis of Seized Drugs. *AH-7921*. Online: Drug Enforcement Administration;2013.
67. Nikolaou, P., et al., *U-47700. An old opioid becomes a recent danger*. Forensic Toxicol., 2017. **35**(1): p. 11-19.
68. Szmuszkowicz, J., *Analgesic N-(2-aminocycloaliphatic)benzamides*, in *Google Patents*, United States Patent and Trademark Office, Editor. 1978, The Upjohn Company: United States.
69. Elliott, S.P., S.D. Brandt, and C. Smith, *The first reported fatality associated with the synthetic opioid 3,4-dichloro-N-[2-(dimethylamino)cyclohexyl]-N-methylbenzamide (U-47700) and implications for forensic analysis*. Drug Test. Anal., 2016. **8**(8): p. 875-879.
70. Cheney, B.V., et al., *Factors affecting binding of trans-N-[2-(methylamino)cyclohexyl]benzamides at the primary morphine receptor*. J. Med. Chem., 1985. **28**(12): p. 1853-1864.
71. World Health Organisation. *U-47700 - Critical Review Report*. Geneva: World Health Organisation;2016.
72. Fabregat-Safont, D., et al., *Updating the list of known opioids through identification and characterization of the new opioid derivative 3,4-dichloro-N-(2-(diethylamino) cyclohexyl)-N-methylbenzamide (U-49900)*. Scientific Reports, 2017. **7**: p. 14.
73. Takemori, A.E. and P.S. Portoghese, *Selective Naltrexone-Derived Opioid Receptor Antagonists*. Annu. Rev. Pharmacool. Toxicol., 1992. **32**: p. 239-269.
74. Fujimoto, R.A., et al., *Synthesis, opioid receptor binding profile, and antinociceptive activity of 1-azaspiro[4.5]decan-10-yl amides*. J. Med. Chem., 1989. **32**(6): p. 1259-1265.
75. Toronto Research Chemicals. *(+/-)-U-50488 Hydrochloride*. 2017 [cited 2019 11th January]; Available from: <https://www.trc-canada.com/product-detail/?U850328>.

76. Cerilliant, *U-51754 HCl - Safety Data Sheet*. 2018, Sigma-Aldrich,: Online.
77. Helander, A., M. Backberg, and O. Beck, *MT-45, a new psychoactive substance associated with hearing loss and unconsciousness*. Clin. Toxicol., 2014. **52**(8): p. 901-904.
78. Baptista-Hon, D.T., et al., *Activation of μ -opioid receptors by MT-45 (1-cyclohexyl-4-(1,2-diphenylethyl)piperazine) and its fluorinated derivatives*. Br. J. Pharmacol., 2020. **177**(15): p. 3436-3448.
79. Blanckaert, P., et al., *Report on a novel emerging class of highly potent benzimidazole NPS opioids: Chemical and in vitro functional characterization of isotonitazene*. Drug Test. Anal., 2020. **12**: p. 422-430.
80. Canadian Community Epidemiology Network on Drug Use. *CCENDU Bulletin: Novel Synthetic Opioids in Counterfeit Pharmaceuticals and Other Illicit Street Drugs*. Online: Canadian Centre on Substance Abuse;2016.
81. Knaus, E.E., B.K. Warren, and T.A. Ondrus, *Analgesic substituted piperidylidene-2-sulfon(cyan)amide derivatives*, in *Google Patents*, United States Patent and Trademark Office, Editor. 1984, Canadian Patents and Development Ltd.: Canada. p. 9.
82. Scientific Working Group for the Analysis of Seized Drugs. *W-15*. Online: Drug Enforcement Administration;2016.
83. Scientific Working Group for the Analysis of Seized Drugs. *W-18*. Online: Drug Enforcement Administration;2016.
84. Krotulski, A.J., et al., *Metabolism of novel opioid agonists U-47700 and U-49900 using human liver microsomes with confirmation in authentic urine specimens from drug users*. Drug Test. Anal., 2018. **10**(1): p. 127-136.
85. Steuer, A.E., et al., *Studies on the metabolism of the fentanyl-derived designer drug butyrfentanyl in human in vitro liver preparations and authentic human samples using liquid chromatography-high resolution mass spectrometry (LC-HRMS)*. Drug Test. Anal., 2017. **9**(7): p. 1085-1092.
86. Wohlfarth, A., et al., *Metabolic characterization of AH-7921, a synthetic opioid designer drug: in vitro metabolic stability assessment and metabolite identification, evaluation of in silico prediction, and in vivo confirmation*. Drug Test. Anal., 2016. **8**(8): p. 779-791.
87. Gupta, P.K., et al., *A convenient one pot synthesis of fentanyl*. J. Chem. Res., 2005. **2005**(7).
88. Higashikawa, Y. and S. Suzuki, *Studies on 1-(2-phenethyl)-4-(N-propionylanilino)piperidine (fentanyl) and its related compounds. VI. Structure-analgesic activity relationship for fentanyl, methyl-substituted fentanyls and other analogues*. Forensic Toxicol., 2008. **26**(1): p. 1-5.
89. Moffat, A.C., et al., *Clarke's Analysis of Drugs and Poisons*. 2011: Pharmaceutical Press.
90. Reynolds, L., et al., *Relative analgesic potency of fentanyl and sufentanil during intermediate-term infusions in patients after long-term opioid treatment for chronic pain*. Pain, 2004. **110**(1-2): p. 182-188.

91. Casale, J.F., J.R. Mallette, and E.M. Guest, *Analysis of Illicit Carfentanil: Emergence of the Death Dragon*. Forensic Chem., 2017. **3**: p. 74-80.
92. Stanley, T.H., *The history and development of the fentanyl series*. J. Pain Symptom Manage., 1992. **7**(3): p. S3-S7.
93. Guerrieri, D., et al., *Acrylfentanyl: Another new psychoactive drug with fatal consequences*. Forensic Sci. Int., 2017. **277**: p. e21-e29.
94. World Health Organisation. *Butyrfentanyl (Butyrylfentanyl) - Critical Review Report*. Geneva: World Health Organisation;2016.
95. Reichle, C.W., et al., *Comparative Analgesic Potency of Heroin and Morphine in Postoperative Patients*. J. Pharmacol. Exp. Ther., 1962. **136**(1): p. 43-46.
96. Backberg, M., et al., *Opioid intoxications involving butyrfentanyl, 4-fluorobutyrfentanyl, and fentanyl from the Swedish STRIDA project*. Clin. Toxicol., 2015. **53**(7): p. 609-617.
97. Pichini, S., et al., *European Drug Users at Risk from Illicit Fentanyls Mix*. Front. Pharmacol., 2017. **8**: p. 2.
98. World Health Organisation. *AH-7921 - Critical Review Report*. Geneva: World Health Organisation;2014.
99. Coppola, M. and R. Mondola, *AH-7921: A new synthetic opioid of abuse*. Drug Alcohol Rev., 2015. **34**(1): p. 109-110.
100. Fels, H., et al., *Two fatalities associated with synthetic opioids: AH-7921 and MT-45*. Forensic Sci. Int., 2017. **277**: p. e30-e35.
101. Zawilska, J.B., *An expanding world of Novel Psychoactive Substances: Opioids*. Front. Psychiatry, 2017. **8**: p. 14.
102. Vandeputte, M.M., A. Cannaert, and C.P. Stove, *In vitro functional characterization of a panel of non-fentanyl opioid new psychoactive substances*. Arch. Toxicol., 2020. **94**(11): p. 3819-3830.
103. Hunger, A., et al., *Benzimidazol-Derivate und verwandte Heterocyclen III. Synthese von 1-Aminoalkyl-2-nenzyl-nitro-benzimidazolen*. Helv. Chim. Acta, 1960. **43**(4): p. 1032-1046.
104. Huang, X.-P., et al., *Pharmacology of W-18 and W-15*. bioRxiv, 2016.
105. Prekupec, M.P., P.A. Mansky, and M.H. Baumann, *Misuse of Novel Synthetic Opioids: A Deadly New Trend*. J Addict Med, 2017. **11**(4): p. 256-265.
106. Matthews, N.S. and G.L. Carroll. *Review of Equine Analgesics and Pain Management*. in *AAEP Proceedings*. 2007. American Association of Equine Practitioners.
107. Tobin, T., *Drugs and the Performance Horse*. 1981, Illinois, USA: Charles C Thomas Publisher, Limited.

108. Kamerling, S., et al., *Dose related effects of the kappa agonist U-50,488H on behaviour, nociception and autonomic response in the horse*. Equine Vet. J., 1988. **20**(2): p. 114-118.
109. Racing Medication & Testing Consortium, *New Synthetic Opioid Designer Drug Found by New York Lab: ARCI 11/9/15*. 2015, Racing Medication & Testing Consortium: Lexington, Ky.
110. Janssen, P.A.J. and J.F. Gardocki, *Method for producing analgesia*, in *Google Patents*, United States Patent and Trademark Office, Editor. 1964, N. V. Research Laboratorium, Dr. C. Janssen: United States. p. 1.
111. Siegfried. *Synthesis of Fentanyl*. 2004 [cited 2017 29th March]; Available from: <https://erowid.org/archive/rhodium/chemistry/fentanyl.html>.
112. Lurie, I.S., et al., *Profiling of illicit fentanyl using UHPLC–MS/MS*. Forensic Sci. Int., 2012. **220**(1–3): p. 191-196.
113. Ulens, C., et al., *Interaction of p-fluorofentanyl on cloned human opioid receptors and exploration of the role of Trp-318 and His-319 in mu-opioid receptor selectivity*. J. Pharmacol. Exp. Ther., 2000. **294**(3): p. 1024-1033.
114. Walz, A.J. and F.-L. Hsu, *Synthesis of 4-anilinopiperidine methyl esters, intermediates in the production of carfentanil, sufentanil, and remifentanil*. Tetrahedron Lett., 2014. **55**(2): p. 501-502.
115. Malaquin, S., et al., *Ugi reaction for the synthesis of 4-aminopiperidine-4-carboxylic acid derivatives. Application to the synthesis of carfentanil and remifentanil*. Tetrahedron Lett., 2010. **51**(22): p. 2983-2985.
116. Zhu, J. and H. Bienayme, eds. *Multicomponent Reactions*. 2006, Wiley-VCH: Weinheim.
117. Akritopoulou-Zanze, I., *Isocyanide-based multicomponent reactions in drug discovery*. Curr. Opin. Chem. Biol., 2008. **12**(3): p. 324-331.
118. Janssen, P.A.J. and P.G.H. Van Daele, *N-(4-piperidiny)-n-phenylamides and -carbamates*, in *Google Patents*, United States Patent and Trademark Office, Editor. 1976, Janssen Pharmaceutica N.V.: United States.
119. Mathew, J. and J.K. Killgore, *Methods for the synthesis of alfentanil, sufentanil, and remifentanil*, in *Google Patents*, United States Patent and Trademark Office, Editor. 2007, Mallinckrodt Inc.: United States.
120. Scientific Working Group for the Analysis of Seized Drugs. *History*. 2017 [cited 2017 21st April]; Available from: <http://www.swgdrug.org/history.htm>.
121. Scientific Working Group for the Analysis of Seized Drugs, *Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) Recommendations Version 7.1*, D.E.A. United States Department of Justice, Editor. 2016.
122. Pasin, D., et al., *Current applications of high-resolution mass spectrometry for the analysis of new psychoactive substances: a critical review*. Anal. Bioanal. Chem., 2017. **409**(25): p. 5821-5836.

123. Knolhoff, A.M. and T.R. Croley, *Non-targeted screening approaches for contaminants and adulterants in food using liquid chromatography hyphenated to high resolution mass spectrometry*. J. Chromatogr. A, 2016. **1428**: p. 86-96.
124. Oberacher, H. and K. Arnhard, *Current status of non-targeted liquid chromatography-tandem mass spectrometry in forensic toxicology*. TrAC, Trends Anal. Chem., 2016. **84**: p. 94-105.
125. Remane, D., D.K. Wissenbach, and F.T. Peters, *Recent advances of liquid chromatography–(tandem) mass spectrometry in clinical and forensic toxicology — An update*. Clin. Biochem., 2016. **49**(13): p. 1051-1071.
126. Pasin, D., et al., *Characterization of hallucinogenic phenethylamines using high-resolution mass spectrometry for non-targeted screening purposes*. Drug Test. Anal., 2017. **9**(10): p. 1620-1629.
127. Dass, C., *Fundamentals of Contemporary Mass Spectrometry*. 2007: Wiley.
128. Demarque, D.P., et al., *Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products*. Nat. Prod. Rep., 2016. **33**(3): p. 432-455.
129. Johnson, A.R. and E.E. Carlson, *Collision-Induced Dissociation Mass Spectrometry: A Powerful Tool for Natural Product Structure Elucidation*. Anal. Chem., 2015. **87**(21): p. 10668-10678.
130. Reid, M.J., et al., *Using biomarkers in wastewater to monitor community drug use: A conceptual approach for dealing with new psychoactive substances*. Sci. Total Environ., 2014. **487**: p. 651-658.
131. Dresen, S., et al., *Detection and identification of 700 drugs by multi-target screening with a 3200 Q TRAP® LC-MS/MS system and library searching*. Anal. Bioanal. Chem., 2010. **396**(7): p. 2425-2434.
132. Zhu, X.C., Y.P. Chen, and R. Subramanian, *Comparison of Information-Dependent Acquisition, SWATH, and MSAll Techniques in Metabolite Identification Study Employing Ultrahigh-Performance Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry*. Anal. Chem., 2014. **86**(2): p. 1202-1209.
133. Wohlfarth, A., et al., *Qualitative Confirmation of 9 Synthetic Cannabinoids and 20 Metabolites in Human Urine Using LC–MS/MS and Library Search*. Anal. Chem., 2013. **85**(7): p. 3730-3738.
134. Oberacher, H., et al., *Detection and identification of drugs and toxicants in human body fluids by liquid chromatography–tandem mass spectrometry under data-dependent acquisition control and automated database search*. Anal. Chim. Acta, 2013. **770**: p. 121-131.
135. Wang, N. and L. Li, *Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis*. Anal. Chem., 2008. **80**(12): p. 4696-4710.
136. Arnhard, K., et al., *Applying 'Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra' (SWATH) for systematic toxicological analysis with liquid chromatography-high-resolution tandem mass spectrometry*. Anal. Bioanal. Chem., 2015. **407**(2): p. 405-414.

137. Roemmelt, A.T., et al., *Liquid Chromatography, in Combination with a Quadrupole Time-of-Flight Instrument (LC QTOF), with Sequential Window Acquisition of All Theoretical Fragment-Ion Spectra (SWATH) Acquisition: Systematic Studies on Its Use for Screenings in Clinical and Forensic Toxicology and Comparison with Information-Dependent Acquisition (IDA)*. Anal. Chem., 2014. **86**(23): p. 11742-11749.
138. Scheidweiler, K.B., M.J.Y. Jarvis, and M.A. Huestis, *Nontargeted SWATH acquisition for identifying 47 synthetic cannabinoid metabolites in human urine by liquid chromatography-high-resolution tandem mass spectrometry*. Anal. Bioanal. Chem., 2015. **407**(3): p. 883-897.
139. Lee, H.K., et al., *Development of a broad toxicological screening technique for urine using ultra-performance liquid chromatography and time-of-flight mass spectrometry*. Anal. Chim. Acta, 2009. **649**(1): p. 80-90.
140. Chindarkar, N.S., et al., *Liquid Chromatography High-Resolution TOF Analysis: Investigation of MSE for Broad-Spectrum Drug Screening*. Clin. Chem., 2014. **60**(8): p. 1115-1125.
141. Pasin, D., S. Bidny, and S.L. Fu, *Analysis of New Designer Drugs in Post-Mortem Blood Using High-Resolution Mass Spectrometry*. J. Anal. Toxicol., 2015. **39**(3): p. 163-171.
142. Gundersen, P.O.M., et al., *Retrospective screening of synthetic cannabinoids, synthetic opioids and designer benzodiazepines in data files from forensic post mortem samples analysed by UHPLC-QTOF-MS from 2014 to 2018*. Forensic Sci. Int., 2020. **311**: p. 110274.
143. Noble, C., et al., *Application of a screening method for fentanyl and its analogues using UHPLC-QTOF-MS with data-independent acquisition (DIA) in MSE mode and retrospective analysis of authentic forensic blood samples*. Drug Test. Anal., 2017. **10**(4): p. 651-662.
144. Hernández, F., et al., *Current use of high-resolution mass spectrometry in the environmental sciences*. Anal. Bioanal. Chem., 2012. **403**(5): p. 1251-1264.
145. Ibáñez, M., et al., *Comprehensive analytical strategies based on high-resolution time-of-flight mass spectrometry to identify new psychoactive substances*. TrAC, Trends Anal. Chem., 2014. **57**: p. 107-117.
146. Royal Society of Chemistry. *ChemSpider*. 2015 [cited 2017 24th July]; Available from: <http://www.chemspider.com/>.
147. Grabenauer, M., et al., *Analysis of Synthetic Cannabinoids Using High-Resolution Mass Spectrometry and Mass Defect Filtering: Implications for Nontargeted Screening of Designer Drugs*. Anal. Chem., 2012. **84**(13): p. 5574-5581.
148. Sleno, L., *The use of mass defect in modern mass spectrometry*. J. Mass Spectrom., 2012. **47**(2): p. 226-236.
149. Zhang, H., et al., *Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry*. J. Mass Spectrom., 2009. **44**(7): p. 999-1016.

150. Broecker, S., et al., *Development and practical application of a library of CID accurate mass spectra of more than 2,500 toxic compounds for systematic toxicological analysis by LC–QTOF-MS with data-dependent acquisition*. Anal. Bioanal. Chem., 2011. **400**(1): p. 101-117.
151. Broecker, S., et al., *Combined use of liquid chromatography–hybrid quadrupole time-of-flight mass spectrometry (LC–QTOF-MS) and high performance liquid chromatography with photodiode array detector (HPLC–DAD) in systematic toxicological analysis*. Forensic Sci. Int., 2011. **212**(1): p. 215-226.
152. Cawley, A., et al., *The potential for complementary targeted/non-targeted screening of novel psychoactive substances in equine urine using liquid chromatography-high resolution accurate mass spectrometry*. Anal. Methods, 2016. **8**(8): p. 1789-1797.
153. Kinyua, J., et al., *A data-independent acquisition workflow for qualitative screening of new psychoactive substances in biological samples*. Anal. Bioanal. Chem., 2015. **407**(29): p. 8773-8785.
154. Kinyua, J., et al., *Qualitative screening of new psychoactive substances in pooled urine samples from Belgium and United Kingdom*. Sci. Total Environ., 2016. **573**: p. 1527-1535.
155. Andrés-Costa, M.J., V. Andreu, and Y. Picó, *Analysis of psychoactive substances in water by information dependent acquisition on a hybrid quadrupole time-of-flight mass spectrometer*. J. Chromatogr. A, 2016. **1461**: p. 98-106.
156. Paul, M., et al., *Analysis of new designer drugs and common drugs of abuse in urine by a combined targeted and untargeted LC-HR-QTOFMS approach*. Anal. Bioanal. Chem., 2014. **406**(18): p. 4425-4441.
157. Hughey, C.A., et al., *Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra*. Anal. Chem., 2001. **73**(19): p. 4676-4681.
158. Anstett, A., et al., *Characterization of 2C-phenethylamines using high-resolution mass spectrometry and Kendrick mass defect filters*. Forensic Chem., 2018. **7**: p. 47-55.
159. Fornal, E., *Identification of substituted cathinones: 3,4-Methylenedioxy derivatives by high performance liquid chromatography–quadrupole time of flight mass spectrometry*. J. Pharm. Biomed. Anal., 2013. **81**: p. 13-19.
160. Fornal, E., *Formation of odd-electron product ions in collision-induced fragmentation of electrospray-generated protonated cathinone derivatives: aryl α -primary amino ketones*. Rapid Commun. Mass Spectrom., 2013. **27**(16): p. 1858-1866.
161. Fornal, E., *Study of collision-induced dissociation of electrospray-generated protonated cathinones*. Drug Test. Anal., 2014. **6**(7-8): p. 705-715.
162. Fornal, E., A. Stachniuk, and A. Wojtyla, *LC-Q/TOF mass spectrometry data driven identification and spectroscopic characterisation of a new 3,4-methylenedioxy-N-benzyl cathinone (BMDP)*. J. Pharm. Biomed. Anal., 2013. **72**: p. 139-144.

163. Strano Rossi, S., et al., *An analytical approach to the forensic identification of different classes of new psychoactive substances (NPSs) in seized materials*. Rapid Commun. Mass Spectrom., 2014. **28**(17): p. 1904-1916.
164. Shevyrin, V., et al., *Synthetic cannabinoid 3-benzyl-5-[1-(2-pyrrolidin-1-ylethyl)-1H-indol-3-yl]-1,2,4-oxadiazole. The first detection in illicit market of new psychoactive substances*. Forensic Sci. Int., 2016. **259**: p. 95-100.
165. Shevyrin, V., et al., *3-Naphthoylindazoles and 2-naphthoylbenzoimidazoles as novel chemical groups of synthetic cannabinoids: Chemical structure elucidation, analytical characteristics and identification of the first representatives in smoke mixtures*. Forensic Sci. Int., 2014. **242**: p. 72-80.
166. Shevyrin, V., et al., *Identification and analytical characteristics of synthetic cannabinoids with an indazole-3-carboxamide structure bearing a N-1-methoxycarbonylalkyl group*. Anal. Bioanal. Chem., 2015. **407**(21): p. 6301-6315.
167. Lobo Vicente, J., et al., *Systematic analytical characterization of new psychoactive substances: A case study*. Forensic Sci. Int., 2016. **265**: p. 107-115.
168. Margagliotti, G. and T. Bollé, *Machine learning & forensic science*. Forensic Sci. Int., 2019. **298**: p. 138-139.
169. MATLAB, *Introducing Machine Learning*, in *Machine Learning with MATLAB*. 2016, MathWorks: Online. 92991v00.
170. Mitchell, T.M., *Machine Learning*. 1997, New York: McGraw-Hill.
171. MATLAB, *Applying Supervised Learning*, in *Machine Learning with MATLAB*. 2016, MathWorks: Online. 80827v00.
172. MATLAB, *Applying Unsupervised Learning*, in *Machine Learning with MATLAB*. 2016, MathWorks: Online. 80823v00.
173. Ekins, S., *Computational Toxicology : Risk Assessment for Chemicals*. 2018, Newark, United States: John Wiley & Sons, Incorporated.
174. Luechtefeld, T., C. Rowlands, and T. Hartung, *Big-data and machine learning to revamp computational toxicology and its use in risk assessment*. Toxicol. Res-UK, 2018. **7**(5): p. 732-744.
175. Yeow, W.L., R. Mahmud, and R.G. Raj, *An application of case-based reasoning with machine learning for forensic autopsy*. Expert Syst. Appl., 2014. **41**(7): p. 3497-3505.
176. Lambert, D., et al., *Combining spectroscopic data in the forensic analysis of paint: Application of a multiblock technique as chemometric tool*. Forensic Sci. Int., 2016. **263**: p. 39-47.
177. Muehlethaler, C., G. Massonnet, and P. Esseiva, *Discrimination and classification of FTIR spectra of red, blue and green spray paints using a multivariate statistical approach*. Forensic Sci. Int., 2014. **244**: p. 170-178.

178. Carriquiry, A., et al., *Machine learning in forensic applications*, in *Significance*. 2019, The Royal Statistical Society. p. 29-35.
179. Pyke, J.S., et al. *Simultaneous Targeted Quantitation and Suspect Screening of Environmental Contaminants in Sewage Sludge*. Online: Agilent Technologies;2019. 5994-0750EN.
180. Miller, T.H., et al., *Prediction of Chromatographic Retention Time in High-Resolution Anti-Doping Screening Data Using Artificial Neural Networks*. *Anal. Chem.*, 2013. **85**(21): p. 10330-10337.
181. Mollerup, C.B., et al., *Prediction of collision cross section and retention time for broad scope screening in gradient reversed-phase liquid chromatography-ion mobility-high resolution accurate mass spectrometry*. *J. Chromatogr. A*, 2018. **1542**: p. 82-88.
182. Groger, T., et al., *Application of two-dimensional gas chromatography combined with pixel-based chemometric processing for the chemical profiling of illicit drug samples*. *J. Chromatogr. A*, 2008. **1200**(1): p. 8-16.
183. Morelato, M., et al., *The use of organic and inorganic impurities found in MDMA police seizures in a drug intelligence perspective*. *Sci. Justice*, 2014. **54**(1): p. 32-41.
184. Lurie, I.S. and S.G. Toske, *Applicability of ultra-performance liquid chromatography-tandem mass spectrometry for heroin profiling*. *J. Chromatogr. A*, 2008. **1188**(2): p. 322-326.
185. Marcelo, M.C.A., et al., *Profiling cocaine by ATR-FTIR*. *Forensic Sci. Int.*, 2015. **246**: p. 65-71.
186. Bumrah, G.S. and R.M. Sharma, *Raman spectroscopy – Basic principle, instrumentation and selected applications for the characterization of drugs of abuse*. *Egypt. J. Forensic Sci.*, 2016. **6**(3): p. 209-215.
187. Penido, C., et al., *Raman spectroscopy in forensic analysis: identification of cocaine and other illegal drugs of abuse*. *J. Raman Spectrosc.*, 2016. **47**(1): p. 28-38.
188. Ryder, A.G., G.M. O'Connor, and T.J. Glynn, *Identifications and quantitative measurements of narcotics in solid mixtures using near-IR Raman spectroscopy and multivariate analysis*. *J. Forensic Sci.*, 1999. **44**(5): p. 1013-1019.
189. West, M.J. and M.J. Went, *Detection of drugs of abuse by Raman spectroscopy*. *Drug Test. Anal.*, 2011. **3**(9): p. 532-538.
190. Morelato, M., et al., *The use of forensic case data in intelligence-led policing: The example of drug profiling*. *Forensic Sci. Int.*, 2013. **226**(1–3): p. 1-9.
191. Collins, M., et al., *Illicit drug profiling: the Australian experience*. *Aust. J. Forensic Sci.*, 2007. **39**(1): p. 25-32.
192. Qi, Y., I.D. Evans, and A. McCluskey, *Australian Federal Police seizures of illicit crystalline methamphetamine ('ice') 1998–2002: Impurity analysis*. *Forensic Sci. Int.*, 2006. **164**(2–3): p. 201-210.

193. Collins, M. and H. Salouros, *A review of some recent studies on the stable isotope profiling of methylamphetamine: Is it a useful adjunct to conventional chemical profiling?* Sci. Justice, 2015. **55**(1): p. 2-9.
194. Stojanovska, N., et al., *A review of impurity profiling and synthetic route of manufacture of methylamphetamine, 3,4-methylenedioxymethylamphetamine, amphetamine, dimethylamphetamine and p-methoxyamphetamine.* Forensic Sci. Int., 2013. **224**(1-3): p. 8-26.
195. Dujourdy, L., et al., *Drug intelligence based on organic impurities in illicit MA samples.* Forensic Sci. Int., 2008. **177**(2-3): p. 153-161.
196. Marclay, F., et al., *Perspectives for Forensic Intelligence in anti-doping: Thinking outside of the box.* Forensic Sci. Int., 2013. **229**(1-3): p. 133-144.
197. Ratcliffe, J.H. *Integrated Intelligence and Crime Analysis: Enhanced Information Management for Law Enforcement Leaders.* Washington D.C.: U.S. Department of Justice, Office of Community Oriented Policing Services (COPS);2007.
198. Morelato, M., et al., *The use of methylamphetamine chemical profiling in an intelligence-led perspective and the observation of inhomogeneity within seizures.* Forensic Sci. Int., 2015. **246**: p. 55-64.
199. Dams, R., et al., *Heroin impurity profiling: trends throughout a decade of experimenting.* Forensic Sci. Int., 2001. **123**(2-3): p. 81-88.
200. Perillo, B.A., R.F.X. Klein, and E.S. Franzosa, *Recent Advances by the US Drug Enforcement Administration in Drug Signature And Comparative Analysis.* Forensic Sci. Int., 1994. **69**(1): p. 1-6.
201. Broseus, J., et al., *Study of common database feeding with results coming from different analytical methods in the framework of illicit drugs chemical profiling.* Forensic Sci. Int., 2013. **230**(1-3): p. 16-28.
202. Esseiva, P., et al., *Chemical profiling and classification of illicit heroin by principal component analysis, calculation of inter sample correlation and artificial neural networks.* Talanta, 2005. **67**(2): p. 360-367.
203. Esseiva, P., et al., *Illicit drug profiling, reflection on statistical comparisons.* Forensic Sci. Int., 2011. **207**(1-3): p. 27-34.
204. Mayer, B.P., et al., *Chemical Attribution of Fentanyl Using Multivariate Statistical Analysis of Orthogonal Mass Spectral Data.* Anal. Chem., 2016. **88**(8): p. 4303-4310.

Chapter 2:
Collision-induced Dissociation
Studies

Chapter 2: Collision-induced Dissociation Studies

2.1 Rationale

Understanding the fragmentation pathways for common synthetic opioids is important to establish a framework from which non-targeted analysis workflows can be developed. The ability of HRMS instruments to operate in DIA modes opens up the possibility for the use of diagnostic product ions to screen for the presence of synthetic opioids. Where novel opioid derivatives contain similar core structures to known analogues, it is possible that these diagnostic product ions will also be present under similar analysis conditions. This chapter focuses on the determination of CID pathways for various subclasses of synthetic opioids in order to identify diagnostic product ions that may be useful in the development of non-targeted screening workflows.

This chapter is taken from a first author publication entitled 'Collision-induced Dissociation Studies of Synthetic Opioids for Non-targeted Analysis', published in *Frontiers in Chemistry*. The experimental work, data analysis and preparation of the initial draft were performed by J. Klingberg, with manuscript edits provided by A. Cawley, R. Shimmon and S. Fu.

Klingberg, J., et al., *Collision-Induced Dissociation Studies of Synthetic Opioids for Non-targeted Analysis*. *Front. Chem.*, 2019. **7**(331).

NOTE: The CID studies performed in this chapter are not exhaustive and were designed solely to identify possible diagnostic product ions belonging to the subclasses of synthetic opioids. The proposal of product ion structures was only to support the selection of diagnostic product ions. The proposed structures may not entirely reflect the actual structures as deuterated standards were not analysed.

Collision-induced Dissociation Studies of Synthetic Opioids for Non-targeted Analysis

Joshua Klingberg¹, Adam Cawley², Ronald Shimmon¹, Shanlin Fu^{1*}

¹Centre for Forensic Science, University of Technology Sydney, Ultimo, NSW, Australia

²Australian Racing Forensic Laboratory, Racing NSW, Sydney, NSW, Australia

*Corresponding Author:

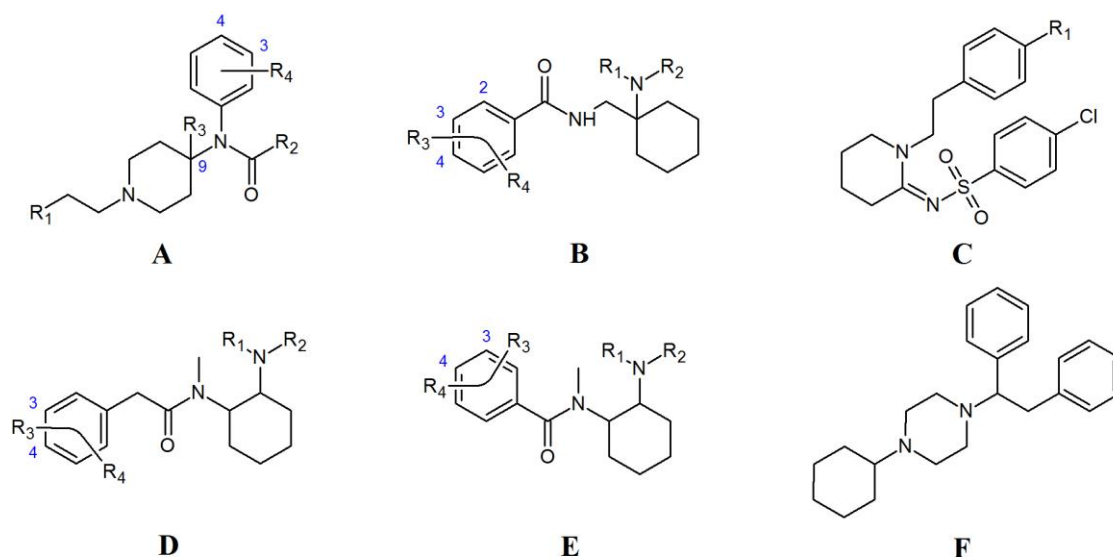
Shanlin Fu

Shanlin.fu@uts.edu.au

2.2 Introduction

The term “opioid” literally means “opium-like substance.” Opium is the dried juice of the opium poppy, *Papaver Somniferum*, which has been used for the relief of pain for thousands of years ^[30]. The actions of the opiates mirror the actions of endogenous opioid peptides, such as endorphins, enkephalins and dynorphins, which exist within the central and peripheral nervous system as neurotransmitters and neuromodulators ^[32]. Synthetic opioids are man-made compounds that act on the opioid receptors in the brain and mimic the effects of natural opiates, such as morphine and codeine ^[29]. While there are a number of synthetic opioids known to forensic toxicologists, such as methadone, levomethorphan and pethidine, two groups have recently come under increased scrutiny, namely NPFs, which are fentanyl derivatives that are obtained or synthesized illicitly, and NSOs. While there can be some overlap between these classifications, NSOs usually refer to the newer classes of opioids, such as the AH and U series’, and pseudo-NSOs, such as the W series, which were originally developed to act as new opioid agonists, but never brought to the market for human use ^[19]. Many of these compounds have recently emerged onto the illicit drug market. The UNODC have identified opioids as the class of compounds which cause the most harm, accounting for 76% of all deaths attributed to drug use worldwide in 2016, with fentanyl derivatives alone accounting for 30% of all deaths in the United States ^[205]. Many of these deaths may come from fentanyl itself, however these statistics still demonstrate the significant threats posed by these compounds. In addition, there is an increasing trend of fentanyl derivatives, and other synthetic opioids, being mixed into illicit drugs, or being sold as prescription opioids. Figure 2-1 shows the chemical structures of the synthetic opioids investigated in this study grouped into subclasses based on their structural features. Individual structures for each of these opioids can be found in Appendix 1.

The continuous introduction of a large number of NPS, along with the lack of CRMs for many of these compounds, highlights the inadequacy of traditional targeted screening methods to detect these substances ^[122]. In particular, many synthetic opioids may not be detected by traditional approaches to illicit drug analysis, and many routine analyses do not screen for these compounds ^[37]. This may be due to numerous factors, including a lack of cross-reactivity with traditional immunoassay screening methods and data for these compounds not being available in the mass spectral libraries that are used for drug screening ^[38]. To combat this, the aim of nontargeted analysis is to identify compounds from samples where the molecular content is unknown ^[123]. Since illicit drug seizures are, by their very nature, unknown entities, this approach to their analysis provides some obvious advantages.



Name	Code	R ₁	R ₂	R ₃	R ₄
Fentanyl	A1	C ₆ H ₅	CH ₂ CH ₃	H	H
Acetyl fentanyl	A2	C ₆ H ₅	CH ₃	H	H
Acryl fentanyl	A3	C ₆ H ₅	CH ₂ CH	H	H
α -methyl fentanyl	A4	CH ₃ C ₆ H ₅	CH ₂ CH ₃	H	H
4-fluoroisobutyl fentanyl	A5	C ₆ H ₅	CH(CH ₃) ₂	H	4-F
4-methoxybutyl fentanyl	A6	C ₆ H ₅	CH ₂ CH ₂ CH ₃	H	4-OCH ₃
Butyl fentanyl	A7	C ₆ H ₅	CH ₂ CH ₂ CH ₃	H	H
Furanyl fentanyl	A8	C ₆ H ₅	C ₄ H ₃ O	H	H
Meta-fluoro fentanyl	A9	C ₆ H ₅	CH ₂ CH ₃	H	3-F
Valeryl fentanyl	A10	C ₆ H ₅	CH ₂ CH ₂ CH ₂ CH ₃	H	H
Alfentanil	A11	CN ₄ OCH ₂ CH ₃	CH ₂ CH ₃	CH ₂ OCH ₂ CH ₃	H
Sufentanil	A12	C ₄ H ₃ S	CH ₂ CH ₃	CH ₂ OCH ₂ CH ₃	H
Carfentanil	A13	C ₆ H ₅	CH ₂ CH ₃	COOCH ₂ CH ₃	H
Remifentanil	A14	COOCH ₃	CH ₂ CH ₃	COOCH ₂ CH ₃	H
AH-7921	B1	CH ₃	CH ₃	3-Cl	4-Cl
AH-8529	B2	CH ₃	CH ₃	4-Cl	H
AH-8532	B3	CH ₃	CH ₃	3-Cl	H
AH-8533	B4	CH ₃	CH ₃	2-Cl	H
AH-8507	B5	C ₄ H ₈ NCH ₃		3-Cl	4-Cl
AH-7959	B6	C ₅ H ₁₀		3-Cl	4-Cl
W-18	C1	NO ₂	-	-	-
W-19	C2	NH ₂	-	-	-
W-15	C3	H	-	-	-
U-50488	D1	C ₄ H ₈		3-Cl	4-Cl
U-51754	D2	CH ₃	CH ₃	3-Cl	4-Cl
U-47700	E1	CH ₃	CH ₃	3-Cl	4-Cl
U-49900	E2	CH ₂ CH ₃	CH ₂ CH ₃	3-Cl	4-Cl
MT-45	F	-	-	-	-

Figure 2-1: General chemical structure of fentanyl (A), AH series opioids (B), W series opioids (C), U series opioids with (D), and without (E) a methylene spacer and MT-45 (F)

The use of HRMS has been investigated for non-targeted analysis due to its ability to be operated in DIA modes in addition to traditional DDA modes [122, 125]. While DIA has been shown to be capable of detecting more compounds at lower concentration than DDA [136], a challenge arises due to the

production of “chimeric spectra.” DIA methods fragment all precursor ions simultaneously, producing a spectrum that includes product ions from all precursor ion subjected to CID ^[122, 124]. This means that the software cannot associate product ions with their correct precursor ions. It has been suggested that deconvolution algorithms may be used to exploit the full potential of DIA data ^[124], however the use of DDA may be useful in situation where rapid identification of product ions related to specific precursor ions is required.

The main advantage of HRMS over other mass spectrometric methods is its ability to differentiate between ions and losses that have the same nominal mass. For example, the loss of an NH₃ molecule or an –OH group would equate to the same nominal mass change of 17 Da. This change may be indistinguishable using lower resolution MS techniques. If an MS technique with a high enough mass resolution was used, however, the accurate masses for those fragments of 17.0265 and 17.0027 Da respectively could be determined and therefore the losses differentiated. The effectiveness of CID data for the development of non-targeted detection strategies has been previously demonstrated for different drug classes ^[126].

While fragmentation patterns have been proposed in the literature for various individual opioids, mainly fentanyl derivatives ^[37, 60, 72, 206-211], there is a need for an overview of the CID pathways for all synthetic opioids to facilitate broad-spectrum opioid screening. Noble et al. ^[143] previously investigated the use of HRMS data collected in DIA mode to identify 50 fentanyl analogues without the use of CRMs. In this study, the identification of common CID pathways is expanded to include both NPFs and NSOs. Additionally, the application of this information to non-targeted screening methods will be discussed.

2.3 Materials and Methods

2.3.1 Solvents and Reagents

All solvents used were LC-MS grade. Acetonitrile, ethyl acetate and methanol were obtained from Merck (Darmstadt, Germany). Ammonium acetate and trichloroacetic acid were obtained from Sigma-Aldrich (Castle Hill, Australia). Acetic acid was obtained from Ajax Chemicals (Sydney, Australia). Ultrapure-grade water was obtained from a Smart2Pure ultra-pure water system (Thermo Scientific, Langenselbold, Hungary).

2.3.2 Opioid Standards

Fentanyl citrate was purchased from Sigma-Aldrich (Castle Hill, NSW). Hydrochloride salts of acetyl fentanyl, 4- fluoroisobutyryl fentanyl, meta-fluoro fentanyl, AH-7921, AH-8529, AH-8533, U-51754, and W-19, along with neat solids of AH-7959, AH-8532, AH-8507, U-49900, and U- 50488, were purchased from Sapphire Bioscience (Redfern, NSW). Hydrochloride salts of furanyl fentanyl, acryl fentanyl, 4-methoxybutyryl fentanyl, valeryl fentanyl, butyryl fentanyl, U-47700, and W-18, along with a neat solid of W-15 and MT-45 dihydrochloride hydrate, were purchased from Chiron Chemicals (Hawthorn, VIC). Remifentanil hydrochloride was purchased from GlaxoSmithKline (Boronia, VIC). Citrate salts of carfentanil, sufentanil and α -methyl fentanyl, along with alfentanil hydrochloride, were purchased from Janssen Pharmaceuticals (North Ryde, NSW).

2.3.3 Sample Preparation

Drug standards were obtained as methanolic standards of varying concentrations. All opioid samples for the CID study were diluted in methanol to a concentration of 10 μ g/mL. Ten microliters of each solution was evaporated to dryness under nitrogen, before being reconstituted in methanol and a 10mM ammonium acetate (pH 4) buffer to give a final concentration of 1 μ g/mL for analysis. All samples were refrigerated until analysis.

A spiked sample of equine plasma was prepared to test the application of the identified diagnostic product ions to a non-targeted screening workflow. A mixed standard containing fentanyl, acetyl fentanyl, AH-7921 and U-50488 was made with a concentration of 200 ng/mL in methanol. Two mL of blank equine plasma was spiked with the mixed standard to give concentrations ranging from 10 to 0.01 ng/mL. A blank sample was also prepared for extraction. The proteins were precipitated out of the samples by the addition of 200 μ L of trichloroacetic acid (10% in H₂O) and the pH of the samples was adjusted to 3– 3.5 using hydrochloric acid. The samples were then centrifuged at 1,500 g for 15min. Following centrifugation, the samples were extracted using XTRACKTR Gravity Flow DAU Extraction Columns (UCT Inc., Bristol, United States). The cartridges were first conditioned with 3mL of methanol, followed by 3mL of purified water. The samples were then loaded, and the cartridges washed with 3mL of 0.1M acetic acid, before being dried under positive pressure. Three mL of methanol was run through the cartridges, which were dried again under positive pressure. Finally, the samples were eluted using 3mL of an elution solvent containing 3% ammonia and 0.5% methanol in ethyl acetate.

One drop of 0.1M methanolic hydrochloric acid was added to the samples before the solvent was evaporated under nitrogen at 60°C. The samples were then reconstituted with one drop of methanol and 100 µL of an ammonium acetate buffer (pH 3.9), before being transferred to a vial for analysis.

2.3.4 Instrumental Analysis

Chromatographic separation was achieved using an Agilent Technologies (Santa Clara, CA, USA) 1290 Infinity II UHPLC, consisting of a high speed pump (G7120A), multisampler (G7167B) and thermostat and column compartment (G1316A, 35°C) coupled to an Agilent 6545 QTOF-MS. All data acquisition was conducted using Agilent MassHunter Workstation (Version B.06.01). A sample volume of 5 µL was injected onto a Phenomenex (Torrance, CA, USA) Gemini 110 Å C18 LC column (2 × 50mm, 5µm particle size) using a gradient elution method with a flow rate of 0.5 mL/min and a total analysis time of 11.5min. Mobile phase A consisted of a 10mM ammonium acetate (pH 9) buffer and mobile phase B consisted of 0.1% acetic acid in acetonitrile. Initial mobile phase composition was 99% A, which was held for 2min before being decreased linearly to 20% A over 6.5min. The mobile phase was then returned to 99% A over 3 min.

The QTOF-MS was operated in positive electrospray ionization mode (ESI+) with capillary and fragmentor voltages of 3,500 and 100V, respectively. An Auto-MS/MS (data dependent) acquisition mode was used with mass ranges of 100–1,000 m/z for MS and 50–500 m/z for MS/MS and a scan rate of 10 spectra/s for both. A preferred precursor ion table was populated with the corresponding $[M + H]^+$ values for each opioid standard and the acquisition method was set to use the preferred list only for MS/MS analysis. CID experiments were performed using collision energies (CE) of 10, 20 and 40 eV in separate experiments, with nitrogen as the collision gas.

For the spiked plasma samples, the QTOF was operated in ESI+ mode with capillary and fragmentor voltages of 3,500 and 100V, respectively. An AllionsMS (DIA) data acquisition mode was used with a mass range of 35–1,000 m/z . Spectra were obtained with an acquisition rate of 10 spectra/s and CEs of 10, 20, and 40 eV were used for CID.

All data acquired was processed using Agilent MassHunter Qualitative Analysis Software (Version B.07.00). The Find by Auto-MS/MS function was used to obtain MS/MS spectra for all precursor ions selected for CID experiments. All MS/MS spectra obtained from the analysed standards can be found in Appendix 2. For the spiked plasma samples, extracted ion chromatograms (EICs) were created for

a number of diagnostic product ions identified from the CID study to screen for possible compounds present within the samples.

2.4 Results and Discussion

2.4.1 Collision-induced Dissociation Pathways

2.4.1.1 *Non-pharmaceutical Fentanyls*

The generic structure of fentanyl (Figure 2-1, structure **A**) shows many different sites that can be altered to produce various analogues. This can lead to different CID pathways being observed depending on the position of substituents. While this can make it difficult to identify common fragments, some overall patterns can be observed for the different sub-groups of NPFs. While a phenylethyl group is the most common substituent found at R₁, some analogues contain different tails, including ester groups and heterocyclic rings. Additionally, substitution at the C9 position of the piperidine ring (R₃) can also be found in some NPFs. Figure 2-2 highlights some common fragment ions that can be found from NPFs containing differing functional groups at these locations.

The main fragments present in the MS² of these compounds may arise from the loss of either the C9 side chain (R₃) or the phenylamide group, resulting in fragments **1a** and **1b**, respectively. Evidently the exact masses of these resultant product ions would be highly dependent on the various functional groups present in the structure, especially the tail group at R₁. The other common product ions found in this subclass of compounds, **1c**, may arise from the inter-ring cleavage of the piperidine ring. Similarly to **1a** and **1b**, the exact mass of this fragment will vary depending on the functional group present at R₁. While the differing masses of the product ions formed is not ideal from a non-targeted screening perspective, understanding the pattern of fragmentation that the molecules undergo is important to help elucidate the structure of possible unknown compounds.

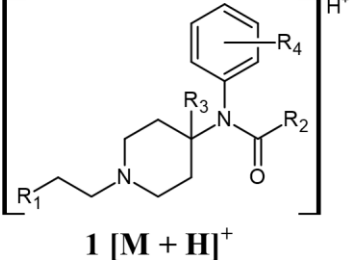
 <p>1 [M + H]⁺</p>		
1a	R ₁ = CN ₄ OCH ₂ CH ₃ R ₂ = CH ₂ CH ₃ R ₄ = H	371.2195
	R ₁ = C ₄ H ₃ S R ₂ = CH ₂ CH ₃ R ₄ = H	343.1844
	R ₁ = C ₆ H ₅ R ₂ = CH ₂ CH ₃ R ₄ = H	335.2123
	R ₁ = COOCH ₃ R ₂ = CH ₂ CH ₃ R ₄ = H	317.1865
1b	R ₁ = CN ₄ OCH ₂ CH ₃ R ₃ = CH ₂ OCH ₂ CH ₃	268.1774
	R ₁ = C ₄ H ₃ S R ₃ = CH ₂ OCH ₂ CH ₃	238.1266
	R ₁ = C ₆ H ₅ R ₃ = COOCH ₂ CH ₃	246.1494
	R ₁ = COOCH ₃ R ₃ = COOCH ₂ CH ₃	228.1236
1c	R ₁ = CN ₄ OCH ₂ CH ₃	170.1042
	R ₁ = C ₄ H ₃ S	140.0534
	R ₁ = C ₆ H ₅	134.0970
	R ₁ = COOCH ₃	116.0712

Figure 2-2: Proposed structures of fragments observed for NPFs containing differing tails and C9 side chains, showing molecular ion (**1**) and common product ions (**1a – 1c**)

As mentioned previously, the most common functional group found at R₁ is an aromatic ring, causing the compound to have a phenylethyl tail off the piperidine ring. Consequently, the most common points of modification between the different derivatives are the R₂ and R₄ groups (Figure 2-1, structure **A**). Figure 2-3 presents the common product ions identified when this core structure is present and there is no additional side chain present on the piperidine ring (R₃). The first common product ion can be found at an *m/z* of 105.0704 which is representative of a methyl-substituted tropylium ion (**2a**). This fragment is very common for alkyl-substituted benzenes and will often break down to form the methyl-substituted aromatic cyclopentadienyl cation, fragment **2b**, with an *m/z* of 79.0548 ^[212]. While these product ions are quite common, and therefore not specific to NPFs, their presence in addition to other common fragments can help confirm the presence of these compounds. Product ions **2c** and **2d** may be formed from the cleavage between the nitrogen and the carbonyl group of the amide side chain. The *m/z* of fragment **2c** may change with substitution of the aromatic ring, however many derivatives have an unsubstituted aromatic ring, leading to a product ion with an *m/z* of 281.2018. The functional groups at R₂ on fragment **2d** are much more varied, leading to different *m/z* values being observed, however the loss between the molecular ion and fragment **2c** can give an indication of the expected *m/z* for fragment **2d**. The further loss of the nitrogen and aromatic ring from fragment **2c** gives rise to fragment **2e**, with an *m/z* of 188.1439. While it has not

been confirmed whether this product ion is formed directly from the fragmentation of the molecular ion, or if further fragmentation of another product ion has occurred, it is not of great importance for non-targeted screening purposes. The important information in this case is that a product ion with the observed m/z is present in the samples analysed. The loss of the phenylethyl tail and the collapse of the piperidine ring leads to the formation of fragment **2f**. Once again, the exact m/z of this fragment will vary due to the presence of the amide side chain, which can vary dramatically. Finally, fragment **2g** can form from the collapse of the piperidine ring and cleavage between the nitrogen and carbonyl group of the amide side chain. Similarly to fragment **2c**, fragment **2g** may vary with substitution of the aromatic ring, however it is most commonly found with an m/z of 146.0970 when no substituent is present. Interestingly, the two fluorine-containing derivatives analysed in this study (4-fluoroisobutyryl fentanyl **A5** and metafluoro fentanyl **A9**, Figure 2-1) both showed fragments at m/z 146.0970. This indicates that fluorine may have been lost before the formation of this fragment, as no corresponding fragment was observed at m/z 165.0954 which would have been expected for the fluorine-substituted product ion.

While all the common product ions identified were not observed for all NPFs, the large number identified meant that a combination of multiple common fragments were present across all derivatives analysed. The presence of a combination of these common fragments in an unknown sample can provide a strong indication that an NPF may be present within the sample. The core structure displayed by these compounds, especially those related to structure **2**, allows for the use of a smaller number of fragments to identify the presence of a broad range of compounds. More importantly, some of the fragments presented in Figure 2-3 have little or no variation in their m/z , regardless of the alterations made to core structure. This means that these product ions could potentially be used to screen for previously unknown derivatives, provided that they contain the same core structure. This is of great importance for non-targeted analysis, since the structures of NPFs are unknown.

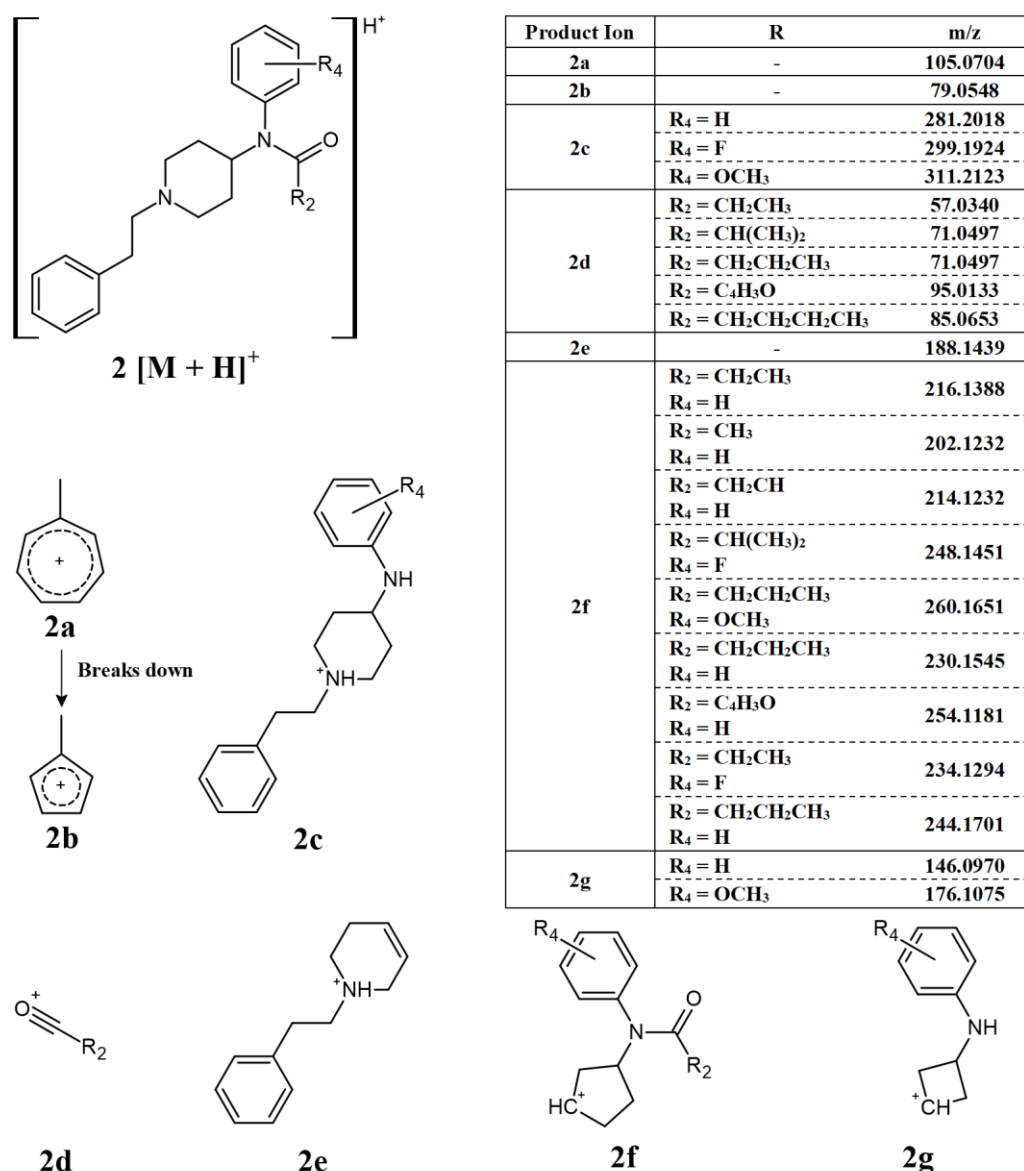
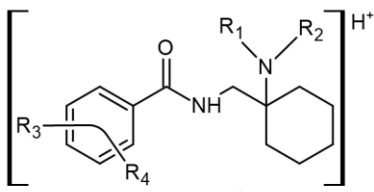


Figure 2-3: Proposed structures of fragments observed for NPFs containing a phenylethyl tail and lacking a C9 side chain, showing molecular ion (**2**) and common product ions (**2a – 2g**)

2.4.1.2 AH Series Opioids

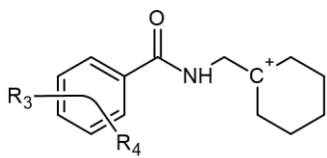
The AH series of synthetic opioids fragment quite predictably, breaking down from the cyclohexane ring toward the aromatic group (Figure 2-4, structure **3**). The first product ion is generated from the loss of the amine side chain attached to the cyclohexane ring to form **3a**. While the mass loss caused by this fragmentation can vary depending on the structure of the amine side chain, most AH series opioids have one or two Cl substituents on the aromatic group, resulting in product ion **3a** having an *m/z* of 250.0999 (³⁵Cl), 252.0969 (³⁷Cl), 284.0609 (³⁵Cl/³⁵Cl), 286.0580 (³⁵Cl/³⁷Cl), or 288.0550

($^{37}\text{Cl}/^{37}\text{Cl}$), respectively. The next major fragments form as a result of the loss of the cyclohexane group and shortening of the amide side chain, giving rise to product ions **3b** and **3c**, respectively. The loss of the amine group and finally the carbonyl group attached to the aromatic ring results in the formation of the final characteristic fragments, **3d** and **3e**, respectively. Fragments **3b** – **3e** may have different m/z values depending on whether they contain one or two Cl substituents, similar to **3a**, however all the compounds within this subclass of synthetic opioids will present a fragment with one of these m/z values. The common m/z produced by these fragments highlight the possibility of using them as targets for screening techniques to detect the presence of this series of NSOs.

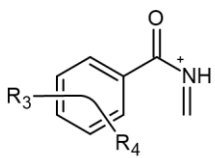


3 $[\text{M} + \text{H}]^+$

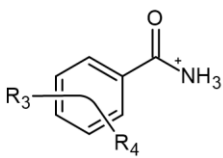
Product Ion	R	m/z
3a	$\text{R}_3 = \text{Cl}$	250.0999 (^{35}Cl)
	$\text{R}_4 = \text{H}$	252.0969 (^{37}Cl)
	$\text{R}_3 = \text{Cl}$	284.0609 ($^{35}\text{Cl}/^{35}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	286.0580 ($^{35}\text{Cl}/^{37}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	288.0550 ($^{37}\text{Cl}/^{37}\text{Cl}$)
3b	$\text{R}_3 = \text{Cl}$	168.0216 (^{35}Cl)
	$\text{R}_4 = \text{H}$	170.0187 (^{37}Cl)
	$\text{R}_3 = \text{Cl}$	201.9827 ($^{35}\text{Cl}/^{35}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	203.9797 ($^{35}\text{Cl}/^{37}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	205.9768 ($^{37}\text{Cl}/^{37}\text{Cl}$)
3c	$\text{R}_3 = \text{Cl}$	156.0216 (^{35}Cl)
	$\text{R}_4 = \text{H}$	158.0187 (^{37}Cl)
	$\text{R}_3 = \text{Cl}$	189.9827 ($^{35}\text{Cl}/^{35}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	191.9797 ($^{35}\text{Cl}/^{37}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	193.9768 ($^{37}\text{Cl}/^{37}\text{Cl}$)
3d	$\text{R}_3 = \text{Cl}$	138.9951 (^{35}Cl)
	$\text{R}_4 = \text{H}$	140.9921 (^{37}Cl)
	$\text{R}_3 = \text{Cl}$	172.9561 ($^{35}\text{Cl}/^{35}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	174.9532 ($^{35}\text{Cl}/^{37}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	176.9502 ($^{37}\text{Cl}/^{37}\text{Cl}$)
3e	$\text{R}_3 = \text{Cl}$	111.0002 (^{35}Cl)
	$\text{R}_4 = \text{H}$	112.9972 (^{37}Cl)
	$\text{R}_3 = \text{Cl}$	144.9612 ($^{35}\text{Cl}/^{35}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	146.9582 ($^{35}\text{Cl}/^{37}\text{Cl}$)
	$\text{R}_4 = \text{Cl}$	148.9553 ($^{37}\text{Cl}/^{37}\text{Cl}$)



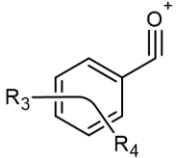
3a



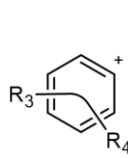
3b



3c



3d



3e

Figure 2-4: Proposed structures of fragments observed for AH series opioids, showing molecular ion (**3**) and common products ions (**3a** – **3e**)

2.4.1.3 U Series Opioids

The U series compounds which do not contain a methylene spacer (Figure 2-1, structure **E**) are constitutional isomers of the compounds within the AH series (Figure 2-1, structure **B**). This means that, while the structures of the fragments are slightly different, they have identical theoretical m/z values. When it comes to non-targeted analysis, this can be a double-edged sword. On the one hand, having the same m/z values means that a smaller number of common product ions can be used to capture a broader range of compounds. Conversely, the presence of multiple fragment ions with the

same m/z reduces the effectiveness of techniques such as class prediction models, making it difficult to determine the identity of these isomers on their fragmentation pattern alone.

The U series compounds which do contain a methylene spacer (Figure 2-1, structure **D**) fragment in a similar manner, however the presence of the spacer changes the m/z of the product ions observed, allowing them to be distinguished from other NSOs.

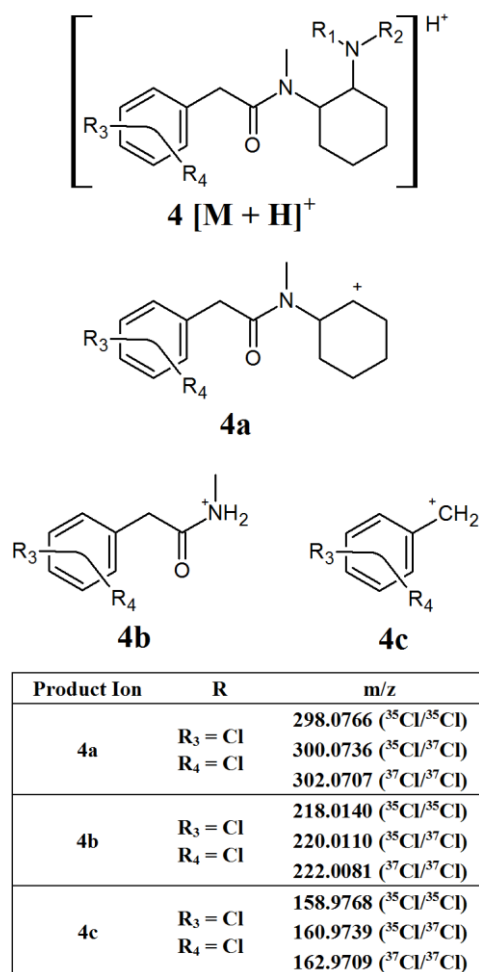


Figure 2-5: Proposed structures of fragments observed for U series opioids containing a methylene spacer, showing molecular ion (**4**) and common product ions (**4a – 4c**)

Figure 2-5 provides an overview of the major common fragments observed from the CID of these compounds. Similarly to the AH series compounds, the loss of the amine group on the cyclohexane ring may produce fragment **4a**. All the compounds analysed in this study containing the additional methylene spacer had two Cl substituents on the aromatic ring, leading to fragment **4a** having a theoretical m/z of 298.0766 ($^{35}Cl/^{35}Cl$), 300.0736 ($^{35}Cl/^{37}Cl$), and 302.0707 ($^{37}Cl/^{37}Cl$). Changing the functional groups on the aromatic ring would lead to different m/z values being observed for this

fragment, as was seen with the AH series compounds, however the two Cl groups appears to be the most common structure found. The subsequent loss of the cyclohexane ring and the amide group produced fragments **4b** and **4c**, respectively. Once again, these fragments may have different m/z values, dependent on the functional groups present on the aromatic ring, however the fragmentation pattern remains consistent.

2.4.1.4 *W Series Pseudo-Opioids*

The W series of pseudo-NSOs do not have an extensive fragmentation pattern, with only a few product ions being detected. Figure 2-6 provides an overview of the common fragment ions proposed. This series of compounds only has one main area of structural variation, meaning that some product ions will have the same m/z regardless of which derivative is analysed. The first product ion identified, **5a**, comes from the loss of one of the aromatic groups attached to the 6-membered heterocyclic ring and contains the differing functional group. This means that the exact m/z of this fragment may vary depending on the substituent attached to the aromatic ring. The remaining two fragments, **5b** and **5c**, respectively, arise from the isolation of the heterocyclic ring and the breaking of the sulfur-nitrogen bond. Given that these fragments do not have any varying functional groups, fragment **5b** will have a consistent m/z of 111.0922. Fragment **5c** will have theoretical m/z values of 174.9621 (^{35}Cl) and 176.9591 (^{37}Cl). While there may not be as many common fragment ions within this series, there are still some masses which may prove useful targets for the development of screening methods.

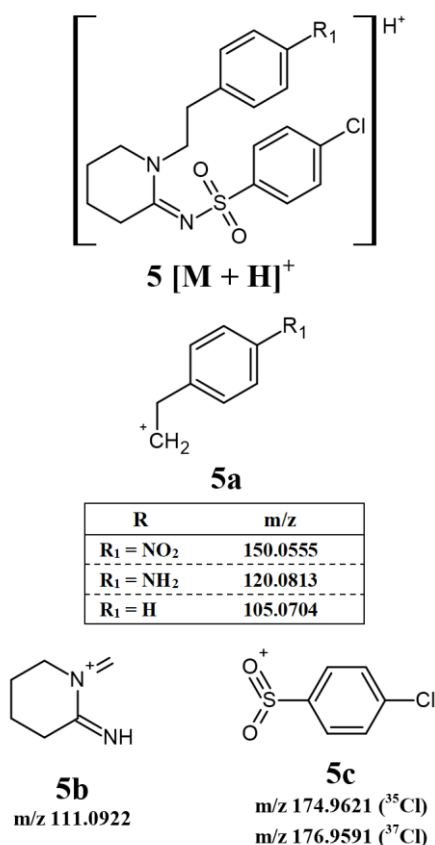


Figure 2-6: Proposed structures of fragments observed for W series opioids, showing molecular ion (**5**) and common product ions (**5a – 5c**)

2.4.1.5 MT-45

MT-45 is a synthetic opioid that does not fit into any of the other defined classes. The fragmentation of this compound was studied, and a number of product ions identified (Figure 2-7). The first fragmentation involves the cleavage of the bond between the piperazine ring and the dibenzene side chain. This gives rise to product ions **6a** and **6b** which can be observed at m/z 181.1017 and 169.1705, respectively. The subsequent cleavage occurs between the piperazine and cyclohexane rings of **6b**, forming product ions **6c** and **6d** with m/z 87.0922 and 83.0861, respectively. While this fragmentation pattern is for a single compound, there is a possibility for the development of additional derivatives based on this core structure. If such derivatives are encountered, the product ions observed can be compared to those presented in Figure 2-7 to identify which ones can be considered diagnostic ions for the new subclass.

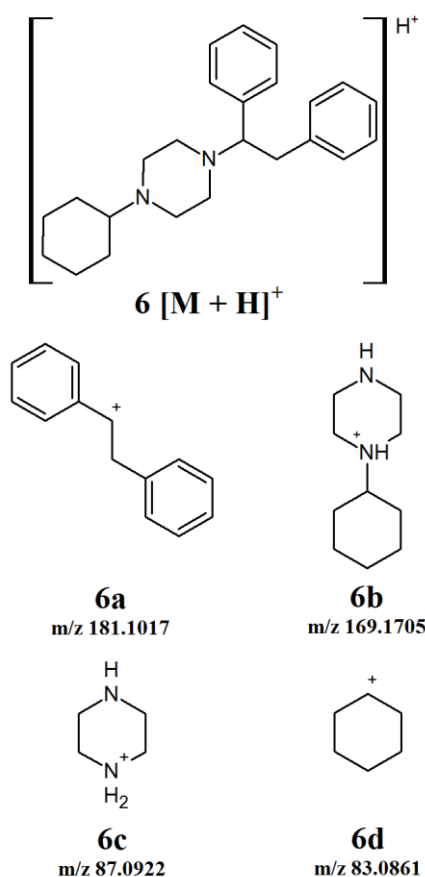


Figure 2-7: Proposed structures of fragments observed for MT-45, showing molecular ion (**6**) and product ions (**6a – 6d**)

2.4.2 Non-targeted Screening

In some ways, the development of new analytical approaches for the detection of NPS is somewhat of an arms race, with forensic toxicologists racing to keep up with the production of novel compounds. From this perspective, methods that take a more holistic approach to the detection of compounds structurally related to known classes of NPS have the distinct advantages over traditional targeted methods. These non-targeted approaches are required for the detection and tentative identification of novel analogues that have not been recorded previously ^[126]. Samples where these kinds of detections arise can then be flagged for further confirmatory analysis when appropriate CRMs are available. In this context, non-targeted screening refers to the methods used for data acquisition, namely DIA and DDA. When DIA is used, all precursor ions that reach the detector are submitted to CID. On the other hand, DDA involves the selection of a limited number of precursor ions for CID, based on a set of user-defined criteria, which is determined at the start of the experiment ^[126]. Generally, this data is compared to either in-house databases, which are populated with

information obtained from the previous analysis of CRMs, or large online databases for compound identification ^[126, 153, 213].

The diagnostic product ions identified in this work for each of the different opioid classes can then be applied to filter through this data to identify possible compounds of interest. A summary of the different diagnostic ions for each opioid class can be found in Table 2-1. This approach relies on the use of MS/MS (sometimes referred to as MS²) data, therefore the effectiveness of these data interpretation methods relies on the selection of an appropriate data acquisition technique ^[126]. Both DIA and DDA techniques can be used for this purpose, however DIA is less likely to result in the loss of data which might be relevant to the compounds of interest. Since DDA methods have a limited number of precursor ions selected for CID, it is possible that the compound/s of interest may not be subjected to CID, especially when there are a number of abundant endogenous compounds in the matrix being analysed. While the number of selected precursor ions can be increased, this can be limited by the scan rate and data processing capabilities of the instrument used ^[126].

Using simple data processing methods, namely EICs and NLF, the identified product ions can be used to detect the presence of different synthetic opioids in complex samples. These techniques can be applied manually, however many data processing software packages allow for the automation of these processes. NLF calculates the mass difference between the product ions identified and the precursor ion. The samples can then be interrogated for compounds that contain these specific mass losses. This technique can be a useful adjunct to EICs in situations where the product ions contain variable R groups, meaning that their *m/z* values will vary, but the neutral molecule which is lost has a consistent mass. For many of the synthetic opioids, this technique has limited value, since many of the neutral losses also contain R groups and, therefore, the exact mass loss will vary. A group where NLF may prove useful are the NPFs that contain side chains at the carbon-9 position of the piperidine ring (Figure 2-2). All of the common product ions identified for these compounds have R groups present, however, the proposed structure for fragment **1a** arises from the loss of the C9 side chain. In all the compounds analysed from this group, the side chain consisted of either an ether or ester group with masses of 59.0497 or 73.0290 Da, respectively. In this case, it may be useful to target these two neutral losses to indicate the presence of one of these compounds.

Table 2-1: Summary of diagnostic product ions for each subclass of synthetic opioids

Opium class	Identifier*	Molecular Formula	Theoretical m/z
Fentanyl (phenethyl tail and no C9 side chain)	2a	$[C_8H_9]^+$	105.0704
	2b	$[C_6H_7]^+$	79.0548
	2e	$[C_{13}H_{18}N]^+$	188.1439
AH	3a	$[C_{14}H_{16}NOR_3R_4]^+$	250.0999 (^{35}Cl) 252.0969 (^{37}Cl) 284.0609 ($^{35}Cl/^{35}Cl$) 286.0580 ($^{35}Cl/^{37}Cl$) 288.0550 ($^{37}Cl/^{37}Cl$)
	3b	$[C_8H_6NOR_3R_4]^+$	168.0216 (^{35}Cl) 170.0187 (^{37}Cl) 201.9827 ($^{35}Cl/^{35}Cl$) 203.9797 ($^{35}Cl/^{37}Cl$) 205.9768 ($^{37}Cl/^{37}Cl$)
	3c	$[C_7H_6NOR_3R_4]^+$	156.0216 (^{35}Cl) 158.0187 (^{37}Cl) 189.9827 ($^{35}Cl/^{35}Cl$) 191.9797 ($^{35}Cl/^{37}Cl$) 193.9768 ($^{37}Cl/^{37}Cl$)
	3d	$[C_7H_3OR_3R_4]^+$	138.9951 (^{35}Cl) 140.9921 (^{37}Cl) 172.9561 ($^{35}Cl/^{35}Cl$) 174.9532 ($^{35}Cl/^{37}Cl$) 176.9502 ($^{37}Cl/^{37}Cl$)
	3e	$[C_6H_3R_3R_4]^+$	111.0002 (^{35}Cl) 112.9972 (^{37}Cl) 144.9612 ($^{35}Cl/^{35}Cl$) 146.9582 ($^{35}Cl/^{37}Cl$) 148.9553 ($^{37}Cl/^{37}Cl$)
U (with methylene spacer)	4a	$[C_{15}H_{18}NOR_3R_4]^+$	298.0766 ($^{35}Cl/^{35}Cl$) 300.0736 ($^{35}Cl/^{37}Cl$) 302.0707 ($^{37}Cl/^{37}Cl$)
	4b	$[C_9H_{10}NOR_3R_4]^+$	218.0140 ($^{35}Cl/^{35}Cl$) 220.0110 ($^{35}Cl/^{37}Cl$) 222.0081 ($^{37}Cl/^{37}Cl$)
	4c	$[C_7H_5R_3R_4]^+$	158.9768 ($^{35}Cl/^{35}Cl$) 160.9739 ($^{35}Cl/^{37}Cl$) 162.9709 ($^{37}Cl/^{37}Cl$)
W	5b	$[C_6H_{11}N_2]^+$	111.0922
	5c	$[C_6H_4SO_2Cl]^+$	174.9621 (^{35}Cl) 176.9591 (^{37}Cl)

*Refer to Figures 2.3 – 2.6

It is important to note, however, that NLF can be more easily applied to data collected using DDA methods, rather than DIA. As reported by Oberacher and Arnhard ^[124], the use of DIA leads to the production of chimeric spectra where the software cannot associate product ions with their correct precursor ions and, therefore, neutral losses cannot be calculated. While deconvolution algorithms may be available to help sift through the data, DDA data does not require this extra processing step and, therefore, may benefit more from the use of NLF. These chimeric spectra do not limit the use of diagnostic product ions, however, since all the product ions created will still be present in the spectra. Since some product ions with consistent m/z values have been identified for the different groups, with the exception of the NPFs discussed previously, the use of EICs to target these diagnostic product ions may be the most effective way to detect their presence in unknown samples.

Figure 2-8 shows the EICs for the diagnostic product ions identified with m/z values of 105.0704 and 188.1439. These product ions were found to be common to NPFs containing a phenethyl tail and lacking any substitution at the C9 position of the piperidine ring (Figure 2-3). It can be seen from these EICs that there are two peaks at 5.5 and 5.85 min which indicate the presence of acetyl fentanyl **A2** and fentanyl **A1** (Figure 2-1), respectively. For the purpose of this proof-of-concept sample, the MS/MS spectra of the identified peaks were observed, and the retention times were compared to a targeted reporting method to confirm the identities of the compounds detected. As mentioned previously, a fragment with an m/z value of 105.0704 is quite common for alkyl-substituted benzenes, therefore, only identifying a peak in this EIC may not be specific to NPFs and requires further confirmation by the presence of other diagnostic product ions or interrogation of the overall MS/MS spectra.

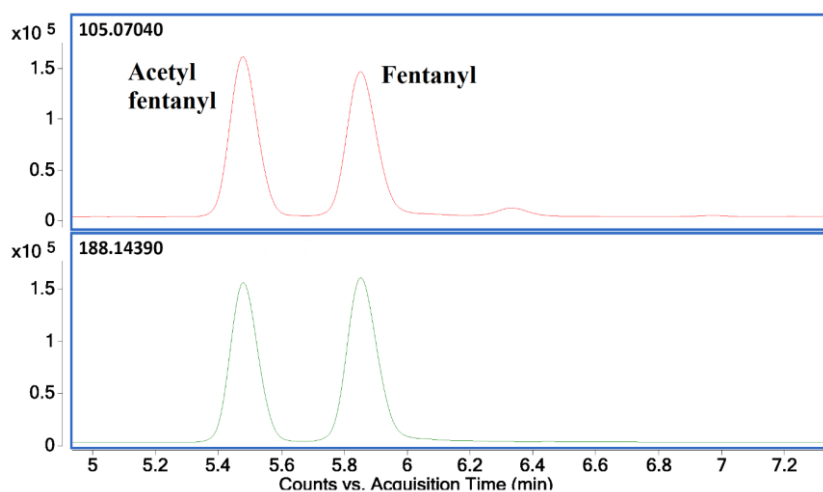


Figure 2-8: Extracted ion chromatograms for diagnostic product ions 105.0704 (**top**) and 188.1439 (**bottom**) showing the identification of acetyl fentanyl and fentanyl in 0.1 ng/mL mixed spike

Figure 2-9 shows additional EICs targeting the diagnostic product ions for the AH opioids (**3a–3e**) and U series opioids with a methylene spacer (**4a–4c**). For this sample, only the m/z values for fragments containing 2 Cl substituents were used. It can be seen that these EICs have indicated the presence of AH-7921 and U-50488 in the sample, with retention times of 5.85 and 6.2 min, respectively. While there are multiple possible m/z values for these fragments due to the different isotopes of Cl which may be present, the detection of any one of these masses may indicate the presence of one of these compounds in the sample being analysed.

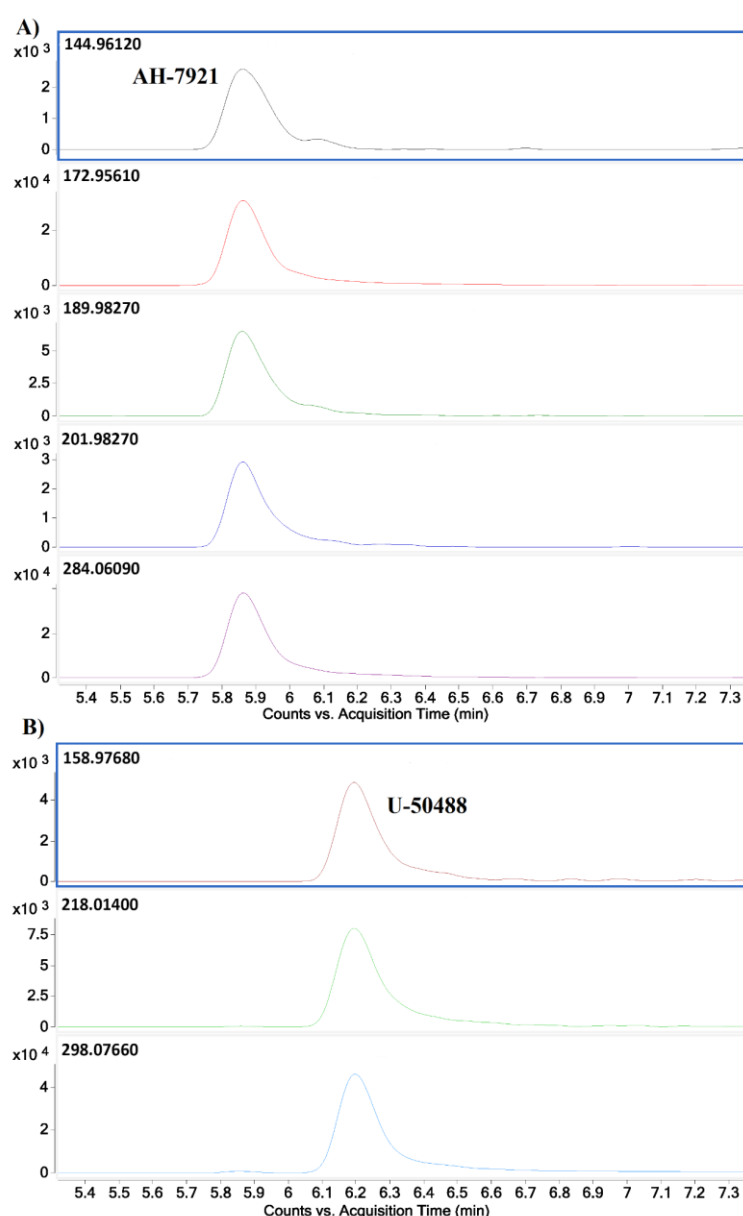


Figure 2-9: Extracted ion chromatograms for diagnostic product ions identified for the AH series opioids: 144.9612, 172.9561, 189.9827, 201.9827, and 284.0609 (A), and for the U series opioids with a methylene spacer: 158.9768, 218.0140 and 298.0766 (B), showing the identification of AH-7921 and U-50488 in an equine plasma sample spiked with a mixture of NPFs and NSOs at 0.1 ng/mL

The concentration range of the spikes was evaluated in order to determine an estimated limit of detection (LOD) for this technique in a relevant biological matrix. It was found that, down to a concentration of 0.05 ng/mL, the identified diagnostic fragments could be clearly seen in the spiked samples. At lower concentrations, some of the product ions no longer presented peaks in their EICs, indicating that their abundance was too low to accurately detect when considering a signal-to noise ratio (S/N) of ≥ 3 . While there were still some product ions that could be detected at lower concentrations, the confidence in the detection of a synthetic opioid would be reduced when fewer diagnostic product ions are identified in a given sample. In the case of fentanyl derivatives containing a phenylethyl tail (Figure 2-3), if the product ion at m/z 188.1439 (Figure 2-8, bottom) is not detected but the fragment at m/z 105.0704 can still be seen, it may cast doubts as to the identity of the detected compound. As previously stated, this fragment is common for any alkyl substituted aromatic compounds, therefore the specificity of identifying this fragment alone would be limited.

The use of EICs to target common product ions for the different groups of synthetic opioids can be useful for the development of a non-targeted screening method. This method does not rely on extensive spectral databases or CRMs to indicate the presence of these compounds. It should be noted that this technique cannot be relied on in isolation and should rather form part of a non-targeted screening workflow. The detection of peaks related to the identified diagnostic product ions can flag samples of interest, which may contain forensically relevant compounds. The MS/MS spectra of the detected peaks can be observed to identify the molecular ion of interest and further, more targeted, analysis can be conducted if necessary. The incorporation of techniques such as this into non-targeted screening workflows can save forensic laboratories time and resources by quickly flagging or eliminating samples and reducing the use of superfluous techniques. This method was applied to 157 race-day equine plasma samples taken during the 2019 Autumn Carnival. While these samples have all shown to be negative for synthetic opioids, it demonstrates the ability of this technique to be incorporated alongside routine analysis.

2.5 Conclusion

Various synthetic opioids, encompassing a number of different subclasses, were subjected to CID using different CEs. A number of different diagnostic product ions were identified for each of the classes analysed. These diagnostic product ions can be incorporated into non-targeted screening strategies to detect and tentatively classify these compounds without relying on comparison to databases or CRMs.

2.6 References

19. Lucyk, S.N. and L.S. Nelson, *Novel Synthetic Opioids: An Opioid Epidemic Within an Opioid Epidemic*. Ann. Emerg. Med., 2017. **69**(1): p. 91-93.
29. World Health Organisation. *Ensuring balance in national policies on controlled substances*. World Health Organisation;2011.
30. Ballantyne, J.C. and K.S. LaForge, *Opioid dependence and addiction during opioid treatment of chronic pain*. Pain, 2007. **129**(3): p. 235-255.
32. Holden, J.E., Y. Jeong, and J.M. Forrest, *The endogenous opioid system and clinical pain management*. AACN clinical issues, 2005. **16**(3): p. 291-301.
37. Breindahl, T., et al., *Identification of a new psychoactive substance in seized material: the synthetic opioid N-phenyl-N- 1-(2-phenethyl)piperidin-4-yl prop-2-enamide (Acrylfentanyl)*. Drug Test. Anal., 2017. **9**(3): p. 415-422.
38. Mohr, A.L.A., et al., *Analysis of Novel Synthetic Opioids U-47700, U-50488 and Furanyl Fentanyl by LC–MS/MS in Postmortem Casework*. J. Anal. Toxicol., 2016. **40**(9): p. 709-717.
60. Rojkiewicz, M., et al., *Identification and physicochemical characterization of 4-fluorobutyrfentanyl (1-((4-fluorophenyl)(1-phenethylpiperidin-4-yl)amino)butan-1-one, 4-FBF) in seized materials and post-mortem biological samples*. Drug Test. Anal., 2017. **9**(3): p. 405-414.
72. Fabregat-Safont, D., et al., *Updating the list of known opioids through identification and characterization of the new opioid derivative 3,4-dichloro-N-(2-(diethylamino) cyclohexyl)-N-methylbenzamide (U-49900)*. Scientific Reports, 2017. **7**: p. 14.
122. Pasin, D., et al., *Current applications of high-resolution mass spectrometry for the analysis of new psychoactive substances: a critical review*. Anal. Bioanal. Chem., 2017. **409**(25): p. 5821-5836.
123. Knolhoff, A.M. and T.R. Croley, *Non-targeted screening approaches for contaminants and adulterants in food using liquid chromatography hyphenated to high resolution mass spectrometry*. J. Chromatogr. A, 2016. **1428**: p. 86-96.
124. Oberacher, H. and K. Arnhard, *Current status of non-targeted liquid chromatography-tandem mass spectrometry in forensic toxicology*. TrAC, Trends Anal. Chem., 2016. **84**: p. 94-105.
125. Remane, D., D.K. Wissenbach, and F.T. Peters, *Recent advances of liquid chromatography–(tandem) mass spectrometry in clinical and forensic toxicology — An update*. Clin. Biochem., 2016. **49**(13): p. 1051-1071.
126. Pasin, D., et al., *Characterization of hallucinogenic phenethylamines using high-resolution mass spectrometry for non-targeted screening purposes*. Drug Test. Anal., 2017. **9**(10): p. 1620-1629.

136. Arnhard, K., et al., *Applying 'Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra' (SWATH) for systematic toxicological analysis with liquid chromatography-high-resolution tandem mass spectrometry*. Anal. Bioanal. Chem., 2015. **407**(2): p. 405-414.
143. Noble, C., et al., *Application of a screening method for fentanyl and its analogues using UHPLC-QTOF-MS with data-independent acquisition (DIA) in MSE mode and retrospective analysis of authentic forensic blood samples*. Drug Test. Anal., 2017. **10**(4): p. 651-662.
153. Kinyua, J., et al., *A data-independent acquisition workflow for qualitative screening of new psychoactive substances in biological samples*. Anal. Bioanal. Chem., 2015. **407**(29): p. 8773-8785.
205. United Nations Office on Drugs and Crime. *World Drug Report 2018*. Vienna: United Nations Office on Drugs and Crime; June 2018 2018.
206. Ohta, H., S. Suzuki, and K. Ogasawara, *Studies on Fentanyl and Related Compounds IV. Chromatographic and Spectrometric Discrimination of Fentanyl and its Derivatives*. J. Anal. Toxicol., 1999. **23**(4): p. 280-285.
207. Lurie, I.S. and R. Iio, *Use of multiple-reaction monitoring ratios for identifying incompletely resolved fentanyl homologs and analogs via ultra-high-pressure liquid chromatography-tandem mass spectrometry*. J. Chromatogr. A, 2009. **1216**(9): p. 1515-1519.
208. Rittgen, J., M. Pütz, and R. Zimmermann, *Identification of fentanyl derivatives at trace levels with nonaqueous capillary electrophoresis-electrospray-tandem mass spectrometry (MSⁿ, n = 2, 3): Analytical method and forensic applications*. Electrophoresis, 2012. **33**(11): p. 1595-1605.
209. Kronstrand, R., et al., *Fatal Intoxications Associated with the Designer Opioid AH-7921*. J. Anal. Toxicol., 2014. **38**(8): p. 599-604.
210. Patton, A.L., et al., *Quantitative Measurement of Acetyl Fentanyl and Acetyl Norfentanyl in Human Urine by LC-MS/MS*. Anal. Chem., 2014. **86**(3): p. 1760-1766.
211. Helander, A., et al., *Intoxications involving acrylfentanyl and other novel designer fentanyls – results from the Swedish STRIDA project*. Clin. Toxicol., 2017: p. 1-11.
212. Pavia, D.L., et al., *Introduction to Spectroscopy*. 5 ed. 2014: Cengage Learning.
213. Baz-Lomba, J.A., M.J. Reid, and K.V. Thomas, *Target and suspect screening of psychoactive substances in sewage-based samples by UHPLC-QTOF*. Anal. Chim. Acta, 2016. **914**: p. 81-90.

Chapter 3:
***Non-Targeted Analysis – Compound
Detection***

Chapter 3: Non-Targeted Analysis – Compound Detection

3.1 Rationale

With the perpetual introduction of new analogues to the illicit drug market, it is important to develop techniques that can screen for new compounds that are chemically and structurally related to known analogues. While there have been a number of papers published which focus on the detection and analysis of new synthetic opioids, the work presented in this chapter is significant for the fact that it focuses more on the ‘back-end’ data analysis and methods that can be utilised alongside traditional screening.

The majority of this chapter is taken from a first author publication entitled ‘Finding the Proverbial Needle: Non-targeted Screening of Synthetic Opioids in Equine Plasma’, accepted for publication in Drug Testing and Analysis. The experimental work, data analysis and preparation of the initial draft were performed by J. Klingberg, with manuscript edits provided by A. Cawley, R. Shimmon, C. Fouracre, D. Pasin and S. Fu.

Klingberg, J., et al., *Finding the Proverbial Needle: Non-targeted Screening of Synthetic Opioids in Equine Plasma*. Drug Test. Anal. 2020. DOI: 10.1002/dta.2893

In addition to the published work, investigation of a molecular feature extraction approach to non-targeted screening is included.

Finding the Proverbial Needle: Non-targeted Screening of Synthetic Opioids in Equine Plasma

Joshua Klingberg¹, Adam Cawley², Ronald Shimmon¹, Chris Fouracre³, Daniel Pasin¹, Shanlin Fu^{1*}

¹Centre for Forensic Science, University of Technology Sydney, Ultimo, NSW, Australia

²Australian Racing Forensic Laboratory, Racing NSW, Sydney, NSW, Australia

³Agilent Technologies, Mulgrave, VIC, Australia

*Corresponding Author:

Shanlin Fu

Shanlin.fu@uts.edu.au

3.2 Introduction

Opioids are compounds that produce analgesic effects through the binding to specific opioid receptors ^[31]. These compounds emulate the effects of endogenous opioid peptides, such as endorphins, enkephalins and dynorphins, which are present in the central and peripheral nervous systems as neurotransmitters and neuromodulators ^[32]. Fentanyl is probably the most well-known synthetic opioid and has been used as a medication for pain management since the 1960s ^[34]. Fentanyl and its analogues, however, have often been misused in place of heroin due to its lower cost ^[214], and there is an increasing trend of fentanyl and its analogues, and other synthetic opioids, being mixed into heroin or sold as prescription opioids ^[205]. Alongside the fentanyl analogues, there are a number of other NSOs emerging onto the illicit drug market. The most prolific of these NSOs are the AH and U series', which were originally developed as new opioid agonists, but were never brought to the market for human use ^[19]. Synthetic opioids have recently emerged as a drug class of particular concern due to the health issues caused by their incredibly high potency. The proliferation of fentanyl analogues, along with other NSOs, also poses a significant challenge for toxicologists, with these compounds accounting for 29% of newly identified NPS in 2017 ^[1].

In addition to the significant human health risks posed by these compounds, it has been found that opioids have a propensity for causing stimulant-like effects in horses, in conjunction with their conventional pain depressing properties ^[107, 215, 216]. This stimulant effect was found to be common to all narcotic analgesics ^[107], however it has been theorized that κ -opioid agonists should provide analgesia without stimulation ^[108]. The effect of the κ -agonist U-50488 on horses was evaluated and it was found that, while there was slight stimulation, it was markedly less than potent μ -agonists, such as fentanyl ^[108]. The stimulant effects have been confirmed to be the result of action at the opioid receptors, as administration of opioid receptor antagonists, such as naloxone, suppress the stimulation ^[107, 217]. This combination of analgesic and stimulant properties makes synthetic opioids a target of importance from an equine anti-doping perspective. Currently in Australia, there is only an established cut-off concentration for Butorphanol at 0.01 ng/mL in plasma. The presence of any other synthetic opioid, regardless of concentration, is prohibited in equine racing. In 2015, the New York Equine Drug Testing Program reported its first positive detection of AH-7921 in a post-race sample ^[109]. This shows that novel synthetic opioids have started to encroach into the racing industry and reinforces the importance of employing up-to-date screening techniques in anti-doping laboratories around the world.

The continual evolution of the illicit drug market necessitates the development of non-targeted screening strategies. While there have been numerous analytical methods developed to detect synthetic opioids [34, 218-223], these methods are often designed using MRM, which needs to be optimised using CRMs for the specific analytes included in the assay. There is often a considerable delay between early use of new substances and the availability of CRMs, therefore these strategies need to be able to function without relying on CRMs or library spectra [122]. Previous CID work has shown that different groups of synthetic opioids demonstrate certain class-specific fragmentation patterns [224], which provides an interesting avenue for the development of non-targeted screening strategies. Noble *et al.* [143] demonstrated that this approach shows promise, through the development of a data-independent screening strategy for 50 fentanyl analogues, taking advantage of these class-specific cleavages. Preliminary work has also been conducted to expand this strategy to the other classes of synthetic opioids [224]. DIA has an advantage over traditional DDA methods in that it can collect the product ions of all present precursor ions in a single scan, providing full scan MS/MS data [122, 143]. While this can lead to the production of ‘chimeric spectra’, where the software cannot associate a product ion to a particular precursor ion [124], the presence of characteristic product ions can be used to identify suspicious samples for further analysis.

In order to interrogate the large amount of data provided by HRMS, various filtering and extraction methods are required to highlight peaks of potential interest. The use of MDF was first reported by Grabenauer *et al.* who applied it to the analysis of synthetic cannabinoids in herbal products and were able to detect a compound that was not visible in the TIC [147]. The mass defect of a particular compound is defined as the difference between its exact mass and its nominal integer mass [148, 149]. Compound classes often have similar mass defects, or specific trends with increasing mass, and therefore MDF can be used to selectively filter out related compounds [148]. A sub-type of MDF, known as KMD, may also be beneficial for the development of non-targeted screening methods. The Kendrick mass scale allows for the recognition of a group of compounds that differ by a specific repeating mass unit [148, 157].

The use of KMD filtering has also been investigated for drug screening. Anstett *et al.* [158] explored the application of KMD filtering to the screening of different phenethylamine classes, including 2C-, aminopropylbenzofuran and 2,5-dimethoxy-*N*-(2-methoxybenzyl) phenethylamines. The authors successfully implemented KMD filters that were able to distinguish between compounds from the different classes of phenethylamines that were analysed [158]. If similar filters could be developed for the different classes of synthetic opioids, this technique could prove a valuable asset in the

development of potentially complementary non-targeted screening methods for the detection of these compounds.

This study presents a proof-of-concept for the development of a toolbox of complementary techniques demonstrating potential for the non-targeted screening of synthetic opioids in equine plasma, which may also be applicable to other biological matrices. These techniques focus on the ‘back end’ data processing and data mining techniques, which allows them to be used in conjunction with routine acquisition methods, without the need for laboratories to develop and validate new instrumental analyses.

3.3 Experimental

3.3.1 Solvents and Reagents

All solvents used were LC-MS grade. Acetonitrile, ethyl acetate and methanol were obtained from Merck (Darmstadt, Germany). Ammonium acetate and trichloroacetic acid were obtained from Sigma-Aldrich (Castle Hill, NSW, Australia). Acetic acid was obtained from Ajax Chemicals (Sydney, NSW Australia). Ultrapure-grade water (18.2 MΩ.cm) was obtained from a Smart2Pure ultrapure water system (Thermo Scientific, Langenselbold, Hungary).

Fentanyl citrate was purchased from Sigma-Aldrich (Castle Hill, NSW, Australia). Hydrochloride salts of acetyl fentanyl and AH-7921, along with a neat solid of U-50488, manufactured by Cayman Chemical (Ann Arbor, MI, USA), were purchased from Sapphire Bioscience (Redfern, NSW, Australia). Carfentanil citrate was purchased from Janssen Pharmaceuticals (North Ryde, NSW, Australia).

3.3.2 Sample Preparation

Blank plasma was obtained from blood samples collected in Lithium Heparin Vacutainers purchased from BD (Mississauga, ON, Canada) from thoroughbred horses following approval of the Racing NSW Animal Care and Ethics Committee (ARA 71).

Spiked equine plasma samples were prepared at varying concentrations to determine the effectiveness of the examined detection methods in an authentic biological matrix. A mixed standard containing fentanyl, acetyl fentanyl, carfentanil, AH-7921 and U-50488 was prepared at a

concentration of 200 ng/mL in methanol. The representative panel of compounds used in this study were chosen to demonstrate the potential of the developed techniques across the range of synthetic opioid classes. AH-7921 and U-50488 were chosen specifically due to their relevance to the racing industry, with AH-7921 being the first NSO reported in an equine plasma sample and U-50488 being previously studied to determine the effects of κ -opioid agonists in horses^[108]. The chemical structures of the different opioids used in this study can be found in Figure 3-1.

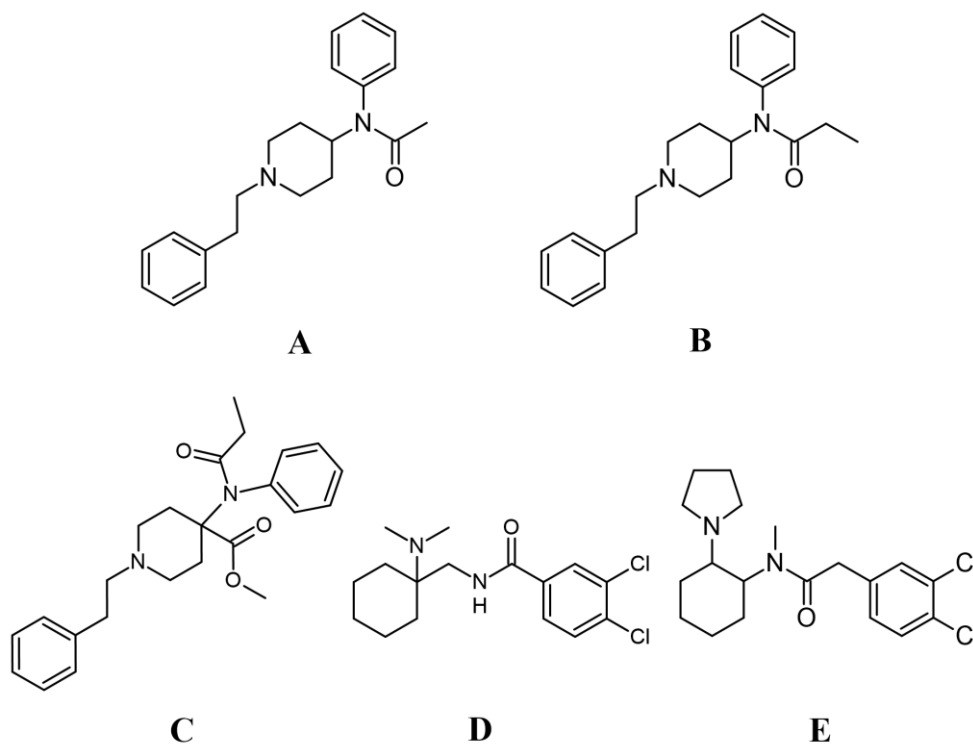


Figure 3-1: Chemical structures of opioids included in the study including acetyl fentanyl (**A**), fentanyl (**B**), carfentanyl (**C**), AH-7921 (**D**) and U-50488 (**E**)

Two millilitres of blank equine plasma was spiked with the mixed standard to produce a concentration range of 0.01 to 10 ng/mL across 13 different samples. This concentration range was used to estimate LODs and screening cut-off levels for the developed techniques. Blank samples were also prepared alongside the spiked samples.

Once an effective screening cut-off level was estimated for each technique, further replicates were prepared to verify the consistency of these cut-offs at concentrations of 0.05 and 0.1 ng/mL. For each concentration, three sets of samples were prepared from a pooled blank plasma source, with seven replicates in each set ($n = 21$). The sample sets were analysed over different days to account for intra-day variability.

Protein precipitation was performed through the addition of 200 μ L of 10% (v/v) trichloroacetic acid to each of the samples. The pH of the samples was then adjusted to 3 – 3.5 using hydrochloric acid and 3 – 3.5 mL of ultrapure water after which they were centrifuged at 1500 x *g* for 10 minutes. SPE was completed using XtrackT[®] Gravity Flow DAU Extraction Columns (3 mL column, 200 mg sorbent, UCT Inc., Bristol, USA). The cartridges were first conditioned with 3 mL of methanol, followed by 3 mL of ultrapure water, after which the samples were loaded (~6 mL). The samples were washed with 3 mL of 0.1 M acetic acid and dried using a positive pressure manifold at 70 psi (UCT Inc., Bristol, PA, USA). The cartridges were washed with 3 mL of methanol and dried under positive pressure. The analytes were eluted from the cartridges using 3 mL of ethyl acetate containing 3% ammonia and 0.5% methanol.

Following SPE, one drop of 0.1 M methanolic hydrochloric acid was added to each of the eluents using a Pasteur pipette (approximately 20 μ L) before the solvent was evaporated under a gentle stream of N₂ at 60 °C. The samples were then reconstituted in one drop of methanol from a Pasteur pipette and 100 μ L of 10 mM ammonium acetate buffer (pH 3.9), before being transferred to vials for LC-HRMS analysis. All samples were stored at 4 °C until analysis.

3.3.3 Instrumental Analysis

Chromatographic separation was achieved on an Agilent Technologies (Santa Clara, CA, USA) 1290 Infinity II UHPLC, consisting of a high-speed pump (G7120A), multisampler (G7167B, 18 °C) and thermostat and column compartment (G1316A, 35 °C) coupled to an Agilent Technologies 6545 QTOF-MS. All data acquisition was performed using Agilent Technologies MassHunter Workstation (Version B.06.01). A sample volume of 1 μ L was injected onto a Phenomenex (Torrance, CA, USA) Gemini 110 Å C₁₈ LC column (2 x 50 mm, 5 μ m particle size) using a gradient elution method with a flow rate of 0.5 mL/min and a total analysis time of 11.5 min. Mobile phase A consisted of a 10 mM ammonium acetate buffer (pH 9) and mobile phase B consisted of 0.1% (v/v) acetic acid in acetonitrile. Initial mobile phase composition was 99% A, which was held for 2 min before being decreased linearly to 20% A over 6.5 min. The mobile phase was then returned to 99% A over 3 min.

The QTOF-MS was operated in ESI+ mode with capillary and fragmentor voltages of 3500 V and 100 V, respectively. An AllionsMS/MS (DIA) data acquisition mode was used with a *m/z* range of 35 – 1000 for both MS and MS/MS spectra. Spectra were obtained with an acquisition speed of 10 spectra/s and CEs of 10, 20 and 40 eV for CID. The chromatographic and MS conditions were chosen

to mimic the routine analytical method employed at the Australian Racing Forensic Laboratory (ARFL) to evaluate the effectiveness of the developed detection techniques for an operational purpose.

3.3.4 Data Analysis

All EICs for the PIS and KMD analysis methods were generated using Agilent Technologies MassHunter Qualitative Analysis Software (Version B.10.0, Build 10.0.10305.0).

3.3.4.1 Targeted Compound Extraction Limit of Detection

The suitability of the analytical method was assessed by performing a targeted extraction on the spiked compounds using MassHunter Quantitative Analysis Software (Version 10.0, Build 10.0.707.0). Spiked samples of decreasing concentrations (10 to 0.01 ng/mL) were evaluated to estimate LODs using the targeted extraction method. Precursor and product ion EICs were assessed within a mass error of ± 10 ppm and a retention time tolerance of ± 0.2 min of their known retention times ^[225]. Where an EIC was assessed for a targeted compound, the S/N was obtained to estimate the LOD where the $S/N > 3$ ^[226]. These estimated values were used for qualitative comparison to the proposed non-targeted screening methods to demonstrate the effectiveness of these methods.

3.3.4.2 Product Ion Searching (PIS)

For the PIS technique, EICs were created from several diagnostic product ions identified in a previous CID study ^[224] to screen for potential synthetic opioids present in the samples within previously observed subclasses. The diagnostic product ions used for each compound class can be found in

Table 3-1. The masses used represent the theoretical exact masses of each diagnostic product ion. All EICs were generated with a mass tolerance of ± 10 ppm. For the purpose of the developed screening method, this mass tolerance was chosen as an arbitrary acceptable cut-off to ensure that no possible compounds were missed during the data analysis. Once a sample has been flagged as potentially containing a compound of interest, further targeted confirmatory tests should be employed as appropriate with more stringent mass tolerance criteria. Product ion masses were extracted at the CE (10, 20 or 40 eV) which gave the greatest peak intensity. This CE was determined automatically by the data analysis software.

Table 3-1: Diagnostic product ions from each compound class used for screening ^[224]

Compound Class	Diagnostic Product Ions (<i>m/z</i>)		
	1	2	3
Fentanyl	105.0704	188.1439	-
AH Series	172.9561	189.9827	284.0609
U Series (with a methylene spacer)	158.9768	218.0140	298.0766

3.3.4.3 Kendrick Mass Defect

When calculating KMD, a normalisation factor is used to ‘correct’ the exact mass of a particular compound and give the Kendrick mass. Most commonly, the mass of the CH₂ group is used as the correction factor. Essentially, the mass of a CH₂ group is set to be exactly 14.00000, allowing the Kendrick mass of a compound to be calculated (Equation 3.1) ^[157].

$$\text{Kendrick mass} = \text{IUPAC mass} \times \frac{14}{14.01565} \quad (3.1)$$

In this study, the KMD values for each compound were calculated as the difference between the Kendrick exact mass and the integer Kendrick mass ^[148]. For example, the Kendrick mass of fentanyl is 336.8509, in this case the KMD is 0.8509, as the integer Kendrick mass is 336.

KMD analysis was conducted using a custom-built program developed by Pasin in the Visual Basic for Applications (VBA) environment for Microsoft Excel (named *DefectDetect*) ^[227]. An averaged mass spectrum over a retention time range (0.5 – 8 min) was created for each sample and an *m/z* and intensity list was generated as a Microsoft Excel comma-separated value (.csv) file from these spectra. This retention time range was chosen as it is the range within which the analytes of interest were expected to elute, based on the representative compounds analysed and the chromatographic conditions. The program then calculated the Kendrick masses and mass defects using a normalisation factor of CH₂ and filtered the results based on a mass range of *m/z* 200 – 500, a KMD tolerance of ± 0.005 Da and mass defect filter of 0.1 – 0.27 Da. Table 4-1 shows the different KMD filters that were used in this study. Examples of the compounds belonging to the different fentanyl groups can be found in Appendix 3 (Figure A3-1). A minimum *m/z* value was populated for each filter and the program only considered masses that varied by a multiple of 14 Da from the minimum *m/z* to account for the difference of CH₂ units. An example of the *DefectDetect* interface can be found in Appendix 3 (Figure A3-2).

Table 3-2: Kendrick mass defect (KMD) filters used for compound detection

Name	KMD	Minimum m/z
Fentanyl (Amide Chain)	0.8510	323
Fentanyl (Cyclic)	0.8380	335
Fentanyl (F Analogues)	0.8215	355
Carfentanil	0.7916	395
AH + U Series	0.7510	329
AH Series	0.8276	295

3.3.4.4 Non-targeted Workflow

For each sample, EICs were extracted for the common product ions identified in

Table 3-1, followed by the extraction of masses identified by the KMD analysis. The generated EICs from each individual technique were first reviewed independently. If a peak of interest was identified by either technique, a comparison was made to the results obtained from the complementary technique to verify concurrence.

3.3.5 Blind Trial Setup

Blind trials were performed to validate the effectiveness of the detection methods under pseudo-operational conditions. Single compound spikes were prepared for fentanyl and acetyl fentanyl at concentrations of 0.05 ng/mL and 0.5 ng/mL ($n = 4$). Routine plasma samples from the ARFL were included to make a batch containing 20 samples. The 'Randbetween' function in Excel was used to randomise the positioning of the spiked samples within the worklist, and this randomisation was completed by a third party to ensure that the analyst was unaware of which samples were spiked. All samples in the worklist were then analysed using the instrument parameters outlined in 3.3.3 and the data was processed using the PIS and KMD analysis methods detailed above. These blind trials were conducted on two separate occasions.

3.4 Results and Discussion

3.4.1 Implementation of Non-targeted Workflow

Once the data from each sample has been processed using the developed techniques, the results from each technique can first be reviewed independently to determine the presence of any possible analytes of interest. The results can then be compared with the complementary technique to provide further evidence of detection. Observation of retention time-matched peaks in MS^1 (KMD) and MS^2 (PIS) may provide strong evidence for the presence of a particular subclass of synthetic opioid. Observation of retention time-matched peaks in MS^2 that is not corroborated by an extracted MS^1 peak may indicate the presence of a synthetic opioid containing the same core structure with novel modifications outside of the determined KMD filters. By implementing screening workflows that contain complimentary techniques, it increases confidence that one technique identifies a feature of interest and provides stronger evidence for detection when both techniques provide corroborative results.

3.4.2 Application of Screening Techniques

The applicability of each developed technique to non-targeted screening was evaluated to ensure that they were fit-for-purpose. The sensitivity of the techniques to the presence of a spiked compound in equine plasma was evaluated to propose an appropriate screening cut-off level. At concentrations below this proposed cut-off, it may be possible to detect an analyte of interest, however the reliability of detection would be reduced. Additionally, the specificity of these techniques was considered to ensure that a sample containing an analyte of interest could easily be distinguished from blank plasma samples.

3.4.2.1 *Product Ion Searching*

In a general screening workflow, detection and putative identification are performed by comparison to a spectral library. In some cases, where the fragmentation patterns of novel compounds are sufficiently similar to known compounds (i.e., where the mass spectrum contains a large number of common product ions), these techniques may be used to detect new analogues. When increased structural diversity is present within groups of analogues, or only a limited number of common product ions are present, simply screening samples against spectral libraries may not be sufficient. To improve the situation, the applicability of PIS for the detection of synthetic opioids in plasma has been demonstrated previously ^[224].

Figure 3-2 displays EICs from the targeted extraction of selected synthetic opioid precursor ions in a spiked plasma sample (0.05 ng/mL), as well as EICs for the common product ions of each opioid subclass in both the spiked plasma sample and a matrix blank. It can be seen from the spiked sample that there are distinct retention time aligned peaks for each set of common product and precursor ions, indicating the presence of an analyte of interest belonging to each of these subclasses. In the matrix blank, however, there are no clear peaks detected, which demonstrates specificity for this technique. The matrix blank displayed for EICs corresponding to the fentanyl analogues showed higher noise compared to the respective EICs for the other two subclasses. This is most likely due to the fact that these masses, belonging to the AH and U series opioids, have mass defects that are indicative of molecules containing halogens, which are less likely to be present in the matrix background than the masses of the common fentanyl product ions. In this study, all EICs were extracted within a mass error of ± 10 ppm. Given that the masses being extracted for the PIS method correspond to the theoretical masses of each ion, this tolerance can be altered on a case-by-case basis depending on the mass accuracy required for specific assays.

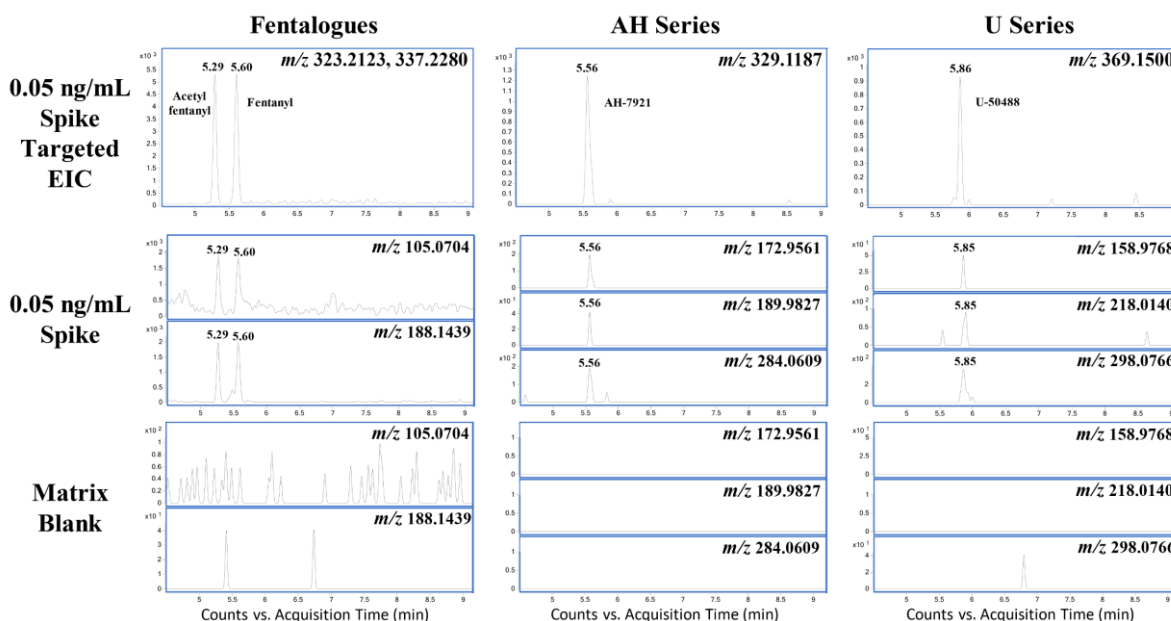


Figure 3-2: EICs showing the targeted extraction of known precursor masses showing the detection of the spiked compounds in a 0.05 ng/mL spike. The common product ions used for the PIS screening also showed detection of the samples peaks and could be easily distinguished from the matrix blank, showing the specificity of the technique.

From a routine screening perspective, the EICs generated from the common product ions can be combined (i.e. a single chromatographic trace that represents multiple m/z values) in MassHunter to reduce the amount of data that the analyst must review. If a peak is detected in a sample however, separate EICs should be created for that sample to ensure that multiple characteristic ions are present. In some instances, a single, high abundance ion can cause a peak to be displayed in the combined EIC. Without the presence of other diagnostic product ions, it cannot be confidently stated whether that ion is present from a compound of interest or if it is extraneous from an unrelated background compound. Alternatively, overlaid EICs for the individual product ions can be viewed for the same purpose. Figure 3-3 shows an example of such a case, where the overlaid EIC for common product ions for the U series compounds containing a methylene spacer show detection of all 3 ions, with decreasing intensities, within the same peak (left). On the other hand, Figure 3-3b shows a large peak at m/z 158.9768, while the EICs at m/z 218.0140 and 298.0766 only show minor peaks at significantly lower abundances barely noticeable on the overlaid chromatogram, which do not have the same retention time and, therefore, cannot be attributed to the same compound. In this case, the lack of corroborating ions means that the peak at m/z 158.9768 cannot be confidently attributed to a U series compound, as there may be a background compound present in the equine plasma with this mass.

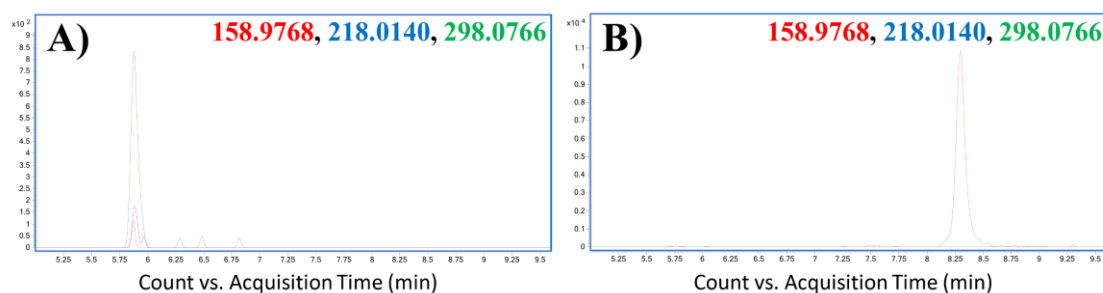


Figure 3-3: Example overlaid EICs for common U series product ions m/z 158.9768 (red), 218.0140 (blue) and 298.0766 (green) from a spiked plasma sample with 0.05 ng/mL U-50488 (A) and a blank sample (B)

The screening cut-off of this technique was found to be 0.05 ng/mL ^[224]. This cut-off level was determined by the concentration at which the common product ions for the majority of analytes in a particular class had a $S/N > 3$. Below this concentration, some product ions may still be detected in a sample, which can give some indication that an opioid may be present. Some product ions, however, were not observed. If the EICs generated from a sample only reveal a single product ion, then a lower confidence in the detection occurs. Some of the product ions that were found to be common to the different subclasses of synthetic opioids are not specific to these compounds alone. For example, the fentanyl product ion at m/z 105.0704 will likely be present in any compound containing a phenylethyl group. While this ion constitutes a good marker for fentanyl analogues in combination with other common product ions, its presence alone would not necessarily be indicative of these compounds. Therefore, care should be taken when interpreting results of samples that may have concentrations below this cut-off, where some common product ions are not detected.

3.4.2.2 Kendrick Mass Defect Analysis

KMD analysis using the *DefectDetect* program operates using an imported mass list containing m/z and intensity values generated from an averaged mass spectrum over a given retention time range. The program will then calculate the Kendrick mass and KMD of each identified m/z value and filter the results based on the user-defined parameters. An example of the results table can be found in Appendix 3 (Figure A3-3). Once the filtered list has been generated, EICs can be created for the selected masses to determine if there is a peak present or if the mass has arisen from background noise. It is possible to include intensity filters in the *DefectDetect* parameters, however it was decided that the absolute intensity can be variable depending on the instrument and analytical method used and, as such, this filter was not included in the analysis.

Figure 3-4 shows example EICs for some of the filtered masses detected by KMD analysis of a plasma sample spiked with 0.1 ng/mL of a mixed standard, as well as the results table generated for this sample. The results tables generated by the *DefectDetect* program are colour coded to correspond to the individual filters used in the analysis. It can be seen in these chromatograms that the masses correlating to the spiked samples show well-defined peaks at relatively high abundances, in the order of 10^3 and 10^4 . On the other hand, the EICs at m/z 323.1914 and 335.2181 have much lower abundances and significantly more noise along the baseline, indicating that they could be present due to the matrix background. Additionally, Figure 3-4 shows the top six masses identified by performing KMD analysis on a blank matrix sample. Once again, the EICs generated are noisy and have quite a low abundance, suggesting that there is no meaningful detection of an analyte of interest. This demonstrates the specificity of KMD analysis for analytes of interest, as they can be easily differentiated from background noise.

A number of different KMD tolerances were trialled in order to determine the best balance between reliability of detection and total number of filtered results. It is important to note that this technique operates using a list of averaged masses for each sample, therefore the accuracy of the calculated KMD values is dependent on the accuracy of the instrument used. In this case, the tolerance used for routine KMD analysis may require adjustment to suit instrument capabilities. For this study, it was determined that a tolerance of ± 0.005 Da was fit for purpose. When optimising the parameters for this detection technique, a reliability of $\geq 95\%$ for each spiked compound at a concentration of 0.1 ng/mL was deemed to be a suitable cut-off. The reliability of detection was evaluated by analysing three sets of spiked samples, each containing seven replicates ($n = 21$), and ensuring that the spiked compounds were being detected by the KMD filters applied. In addition, an arbitrary threshold of 20

filtered results was decided upon to limit the amount of data that needed review. Where a sample returned more than 20 filtered results after KMD analysis, the peak abundance was used to determine which masses were reviewed. It is important to note that, for the purpose of this proof-of-concept study, the retention times of the peaks identified by the KMD analysis were compared to the known retention times of the spiked compounds (within a tolerance of ± 0.2 min ^[225]) to ensure that the suspected compound was being correctly identified.

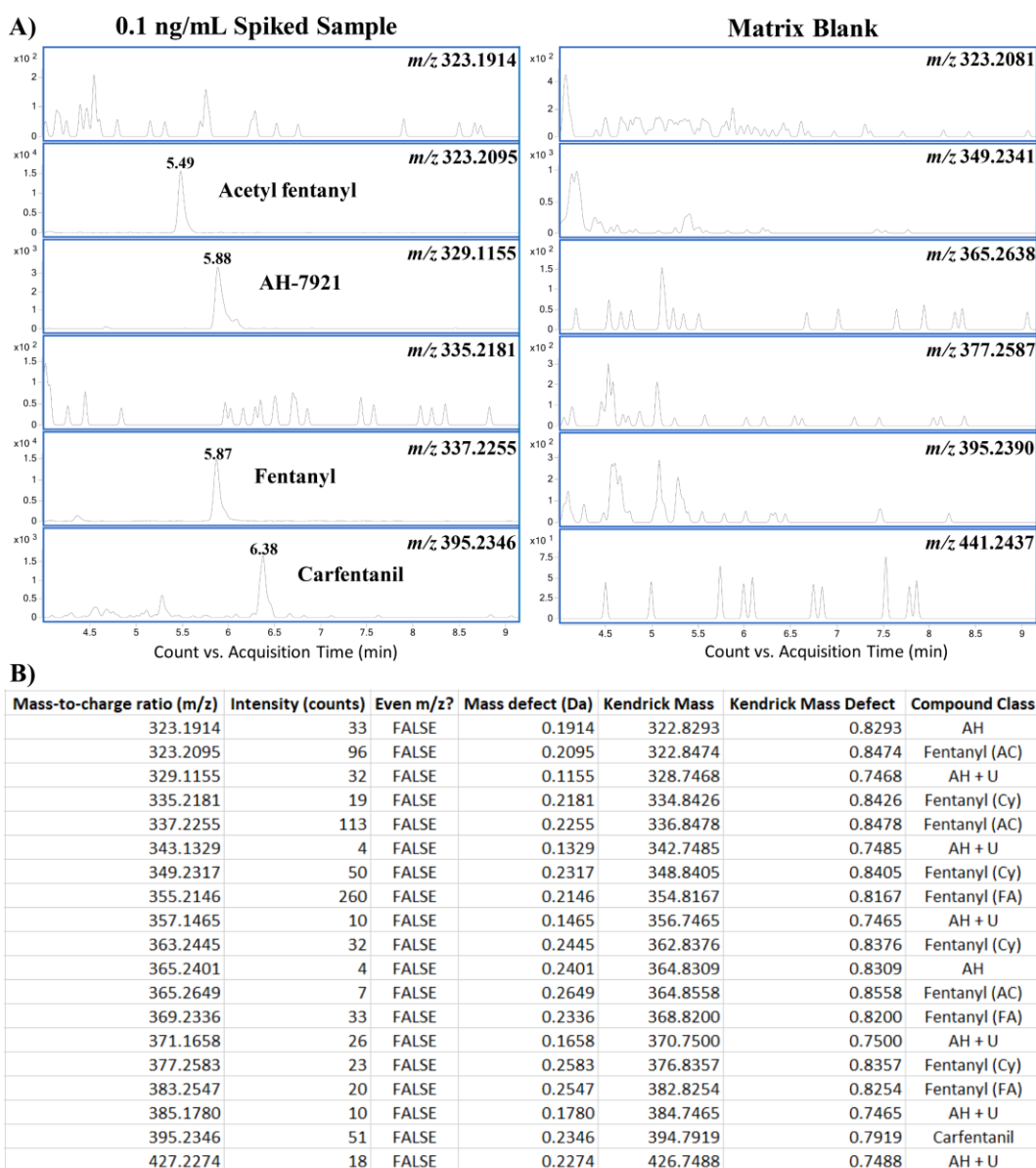


Figure 3-4: EICs for some of the identified masses from DefectDetect showing selectivity to the spiked compounds in comparison to the blank (A) and the results table given by DefectDetect for the spiked sample (B). The detected class for each mass in the results table is given according to the filters listed in Table 3-2. AC: Amide Chain, Cy: Cyclic, FA: F analogues

In addition to the KMD filters, an MDF can be applied to further reduce the number of filtered results. The mass defect range for a large number of synthetic opioids fell within a range of 0.10 – 0.27 Da, allowing some variation either side to account for mass tolerance and different structures. This range is broad, and by itself would not mean much. When it was applied on top of the other filters that were already in place, however, it significantly reduced the number of results that were obtained for each sample (50 – 63% reduction), while not affecting the reliability of the detection of the spiked compounds. Table 3-3 provides a summary of one set of the spiked samples showing the reliability and variation in the number of filtered results obtained. These results demonstrate that KMD analysis can reliably detect the presence of an analyte of interest, while still returning a reasonable number of results that need to be reviewed by generating EICs in an operational environment. Across all the 21 samples analysed, the spiked compounds within the defined KMD filters were detected in every sample.

Table 3-3: Kendrick mass defect (KMD) analysis of one set of 0.1 ng/mL samples, showing consistent detection of analytes of interest and total number of filtered results for each sample

Sample	Acetyl fentanyl	Fentanyl	AH-7921	Carfentanil	Filtered Results
A	✓	✓	✓	✓	11
B	✓	✓	✓	✓	15
C	✓	✓	✓	✓	13
D	✓	✓	✓	✓	13
E	✓	✓	✓	✓	14
F	✓	✓	✓	✓	16
G	✓	✓	✓	✓	12

3.4.3 Blind Trial

In order to evaluate the effectiveness of these techniques under more realistic conditions, the blind trial was completed on two different occasions using the PIS and KMD analysis methods to provide complementary detection. On both of these occasions, the higher concentration spikes (0.5 ng/mL) were quite clearly detectable within the dataset. Figure 3-5 shows the PIS results for the acetyl fentanyl spikes, along with a negative sample and the matrix blank.

There are distinct peaks present at relatively high abundances for both the m/z 105.0704 and 188.1439 product ions expected from a fentanyl derivative. Importantly, the peaks present in the

0.5 ng/mL spike can be clearly differentiated from both the matrix blank and the negative sample presented in Figure 3-5. While these peaks could not specifically identify acetyl fentanyl in a routine sample, they could be used to identify the sample as possibly containing a fentanyl derivative for further putative identification.

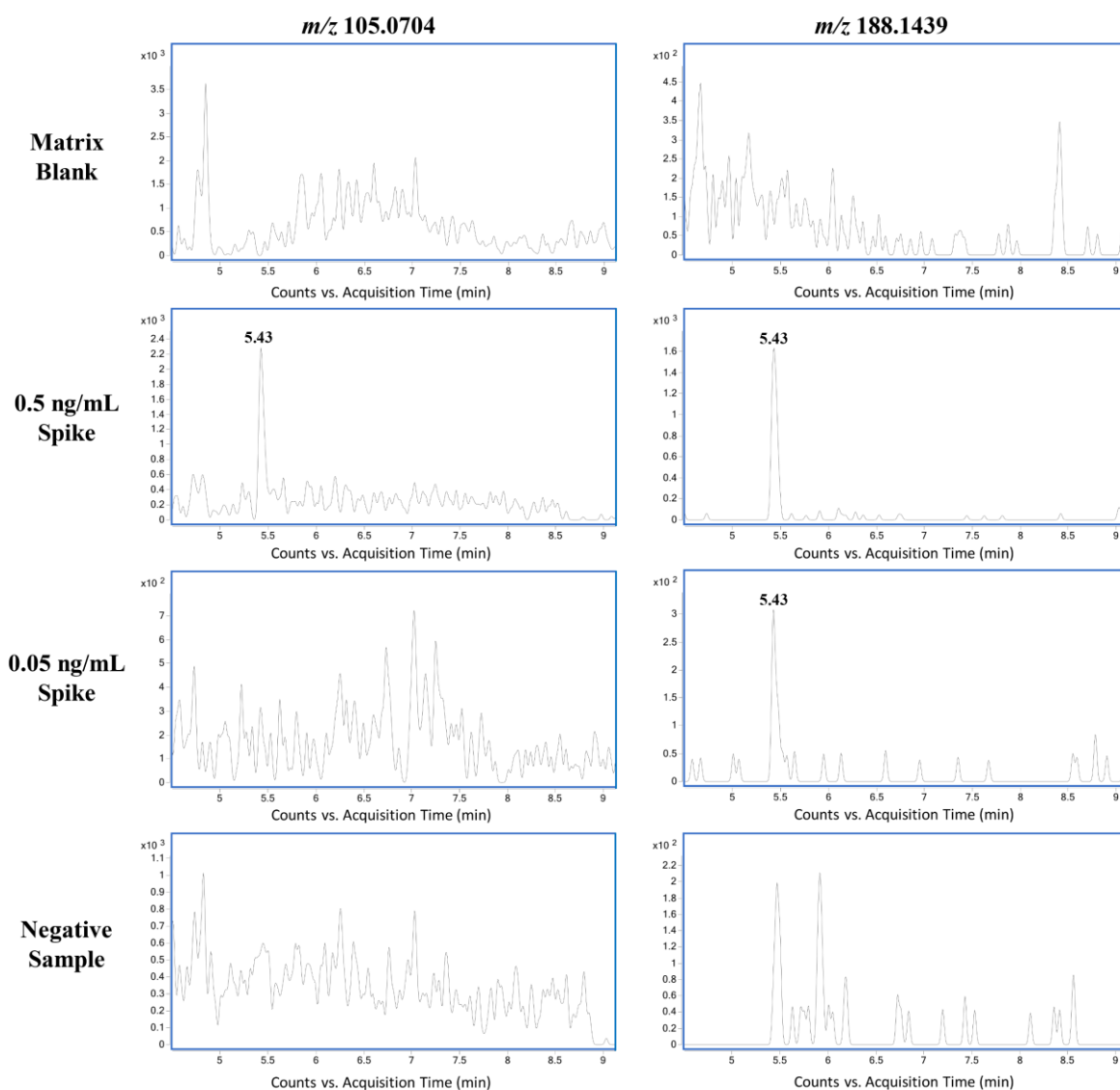


Figure 3-5: PIS results obtained from the blind trail showing the detection of both monitored ions in the 0.5 ng/mL spike and only detection of the m/z 188.1439 ion in the 0.05 ng/mL spike

In the 0.05 ng/mL acetyl fentanyl spike (Figure 3-5), one of the common product ions at m/z 105.0704 was not detected in the sample. Given that the m/z 105.0704 ion is common for organic molecules containing a phenethyl moiety, it is reasonable to suggest that this ion may be more commonly present in the background of biological matrices. This is further demonstrated by the fact that EICs generated for m/z 105.0704 tend to have a noisier baseline, as seen in Figure 3-2 and Figure 3-5. At

lower concentrations, therefore, it is possible that the presence of this common fentanyl product ion is masked by the background evident from the sample matrix. This can affect the confidence with which the presence of a compound of interest can be indicated. In this blind trial, there were no peaks identified in the negative samples that would suggest that they might have contained one of the compounds of interest, however it is possible for false positives to arise due to the variability of biological matrices. This again highlights the benefit of using complementary techniques. In addition to one technique being able to detect compounds that are missed by the other, the results from one detection technique can be used to corroborate the findings from another.

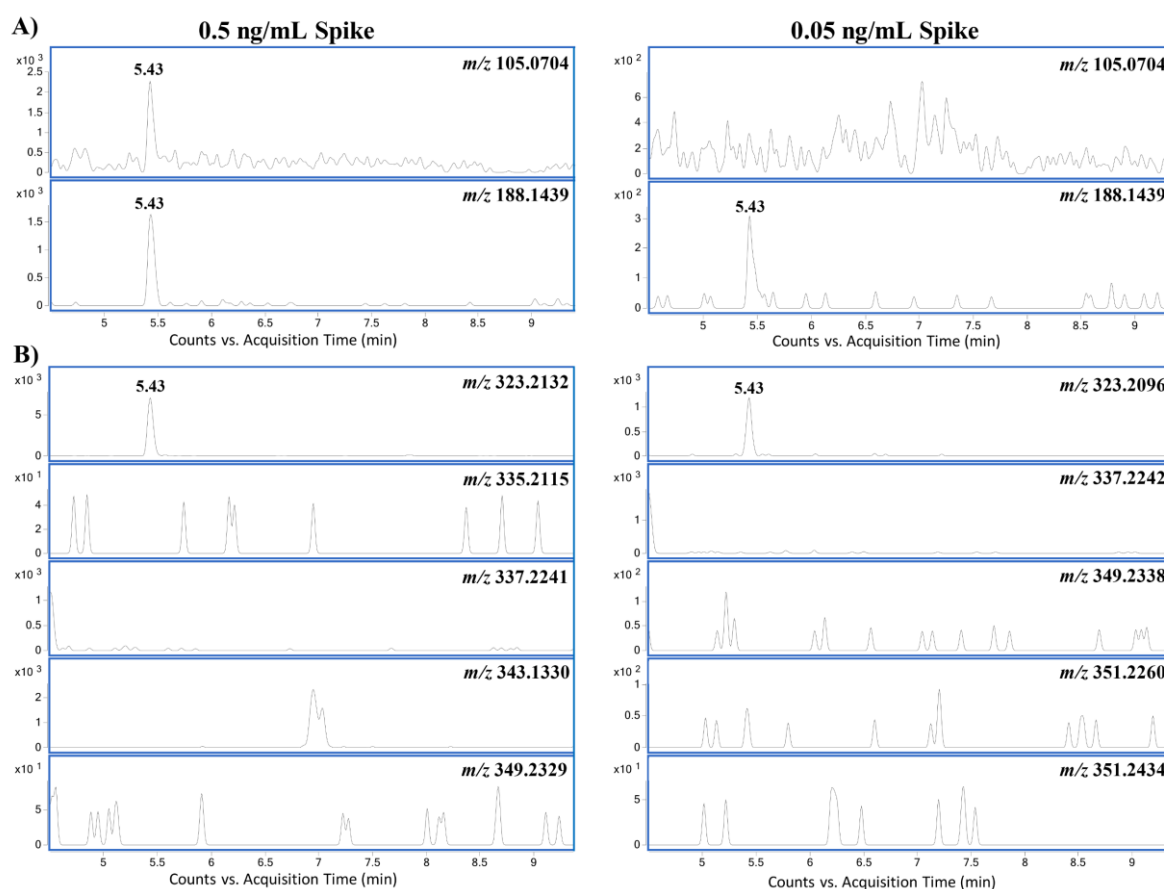


Figure 3-6: PIS results for the spiked acetyl fentanyl standards (A) compared with the associated KMD results for each sample (B)

Figure 3-6 presents the KMD results for each of the spiked acetyl fentanyl standards compared to their PIS results. In both spiked samples, the KMD analysis identified a peak with comparable retention time at the expected molecular ion mass of acetyl fentanyl (within a mass error of ± 10 ppm). In this blind trial, the molecular ion for the compounds of interest fell within the list of 20 filtered results proposed for the KMD technique, as seen in Figure 3-6 above. From these results, it

can also be seen how the other masses identified by the KMD analysis can be eliminated by reviewing their EICs and comparing with the PIS results. While the detection of the common product ions at m/z 105.0704 and 188.1439 cannot indicate the specific identity of the compound detected, the supporting KMD results identify the mass of the suspected molecular ion, which can be used for further putative identification work.

Interestingly, in the lower concentration (0.05 ng/mL) samples used for the blind trial, the KMD analysis was able to show the acetyl fentanyl peak even though this sample fell below the proposed screening cut-off. A cut-off of 0.1 ng/mL was proposed as this is the concentration where the reliability of detection with this technique was > 95%. It is, of course, possible that detections may be made below this concentration, as demonstrated in this blind trial, however, the reliability of identifying all samples with compounds of interest at lower concentrations would be affected. The spiked fentanyl samples at both concentrations (0.05 and 0.5 ng/mL) were also identified among the dataset and it was found that they followed the same trends highlighted by the acetyl fentanyl results presented (Figures A3-4 and A3-5).

While the higher concentration samples (0.5 ng/mL) were clearly identified with corroborative evidence from both complementary techniques, the lower concentration samples (0.05 ng/mL) lacked the identification of the common product ion at m/z 105.0704. In practice, this can be used to create 'confidence levels' when it comes to the detection of compounds of interest. When all components of a detection workflow are in agreement, as seen in Figure 3-6 for the 0.5 ng/mL spike, a high degree of confidence can be placed in the detection of a given compound. On the other hand, when no indication of an exogenous substance is provided by either technique, it can be said that the compounds of interest are not present. A compromise position can also exist, where some components of the detection workflow indicate the presence of a compound, but others do not, as seen in Figure 3-6 for the 0.05 ng/mL spike. In these cases, laboratories can set out a standard operating procedure about what samples should be subject to further analysis.

3.4.4 Comparison of Techniques

The major limitation of both techniques is the structural diversity of the opioids themselves. While the identified diagnostic product ions cover a large majority of the known structures of synthetic opioids, it is always possible that new structures will arise that will not lead to the formation of these ions. As has been demonstrated previously ^[224], certain fentanyl analogues that contain additional substitutions on their piperidine ring, such as carfentanil, do not give rise to the same diagnostic

product ions as the majority of fentanyl analogues. Carfentanil, in particular, presents the common product ion at m/z 105.0704 similar to other fentanyl derivatives (Figure 3-7), albeit at a lower abundance. This compound, however, does not display the other common product ion at m/z 188.1439. This means that screening using the proposed PIS method may not identify samples where this compound is present. It is possible to include further diagnostic product ions that are more specific to different compounds, however this begins to create a large list of possible masses for extraction, comparable to targeting a wide array of theoretical opioid analogue precursor m/z values, which introduces difficulties for high-throughput screening. The benefit of using a smaller number of product ions to encompass a larger range of compounds is that there is much less data for the analyst to review, making the screening easier from an operational perspective. While there are certainly limitations to the PIS technique, the broad array of compounds that it does encompass, along with the opportunity for a similar technique to be applied to different compound classes ^[126], demonstrates the applicability of this approach to a non-targeted screening workflow.

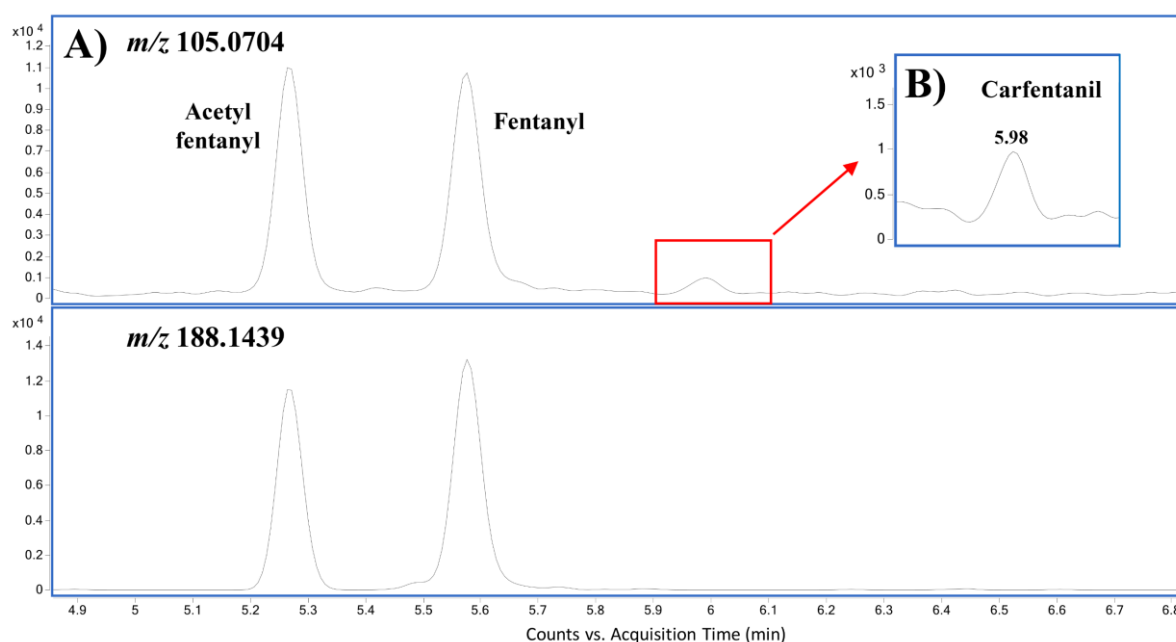


Figure 3-7: EICs showing common fentanyl product ions at m/z 105.0704 and 188.1439 (A) showing a peak for carfentanil at a lower abundance at m/z 105.0704 (B) but not m/z 188.1439

The structural diversity of the opioids also presents challenges for KMD analysis, which is particularly evident for the fentanyl analogues. This technique works well for compounds that contain repeating units, however if multiple different changes are present in the chemical structure, it will change the KMD value, meaning that it may no longer fall within the filter parameters. This drawback has been somewhat mitigated by the inclusion of multiple filters for different subgroups of fentanyl analogues,

which can then capture a representative group of the known compounds in this group. This means that a balance needs to be maintained between targeted and non-targeted approaches by selecting the right number of filters appropriate to a specific application. When screening for high priority compounds, such as carfentanil, specific filters can be added to detect these compounds, however, for the purpose of non-targeted screening, this should be used sparingly. In comparison to the fentanyl analogues, a larger portion of AH and U opioids are covered by a small number of filters, however there are still some compounds, such as U-50488, that may fall outside of these filters.

These drawbacks should not detract from the application of this technique. Given that *DefectDetect* operates using an averaged mass list, it is vendor software agnostic. This means that it can be implemented into a routine workflow regardless of the type of instrument used in a laboratory, so long as an averaged mass list can be generated. In addition, this technique has been previously shown to be useful for other drug classes, namely cathinones and phenethylamines ^[158, 227]. The filters used for the KMD analysis are user-defined, meaning that multiple different drug classes can be combined into a single output, depending on the priorities of the specific laboratory. In this study, KMD values were corrected for a repeating CH₂ unit, which was found to be most appropriate for synthetic opioids, however different correction factors may be applied to more efficiently detect different drug classes. It is important when applying this technique to new compound groups that common structural features are identified and different correction factors are evaluated to determine the optimal filters to use, which encompass the broadest range of compounds.

These non-targeted screening methods were applied to 157 race-day equine plasma samples obtained by the ARFL during the 2019 Autumn Carnival. These samples were analysed using the routine screening method in place at ARFL before being subjected to the non-targeted workflow. While these samples did not return any positive detections with the applied methods, it did allow for both PIS and KMD analysis to be evaluated alongside routine analysis in terms of their ease of use. The PIS method was found to be easy to append to previously employed data analysis workflows. Having an established list of product ions to be extracted made it simple to apply to routine batches containing 30 to 50 samples. KMD analysis, while still easy to perform, proved slightly more time consuming in the generation of EICs for each sample based on the results table provided by *DefectDetect*. The implementation of both PIS and KMD techniques alongside routine workflows can be completed in approximately 2 to 3 hours (in addition to normal data checking procedures) for a routine sequence containing 30 to 50 samples.

Table 3-4 provides a summary of the performance of the PIS and KMD techniques. While quantitative validation was not completed for this proof-of-concept study, the screening cut-off levels show promise for the application of these techniques to routine screening. The estimated cut-off levels indicate that PIS can detect analytes close to their estimated LODs from validation performed by the ARFL for conventional targeted screening. KMD analysis resulted in a higher screening cut-off level being estimated. This may be due to the averaging of mass spectral data, specifically the averaging of ion intensities, potentially causing low concentration analytes to fall outside of the defined filters.

Table 3-4: Overview of the estimated limits of detection (LODs) and screening cut-offs for the different detection techniques applied

Compound	Estimated Targeted LOD (ng/mL) ^[226]	Screening Cut-off (ng/mL)	
		PIS	KMD
Acetyl fentanyl	0.01	0.05	0.1
Carfentanil	0.05	-	0.1
Fentanyl	0.01	0.05	0.1
AH-7921	0.01	0.05	0.1
U-50488	0.05	0.05	-
Acetyl fentanyl	0.01	0.05	0.1

*PIS: Product Ion Searching; KMD: Kendrick Mass Defect

Given that these techniques are focused on the ‘back-end’ data analysis methods, their sensitivity can also be affected by the instrument and analytical methods used. For both techniques, it is possible that compounds at concentrations lower than the estimated cut-off levels will be detected, however the reliability of detection will be lower.

Due to the structural diversity of the synthetic opioids, there is no one technique that is best for all the compounds analysed. The compounds used for this study were chosen to provide a representative panel of compounds that belong to each group of the synthetic opioids. While there might be other compounds that do not follow the patterns observed in this study, the results obtained demonstrate the importance of using a workflow containing multiple, complimentary techniques. In this way, the drawbacks of one technique can be made up for by another, providing a more holistic screening for compounds that might be present in a given sample.

This study has demonstrated the potential of the of PIS and KMD analysis for non-targeted screening of synthetic opioids. There is scope for future work to continue to expand the usefulness of these

techniques to further derivatives to increase their effectiveness. Newer synthetic opioid classes, such as the benzimidazole opioids recently reported by Blanckaert *et al.* ^[79], can also be investigated for inclusion in non-targeted screening workflows implementing these techniques. From an operational perspective, further work into the automation of the developed techniques may prove valuable to assist their application to routine analysis.

3.5 Molecular Feature Extraction

3.5.1 Background

In addition to the work presented in the manuscript above, molecular feature extraction (MFE) was attempted using the Agilent Technologies *MassHunter Profinder* software. MFE algorithms locate individual sample components (molecular features) from within complex chromatograms ^[228]. The algorithms can use the accuracy of the mass measurements to assign multiple species that are related to the same neutral molecule (such as different charge states or adducts) to a single compound group ^[228]. This approach can allow for the simplification of complex matrices into the different compounds that are present within a given sample.

3.5.2 Experimental

Data files submitted to *Profinder* underwent a Batch Recursive Feature Extraction workflow. This workflow involves a two-stage process to extract and collate the identified entities ^[229]. The data files were first analysed using an MFE algorithm and the information extracted by this algorithm was then used for a targeted Find by Ion (Fbi) extraction process. The input data range for the MFE was set to only include features with a retention time between 3.5–7 min and an m/z range of 200–500. A peak filter was added so that only peaks with a height of ≥ 750 counts would be included during feature extraction. Only the $[M+H]^+$, $[M+Na]^+$ and $[M+NH_4]^+$ ion species were considered in the feature extraction. A common organic molecules isotope model was used for analysis. This model includes isotopes of C, N, O, P, H, S, Cl and Br. Only molecules with a charge state of +1 were included in analysis. A mass defect filter was used during the MFE process with a range of 0–0.27 Da and an ion count threshold of two or more ions was added.

Compound binning and alignment allows the algorithm to determine which features are the same across the dataset. Features within a retention time tolerance of ± 0.3 min and a mass tolerance of

$\pm(50 \text{ ppm}+3 \text{ mDa})$ were assigned to the same bin. These tolerance parameters are specific to the entity extraction within *Profinder* and do not affect the accuracy of the instrument used. In order for an entity to be isolated, it would need to fulfil mass, retention time and isotopic abundance parameters, therefore allowing a slightly larger mass tolerance can account for any mass drift observed. In addition, this larger tolerance can correct for compounds with a lower ionisation ability and allow for batch-to-batch data to be collated into a single *Profinder* extraction. These aligned features are referred to as ‘compound groups’ in *Profinder*. An MFE post-processing filter was set with an absolute peak height of ≥ 750 counts and features needed to satisfy the filter conditions in at least 1 file across all sample files to be included as a compound group.

Once the MFE algorithm finished processing the sample files, a consensus library of target ions was compiled. The Fbl algorithm then used that consensus library to re-extract the target ions from the raw data files. The match tolerances for the re-extraction of the target formulae were set to ± 20 ppm for mass tolerance and ± 0.2 min for retention time tolerance. The extraction window used for the creation of EICs had an m/z range of ± 20 ppm and retention time range of ± 1 min. Target ions that failed the match tolerance criteria, but fell within the specified extraction window were still seen in the EIC, however they were flagged as ‘**RT tolerance**’ or ‘ **m/z tolerance**’. The overall score for each target compound was made up of 3 different component scores, each with a different relative weighting. The relative weightings were 100 for the mass score, 60 for the isotope abundance score and 50 for the isotope spacing score. Expected data variation parameters were defined which affect the extent of the scoring penalties applied for variations in the mass and retention time. The expected variation from the MS mass used by the Fbl algorithm was set to $\pm(2 \text{ mDa}+10 \text{ ppm})$ with an isotope abundance variation of 7.5%. The MS/MS mass variation was set to $\pm(5 \text{ mDa}+10 \text{ ppm})$ and an expected retention time variation of ± 0.115 min. Compound groups that returned a score of lower than 75 were flagged with a ‘**low score**’ warning. Additionally, a second warning flag was included when a second ion was expected from the molecular formula with an abundance of greater than 50 counts but was not observed in the Fbl extraction.

An Agile 2 integrator was used for matched peaks and an absolute height filter of ≥ 750 counts was used for integrated peaks. This filter reduces the probability that a low abundance peak may score better than the targeted peak by chance and, therefore, be preferentially selected. The mass spectra that are displayed for each compound group are averaged spectra for each extracted peak. Thresholds were set to determine which peak regions were used when determining the average peak spectrum. In this analysis, the algorithm was set to include average scans that were $>10\%$ of the peak height and exclude any spectra that were above 20% of the saturation value of the detector. Finally,

a post-processing filter was applied to the Fbl extraction. Compound groups were required to return a target score of ≥ 50 in at least 1 file across all sample files in order to be included in the results.

3.5.3 Results and Discussion

In comparison to PIS and KMD analysis, recursive feature extraction (RFE) using *Profinder* is an unbiased detection technique. The technique doesn't rely on a targeted analyte showing similarity to known drugs ^[122, 145], rather it will extract any compound that passes the user-defined criteria. While this may enable the technique to be applicable to a much broader range of compounds than the other techniques, limitations can be found with respect to the relevance of the extracted compounds. In a complex biological matrix, such as plasma, there are many background compounds that will be present in the sample. While the extraction parameters can be optimised for the targeted analyte(s), some of these background compounds will be extracted. This can make it difficult for the analyst to review the data and identify exogenous compounds that are of interest from a forensic toxicology perspective.

Figure 3-8 shows the layout of the *Profinder* program once the RFE algorithm has processed a batch of samples. The first window (Figure 3-8A) provides a list of the compound groups extracted from the batch of samples. This window also summarises the major information for the compound group, including retention time and mass data, as well as how many times within the data set the compound was found and the score given by the MFE algorithm, among others. When an individual compound group is selected, the second window (Figure 3-8B) shows the information for the compound group for each sample within the data set. If any of the samples were flagged during the RFE workflow, whether it be for RT tolerance, low MFE score, or any other defined criteria, these flags will be displayed in this window and colour coded to stand out to an analyst. The third window (Figure 3-8C) shows EICs for the selected compound group across all the samples in the data set. This allows an analyst to review the extracted peaks and manually extract compounds in any sample where the algorithm has not identified a compound, but where a peak is still present. The information for this manually extracted compound will then be updated in the compound information window (Figure 3-8B), where it can be reviewed by an analyst. Conversely, compounds can manually be deleted from the compound group where it is judged that the algorithm has extracted an erroneous or background peak. The final window (Figure 3-8D) displays the Find by Formula mass spectra for the selected compound group. These spectra display the molecular ion for each sample, along with any adducts detected, which can be reviewed by an analyst. Close-up images of the different *Profinder* panels can be found in Appendix 3 (Figures A3-6 – A3-8).

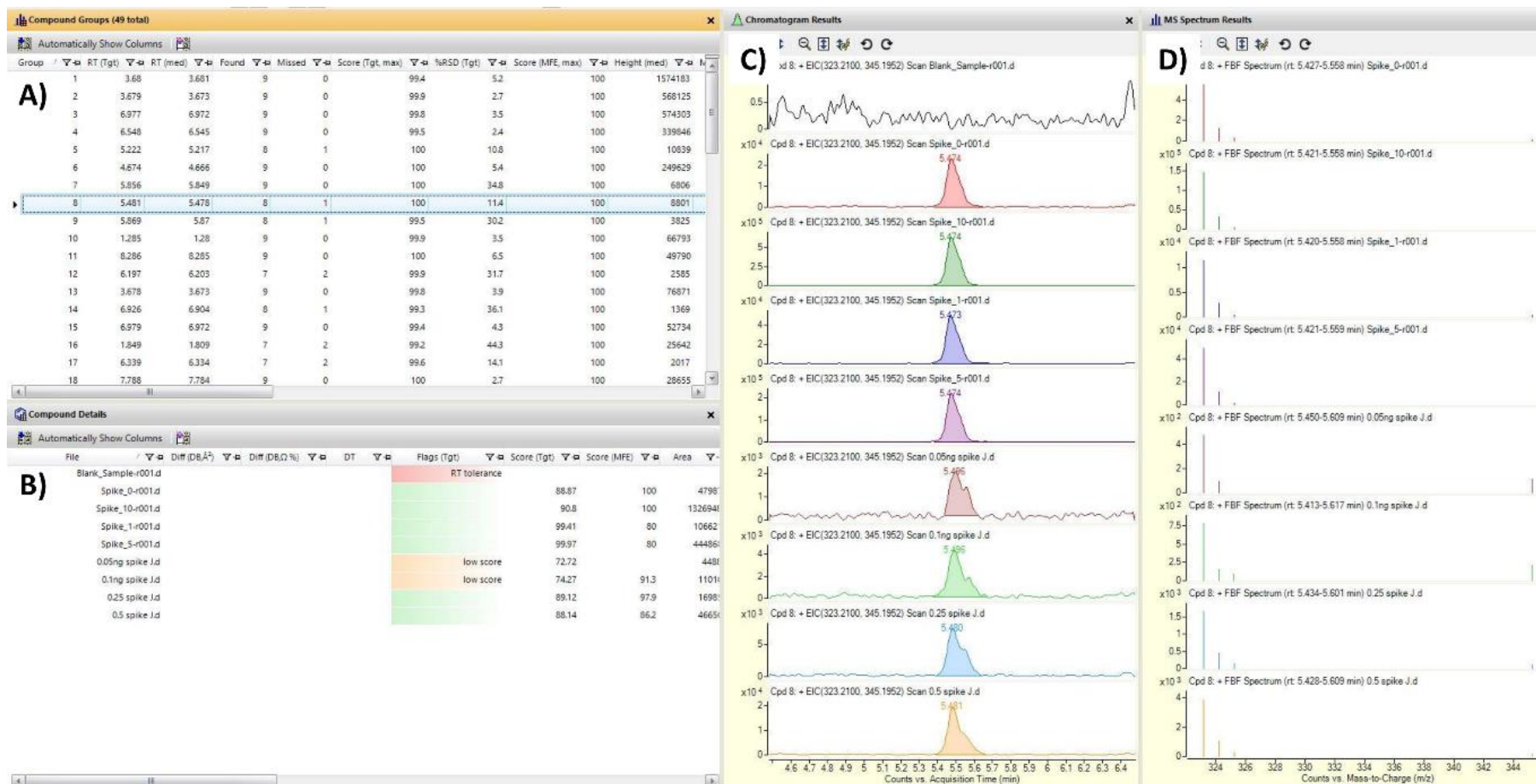


Figure 3-8: Layout of the Profinder outputs showing: (A) list of compound groups; (B) compound information within each sample; (C) EICs for selected compound group; (D) Find by Formula mass spectra for selected compound group

Where this technique excels is its sensitivity when extracting the compounds. In all the spiked samples that were tested, the RFE algorithm extracted the targeted compounds down to concentrations between 0.01 and 0.05 ng/mL. For some of these compounds, this was comparable with the targeted extraction sensitivity, without being given any prior information about the compounds, such as masses or expected retention times. On the other hand, the optimised parameters developed in this study extracted a total of 437 compounds from the samples. This makes it difficult to determine which compounds might be of interest and which are present because of the matrix. For the purpose of this study, 170 routine out of competition plasma samples were extracted using the same parameters as a background population. The list of compounds generated from this background population (478 compounds) was used as a mass exclusion list when analysing the spiked samples. While this did reduce the number of compounds that were extracted, resulting in the extraction of 236 compounds from the spiked samples (approximately 46% reduction), it was not significant enough to clearly differentiate the spiked compounds from the background.

Mass Profiler Professional (Version 15.1) was used to compare the spiked sample set to the background population to identify entities (compounds) that were unique to each sample set. This reduced the number of extracted compounds by approximately 58%, however there were still 183 different entities produced by this analysis. Additionally, this analysis requires the identified compounds to be present in a user-defined proportion of each sample set. For example, the parameters can be set so that a specific compound must be present in every sample of one of the sample sets before it is considered to be a unique entity. This is not appropriate from a drug screening perspective, as there may only be one sample in a sequence that contains an analyte of interest. While the proportion of samples that the compound must be present in can be reduced, this will naturally increase the number of unique entities that are identified, which defeats the purpose of applying this statistical technique. Figure 3-9 provides an overview of the effects of the different refinement techniques applied to the RFE process.

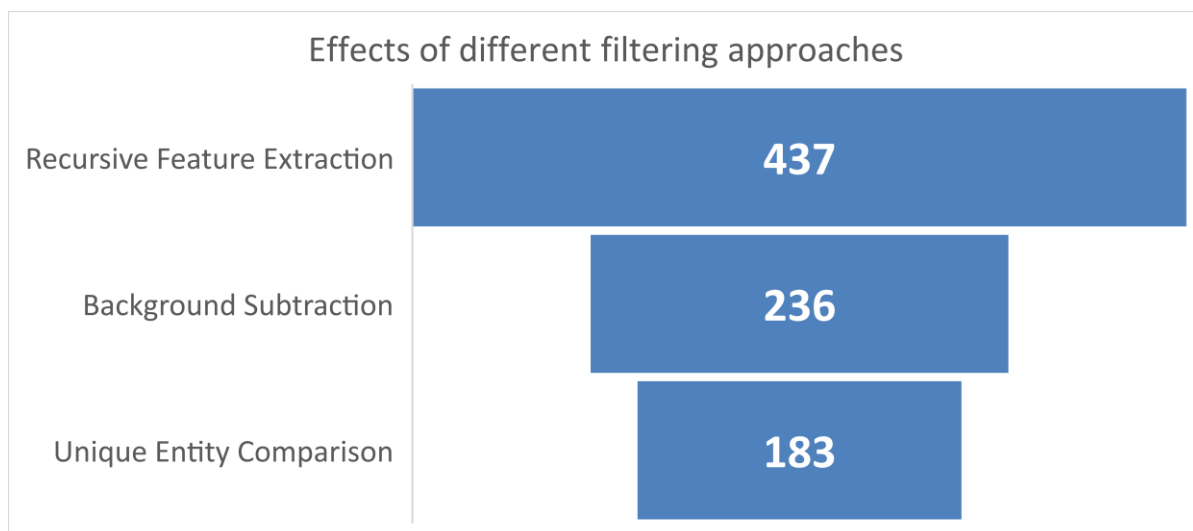


Figure 3-9: Overview of the effects of different filtering processes on the number of compounds extracted by the recursive feature extraction approach

3.6 Conclusion

Non-targeted screening of synthetic opioids within established structural classes can be performed by exploiting common fragmentation patterns within the different subclasses of opioids and trends within the KMD of the compounds. The blind trials demonstrated the potential of the PIS and KMD analysis techniques in a realistic, pseudo-operational context to enhance the screening capabilities for fentanyl analogues. In addition, they demonstrated the importance of using the detection techniques as part of a complementary workflow to increase the confidence in the detection of compounds of interest.

In addition, when considering the molecular feature extraction, it appears that RFE is not fit-for-purpose for first-pass drug screening. While the sensitivity of the technique is excellent, it is difficult to differentiate between analytes of interest and background compounds. This technique is more suited to a metabolomics-type approach, where all identified compounds may be relevant to the analysis.

3.7 References

1. United Nations Office on Drugs and Crime. *World Drug Report 2019*. Vienna: United Nations Office on Drugs and Crime; June 2019.

19. Lucyk, S.N. and L.S. Nelson, *Novel Synthetic Opioids: An Opioid Epidemic Within an Opioid Epidemic*. Ann. Emerg. Med., 2017. **69**(1): p. 91-93.
31. Sandilands, E.A. and D.N. Bateman, *Opioids*. Medicine, 2016. **44**(3): p. 187-189.
32. Holden, J.E., Y. Jeong, and J.M. Forrest, *The endogenous opioid system and clinical pain management*. AACN clinical issues, 2005. **16**(3): p. 291-301.
34. Busardò, F.P., et al., *Ultra-High-Performance Liquid Chromatography-Tandem Mass Spectrometry Assay for Quantifying Fentanyl and 22 Analogs and Metabolites in Whole Blood, Urine, and Hair*. Front. Chem., 2019. **7**: p. 184-184.
79. Blanckaert, P., et al., *Report on a novel emerging class of highly potent benzimidazole NPS opioids: Chemical and in vitro functional characterization of isotonitazene*. Drug Test. Anal., 2020. **12**: p. 422-430.
107. Tobin, T., *Drugs and the Performance Horse*. 1981, Illinois, USA: Charles C Thomas Publisher, Limited.
108. Kamerling, S., et al., *Dose related effects of the kappa agonist U-50,488H on behaviour, nociception and autonomic response in the horse*. Equine Vet. J., 1988. **20**(2): p. 114-118.
109. Racing Medication & Testing Consortium, *New Synthetic Opioid Designer Drug Found by New York Lab: ARCI 11/9/15*. 2015, Racing Medication & Testing Consortium: Lexington, Ky.
122. Pasin, D., et al., *Current applications of high-resolution mass spectrometry for the analysis of new psychoactive substances: a critical review*. Anal. Bioanal. Chem., 2017. **409**(25): p. 5821-5836.
124. Oberacher, H. and K. Arnhard, *Current status of non-targeted liquid chromatography-tandem mass spectrometry in forensic toxicology*. TrAC, Trends Anal. Chem., 2016. **84**: p. 94-105.
126. Pasin, D., et al., *Characterization of hallucinogenic phenethylamines using high-resolution mass spectrometry for non-targeted screening purposes*. Drug Test. Anal., 2017. **9**(10): p. 1620-1629.
143. Noble, C., et al., *Application of a screening method for fentanyl and its analogues using UHPLC-QTOF-MS with data-independent acquisition (DIA) in MSE mode and retrospective analysis of authentic forensic blood samples*. Drug Test. Anal., 2017. **10**(4): p. 651-662.
145. Ibáñez, M., et al., *Comprehensive analytical strategies based on high-resolution time-of-flight mass spectrometry to identify new psychoactive substances*. TrAC, Trends Anal. Chem., 2014. **57**: p. 107-117.
147. Grabenauer, M., et al., *Analysis of Synthetic Cannabinoids Using High-Resolution Mass Spectrometry and Mass Defect Filtering: Implications for Nontargeted Screening of Designer Drugs*. Anal. Chem., 2012. **84**(13): p. 5574-5581.
148. Sleno, L., *The use of mass defect in modern mass spectrometry*. J. Mass Spectrom., 2012. **47**(2): p. 226-236.

149. Zhang, H., et al., *Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry*. J. Mass Spectrom., 2009. **44**(7): p. 999-1016.
157. Hughey, C.A., et al., *Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra*. Anal. Chem., 2001. **73**(19): p. 4676-4681.
158. Anstett, A., et al., *Characterization of 2C-phenethylamines using high-resolution mass spectrometry and Kendrick mass defect filters*. Forensic Chem., 2018. **7**: p. 47-55.
205. United Nations Office on Drugs and Crime. *World Drug Report 2018*. Vienna: United Nations Office on Drugs and Crime; June 2018 2018.
214. Pichini, S., et al., *Acute Intoxications and Fatalities From Illicit Fentanyl and Analogues: An Update*. Ther Drug Monit, 2018. **40**(1): p. 38-51.
215. Knych, H.K., et al., *Pharmacokinetics and pharmacodynamics of butorphanol following intravenous administration to the horse*. J. Vet. Pharmacol. Ther., 2013. **36**(1): p. 21-30.
216. Kamerling, S.G., et al., *Dose-related effects of fentanyl on autonomic and behavioral responses in performance horses*. Gen. Pharmacol.: Vasc. S., 1985. **16**(3): p. 253-258.
217. Combie, J., et al., *Pharmacology of narcotic analgesics in the horse: selective blockade of narcotic-induced locomotor activity*. Am J Vet Res, 1981. **42**(5): p. 716-721.
218. Marchei, E., et al., *New synthetic opioids in biological and non-biological matrices: A review of current analytical methods*. TrAC, Trends Anal. Chem., 2018. **102**: p. 1-15.
219. Cooreman, S., et al., *A comprehensive LC-MS-based quantitative analysis of fentanyl-like drugs in plasma and urine*. J. Sep. Sci., 2010. **33**(17-18): p. 2654-2662.
220. Sofalvi, S., et al., *An LC-MS-MS Method for the Analysis of Carfentanil, 3-Methylfentanyl, 2-Furanyl Fentanyl, Acetyl Fentanyl, Fentanyl and Norfentanyl in Postmortem and Impaired-Driving Cases*. J. Anal. Toxicol., 2017. **41**(6): p. 473-483.
221. Wang, L. and J.T. Bernert, *Analysis of 13 Fentanils, Including Sufentanil and Carfentanil, in Human Urine by Liquid Chromatography-Atmospheric-Pressure Ionization-Tandem Mass Spectrometry*. J. Anal. Toxicol., 2006. **30**(5): p. 335-341.
222. Gerace, E., A. Salomone, and M. Vincenti, *Analytical Approaches in Fatal Intoxication Cases Involving New Synthetic Opioids*. Curr. Pharm. Biotechnol., 2018. **19**(2): p. 113-123.
223. Eckart, K., et al., *Development of a new multi-analyte assay for the simultaneous detection of opioids in serum and other body fluids using liquid chromatography-tandem mass spectrometry*. J. Chromatogr. B, 2015. **1001**: p. 1-8.
224. Klingberg, J., et al., *Collision-Induced Dissociation Studies of Synthetic Opioids for Non-targeted Analysis*. Front. Chem., 2019. **7**(331).

225. Association of Official Racing Chemists. *AORC Guidelines for the Minimum Criteria for Identification by Chromatography and Mass Spectrometry*. Online 2016. <http://www.aorc-online.org/documents/aorc-ms-criteria-modified-23-aug-16/>.
226. Thompson, M., S.L.R. Ellison, and R. Wood, *Harmonized guidelines for single-laboratory validation of methods of analysis - (IUPAC technical report)*. Pure Appl. Chem., 2002. **74**(5): p. 835-855.
227. Pasin, D., *Non-targeted analysis of new psychoactive substances using mass spectrometric techniques*, in Centre for Forensic Science. 2018, University of Technology Sydney: Online.
228. Sana, T.R., et al., *Molecular formula and METLIN Personal Metabolite Database matching applied to the identification of compounds generated by LC/TOF-MS*. J Biomol Tech, 2008. **19**(4): p. 258-266.
229. Agilent Technologies. *Agilent Mass Profiler Professional Operation for Chemometric Analysis*. United States of America: Agilent Technologies; 2019. SW-MPP-3101c.

Chapter 4:

***Non-Targeted Analysis – Compound
Identification***

Chapter 4: Non-Targeted Analysis – Compound Identification

4.1 Rationale

Once an unknown compound has been detected in a non-targeted workflow, the identity of the compound must be established. While definitive compound identification must be performed using a CRM, it is useful to be able to putatively determine the identity of an unknown before time and/or resources are invested into confirmation against a reference standard. This chapter focuses on the use of machine learning and statistical modelling approaches to propose putative identities of unknown compounds within biological samples.

This chapter is taken from a first author manuscript entitled ‘Towards Compound Identification in Non-targeted Screening Using Machine Learning Techniques’, submitted for publication in Drug Testing and Analysis. The experimental work, data analysis and preparation of the initial draft were performed by J. Klingberg, with manuscript edits provided by A. Cawley, R. Shimmon, and S. Fu.

Towards Compound Identification in Non-targeted Screening Using Machine Learning Techniques

Joshua Klingberg¹, Adam Cawley², Ronald Shimmon¹, Shanlin Fu^{1*}

¹Centre for Forensic Science, University of Technology Sydney, Ultimo, NSW, Australia

²Australian Racing Forensic Laboratory, Racing NSW, Sydney, NSW, Australia

*Corresponding Author:

Shanlin Fu

Shanlin.fu@uts.edu.au

4.2 Introduction

Recently, there has been an increasing focus on non-targeted analysis methods in forensic toxicology to detect illicit drugs in biological matrices for high-throughput screening applications, without needing to rely on CRMs or library databases ^[122]. Previous studies have investigated the use of product ion searching to detect novel analogues of known drug classes, exploiting class-specific cleavages giving rise to common product ions ^[126, 143, 224, 230]. Other studies have looked at more advanced data analysis methods, incorporating mass filtering techniques such as Kendrick Mass Defect analysis to detect structurally related compounds within a complex matrix ^[158, 230]. While these data analysis techniques are useful for detecting the presence of an unknown compound within a sample, further investigation is required to determine the identity of the analyte. Furthermore, while many vendor software packages, such as *Molecular Structure Correlator* (Agilent Technologies) and *Compound Discoverer* (Thermo Fischer Scientific), can help with this process, having a basic understanding of the general classification of an unknown compound can assist with their timely identification.

With the increasing volume of data generated by HRMS and the availability of cheaper and more powerful computational processing ^[168], the use of artificial intelligence approaches, such as machine learning, for toxicological applications has become viable. In general, machine learning algorithms enable a computer to ‘learn’ information directly from a data set without needing a predetermined equation to use as a model ^[169]. In this way, the algorithm can adaptively improve its performance over time with experience ^[170]. Machine learning can be further divided into supervised and unsupervised learning. Supervised learning involves using a known set of input data (called the training set) with known responses to that data (the output) to train a model that has the ability to predict the response for new input data ^[168, 171]. Supervised learning models are categorised as either ‘classification’ or ‘regression’ models. Classification models aim to predict a discrete output, whereas regression models attempt to predict an output within a continuous space, such as a temperature or time range ^[168]. Unsupervised learning, on the other hand, is useful when attempting to explore data without specific goals or previous knowledge of the information contained within the data ^[168, 172].

Machine learning approaches have been used in the field of clinical toxicology previously, notably within the context of determining structure activity relationships in drug discovery ^[173, 174]. Retention time prediction under a specific chromatographic system has also been explored using regression models from both an environmental ^[179] and drug analysis perspective ^[180, 181].

This study aimed to develop a machine learning approach to assist with the identification of unknown compounds indicated from non-targeted screening workflows such as those presented by our group [230]. The development of a classification model to predict the subclass of synthetic opioids based on MS² dissociation data is presented, alongside the development of a regression model to predict the retention time of suspected compounds from calculated molecular features. CRMs of novel drug analogues can be difficult and costly to procure. If an analyst has evidence supporting the putative identification of an unknown compound prior to procuring a CRM, it may save time and money for subsequent confirmation of the drug.

4.3 Experimental

4.3.1 Solvents and Reagents

All solvents used were LC-MS grade. Acetonitrile, ethyl acetate and methanol were obtained from Merck (Darmstadt, Germany). Ammonium acetate and trichloroacetic acid were obtained from Sigma-Aldrich (Castle Hill, NSW, Australia). Acetic acid and hydrochloric acid were obtained from Ajax Chemicals (Sydney, NSW, Australia). Ultrapure-grade water was obtained from a Smart2Pure ultra-pure water system (Thermo Scientific, Langenselbold, Hungary).

4.3.2 Drug Standards

Fentanyl citrate was purchased from Sigma-Aldrich (Castle Hill, NSW, Australia). Hydrochloride salts of 3,4-ethylenedioxy U-47700, 3,4-ethylenedioxy U-51754, 3,4-methylenedioxy U-47700, 4-fluoroisobutyl fentanyl, 4-chloroisobutyl fentanyl acetyl fentanyl, AH-7921, AH-8529, AH-8533, meta-fluoro fentanyl, para-fluoro methoxyacetyl fentanyl, phenyl fentanyl, tetrahydrofuran fentanyl, U-48800, U-51754, UF-17 and β -hydroxythio fentanyl, manufactured by Cayman Chemical (Ann Arbor, MI, USA), were purchased from Sapphire Bioscience (Redfern, NSW, Australia). U-62066 mesylate, along with free base standards of AH-7563, AH-7959, AH-8507, AH-8532, crotonyl fentanyl, isopropyl U-47700, meta-fluoro methoxyacetyl fentanyl, methacryl fentanyl, N-methyl U-47931E, propyl U-47700, seneciyl fentanyl, thiofentanyl, U-47109, U-47931E, U-48520, U-49900 and U-50488, manufactured by Cayman Chemical, were also purchased from Sapphire Bioscience. Hydrochloride salts of 4-methoxybutyl fentanyl, acryl fentanyl, butyl fentanyl, furanyl fentanyl, U-47700, valeryl fentanyl, 4-fluorobutyl fentanyl, benzyl fentanyl, N-methyl carfentanil, ocfentanil and ortho-fluoro fentanyl, along with MT-45 dihydrochloride hydrate and neat standards of

despropionyl para-fluoro fentanyl, benzodioxole fentanyl, cyclopropyl fentanyl, cyclopentyl fentanyl, phenylpropionyl fentanyl and norfentanyl, manufactured by Chiron Chemicals (Hawthorn, VIC, Australia), were purchased from PM Separations (Capalaba, QLD, Australia). Remifentanyl hydrochloride was purchased from GlaxoSmithKline (Boronia, VIC, Australia). Citrate salts of carfentanyl, sufentanyl and α -methyl fentanyl, along with alfentanyl hydrochloride, were purchased from Janssen Pharmaceuticals (North Ryde, NSW, Australia). Desipramine- d_3 was purchased from Grace (Columbia, MD, USA). The chemical information for all opioid standards used in this study can be found in Appendix 4 (Table A4-1).

4.3.3 Sample Preparation

4.3.3.1 Class Prediction Samples

Drug standards were obtained as methanolic standards of varying concentrations. Neat standards of all opioids were diluted in methanol to a concentration of 10 $\mu\text{g/mL}$. Ten microliters of each solution was evaporated to dryness under nitrogen, before being reconstituted in one drop of methanol using a Pasteur pipette (approximately 20 μL) and 100 μL of 10 mM ammonium acetate (pH 4) buffer to give a final concentration of 1 $\mu\text{g/mL}$ for analysis. All samples were stored at 4 $^{\circ}\text{C}$ until analysis.

4.3.3.2 Retention Time Repeatability Studies

Neat mixed standards containing 4-chloroisobutyryl fentanyl, 4-fluoroisobutyryl fentanyl, 4-methoxybutyryl fentanyl, acetyl fentanyl, acryl fentanyl, AH-7563, AH-7921, U-47700, AH-7959, AH-8507, AH-8529, AH-8533, α -methyl fentanyl, butyryl fentanyl, carfentanyl, cyclopentyl fentanyl, cyclopropyl fentanyl, fentanyl, furanyl fentanyl, meta-fluoro fentanyl, MT-45, remifentanyl, sufentanyl, U-50488, U-51754 and valeryl fentanyl were prepared in methanol to give a concentration of 1 $\mu\text{g/mL}$. Sets of 7 samples were prepared, and the study was repeated 7 times ($n = 49$) to evaluate the repeatability of absolute retention times without the possibility of matrix effects.

Blank plasma was obtained from blood samples collected in Lithium Heparin Vacutainers purchased from BD (Mississauga, ON, Canada) from four thoroughbred research horses following approval of the Racing NSW Animal Care and Ethics Committee (ARA 71).

Spiked equine plasma samples (2 mL) were prepared to determine the precision of retention time measurements in a relevant biological matrix. Mixed standards containing the same panel of opioids

used in the neat standards were spiked into the plasma samples at a concentration of 10 ng/mL. Spiked samples were prepared in sets of 7, and the extraction was repeated 7 times ($n = 49$) to determine the repeatability of measured retention times. Desipramine- d_3 (10 ng/mL) was added as an internal standard to compare the repeatability of the absolute retention times (RT) to relative retention times (RRT) in comparison to the internal standard. The RRT was calculated by dividing the RT of the opioid standard by the RT of the internal standard.

4.3.3.3 Retention Time Prediction Samples

Once the retention time repeatability had been determined, separate plasma samples were spiked with each of 59 different opioid standards at a concentration of 10 ng/mL. Desipramine- d_3 was again added as an IS at a concentration of 10 ng/mL. Triplicate injections were completed from each sample and the extraction was repeated 3 times for a total of 9 measurements for each standard.

4.3.3.4 Plasma Extraction Method

Protein precipitation was completed through the addition of 200 μ L of trichloroacetic acid (10% in H_2O) to 2 mL samples. The pH of the samples was then adjusted to 3 – 3.5 using hydrochloric acid after which they were centrifuged at 1500 g for 10 minutes. Solid phase extraction (SPE) was completed using XtrackT[®] Gravity Flow DAU Extraction Columns (UCT Inc., Bristol, PA, USA). The cartridges were first conditioned with 3 mL of methanol, followed by 3 mL of purified water, after which the samples were loaded. The samples were washed with 3 mL of 0.1 M acetic acid and dried under positive pressure. The cartridges were again conditioned with 3 mL of methanol and dried under positive pressure. The analytes were eluted from the cartridges using 3 mL of solvent containing 3% ammonia and 0.5% methanol in ethyl acetate.

Following SPE, one drop of 0.1 M methanolic hydrochloric acid was added to each of the samples using a Pasteur pipette before the solvent was evaporated under a gentle stream of N_2 at 60 °C. The samples were then reconstituted in one drop of methanol using a Pasteur pipette and 100 μ L of an ammonium acetate buffer (pH 3.9), before being transferred to vials for analysis. All samples were stored at 4 °C until analysis.

4.3.4 Instrumental Analysis

Chromatographic separation was achieved on an Agilent Technologies (Santa Clara, CA, USA) 1290 Infinity II UHPLC, consisting of a high-speed pump (G7120A), multisampler (G7167B) and thermostat and column compartment (G1316A, 35 °C) coupled to an Agilent Technologies 6545 QTOF mass spectrometer. All data acquisition was conducted using Agilent Technologies MassHunter Workstation (Version B.06.01). A sample volume of 1 µL was injected onto an Agilent Technologies Poroshell 120 EC-C18 LC column (2.1 x 75 mm, 2.7 µm particle size) using a gradient elution with a flow rate of 0.4 mL/min and a total analysis time of 12 min. Mobile phase A consisted of 10 mM ammonium acetate buffer (pH 9) and mobile phase B consisted of 0.1% acetic acid in acetonitrile. Initial mobile phase composition was 75% A, which was held for 0.5 min before being decreased linearly to 67% A over 2 min, before being held for another 1 min. The mobile phase was then further decreased to 55% A over 6 min, before being held for 1 min and returned to 75% A over 0.5 min, before a final 1 min hold. The column was allowed to equilibrate for 2 min before the next sample was injected.

The QTOF was operated in ESI+ mode with capillary and fragmentor voltages of 3500 V and 100 V, respectively. The QTOF was calibrated each day before use with an average resolution of 16,679 (FWHM) over the period of analysis at a reference mass of m/z 322.048121. The calibration solution was made up using 10 mL of the Agilent Technologies ESI-L tuning mix, diluted with 85.5 mL of acetonitrile and 4.5 mL of ultrapure water. Three microliter of 0.1 mM reference mass solution containing HP-0321 was added to the calibration solution. An AutoMS-MS data acquisition mode was used with mass ranges of 100 – 1000 m/z for MS and 50 – 850 m/z for MS/MS. Spectra were obtained with an acquisition speed of 10 spectra/s for both MS and MS/MS and CEs of 10, 20 and 40 eV were used for CID. A maximum of 5 precursors from the MS scan were selected for CID per cycle, with an abundance threshold of 5000 counts. Active exclusion was used, with precursor being excluded after 2 spectra and released after 0.1 min.

4.3.5 Data Analysis

All data files were analysed using Agilent Technologies MassHunter Qualitative Analysis Software (Version B.10.0, Build 10.0.10305.0) to generate EICs and mass spectra for use in the statistical analysis software.

4.3.5.1 *Class Prediction Samples*

MT-45 and UF-17 were excluded from the classification model as they fall within their own classes and the classification model could not reliably account for these compounds when there is only 1 sample in the training set.

The Find by AutoMS-MS function in the qualitative analysis software was used to extract MS spectra from the neat opioid standards at all 3 CEs (10, 20 and 40 eV). The spectra were then exported, along with a list of the top 10 most abundant masses, for use in building the class prediction model. The most abundant ions for each synthetic opioid standard were then populated into an Excel spreadsheet from highest to lowest abundance.

4.3.5.2 *Retention Time Prediction Samples*

Compound and spectral information for all opioid standards analysed was curated into a database using Agilent Technologies MassHunter Personal Compound Database and Library (PCDL) Manager (Version B.08.00, Build 8209.7 SP1). This PCDL was used in conjunction with the Find by Formula (FbF) function in the qualitative analysis software to generate EICs for all the spiked plasma samples analysed.

For the repeatability studies, absolute RT and RRT were collected and populated within a spreadsheet in Microsoft Excel. The average, standard deviation (SD) and relative standard deviation (%RSD) was calculated within each set of 7 samples and between the 7 sets of samples analysed (n = 49).

The individual spiked samples were analysed and the average RRT across all 9 measurements was determined for use in training the retention time prediction model.

4.3.6 **Molecular Features**

Thirteen molecular features were used in the retention time prediction model. Features were calculated using online software, including the Chemicalize tool (ChemAxon, San Diego, CA, USA) ^[231] and Pharmaceutical Data Exploration Laboratory ^[232], or predefined equations. A list of all the molecular features used can be found in Table 4-1. A pH of 9 was chosen for the determination of logD and logS values as this reflected the pH of the mobile phase used for analysis.

The DBE was calculated using Equation 4.1, where a is number of carbons, b is the number of hydrogens, c is the number of nitrogens and f is the number of halogens in the molecule.

$$DBE = (a + 1) \frac{b - c + f}{2} \quad (4.1)^{[233]}$$

Table 4-1: Molecular features used for retention time prediction

Feature	Abbreviation	Source
Strongest basic pKa	pKa	Chemicalize ^[231]
Number of carbon atoms	nC	Chemicalize ^[231]
Number of oxygen atoms	nO	Chemicalize ^[231]
Number of double bonds	nDB	Chemicalize ^[231]
Number of 6-membered rings	nR06	Chemicalize ^[231]
Number of benzene rings	nBNZ	Chemicalize ^[231]
Partition coefficient	logP	Chemicalize ^[231]
Distribution coefficient (pH 9)	logD (pH 9)	Chemicalize ^[231]
Aqueous solubility (pH 9)	logS (pH 9)	Chemicalize ^[231]
Ghose-Crippen logP	AlogP	Pharmaceutical Data Exploration Laboratory ^[232]
Moriguchi logP	MlogP	Pharmaceutical Data Exploration Laboratory ^[232]
Double bond equivalents	DBE	Equation 1 ^[233]
Hydrophilic factor	Hy	Equation 2 ^[234]

The hydrophilic factor (Hy) was calculated using the Equation 4.2. N_{Hy} is the number of hydrophilic groups in the molecule (or the total number of hydrogen atoms attached to oxygen, nitrogen or sulfur atoms), N_c is the number of carbon atoms in the molecule and A is the number of non-hydrogen atoms in the molecule ^[234].

$$Hy = \frac{(1+N_{Hy}) \log_2(1+N_{Hy}) + N_c \left(\frac{1}{A} \log_2 \frac{1}{A} \right) + \sqrt{\frac{N_{Hy}}{A^2}}}{\log_2(1+A)} \quad (4.2)^{[234]}$$

4.3.7 Statistical Analysis

All statistical analysis was conducted using Microsoft Excel (Version 16.0.12624.20382) and MathWorks MATLAB® (Version R2019b, 9.7.0.1319299) equipped with the Statistical and Machine Learning Toolbox (Version 11.6). The machine learning algorithms evaluated for both the class prediction modelling and retention time prediction reflect the common supervised learning

algorithms available for both classification and regression problems ^[235, 236]. In most machine learning applications, the decision of which model architecture to use is based on a process of trial and error, meaning it is important to evaluate the performance of a number of different model types in order to choose the best approach ^[169]. For this study, ANNs were not used as they require large amounts of data to operate effectively and therefore were not suitable for the intended dataset ^[237].

4.3.7.1 *Class Prediction Modelling*

Product ion data from all 3 CEs (10, 20 and 40 eV) were compared and it was found that the 40 eV spectra produced the most accurate models. The compiled data can be found in Appendix 4 (Table A4-2).

The most abundant ion data was imported into the Classification Learner app within MATLAB. The model response variable was set to be the compound class and *k*-fold cross-validation was used with 50 folds. This method of validation splits the data into a finite number of groups, or folds. When training the model, each iteration uses *k*-1 folds as the training data and remaining group as the test data. This process is then repeated for the remaining groups and an average of the accuracies of each iteration is given at the end ^[238]. This method of validation provides a suitable estimate of the model accuracy ^[238].

Several different model types, namely decision trees, discriminant analysis, naïve Bayes classifiers, support vector machines, nearest neighbour classifiers and ensemble classifiers, were investigated in order to determine the most accurate model. Hyperparameter tuning was used to optimise the models and provide the highest accuracy possible. The hyperparameters of a model are the variables that control the training process itself. These hyperparameters are set before the model is trained and can be considered the model settings that need to be optimised ^[239]. The aim of hyperparameter tuning, therefore, is to find the settings that return the best model performance.

Bayesian optimisation was used to train the model with an acquisition function of 'expected improvement per second plus'. The objective of Bayesian optimisation is to build a probability model of the objective function, in this case the classification error of the model, and use it to select the most promising hyperparameters to train the classification model ^[239]. This optimisation methods learns with each iteration and uses the results from previous trials to determine the best set of hyperparameters to use for the next trial.

The model was trained and optimised over 30 iterations and the Classification Learner app returned the most accurate model from those iterations. In addition to the overall model accuracy, the F1 score and Matthew's correlation coefficient (MCC) were calculated for the optimised model. The code for the trained model could then be exported for use in classifying new samples.

4.3.7.2 Retention Time Prediction

The calculated molecular feature data for each synthetic opioid standard was populated into an Excel spreadsheet, along with the average RRT calculated for each compound. The compiled data can be found in Appendix 4 (Table A4-3). All the compound data was then imported into the Regression Learner app within MATLAB. The model response variable was set to be the RRT and *k*-fold cross validation was again used with 50 folds.

Several regression models were also investigated to determine the most accurate model, namely regression trees, support vector machines and gaussian process regression models. Hyperparameter tuning was also used for the regression models with the same parameters as outlined in Section 4.3.7.1. To assess the influence of individual predictors, each molecular feature was sequentially removed, and the model retrained without that predictor. This process was completed 3 times for each feature and an average RMSE value was determined for each feature. The code was again exported so that it could be used to classify new samples.

4.4 Results and Discussion

4.4.1 Class Prediction Modelling

Previous work has shown that different subclasses of synthetic opioids display class specific fragmentation that can be used for non-targeted screening methods ^[143, 224, 230]. By exploiting this same phenomenon, class prediction modelling can be attempted using MS² data and the presence of abundant product ions within the resultant spectra. The generic structures of the 3 opioid subclasses included in the class prediction models, namely fentanyl analogues, AH series and U series opioids, can be found in Figure 4-1 below.

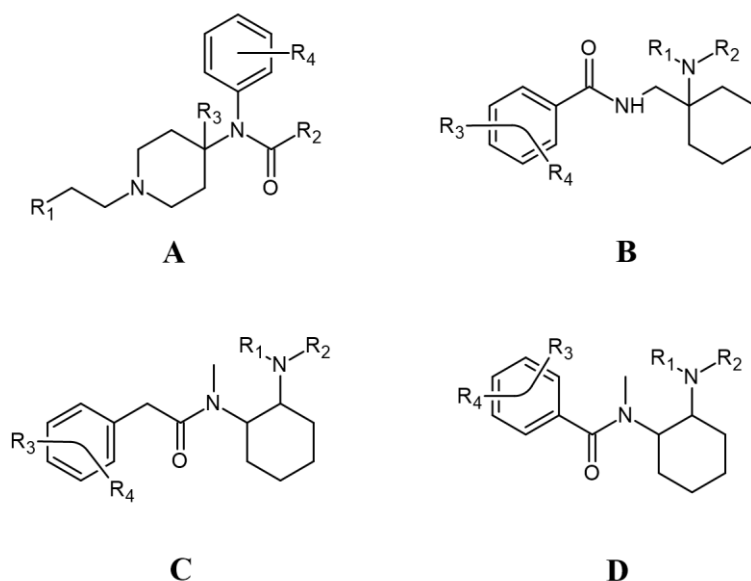


Figure 4-1: Generic structures of the opioid subclasses used in the class prediction model, including fentanyl analogues (A), AH series (B), and U series with (C), and without (D) a methylene spacer

For this study, product ion data from all 3 CEs used for analysis was compared to determine the best input data for training the models. It was found that the MS² spectra using a CE of 40 eV gave the highest accuracy when developing the class prediction models. This is likely due to the fact that the higher collision energy caused greater dissociation of the compound structure, breaking it down into smaller units which are more closely related to the core structures of the compounds. The common fragmentation patterns of the opioid subclasses have been presented previously ^[224].

The use of *k*-fold cross validation is beneficial where smaller data sets are being examined. Using the more simplistic hold-out method requires separating the data set into a training and validation sets, which results in a large proportion of the data not being able to be used for the training of the model ^[240]. The *k*-fold method, on the other hand, allows the use of all data in the training of the model. By randomly separating the training data into different groups, or folds, training and validation can both be performed on all samples within the data set. The choice for how many folds to use in cross-validation is often determined by the computational power available and the variance allowed in the test error calculated ^[240]. When the number of folds is close to, or equal to, the number of samples in the data set, there is less bias in the test error that is calculated ^[240]. This occurs because the model is essentially being trained and tested against each individual sample, which will result in the best estimation of the overall model accuracy. On the other hand, if the number of folds is much smaller than the number of samples, there will tend to be some variance in reported accuracy of the model each time training is completed as the algorithm will more randomly assign samples to each group. If

the composition of each validation group changes, the training process and reported accuracy of the model can change as well.

The downside of using a larger number of folds, however, is that it can slow down the training of the model itself. When using k -fold cross-validation, the model must be trained $k-1$ times. The more folds that are used, the more training iterations the model must go through, which can significantly increase the time taken to train the model, especially when large data sets are being used. Therefore, a compromise is made between the variance allowed in the reported error and the computational power available for training the model. For the model applied in this study, a data set containing 57 samples was used. In comparison to many machine learning problems looking at 'big data' this is a rather small sample size. Therefore, the use of 50-folds for cross validation (the most allowed by the classification learner app) provided an unbiased measure of the overall model error, as k was close to the total number of sample present meaning that bias would not be introduced as a result of the random assignment of training data to the different folds, while still being able to train the model in a reasonable timeframe (within 5-10 minutes).

The classification accuracies for the different models trained can be found in Table 4-2. The overall accuracy of the model can be defined as the total number of correct prediction (i.e. total true positives and true negatives) in relation to the total number of samples in the training set ^[241]. After investigating several different types of classification models available in the MATLAB classification learner app, it was found that the naïve Bayes model provided the best option for class prediction. While the ensemble model provided a slightly higher overall accuracy, it resulted in a loss of accuracy for the fentanyl analogues and no change in the prediction accuracy of the AH series, with the increase in overall accuracy coming from the U series compounds. In addition, an ensemble model is a more complicated model, being comprised of multiple different classification models used in tandem, and took significantly longer to train than the naïve Bayes. The compromise between a slightly lower overall accuracy and simpler/faster model to use and re-train as new data is obtained means that a naïve Bayes model may be more suitable for routine use.

As the name suggests, a naïve Bayes classifier uses Bayes theorem to determine the appropriate output based on the given data ^[242]. Bayes theorem is shown in Equation 4.3 below, where A is the desired output, in this case the class of the compound, and B is the input data that we are using, in this case the most abundant product ions.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (4.3)$$

Table 4-2: Accuracy of each class prediction model trained

Model	Overall Accuracy (%)	Class Accuracy (%)		
		Fentanyl	AH Series	U Series
Decision Tree	87.7	94.3	57.1	86.7
Discriminant Analysis	86.0	85.7	71.4	93.3
Naïve Bayes	89.5	100.0	57.1	80.0
Support Vector Machine	84.2	97.1	71.4	60.0
Nearest Neighbour	87.7	100.0	71.4	66.7
Ensemble	91.2	97.1	57.1	93.3

A naïve Bayes classifier also involves several assumptions. The first of these assumptions is that the input variables (predictors) are independent of each other, i.e. the presence of one feature does not affect the others ^[243]. Secondly, this model type assumes that all the predictors have an equal effect on the outcome ^[243]. While these assumptions are not always true for every data set, naïve Bayes classifiers still provide high classification accuracy ^[244]. In the case of the CID data used in this study, it is reasonable to assume that some of the predictors would not be independent of each other. When considering fragmentation pathways, it is important to take into account the presence of ion formed through secondary fragmentation (i.e., the fragmentation of a product ion to form further product ions). In this case, the presence of the secondary fragment ions relies on the production of their parent ion. While this may contribute to the higher prediction accuracy achieved by the Naïve Bayes model, it is important to consider that compounds that fall within the same chemical class will often produce common diagnostic ions (as discussed in previous chapters). Therefore, these secondary fragmentation relationships would be consistently present across the different subclasses, meaning that the training data submitted to the model would still accurately reflect the subsequent unknown data that is being predicted. As stated by Webb ^[244], Naïve Bayes algorithms can still produce high accuracy models when applied to data sets that do not meet the assumptions.

The confusion matrix produced from the trained model is displayed in Figure 4-2. Using the hyperparameter tuning incorporated in the classification learner app, it was found that a kernel Naïve Bayes structure performed the best with a gaussian kernel type. Kernel distributions are appropriate when the predictors have a continuous distribution ^[245]. The main advantage of kernel distributions is that they can be applied to numerical predictors, such as the fragment ion masses used in this

study, rather than categorical predictors ^[246]. The minimum classification error plot showing the change in error with each training iteration can be found in Appendix 4 (Figure A4-1). The overall accuracy of the trained model was found to be 89.5% and it can be seen in Figure 4-2 that the model was most accurate for the fentanyl analogues. The confusion matrix shows that there is also some bias towards the fentanyl analogues, with the model preferentially classifying this group over the others. This is likely due to the fact that there were significantly more fentanyl analogue standards used in the training of the model than the other two subclasses. This is somewhat unavoidable given that this group tends to be much more prevalent and diverse than the AH and U series opioids. In the context of routine drug screening, having a model that is more accurate for the detection of fentanyl analogues may, in fact, be beneficial. These compounds are the most prevalent of the synthetic opioids with the United Nations Office on Drugs and Crime reporting that the majority of the 22 new opioids reported to their early warning advisory as of December 2019 were fentanyl analogues ^[1]. Additionally, in Europe, until the end of 2018 they accounted for the majority of the novel opioid analogues being reported to the European Monitoring Centre for Drugs and Drug Addiction ^[79]. This suggests that the inherent biases present in the model may reflect the current landscape of the illicit opioid market.

Experimental Class	AH Series	4	3	
	Fentanyl		35	
	U Series		3	12
		AH Series	Fentanyl	U Series
		Predicted Class		

Figure 4-2: Confusion matrix showing the prediction accuracy of the developed Naïve Bayes model for each opioid subclass

It is possible, however, to artificially introduce biases into the training of the model using a cost matrix. When using a cost matrix, a particular penalty is associated with incorrect predictions ^[247]. In the case of the data used in this study, an increased cost could be applied to the misclassification of AH and U series opioids as fentanyl analogues. In a cost-sensitive learning approach, the algorithm will attempt to minimise the cost resulting from the trained model ^[247], meaning that true positive predictions of AH and U series opioids may be favoured over false negative fentanyl analogue predictions. While this may reduce the overall accuracy of the model, it may help to balance the false classifications from the AH and U series compounds. This compromise should therefore be evaluated on a case-by-case basis, depending on the intended application of the model. As the drug market continues to develop with the production of new compounds, it may be possible to further refine the model if more AH and U series opioids are developed which can be incorporated into the training of the model.

In addition to the overall accuracy of the model, the F1 score and MCC were calculated based on the confusion matrix presented in Figure 4-2. It has been suggested that, when a dataset is unbalanced (i.e. there are more samples belonging to one class than the others), accuracy alone may not be a reliable measure of model performance ^[248, 249]. This is because the accuracy measurement may provide an overoptimistic estimation based on the classification ability of the majority class ^[248, 249]. The F1 score, however, is a better metric when the classes are unbalanced and takes into account both the precision (also known as the positive predictive value) and recall (also known as sensitivity) of the model ^[241]. The MCC can also be a reliable metric as it will only produce a high score if the model prediction obtained good results for all four of the confusion matrix categories, namely true positive, true negative, false positive and false negative ^[249]. While both the F1 score and MCC are designed for use on binary datasets (i.e., where there are only two classes present), they can be applied to multiclass models through the use of ‘micro averaging’ and weighted averages. Micro averaging treats the entire data set as an aggregate result and does not consider the individual classes. In this case, all of the confusion matrix categories mentioned above would be added together for each class and used to calculate the F1 score and MCC. On the other hand, the F1 score and MCC can be calculated for each class individually and a weighted average taken, taking into account any class imbalance. Table 4-3 shows the F1 score and MCC values calculated for the optimised model using both calculation methods. A score of 1 for both of these metrics would indicate a perfect correlation. It can be seen from these scores that they all return high values, with the micro averaged F1 score equalling the overall accuracy of the model, which supports the high accuracy and suitability of the trained model.

Table 4-3: F1 scores and Matthew's Correlation Coefficients calculated from the optimised classification model

Metric	Micro Average	Weighted Average
F1 Score	0.895	0.889
Matthew's Correlation Coefficient	0.842	0.801

As new opioids belonging to any of the 3 classes are identified, their MS² data can be collected and incorporated into the model, ideally leading to an increase in prediction accuracy over time as new data is generated. Additionally, if new synthetic opioid subclasses are identified, such as the benzimidazole class recently reported by Blanckaert *et al.* [79], product ion information can be collated and the model retrained to expand the scope of compounds covered. In this way, the classification model can continue to be developed and adapted over time to stay up to date with developments in the illicit drug market.

4.4.2 Retention Time Repeatability Studies

Before retention time prediction can be attempted, it is important to establish the precision of RT values under the applied chromatographic conditions. If significant variation in the RT of a given compound is observed, the efficacy of a retention time prediction algorithm would be severely limited.

In order to first determine the repeatability of RT values without the possibility of matrix effects, mixed neat standards were analysed using the developed UHPLC chromatographic method. A representative panel of 26 different synthetic opioids, including compounds from each of the different subclasses, were chosen for the repeatability study. Sets of samples were analysed across multiple days to account for both intra and inter-day variability, including different batches of the mobile phase, and the average RT, SD, and %RSD were calculated.

Table 4-4 provides an overview of the average SD and %RSD values calculated across all the neat standards analysed. It was found that the %RSD for all the representative compounds was within 4%, suggesting a high degree of precision. Importantly, the absolute SD of all the compounds was ≤ 0.120 min. This falls within ± 0.2 min required by the Association of Official Racing Chemists (AORC) for the identification of compounds [225].

Table 4-4: Overview of the average standard deviation (SD) and relative standard deviation (%RSD) values for the repeatability studies conducted on neat standards

Compound	SD (min)	%RSD	Compound	SD (min)	%RSD
4-chloroisobuteryl fentanyl	0.082	1.3	Carfentanil	0.120	2.4
4-fluoroisobuteryl fentanyl	0.089	1.8	Cyclopentyl fentanyl	0.079	1.2
4-methoxybutyl fentanyl	0.080	1.7	Cyclopropyl fentanyl	0.084	2.2
Acetyl fentanyl	0.095	3.6	Fentanyl	0.077	2.3
Acryl fentanyl	0.084	2.5	Furanyl fentanyl	0.095	2.4
AH-7563	0.019	1.8	m-fluoro fentanyl	0.089	2.3
AH-7921/U-47700	0.066	2.1	MT-45	0.092	1.4
AH-7959	0.060	1.1	Remifentanil	0.106	3.7
AH-8507	0.056	1.1	Sufentanil	0.112	1.9
AH-8529	0.035	1.7	U-50488	0.052	1.3
AH-8533	0.050	4.0	U-51754	0.031	0.8
α -methyl fentanyl	0.073	2.0	Valeryl fentanyl	0.076	1.2
Butyl fentanyl	0.091	2.0			

While the precision of the RT values in neat standards shows promise, in order to determine the applicability to realistic samples, spiked plasma samples were analysed to determine if matrix effects would affect this. These samples were also prepared at a lower concentration (10 ng/mL) to further simulate a more realistic case scenario. In addition, an internal standard of desipramine-d₃ was included to evaluate if the use of RRT values could further improve the precision of measurements. Table 4-5 provides an overview of the average SD and %RSD values for both the absolute RT and RRT measurements. It was found that there is no significant difference in the absolute RT measurements between the neat and spiked samples, with most compounds still displaying %RSD values within 4%. When comparing the %RSD values for the absolute RT and RRT measurements, the RRT measurements showed improved precision between 0.1% and 0.5%. Based on these results, it is suggested that an internal standard is included when developing retention time prediction models, so that RRT values can be used. While the internal standard used in this study was desipramine-d₃, in practice, any compound with a suitable retention time could be used to determine RRT values for training and implementing the prediction model.

Table 4-5: Overview of the average standard deviation (SD) and relative standard deviation (%RSD) values for both the absolute retention time (RT) and relative retention time (RRT) of the spiked plasma samples (n = 49)

Compound	RT		RRT		Compound	RT		RRT	
	SD (min)	%RSD	SD	%RSD		SD (min)	%RSD	SD	%RSD
4-chloroisobutyl fentanyl	0.093	1.4	0.021	1.4	Carfentanil	0.128	2.6	0.030	2.5
4-fluoroisobutyl fentanyl	0.098	2.0	0.022	1.9	Cyclopentyl fentanyl	0.086	1.3	0.018	1.2
4-methoxybutyl fentanyl	0.087	1.8	0.019	1.7	Cyclopropyl fentanyl	0.081	2.1	0.018	2.0
Acetyl fentanyl	0.093	3.6	0.022	3.5	Fentanyl	0.072	2.1	0.016	2.0
Acryl fentanyl	0.080	2.4	0.018	2.3	Furanyl fentanyl	0.094	2.4	0.022	2.3
AH-7563	0.017	1.7	0.003	1.3	m-fluoro fentanyl	0.087	2.3	0.020	2.2
AH-7921/U-47700	0.058	1.8	0.014	1.7	MT-45	0.079	1.2	0.011	0.7
AH-7959	0.047	0.9	0.012	0.9	Remifentanil	0.112	4.0	0.026	3.9
AH-8507	0.045	0.9	0.006	0.5	Sufentanil	0.121	2.1	0.028	2.0
AH-8529	0.034	1.7	0.007	1.4	U-50488	0.036	0.9	0.005	0.5
AH-8533	0.034	2.7	0.007	2.5	U-51754	0.028	0.8	0.006	0.6
α -methyl fentanyl	0.069	1.9	0.016	1.8	Valeryl fentanyl	0.091	1.5	0.020	1.4
Butyl fentanyl	0.097	2.2	0.022	2.0					

4.4.3 Retention Time Prediction

When training a regression model, the accuracy of the model can be evaluated by considering the root mean square error (RMSE) of the output variables. The mean square error is average square of the difference between the predicted and actual target variables ^[250]. This metric, however, is measured in units that are the square of the target output, therefore the RMSE is often preferred as it is easier to interpret in the context of the developed model ^[238, 250]. Put simply, the RMSE measures the standard deviation present in the residuals of the model. The residuals display the difference (error) between the experimental value and the predicted value given by the model. The less variation there is between the experimental and predicted values, the smaller the residuals will be and, therefore, a more accurate model will give a lower RMSE.

In this study, four different regression models were evaluated to determine the best model architecture for this application. The observed RMSE values for each model type are presented in Table 4-6. Since regression models are predicting values in a continuous space, variation can be seen between the RMSE values for the same model type over different training periods. Therefore, the RMSE values presented in Table 4-6 show an average over 3 models trained on the same data.

Table 4-6: Accuracy of each retention time prediction model trained, measured by the root mean square error (RMSE)

Model	RMSE (mean \pm range)
Regression Tree	0.177630 \pm 0.0000
Support Vector Machine	0.123920 \pm 0.01810
Ensemble of Trees	0.158627 \pm 0.01540
Gaussian Process Regression	0.096608 \pm 0.018173
Gaussian Process Regression (omitted features)	0.084348 \pm 0.00000

The Gaussian Process Regression (GPR) model provided the best accuracy for the RT prediction. GPR models are non-parametric models, which use a Bayesian approach to regression problems ^[251]. One advantage of GPR approaches is that the prediction is probabilistic, meaning that the estimate for a given point contains uncertainty information as well ^[252]. It has been suggested that this type of model is suitable for complex regression problems, with high-dimensional data (i.e. large number of predictor variables) and small samples sizes ^[253].

Once evaluation of the different model types had been completed, the effect of individual predictors on the overall accuracy of the model was explored. These accuracies were then plotted against the

overall accuracy of a model trained using all predictors (Figure 4-3). It can be seen from Figure 4-3 that 5 of the predictors resulted in a large increase in RMSE when removed, indicating that they are important predictors for determining retention time. A further 3 indicators showed a smaller increase in RMSE, or no change, indicating that they do not have as large of an effect, but can still influence the accuracy of the model. For 5 of the predictors, however, namely nO, nDB, nR06, logS and Hy, it was found that the accuracy of the model improves when those features are excluded. This indicates that these features are not good predictors of retention time and may harm the overall accuracy of the model. Following this determination, a new GPR model was trained, which omitted these 5 predictors, and was determined to have an average RMSE of 0.084348. Interestingly, when training the adjusted model without these predictors, no variation in the RMSE was observed between training attempts, unlike the other models. The difference between using RRT and absolute RT values for the model output was also evaluated for the optimised model. The model returned an RMSE value of 0.4432 using the absolute RT values. The difference in the error can be explained by the absolute RT values being larger than the RRT, resulting in a larger standard deviation being measured for the residuals. When the trained model was evaluated, however, a greater variance was seen between the predicted and experimental RT values. The outputs from the model trained using absolute RT values can be found in Appendix 4 (Figures A4-2 and A4-3).

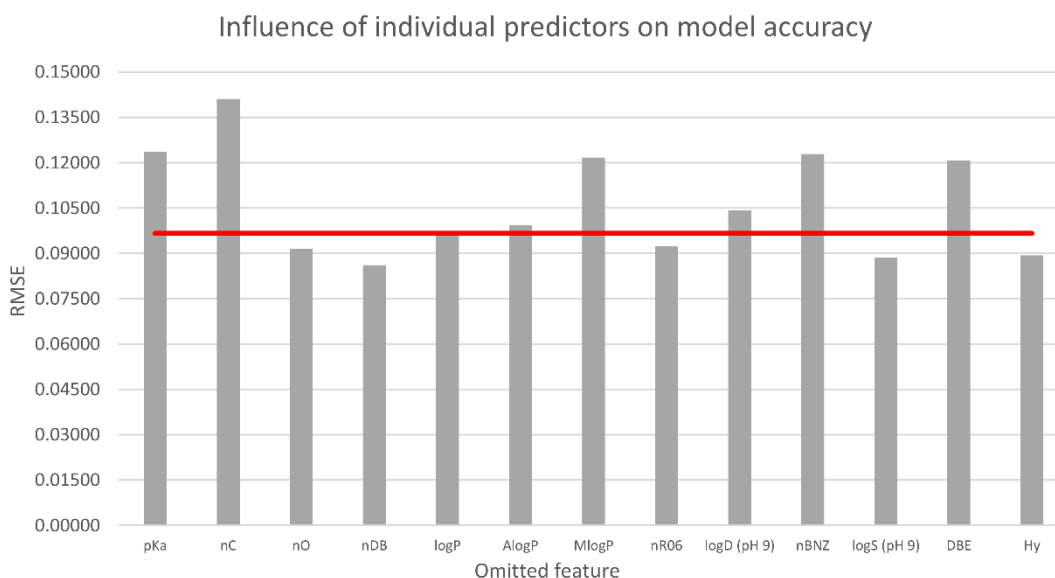


Figure 4-3: Model accuracy following removal of individual predictors in comparison to the accuracy of a model trained with all features (red). A lower RMSE indicates a higher model accuracy

Figure 4-4 presents the predicted response (in this case RRT) compared to the experimentally determined 'true' RRT to provide a visual representation of the model accuracy. The majority of the

data points are in good agreement with the regression line. The hyperparameter optimisation completed on the GPR model found that a basis function of zero and a nonisotropic squared exponential kernel function provided the best model accuracy. The sigma value used in the Gaussian processes varied slightly between the different training attempts, however the overall RMSE produced was the same. The specific hyperparameters used in the optimised model were determined automatically by the training algorithm. While these hyperparameters can be useful for comparison between similar models, they do not impact the routine implementation of the model once trained. The minimum mean square error (MSE) plot showing the change in the model error over the sequential training iterations can be found in Appendix 4 (Figure A4-4).

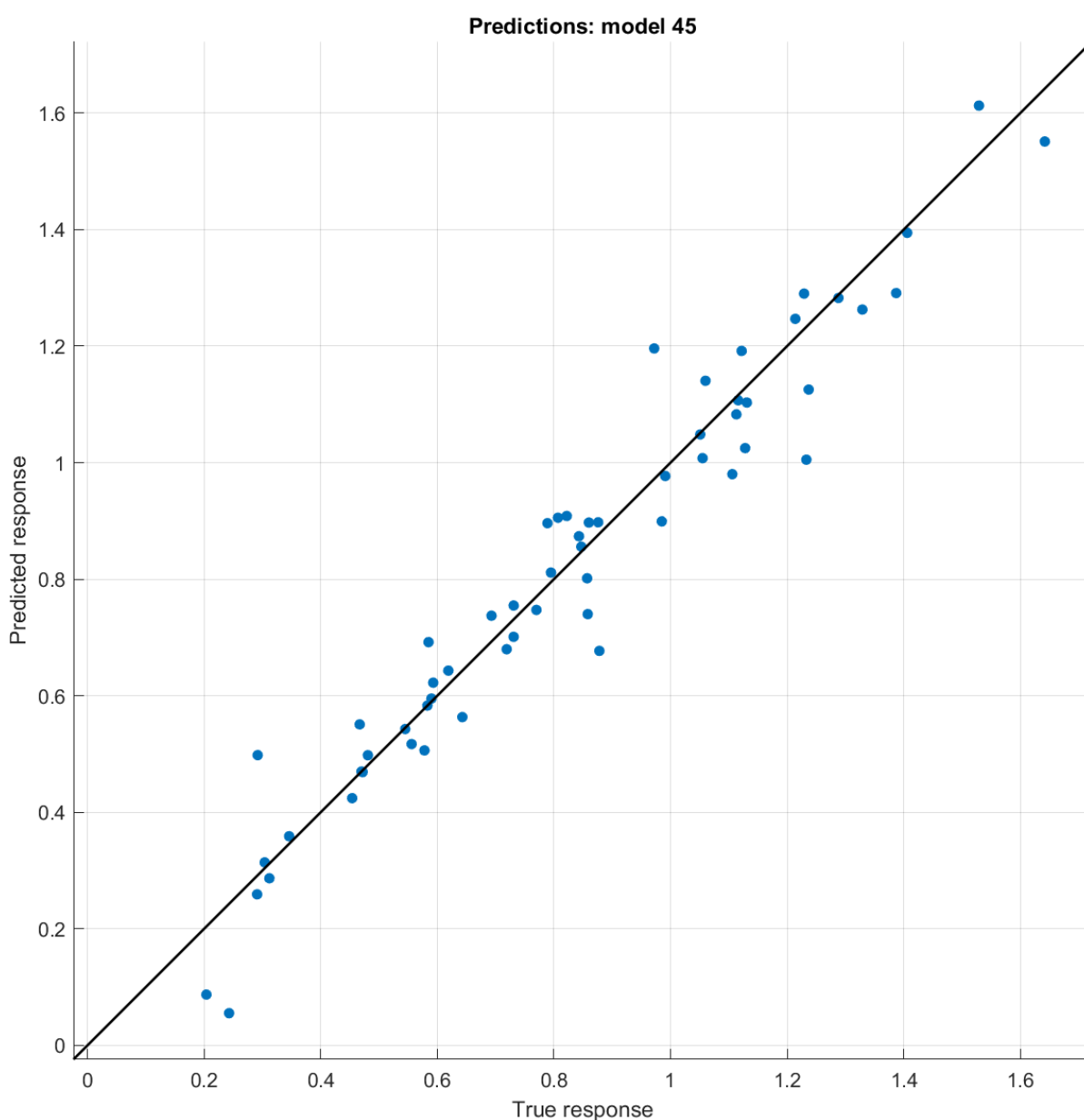


Figure 4-4: Predicted (RRT_p) vs. experimental (RRT_e) for the developed Gaussian Process Regression model

Another visual method that can be used to examine the errors present within the model is by examining the residuals produced from the predictions (Figure 4-5). When the residuals are clustered around the lower end of the y-axis, it indicates a higher model accuracy, as a residual of 0 means that the model has correctly predicted the RRT of the given sample. In the case of this model, 79.7% of the samples produced residuals within ± 0.1 . The RMSE of 0.084348 calculated for this model refers to the standard deviation of the residuals shown in Figure 4-5. This means that samples predicted within approximately ± 0.1 fall within the expected variation of the model. This further reinforces the suitability of the trained model to the application of retention time prediction. It can be seen from Figure 4-5 that some samples fall outside of the ± 0.1 threshold, indicating that the algorithm used in this model cannot correlate the RRT of the compound with the molecular features inputted. In some cases, slight structural differences may cause an increase in error, especially in circumstances where there are not many occurrences of these structural features within the training set, such as in the case of UF-17 and N-methyl carfentanil. It is therefore important that each individual laboratory assess what threshold is appropriate for their intended purpose.

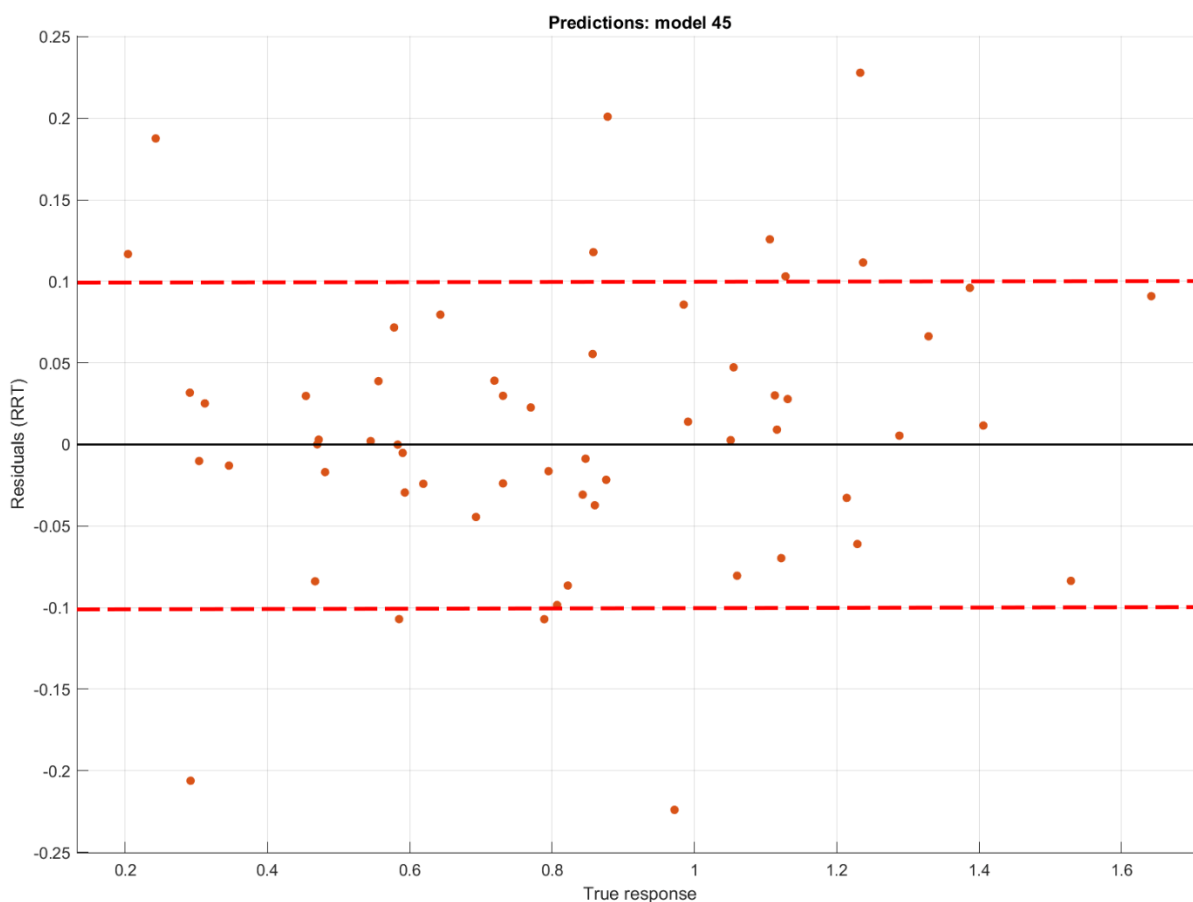


Figure 4-5: Residuals produced from the retention time prediction model

Figure 4-5 also shows that the residuals are relatively symmetrical and there are no clear patterns present. This indicates that there is no significant bias in the model. If an observable trend were displayed in the residual plot, such as an increase in error as the experimental RRT increased, it would indicate that there was a problem with the algorithm the model was using for prediction.

In the same way as the classification model, the GPR model can be re-trained with new data to expand it to include more compounds. It is important, however, that any new samples which are to be subjected to the prediction model, or used to re-train the model, are analysed using the same chromatographic conditions as the original training data. This means that if a laboratory incorporates retention time prediction into their data analysis workflow, the model used needs to be trained and optimised to suit the analytical methods being used.

It is important to note that the accuracy of the predicted retention time values will depend on the accuracy of the molecular features that are calculated. It is possible that different software packages and calculation approaches may result in slight changes in the values that are generated for the different molecular features, and this will affect the predictions generated by the model. It is necessary, therefore, for laboratories looking to implement retention time prediction into their workflows to assess the different molecular feature generation options to determine the best approach for modelling. Future studies may also evaluate the extent of the difference between various software packages and the effects that this has on the prediction modelling.

A limitation of retention time prediction modelling using molecular features is that a suspected compound identity must be known before the model can be applied. In order to generate the predictors that are used by the model, a suspected structure is needed. This does not diminish the usefulness of the model when incorporated into a rigorous compound identification workflow, however. A classification model, such as the one presented in Section 4.4.1, can be used to give a general indication of the type of compound present in the suspected sample. This can help inform further identification processes, such as the use of vendor software like *Molecular Structure Correlator* or *Compound Discoverer*, in order to putatively identify an unknown compound. Once this putative identification has been achieved, the required molecular features for the compound can be generated and input to the retention time prediction model. The predicted RRT can then be compared to the experimental RRT to provide further evidence to support the identity of the unknown. In cases where multiple possible identities are determined for a specific unknown, the molecular features of all possible identities can be generated, and the predicted RRT values compared to determine the most likely identity of the unknown compound. By using models such as these in

addition to knowing the accurate mass of the parent molecule, laboratories can perform a putative identification of an unknown compound with a higher degree of confidence, which can save time and money by limiting the purchase and/or production of erroneous CRMs.

A limitation of the models presented in this study is the lack of authentic samples with which to validate the applicability of the models. Future studies should strive to demonstrate the accuracy of models such as these on authentic administration samples of synthetic opioids. This limitation does not preclude the inclusion of these models in a complementary targeted/non-targeted screening workflow, as the intention of models such as these is to provide preliminary intelligence to assist an analyst in the identification of an unknown component within a sample.

4.5 Conclusion

The use of machine learning to assist with the identification of unknown compounds has shown significant potential. The classification model developed in this study showed a high degree of accuracy for the prediction of opioid subclasses. This model can be further developed and refined as new compounds are produced to encompass a broad spectrum of compounds. While the developed model showed some bias towards the classification of fentanyl analogues, cost matrices can be introduced to counteract this bias. These can be applied on a case-by-case basis depending on the priorities of the laboratory. Additionally, the retention time prediction model showed good correlation between predicted and experimental RRT values. The use of an internal standard to correct for any intra- and inter-day variations resulted in improved precision compared to absolute RT values alone. The developed models can be incorporated into a compound identification workflow and expanded and optimised based on the requirements of an individual laboratory.

4.6 References

1. United Nations Office on Drugs and Crime. *World Drug Report 2019*. Vienna: United Nations Office on Drugs and Crime; June 2019.
79. Blanckaert, P., et al., *Report on a novel emerging class of highly potent benzimidazole NPS opioids: Chemical and in vitro functional characterization of isotonitazene*. Drug Test. Anal., 2020. **12**: p. 422-430.
122. Pasin, D., et al., *Current applications of high-resolution mass spectrometry for the analysis of new psychoactive substances: a critical review*. Anal. Bioanal. Chem., 2017. **409**(25): p. 5821-5836.

126. Pasin, D., et al., *Characterization of hallucinogenic phenethylamines using high-resolution mass spectrometry for non-targeted screening purposes*. Drug Test. Anal., 2017. **9**(10): p. 1620-1629.
143. Noble, C., et al., *Application of a screening method for fentanyl and its analogues using UHPLC-QTOF-MS with data-independent acquisition (DIA) in MSE mode and retrospective analysis of authentic forensic blood samples*. Drug Test. Anal., 2017. **10**(4): p. 651-662.
158. Anstett, A., et al., *Characterization of 2C-phenethylamines using high-resolution mass spectrometry and Kendrick mass defect filters*. Forensic Chem., 2018. **7**: p. 47-55.
168. Margagliotti, G. and T. Bollé, *Machine learning & forensic science*. Forensic Sci. Int., 2019. **298**: p. 138-139.
169. MATLAB, *Introducing Machine Learning*, in *Machine Learning with MATLAB*. 2016, MathWorks: Online. 92991v00.
170. Mitchell, T.M., *Machine Learning*. 1997, New York: McGraw-Hill.
171. MATLAB, *Applying Supervised Learning*, in *Machine Learning with MATLAB*. 2016, MathWorks: Online. 80827v00.
172. MATLAB, *Applying Unsupervised Learning*, in *Machine Learning with MATLAB*. 2016, MathWorks: Online. 80823v00.
173. Ekins, S., *Computational Toxicology : Risk Assessment for Chemicals*. 2018, Newark, United States: John Wiley & Sons, Incorporated.
174. Luechtefeld, T., C. Rowlands, and T. Hartung, *Big-data and machine learning to revamp computational toxicology and its use in risk assessment*. Toxicol. Res-UK, 2018. **7**(5): p. 732-744.
179. Pyke, J.S., et al. *Simultaneous Targeted Quantitation and Suspect Screening of Environmental Contaminants in Sewage Sludge*. Online: Agilent Technologies;2019. 5994-0750EN.
180. Miller, T.H., et al., *Prediction of Chromatographic Retention Time in High-Resolution Anti-Doping Screening Data Using Artificial Neural Networks*. Anal. Chem., 2013. **85**(21): p. 10330-10337.
181. Mollerup, C.B., et al., *Prediction of collision cross section and retention time for broad scope screening in gradient reversed-phase liquid chromatography-ion mobility-high resolution accurate mass spectrometry*. J. Chromatogr. A, 2018. **1542**: p. 82-88.
224. Klingberg, J., et al., *Collision-Induced Dissociation Studies of Synthetic Opioids for Non-targeted Analysis*. Front. Chem., 2019. **7**(331).
225. Association of Official Racing Chemists. *AORC Guidelines for the Minimum Criteria for Identification by Chromatography and Mass Spectrometry*. Online2016. <http://www.aorc-online.org/documents/aorc-ms-criteria-modified-23-aug-16/>.

230. Klingberg, J., et al., *Finding the Proverbial Needle: Non-targeted Screening of Synthetic Opioids in Equine Plasma*. Drug Test Anal, Accepted 29th June 2020.
231. ChemAxon. *Chemicalize*. 2020; Available from: <https://chemicalize.com/app>.
232. Wei, Y.C. *PaDEL-Descriptor*. 2014; Available from: <http://www.yapcsoft.com/dd/padeldescriptor/>.
233. Lambert, J.B., *Organic structural spectroscopy*. 2nd ed. ed. 2011, Upper Saddle River, N.J: Pearson Prentice Hall.
234. *Molecular Descriptors Guide*, U.S. Environmental Protection Agency, Editor. 2008: Online.
235. Brownlee, J., *Master Machine Learning Algorithms: Discover How They Work and Implement Them From Scratch*. 2016: Jason Brownlee.
236. Bonaccorso, G., *Machine Learning Algorithms*. 2017: Packt Publishing.
237. Priddy, K.L. and P.E. Keller, *Artificial Neural Networks: An Introduction*. 2005: SPIE Press.
238. Ozdemir, S., *Principles of Data Science*. 2016: Packt Publishing.
239. Koehrsen, W. *A Conceptual Explanation of Bayesian Hyperparameter Optimization for Machine Learning*. Towards Data Science 2018 [cited 2020 24th April]; Available from: <https://towardsdatascience.com/a-conceptual-explanation-of-bayesian-model-based-hyperparameter-optimization-for-machine-learning-b8172278050f>.
240. James, G., et al., *Resampling Methods*, in *An Introduction to Statistical Learning: with Applications in R*. 2013, Springer New York: New York, NY. p. 175-201. 978-1-4614-7138-7.
241. Huilgol, P. *Accuracy vs. F1 Score*. Analytics Vidhya 2019 [cited 2020 9th July]; Available from: <https://medium.com/analytics-vidhya/accuracy-vs-f1-score-6258237beca2>.
242. Cichosz, P., *Naive Bayes classifier*, in *Data Mining Algorithms : Explained Using R*. 2015, John Wiley & Sons, Incorporated: Somerset, United Kingdom. 9781118950807.
243. Gandhi, R. *Naive Bayes Classifier*. Towards Data Science 2018 [cited 2020 9th April]; Available from: <https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c>.
244. Webb, G.I., *Naïve Bayes*, in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G.I. Webb, Editors. 2017, Springer US: Boston, MA. p. 895-896. 978-1-4899-7687-1.
245. MathWorks. *Naive Bayes Classification*. n.d. [cited 2021 31st January]; Available from: <https://au.mathworks.com/help/stats/naive-bayes-classification.html>.
246. RapidMiner. *Naive Bayes (Kernel)*. RapidMiner Documentation 2021 [cited 2021 31st January]; Available from: https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/bayesian/naive_bayes_kernel.html#:~:text=In%20contrast%20to%20the%20Naive,be%20applied%20on%20numerical%20attributes.&text=Kernels%20are%20used%20in%20kernel,expectation%20of%20a%20random%20variable.

247. Brownlee, J. *Cost-Sensitive Learning for Imbalanced Classification*. Machine Learning Mastery 2020 [cited 2021 27th January]; Available from: <https://machinelearningmastery.com/cost-sensitive-learning-for-imbalanced-classification/>.
248. Chicco, D. and G. Jurman, *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation*. BMC Genomics, 2020. **21**(1): p. 6.
249. Bekkar, M., H.K. Djemaa, and T.A. Alitouche, *Evaluation measures for models assessment over imbalanced data sets*. J. Informa. Eng. Appl., 2013. **3**(10).
250. Rajdeep, D., G. Manpreet Singh, and P. Nick, *Mean Squared Error and Root Mean Squared Error*, in *Machine Learning with Spark - Second Edition*. 2017, Packt Publishing. 1785889931.
251. Schulz, E., M. Speekenbrink, and A. Krause, *A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions*. J Math. Psychol., 2018. **85**: p. 1-16.
252. Rasmussen, C.E., *Gaussian processes for machine learning*. Adaptive computation and machine learning., ed. C.K.I. Williams. 2006, Cambridge, Mass: MIT Press.
253. Yang, D., et al., *A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve*. J. Power Sources, 2018. **384**: p. 387-395.

Chapter 5:

Non-targeted Screening Workflow

Chapter 5: Non-targeted Screening Workflow

5.1 Rationale

The techniques developed and discussed in the previous chapters should not be used in isolation, but rather should form a complementary non-targeted screening workflow. In this chapter a proposed non-targeted screening workflow using the developed methods will be discussed for the detection of unknown compounds and proposal of putative structures where no CRMs are available to the particular laboratory carrying out the testing and no publicly available spectra based on authentic samples can be found. The decisions that need to be made on a laboratory level will also be outlined. While the developed methods and workflow presented in this chapter are tailored to synthetic opioids, they could easily be adapted to suit the detection and identification of a broader range of drug classes in future.

5.2 Implementation of Non-targeted Screening Processes in Forensic Toxicology

5.2.1 Compound Detection

The use of a non-targeted screening approach for biological matrices should not preclude routine targeted screening and should instead be viewed as a complementary detection approach. Obviously, if a compound is detected using targeted screening, the use of the proposed non-targeted screening methods is not required. It is still recommended that laboratories update their targeted libraries with emerging synthetic opioids as they become more prevalent and CRMs become more readily available. For novel analogues, the production of CRMs and inclusion of spectral data in libraries often lags behind the preliminary identification of these compounds ^[122], meaning that alternative techniques are needed that do not rely on this information.

Even where CRMs are available, the decision of which compounds to include in a targeted screening method needs to be made by the individual laboratory, and this can be based on specific compounds of concern within a given context. Currently, the ARFL monitor fentanyl, acetyl fentanyl, carfentanil, AH-7921/U-47700 (grouped due to them being structural isomers) and U-50488, among other related compounds, as part of their routine targeted analysis. The fentanyl analogues are some of the most commonly encountered synthetic opioids and carfentanil is a compound of particular concern to the racing industry because of its ultra-high potency (approximately 100x that of fentanyl ^[91], EC₅₀ 72 nM in rat brain membrane ^[254]). The inclusion of AH-7921 and U-50488 can also be traced to their importance to the racing industry. AH-7921 was the first and, to date, only known detection of an NSO in a post-race sample, reported by the New York Equine Drug Testing Program in 2015 ^[109]. U-50488, on the other hand, has not been reported in an authentic sample, however it was previously studied to determine the effects of κ -opioid agonists in horses ^[108]. In a similar manner, the inclusion of specific derivatives into a targeted screening method should be adapted based on the specific needs of a particular laboratory.

The use of the developed non-targeted detection methods can then be used as an orthogonal approach, ideally detecting any synthetic opioid derivatives not included in the targeted screen. Figure 5-1 outlines a proposed workflow that can be used for the non-targeted detection of opioids. Crucially, the techniques and steps outlined in the workflow focus on the 'back end' data analysis,

meaning that they can be implemented alongside currently applied analytical approaches used in toxicological laboratories.

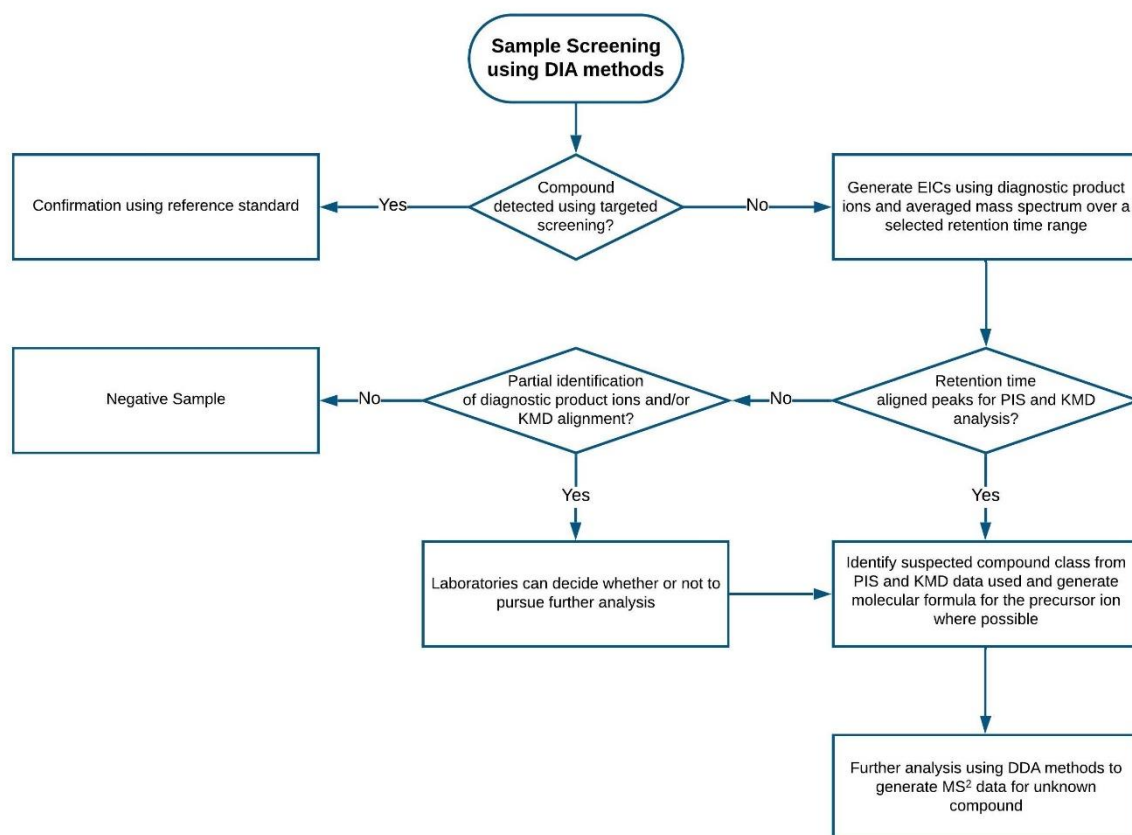


Figure 5-1: Proposed non-targeted screening workflow for detection of unknown compounds

When implementing the proposed non-targeted screening approach, EICs can be generated for the diagnostic product ions identified in the PIS method presented in Chapter 3 (Section 3.3.4.2). Further, an averaged mass spectrum can be generated for each sample. The retention time range over which this spectrum is generated can be determined depending on the chromatographic conditions used for analysis. The averaged mass spectrum can then be exported as a CSV file for submission to the *DefectDetect* program. Once a results table has been obtained from the KMD analysis, EICs of the identified masses can be generated for review. Once again, an individual laboratory can determine how many results to review, focusing on a defined number of most abundant masses in each sample or reviewing all filtered masses. Once EICs have been generated for both the PIS and KMD analysis techniques, the results from these complementary approaches can be compared. As discussed in Chapter 3, the presence of retention time aligned peaks can be used to select a sample for further analysis and identification of the unknown compound present. On the other hand, it is possible that

some compounds detected by the PIS method may fall outside of the KMD filters, such as U-50488. Alternatively, there may be non-detection of one or more of the diagnostic product ions, although the product ion/s that are detected may be supported by the KMD analysis. In cases where partial detections such as these are present, individual laboratories should determine selection criteria for which samples will be selected for further analysis.

Once the sample has been flagged for further analysis, a suspected compound class may be determined from the PIS results and KMD filter used. In some cases, it may be possible to determine the precursor ion for the unknown peak. One of the major drawbacks of using a DIA method is the production of chimeric spectra, where the software cannot directly correlate product ions to their respective precursor ions ^[122, 124]. In these cases, peaks detected using the PIS method may not have a clearly identifiable precursor ion. Where a retention time aligned peak is detected by the KMD analysis, the mass identified by *DefectDetect* should correlate to the molecular mass of the precursor ion. In situations where a precursor ion has been identified, a molecular formula can be generated from the precursor mass to assist with compound identification. Regardless of whether a precursor ion can be identified in the preliminary screening results, any selected samples should be submitted to further analysis using a DDA method to further study the precursor/product ion relationship and attempt putative identification of the unknown compound.

One of the major advantages that comes from these screening techniques is that they are vendor software agnostic. This means that, regardless of the instrument and software package that is being used by a laboratory, these methods can be applied alongside routine analysis. All data analysis software packages can be used to generate EICs, meaning that the PIS method can be implemented, provided that the instrument is capable of data-independent acquisition. In addition, the KMD analysis approach simply works off a generated mass list from a particular sample. Once again, this means that any data analysis software suite can be used. As mentioned in Section 3.4.4, these approaches are easy to implement and can be completed within 2-3 hours, depending on the size of the routine batches being analysed.

5.2.2 Compound Identification

Once again, the compound identification methods developed focus mostly on the 'back end' data analysis, however they do have some considerations that need to be taken into account when implementing the workflow. Firstly, the class prediction model requires the suspected sample to be analysed using a DDA method. This is due to the fact that it relies on the production of MS² data

specific to the unknown compound, rather than the simultaneous fragmentation of all molecules that reach the detector, as would be the case in a DIA method. Additionally, the retention time prediction model needs to be trained and optimised for the chromatographic conditions used in each laboratory. This does not necessarily mean that a new method needs to be developed and validated, but rather training data acquired under the routinely employed chromatographic conditions must be used to adapt the model before unknown samples can be submitted. In this study, the retention time prediction model was trained and optimised using a range of synthetic opioid standards, however this model is not class specific. This means that the model could potentially be trained using any number of standards across a range of different compound classes, further expanding its applicability.

Figure 5-2 presents a workflow for the proposal of putative identities for unknown compounds. When a suspected sample is analysed using a DDA method, the MS² spectra for the unknown compound can be generated. It was found during the development of the class prediction model that the 40 eV spectra provided the best classification accuracy for the optimised model. It is important, therefore, that whatever routine method is used includes this CE or the class prediction model is retrained using MS² data from an alternate CE. Once these spectra have been generated, the top 10 most abundant ions can be tabulated in order of abundance and submitted to the class prediction model. The model will then classify the sample based on the trained opioid subclasses, as discussed in Chapter 4 (Section 4.4.1). It is possible that a suspected class may already be known based on the PIS and KMD results obtained in the initial screening, however this can be further supported by the results obtained from the class prediction model. Once the predicted class is identified, a molecular formula can be generated for the precursor ion of the unknown compound and possible compound structures can be assessed with the help of software, such as *Molecular Structure Correlator* (Agilent Technologies) and *Compound Discoverer* (Thermo Fischer Scientific). While there may be numerous possible structures for a given molecular formula, having a suspected compound class can help narrow down the possibilities. Possible structures can also be compared to the fragmentation pathways displayed in the MS² data of the unknown compound to help determine a putative identity. If a possible structure that fits the defined criteria (i.e., predicted compound class) cannot be identified, the unknown compound may belong to a novel opioid subclass that has not been included in the prediction model. Alternatively, the unknown compound may not be an opioid at all, but shares some common structural features, leading to it being detected using the initial screening workflow. In either case, further analysis may be required in order to determine the identity of the unknown compound.

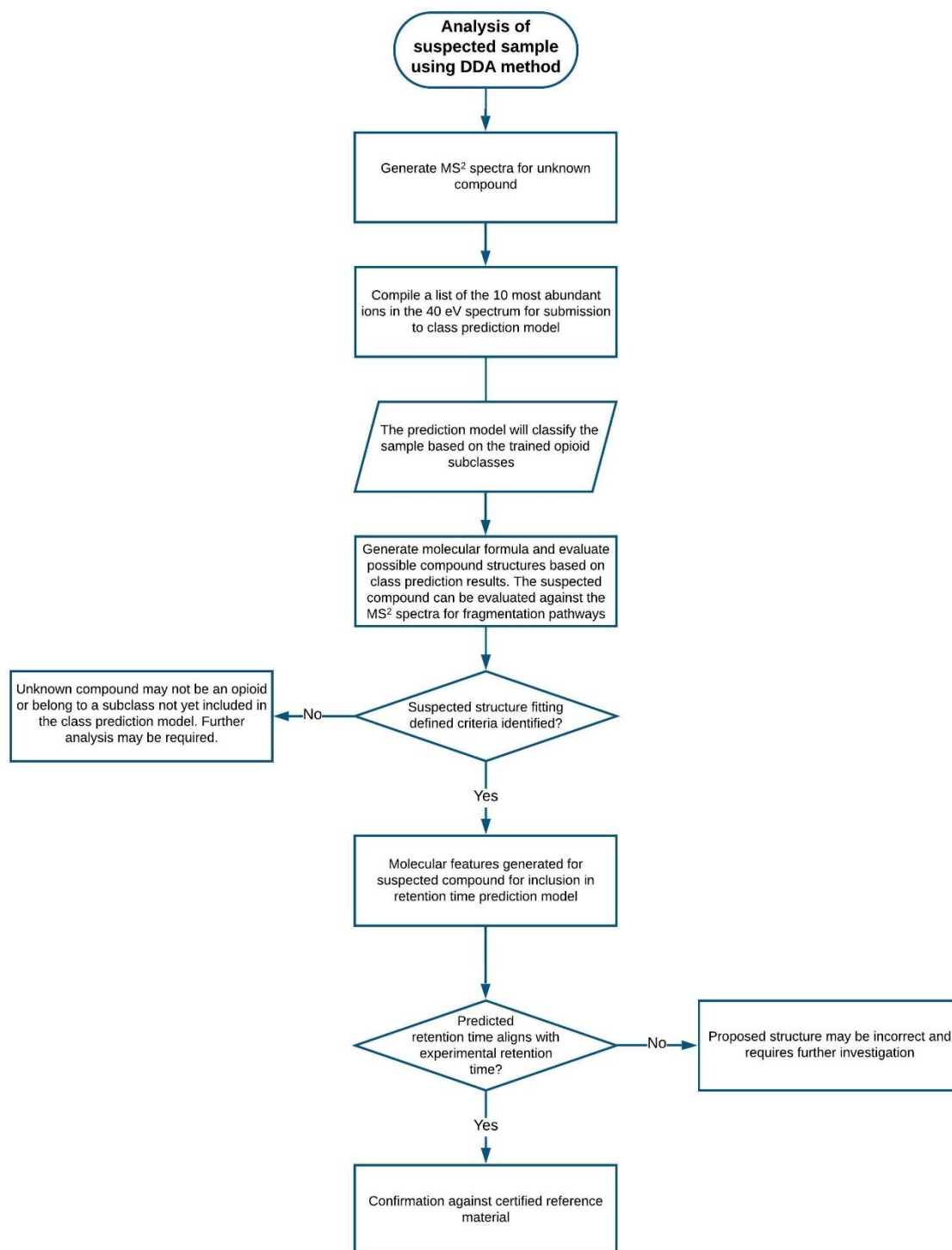


Figure 5-2: Non-targeted screening workflow for proposing putative identifications of detected unknown compounds

Once a putative identity has been determined, the molecular features identified in Chapter 4 (Section 4.3.6) can be calculated for submission to the retention time prediction model. The predicted RRT given by the model can then be compared to the experimentally determined RRT. Again, laboratories

should determine an appropriate detection window for retention time alignment. It is important to note that this model uses RRT values, and therefore it is imperative that the same internal standard is used for the training of the model and routine analyses. If there is alignment between the predicted RRT and experimental RRT, it provides further confidence in the putative identity of the unknown compound and an appropriate reference standard can be obtained for final confirmation. If, on the other hand, the predicted RRT shows a significant difference from the experimental RRT, it may indicate that the proposed putative identity is incorrect and further investigation is required. If multiple possible structures are identified following evaluation of the molecular formula, molecular features for all possible structures can be calculated and submitted to the prediction model to determine which possible compound aligns most closely with the experimental RRT. To the best of the author's knowledge, retention time prediction is a function that is not provided in any vendor-specific data analysis software packages.

Once the prediction models have been trained, their application to new data is a quick and straightforward process. The most time-consuming element of this process is the gathering of the product ion and molecular feature data for input into the models. Once the data has been tabulated, it can be submitted to the different prediction models and a result can be obtained immediately.

5.3 Conclusion

The use of a non-targeted screening workflow alongside traditional targeted analysis can enhance the capabilities of a laboratory for the detection of unknown compounds in questioned samples. This chapter proposed workflows that could be used to assist both the detection and identification of synthetic opioids, utilising methods developed throughout this thesis. While the developed methods, and hence the proposed workflows, were designed for synthetic opioids, similar approaches could be used by laboratories to improve the detection and identification of other drug classes.

5.4 References

91. Casale, J.F., J.R. Mallette, and E.M. Guest, *Analysis of Illicit Carfentanil: Emergence of the Death Dragon*. Forensic Chem., 2017. **3**: p. 74-80.
108. Kamerling, S., et al., *Dose related effects of the kappa agonist U-50,488H on behaviour, nociception and autonomic response in the horse*. Equine Vet. J., 1988. **20**(2): p. 114-118.

109. Racing Medication & Testing Consortium, *New Synthetic Opioid Designer Drug Found by New York Lab: ARCI 11/9/15*. 2015, Racing Medication & Testing Consortium: Lexington, Ky.
122. Pasin, D., et al., *Current applications of high-resolution mass spectrometry for the analysis of new psychoactive substances: a critical review*. Anal. Bioanal. Chem., 2017. **409**(25): p. 5821-5836.
124. Oberacher, H. and K. Arnhard, *Current status of non-targeted liquid chromatography-tandem mass spectrometry in forensic toxicology*. TrAC, Trends Anal. Chem., 2016. **84**: p. 94-105.
254. European Monitoring Centre for Drugs and Drug Addiction. *Report on the risk assessment of methyl 1-(2-phenylethyl)-4-[phenyl(propanoyl)amino]piperidine-4-carboxylate (carfentanil) in the framework of the Council Decision on new psychoactive substances*. Luxembourg: Publication Office of the European Union;2018.

Chapter 6:
Synthetic Route Profiling

Chapter 6: Synthetic Route Profiling

6.1 Rationale

During the production of illicit drugs, the synthetic procedure can lead to the formation of impurities and by-products. The specific impurities that are present in a given sample can be affected by the synthetic approach used in the production of the compound. By analysing the impurities found in drug samples synthesised through different routes, a chemical profile can be compiled to give analysts an indication of the method used for the synthesis of seized samples. This chapter presents a preliminary investigation into the synthetic route profiling of acetyl fentanyl. While this study considered a 'best case' scenario where purification procedures were not completed, it provides the foundations for future work to expand and refine the profiles.

All experimental work, data analysis and writing of the chapter were performed by J. Klingberg, with edits provided by A. Cawley, R. Shimmer and S. Fu.

6.2 Introduction

The use of drug profiling has been demonstrated in Australia through the AIDIP. This program, developed in 2003, was a collaboration between the AFP, NMI and the US DEA Special Testing and Research Laboratory (STRL) ^[255]. At the time, the cornerstones of the AIDIP were the heroin signature program and the cocaine signature program, which were capable of determining the geographic origin of drug samples ^[191]. In order to do this, these signature programs analysed the manufacture by-products, alkaloid impurities, solvents and even isotope ratios for cocaine ^[191]. More recently, the AIDIP has expanded into the profiling of more high volume synthetic drugs, such as methamphetamine ^[255]. The methamphetamine profiling program focuses on purity determination of samples, identification of manufacturing by-products, determination of enantiomeric composition and isotope ratio analysis to provide precursor information ^[255]. The goal of illicit drug profiling in general is to provide strategic intelligence for investigation, whether that be to localise the origin of a drug or provide links between separate seizures ^[255].

The profiling of fentanyl has been demonstrated previously by Lurie *et al.* ^[112] and Mayer *et al.* ^[204]. Lurie *et al.* ^[112] developed profiles for fentanyl samples synthesised by the original method patented by Dr. Janssen ^[110] and a method available on drug user forums, the 'Siegfried' method ^[111]. This study identified several processing impurities, including some that were specific to either route, and went on to analyse a number of seized samples and classify them by synthetic route ^[112]. Following this study, Mayer *et al.* investigated the CAS for fentanyl samples synthesised by the Siegfried method, as well as methods published by Valdez *et al.* ^[43] and Gupta *et al.* ^[87] and some hybrid methods ^[204]. This study utilised multiple, orthogonal mass spectral techniques for the analysis of the samples and introduced the use of statistical approaches to assist with the classification of samples ^[204]. Both of these studies demonstrate the potential for fentanyl standards to be differentiated based on the synthetic approach used by exploiting the differences in the manufacturing by-products that are introduced.

This work aimed to build upon the previously conducted studies and provide a preliminary investigation into the profiling of a fentanyl derivative, namely acetyl fentanyl, utilising a number of different synthetic approaches. Given the structural similarity between fentanyl and acetyl fentanyl, the impurities produced during the synthesis of the latter compound can be compared to previous work on fentanyl to determine if a combined model for the classification of both compounds is feasible.

6.3 Experimental

6.3.1 Synthetic Strategy

The synthesis of NPFs dates back to the 1960s, with the production of fentanyl itself originally patented in 1964 by Dr. Paul Janssen ^[110]. This method, however, employs the use of complicated and expensive techniques, such as hydrogenation using a precious metal catalyst. Therefore, it has been speculated that this synthetic approach is too complicated and expensive to realistically be used in a clandestine setting ^[204]. It is likely that the clandestine synthesis of NPFs would proceed through the use of an NPP intermediate. While NPP is commercially available, it has been controlled by some agencies (such as the DEA) ^[204], meaning that in-house synthesis of NPP using readily available, smaller compounds is preferred.

Three different synthetic methods were sourced from literature for the synthesis of fentanyl. Method 1 was published by Valdez *et al.* as an efficient method providing high yields ^[43]. Method 2 was described as a convenient 'One Pot' method in the original publication by Gupta *et al.* ^[87]. Finally, method 3 was sourced from a drug enthusiasts website (Erowid) and is commonly referred to as the 'Siegfried' method due to the author of the procedure ^[111]. These methods were used with minor modifications for this research. The purification procedures employed by the different synthetic routes, including column chromatography and formation of salts, were not completed to simulate a rougher, clandestine synthesis of these compounds.

The chosen synthetic procedures can be easily modified to provide other NPFs by the substitution of various reagents at different steps in the synthesis. The work published by Valdez *et al.* includes the synthesis of acetyl fentanyl in addition to fentanyl, where the product was obtained by the substitution of acetic anhydride with the propionyl chloride used in the synthesis of fentanyl ^[43]. While the Siegfried method is published for fentanyl only, the author has suggested that common analogues, such as α -methylfentanyl and 4-fluorofentanyl, can be produced by the substitution of different reagents at various stages of the synthesis ^[111].

A minor modification was made to the Siegfried method ^[111] as suggested by Mayer *et al.* ^[204]. During the second step in the reaction, the formation of the NPP N-phenyl-imine was conducted using one molar equivalent of aniline dissolved in dichloromethane (DCM), rather than using aniline as the reaction solvent.

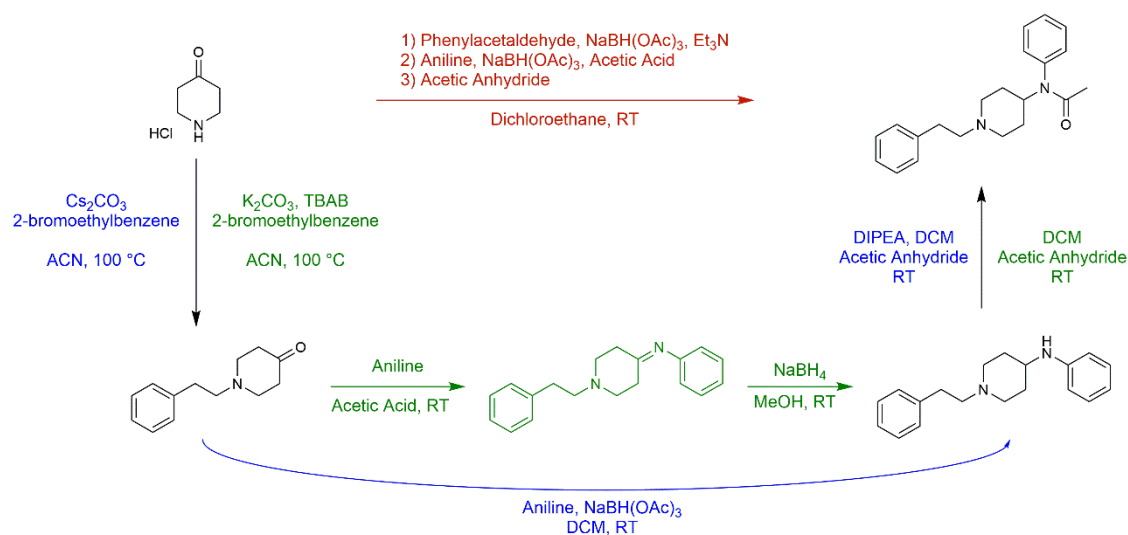


Figure 6-1: Overall synthetic route for the production of acetylfentanyl. Compounds in black are those specific to the majority of synthetic methods. Compounds specific to each method are shown in blue, Valdez^[43] (Method 1); brown, One Pot^[87] (Method 2); and green, Siegfried^[111] (Method 3). TBAB: tetra-*N*-butylammonium bromide; ACN: acetonitrile; DCM: dichloromethane; DIPEA: diisopropylethylamine.

The synthesis of acetyl fentanyl via the 3 identified methods was replicated over multiple distinct batches. These batches were prepared on different days to account for inter-day variation and incorporated different lots of the 4-piperidone monohydrate hydrochloride starting material.

6.3.1.1 Materials

Unless otherwise specified, all reagents and solvents were obtained from commercial suppliers (Sigma-Aldrich, Castle Hill, NSW) and used as received.

6.3.2 Instrumentation

6.3.2.1 Gas Chromatography – Mass Spectrometry

Samples of the different products obtained were dissolved in methanol at varying concentrations for GC-MS analysis. An Agilent Technologies (Santa Clara, CA, USA) 6890N Network GC System equipped with an Agilent 19091P-M15 column (50 m x 320 μm x 8 μm) was used for chromatographic separation. A helium carrier gas was used, with a flow rate of 10 mL/min at the inlet and 1.2 mL/min on the column. One microliter of the sample was manually injected into the injection port held at 250 °C and a 5:1 split injection used. The oven temperature was held at 50 °C for 3 min, then ramped at 10 °C/min to 250 °C where it was held for 3 min before being ramped to 280 °C at

10 °C/min and held for another 5 min. Detection was achieved with an Agilent 5973 Network Mass Selective Detector and operated in EI mode. A 3 min solvent delay was implemented, and the detector was operated in scan mode (m/z 40 – 500, 5.92 scan/s). The ionisation source and mass analyser were held at 230 °C and 150 °C, respectively.

All data acquired was processed using Agilent ChemStation software (Version E.02.02.1431). Compounds were identified based on spectral comparisons to the National Institute of Standards and Technology (NIST) Mass Spectral Library (NIST 08) with a similarity match cut-off of 85%.

6.3.2.2 *Liquid Chromatography – Mass Spectrometry*

Samples of the different synthesis products obtained were dissolved in LC-MS grade acetonitrile (ACN) to yield 1 µg/mL solutions. The samples were analysed by direct injection into an Agilent Technologies (Santa Clara, CA, USA) 1290 Infinity LC system, coupled to a 6510 QTOF-MS. All data acquisition was conducted using Agilent MassHunter Workstation (Version B.06.01).

A sample volume of 10 µL was injected onto an Agilent Technologies Poroshell 120 EC-C18 LC column (2.1 x 75 mm, 2.7 µm particle size), held at a column temperature of 60 °C, using a gradient elution with a flow rate of 0.3 mL/min and a total analysis time of 13.5 min. Mobile phase A consisted of 0.1% formic acid in ultrapure water and mobile phase B consisted of 0.1% formic acid in acetonitrile. The gradient elution method was taken from literature ^[219]. Initial mobile phase composition was 90% A, which was held for 0.5 min, before being decreased to 73% A and held for 8 min. The mobile phase composition was then reverted to the initial composition over 0.1 min with a final hold of 5 min.

The QTOF-MS was operated in ESI+ mode with capillary and fragmentor voltages of 3500 V and 175 V, respectively. An extended dynamic range of 2 GHz was used for analysis. Mass spectra were acquired using Full Scan MS, with a mass range of 50 – 1100 m/z and an acquisition rate of 1 spectra/s.

All data acquired was processed using Agilent MassHunter Qualitative Analysis Software (Version B.06.00).

6.3.3 Recursive Feature Extraction

6.3.3.1 Gas Chromatography – Mass Spectrometry

A Batch Recursive Feature Extraction workflow was conducted using Agilent Technologies MassHunter *Profiler* software. The purpose of each stage of this extraction process is explained in Chapter 3 (Section 3.5.2).

The input data range for the GC-MS data was set to retention times between 3–34 min and an m/z range of 100–750. A peak filter was added so that only peaks with a height of ≥ 100 counts would be included during feature extraction and relative peak heights of $\geq 0.1\%$ of the largest peak were required. Compounds were only extracted if they contained ≥ 5 peaks in their MS. A retention time tolerance of ± 0.3 min was allowed for compound binning. MFE post-processing filters were set with a relative peak height of $\geq 1\%$ and an absolute peak height of ≥ 2000 counts, and features needed to satisfy the filter conditions in at least 2 files across all sample files to be included as a compound group.

The extraction window used for the creation of EICs had an m/z range of ± 35 ppm and retention time range of ± 1.5 min. The relative weightings of the overall compound score were 60 for the isotope abundance score and 50 for the retention time. An expected retention time variation of ± 0.115 min was used for the Fbl algorithm. Compound groups that returned a score of lower than 75 were flagged with a '**low score**' warning. Additionally, a second warning flag was included when a second ion was expected from the molecular formula with an abundance of greater than 50 counts but was not observed in the Fbl extraction.

An Agile 2 integrator was used for matched peaks and a Gaussian smoothing function was applied to the EICs before integration. In this analysis, the algorithm was set to include average scans that were $>10\%$ of the peak height. Finally, a post-processing filter was applied to the Fbl extraction. Compound groups were required to return a target score of ≥ 50 in at least 2 files across all sample files in order to be included in the results.

6.3.3.2 Liquid Chromatography – Mass Spectrometry

Profiler cannot process data that is acquired using different ionisation techniques simultaneously. Therefore, the LC-MS data needed to be treated separately. This data was extracted using a similar

RFE workflow to that presented above, with some changes. The input data range for retention time was set to 0 – 13 min and the mass range was set to 50 – 750 m/z . The retention time tolerance for compound binning was set to ± 0.1 min and a mass tolerance was included at $\pm(50 \text{ ppm} + 2 \text{ mDa})$. The relative weightings for the MFE score were set at 100 for the mass score, 60 for the isotope abundance score and 50 for the isotope spacing score. In addition to the expected retention time variation, expected variations of $\pm(2 \text{ mDa} + 5.6 \text{ ppm})$ for the MS mass, 7.5% for the MS isotope abundance were included. As discussed in Section 3.5.2, these expected variation affect the score penalty that is applied to compounds that fall outside of these parameters. Finally, in addition to the inclusion of scans that were $>10\%$ of the peak height for the integration, a parameter was set to exclude any TOF spectra above 20% of the saturation value of the detector.

6.3.4 Statistical Modelling

The extracted data was analysed using Agilent Technologies *Mass Profiler Professional* (MPP, Version 14.0). This software package allows for differential analysis of multiple sample sets using a variety of statistical tools which can highlight the differences between samples ^[256]. For all trials run in this preliminary study, the experiment type was set to unidentified, as prior identification of all compound groups extracted by *Profinder* was not completed. When MPP is used in conjunction with data that has been extracted using a *Profinder* workflow, the statistics consider each entity as a whole. This means that all parameters identified by the RFE workflow, including retention time, mass, peak abundance, MFE scores etc., will be used to differentiate between the submitted samples. This approach allows for a more robust analysis of the differences between the overall profiles of each synthesised sample. The data processing that was completed in MPP can be broken down into 2 basic stages.

The first stage involves the normalisation of the data to be used. In this study, a \log_2 normalisation was applied across all entities submitted for statistical analysis. This step is designed to reduce any redundant data in the sample sets by grouping certain data attributes to form entities ^[257]. In this study, compounds were aligned based on RT and mass, with a RT window of ± 0.15 min and a mass window of $15 \text{ ppm} \pm 2 \text{ mDa}$. After compound alignment, the data was baselined to the median of all samples. This means that the data was centred around the median abundance of each extracted compound from all data files, and then subtracted from the same compound in each individual data file ^[256]. This baselining process ensures that compound with large abundances cannot exert a disproportionate influence on the statistical models developed ^[256].

Once the data had been aligned and baselined, it could then be viewed using a PCA plot. This plot allowed for the data spread to be viewed in a 3D space, in order to easily visualise the differences in the data sets. These PCAs were visually inspected to give a primary indication of the separation achieved between the different synthetic pathways. If all the samples were clustered together, it would not be realistically possible to predict the synthetic route based on the acquired data.

Following review of the PCAs, some classification models were created to determine the possibility of using statistical approaches to differentiate between different synthetic pathways. The MPP software suite contains 5 different prediction model archetypes that can be trained simultaneously, namely PLS-DA, SVM, Naïve Bayes, Decision Tree and ANN. These models are created and optimised by the software based on the pre-processed data and require very little user input in their initial training. Once the models were created in MPP, the prediction accuracies could be evaluated.

6.4 Results and Discussion

6.4.1 Identification of Major Impurities

While the identification of all compounds present in the samples is not required for the purpose of statistical analysis, identifying the major impurities present within the different samples may allow for some preliminary indication of synthetic route. The identification of these impurities and rationalisation by comparison to the synthetic procedure can allow for a 'first principles' classification of the samples. By identifying the major impurities present, a comparison can also be made to previous profiling studies, which have investigated the synthetic impurities in fentanyl samples, in order to identify similarities between the two analogues. For this study, major impurities were considered to be those with a high abundance in the overall sample TIC and a S/N ratio of >3 . The mass spectra library comparison for all identified impurities can be found in Appendix 5. The impurities discussed in the following sections were identified by both GC-MS and LC-MS/MS methods.

6.4.1.1 Common Synthetic Impurities

Although the identification of route-specific impurities is important for classification purposes, some compounds were identified in samples synthesised by different methods. This is not unexpected as the different synthetic approaches share many similarities, as seen in Figure 6-1. While these compounds are not particularly useful to differentiate between the different synthetic routes, it is

still important to identify these major impurities. If there is a significant impurity present in all samples, it may warrant inclusion into a targeted drug screen, such as that discussed in Chapter 5, as it may be considered a marker for the drug of interest. The structures of some of the common impurities identified can be found in Figure 6-2 below.

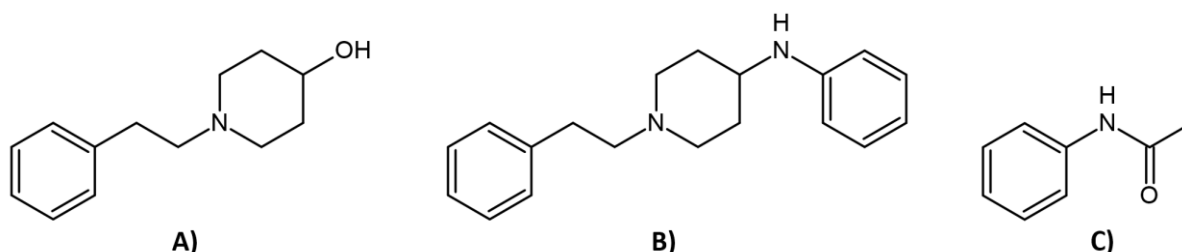


Figure 6-2: Chemical structures of some of the common impurities identified. (A) 1-phenethyl-piperidine-4-ol; (B) 4-anilino-N-phenethylpiperidine; (C) acetanilide

The first common impurity identified, 1-phenethyl-piperidine-4-ol (Figure 6-2A), arises from the reduction of the NPP intermediate formed during the first step of the synthesis. All 3 synthetic pathways used go through this intermediate, so it is not surprising that an impurity related to NPP may be found in the final product. The second step of the synthesis involves the formation of an imine before it is reduced to give the second ANPP intermediate. Therefore, it is possible that the ketone group of NPP is reduced to the alcohol found in the impurity but does not continue on to react with the aniline, resulting in the identified impurity. Interestingly, NPP itself was not a major impurity identified in the final products analysed. This indicates that the reaction with the aniline is the limiting step in this reaction, as the reduction of the ketone group is quite efficient, leaving the alcohol as an impurity instead of the NPP.

The second common impurity identified was ANPP (Figure 6-2B). As previously mentioned, this is the second intermediate formed during the synthesis of acetyl fentanyl and indicates that the final reaction step has not gone to completion. Again, since all the chosen synthetic pathways involve the formation of this intermediate, its detection in all analysed samples is unsurprising. ANPP may therefore be a compound of specific interest for including in a targeted screening method. In addition to being a common synthetic impurity, it has also been suggested that this compound can be formed as a metabolite of fentanyl ^[258]. This means that ANPP may be an important fentanyl marker from both a bulk drug and toxicological screening perspective.

The final major impurity common to all synthetic pathways was acetanilide (Figure 6-2C). This impurity can be formed from the reaction of any unreacted aniline from the imine formation with the acetic anhydride used in the final reaction step. Both the 1-phenethyl-piperidine-4-ol (Figure 6-2A) and ANPP (Figure 6-2B) were identified in the previously conducted profiling studies of fentanyl^[112, 204]. This is understandable as the only difference between the synthesis of fentanyl and acetyl fentanyl is the reagent used for the acetylation of ANPP in the final reaction step. Interestingly, the study conducted by Mayer *et al.* also identified the presence of acetanilide in addition to propionanilide^[204], which would be expected from the reaction of aniline with propionyl chloride. This indicates that acetanilide may not be a marker specific to the synthesis of acetyl fentanyl and may again be more generic to fentanyl derivatives synthesised by these reaction pathways.

Another impurity that was found in samples synthesised by methods 1 and 2 was N-ethylacetanilide (Figure 6-3). This compound is a derivative of acetanilide containing an additional ethyl moiety on the amine group. While the specific origin of this impurity has not been clearly identified, the study conducted by Mayer *et al.* identified a similar impurity, N-ethylpropionanilide, which contains an extra carbon in the acetyl chain, which would be expected when propionyl chloride is used for the synthesis of fentanyl^[204]. This provides evidence that this impurity has actually arisen from the synthetic procedure, rather than being an unrelated artefact in this analysis. Interestingly, Mayer's study also found that this impurity was most abundant in samples synthesised by the Valdez (method 1) and Gupta (method 2) methods, with only very trace amounts being found in samples synthesised by the Siegfried method. This presents an interesting opportunity for the classification of samples, where the lack of a particular common impurity may indicate the synthetic route used, rather than the presence of a route-specific impurity.

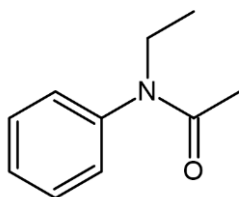


Figure 6-3: Chemical structures of N-ethylacetanilide

Two impurities common to methods 1 and 3 were 2-CEB and 2-BEB (Figure 6-4). The starting material used in both of these methods was 2-BEB, which accounts for its presence in the final product. The presence of 2-CEB can be explained by a halide exchange. The other starting material, 4-piperidone, is used as a hydrochloride salt. Since chlorine has a higher electronegativity than bromine^[259], an

exchange can occur between the bromine in the starting material and the hydrochloride salt, giving rise to 2-CEB. These findings were echoed by Mayer *et al.* [204], however the study conducted by Lurie *et al.*, which analysed samples synthesised by the Siegfried method (method 3), did not report either of these impurities [112]. If the first reaction step goes to completion or purification steps are employed, it is possible that these impurities would be removed, however it is expected that at least trace amounts would be present. Similar to N-ethylacetanilide, it is possible that the lack of either of these impurities may help provide an indication that the sample was synthesised using method 2.

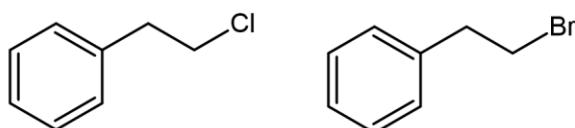


Figure 6-4: Chemical structures of (2-chloroethyl)benzene (**left**) and (2-bromoethyl)benzene (**right**)

6.4.1.2 Route-specific Impurities

In addition to the common synthetic impurities that were found in samples synthesised by multiple methods, several impurities were found to be route-specific. Figure 6-5 displays the chemical structures of some impurities found specific to method 1, the Valdez method.

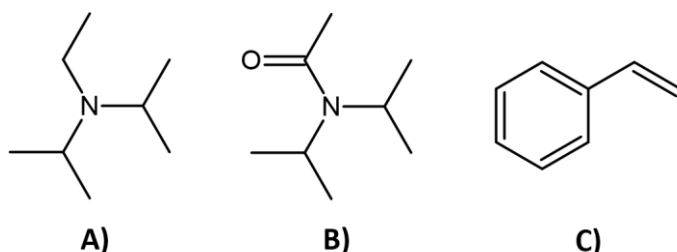


Figure 6-5: Chemical structures of some impurities specific to the Valdez method. (**A**) diisopropylethylamine (DIPEA); (**B**) N,N-diisopropylacetamide; (**C**) styrene

The first impurity identified is diisopropylethylamine (DIPEA, Figure 6-5A). This is a reagent used in the final step of the Valdez synthesis and, as such, its presence may be explained by incomplete removal during the extraction of the final acetyl fentanyl product. This particular reagent is only used in the Valdez synthesis, so it is not surprising that this impurity is specific to this method. The second impurity is an amide derivative of DIPEA, N,N-diisopropylacetamide (Figure 6-5B). While the specific source of this impurity is not categorically known, it may be formed from the oxidation of DIPEA or a

reaction between DIPEA and the acetic anhydride used for the acetylation of ANPP. This impurity was also found by Mayer *et al.* ^[204], suggesting that the oxidation reaction may be more likely as this study used propionyl chloride, rather than acetic anhydride, for the formation of fentanyl. Either way, the presence of this impurity in previously conducted studies indicates that it is indeed likely to have originated from the synthetic process rather than being present as an analytical artefact. Comparable to the results obtained in this study, Mayer *et al.* reported that N,N-diisopropylacetamide was found predominantly in samples synthesised by the Valdez route, with only a very small percentage (0.327%) found in samples synthesised by the one-pot method (Method 2).

The final route-specific impurity found for the Valdez method was styrene (Figure 6-5C). It is possible that this impurity was formed from the loss of the halogen in the starting material (2-BEB) causing the formation of the double bond. It would have been expected that this impurity would have been found for the Siegfried method as well, since both methods use the same starting material, however this was not the case. This suggests that the conditions present during the initial step of the Valdez reaction favoured the formation of this compound. Neither of the previous profiling studies identified this impurity however ^[112, 204], indicating that its formation may have been specific to the reactions conducted in this study. Styrene was found in all samples synthesised by the Valdez method, however, suggesting that its presence was not simply an artefact of the analysis. An increased sample size and further replication of the synthetic process in future studies could therefore be useful in confirming the importance of this impurity.

Figure 6-6 outlines the major route-specific impurities identified for method 2, the one pot method. The specific impurities found are all related to the starting material. The first impurity identified, phenacetaldehyde (Figure 6-6A), is the starting material itself. Any unreacted material from the incomplete first reaction step has carried through to the final product, accounting for its presence.

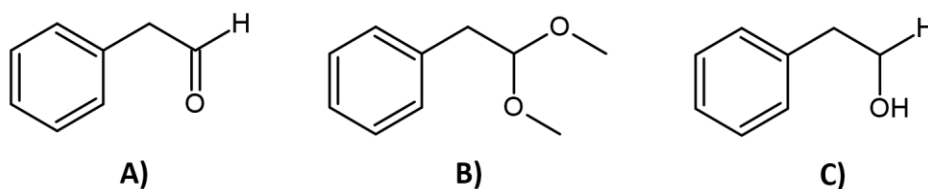


Figure 6-6: Chemical structures of some impurities specific to the One-pot method. (A) phenacetaldehyde; (B) phenacetaldehyde dimethyl acetal; (C) phenethyl alcohol

Phenacetaldehyde itself is not particularly stable for long-term storage, however, with its high reactivity often causing polymerisation or autocondensation ^[260]. Therefore, it is often obtained with protecting groups in place of the aldehyde functional group. In this study, phenacetaldehyde dimethyl acetal (Figure 6-6B) was used and the dimethyl acetal protecting groups were removed prior to the synthesis by dissolution in n-pentane before being treated with 85% formic acid. Any residual starting material from this reaction may carry through with the phenacetaldehyde and be present in the final product. The protecting groups on this compound are specifically designed to be non-reactive, so it is unsurprising that the compound would make it through the synthesis of acetyl fentanyl unchanged. Obviously, if a different starting material was used to generate the phenacetaldehyde used in this reaction, this impurity would not be present. This means that this particular route-specific impurity is somewhat context-specific, with this particular impurity only being present in cases where phenacetaldehyde has been generated from the dimethyl acetal starting material. The final major route-specific impurity identified for this synthetic pathway was phenethyl alcohol (Figure 6-6C). This impurity arises from the reduction of the phenacetaldehyde starting material, which has then not gone on to react with 4-piperidone to form the initial NPP intermediate.

Interestingly, no major route-specific impurities were identified from method 3, the Siegfried method. This does not necessarily mean that samples synthesised by this method cannot be differentiated from the other two methods. It is possible that there are trace impurities, which were not putatively identified in this preliminary study, but were present in the overall profile. These unidentified impurities may contribute to the overall variation between the different methods when statistical approaches are taken. Alternatively, the Siegfried method may contain a combination of impurities, which are individually not route-specific, but are nonetheless unique to this particular pathway when viewed together. Either way, the use of statistical analysis methods will allow these differences to be highlighted.

It is important to note that, for the purpose of this preliminary proof-of-concept study, no purification methods were employed during the synthesis of the acetyl fentanyl samples and salt formation was not completed. This was done in order to achieve a 'best-case' scenario, with respect to the impurity profile, and simulate a more rudimentary clandestine-based synthesis. Future work should build upon the foundations laid by this study and evaluate the effects that purification and salt formation have on the obtained impurity profile. Additionally, the synthesis should be carried out in replicate over a number of different days to determine the reproducibility of the impurity profile. Recently, seizures of fentanyl, especially those originating from laboratories in China, have been found to have quite high purities ^[1], with profiling programs in the United States finding some seizures to have purities

between 90 and 100% ^[261]. The higher purity of these opioid seizures originating from China, in comparison to seizures of similar drugs originating from other sources, such as Mexico ^[261], may stem from the more sophisticated clandestine laboratory setups that can be found there. If this trend of increasing purity continues, it may limit the usefulness of organic impurity profiling for providing strategic intelligence. In such a case, future research may focus more on stable isotope ratio analysis, as used in the methamphetamine profiling program as part of the AIDIP ^[255], in order to try and ascertain information about the starting materials used in the synthesis. Ultimately, it is likely that a combination of these two different approaches will provide the most complete solution to the impurity profiling of bulk drug seizures.

6.4.2 Statistical Approaches to Route Classification

While the identification of some major impurities found in the synthesised samples can assist with the classification of synthetic pathways, a statistical approach can provide a more transparent qualitative approach and provide clearer visualisation of the impurity profile differences between synthetic batches prepared using the different synthetic routes. Since *Profinder* cannot simultaneously process data that is acquired using different ionisation methods, the data from the GC-MS and LC-MS methods was treated separately. The combination of multiple techniques into a single analysis may provide a more comprehensive overview of the impurity profile. Viewing the data produced by GC-MS alone, however, can be useful as this technique is the most commonly used method for bulk drug analysis in forensic laboratories. Additionally, by utilising the orthogonal techniques individually, they can be used as a truly complementary process, where the results from one analysis can be supported by a separate technique. This may lead to increased confidence in the identified synthetic route. Specific entity lists for both the GC and LC data can be found in Appendix 5 (Table A5-1 and Table A5-2). These lists show all the entities that were used by MPP for the generation of PCA plots and construction of statistical models. It can be seen from these lists that different entities were included for each technique, meaning that they can effectively be used as orthogonal approaches.

Figure 6-7 shows the PCA results related to the GC-MS analysis of samples synthesised using different synthetic routes. It can be seen that the Valdez samples are the most clustered, while there is some significant spread between the Siegfried samples. This may be due to the fact that no major route-specific impurities were identified in the Siegfried samples and differences in the abundances of other common impurities may have caused the spread observed. It is important to note, however, that regardless of the observed spread, there is clear separation in multivariate space between the

samples synthesised by the three different methods. This is a promising initial outcome from a route classification perspective but requires further within and between batch replication with different reagent sources to further understand the extent of the variation.

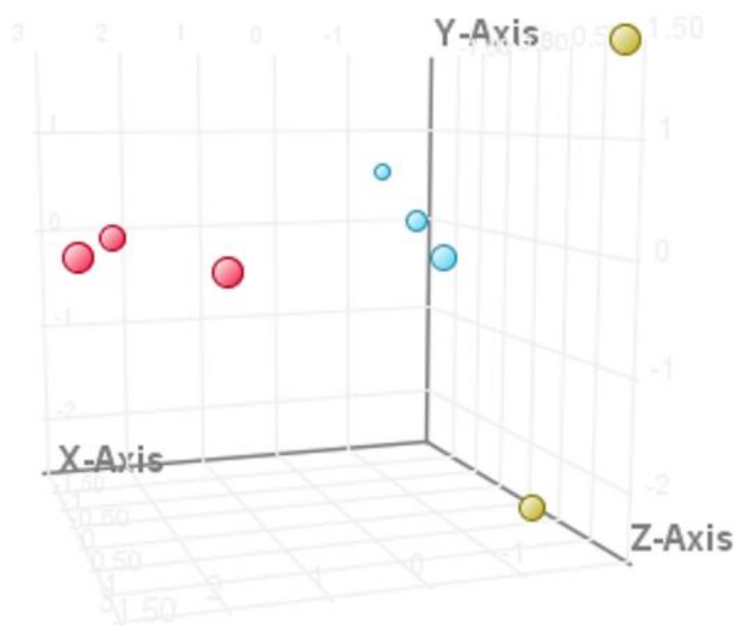
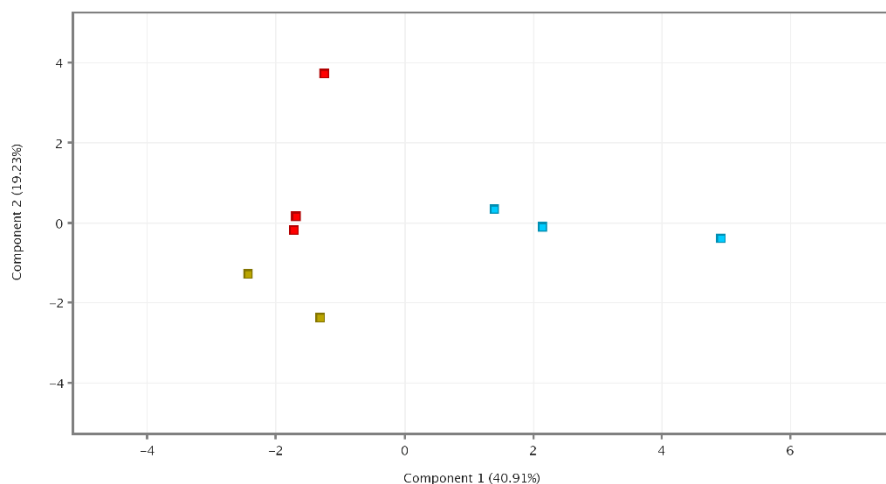


Figure 6-7: Principal component analysis showing the separation between different synthetic pathways based on GC-MS analysis for 2 components (**top**) and 3 components (**bottom**). Red: Valdez (Method 1); Blue: One-pot (Method 2); Yellow: Siegfried (Method 3). Each data point represents samples prepared in separate synthesis batches. The first 3 components accounted for 76.81% of the variance present.

The PCA loading plot for the GC-MS analysis can be found in Figure 6-8. Comparing the entities present in the loading plot with the identified impurities presented in Section 6.4.1, it can be seen

that there is a correlation present between the identified common and route-specific impurities with the separation achieved by the PCA plots, as can be expected. The three main impurities that were common to all 3 synthetic routes, namely 1-phenethyl-piperidine-4-ol, 4-anilino-N-phenethylpiperidine and acetanilide, can be seen clustered roughly around the centre of the plot as the light purple (114.2), dark blue (280.3) and dark green (135.2) spots, respectively. This is unsurprising, as these common impurities would contribute to all the synthetic routes, and therefore would not be expected to trend closer to any of the clusters. The pink spot (163.2) representing the identified N-ethylacetanilide can be found between where the Valdez and One-pot methods have grouped, meaning that this impurity may contribute to the separation of these two clusters from the Siegfried samples. Similarly, the yellow (140.1) and grey (186.1) spots, representing the (2-chloroethyl)benzene and (2-bromoethyl)benzene, respectively, indicating that they can be used to separate the Valdez and Siegfried samples from the One-pot method.

In addition to these common synthetic impurities, many of the route-specific impurities identified in the previous section can be seen on the PCA loading plot clustering around where each of the synthetic route samples clustered in Figure 6-7. In particular, the DIPEA (129.3), N,N-diisopropylacetamide (143.2) and styrene (104.2) can be found in the top left of the loading plot, and phenylacetaldehyde (120.2) and phenylethyl alcohol (122.2) can be found on the far right. These locations correlate with the clusters for the Valdez and One-pot methods, respectively, which further supports their presence as route-specific impurities. Some of the entities present in Figure 6-8 were not putatively identified as major impurities in this preliminary study, however, may have had some effect on the separation of the different samples. A similar theory was applied to the PCA loading plot obtained from the LC-MS analysis (Figure A5-13), however this plot contained a lot more minor entities that were not identified as a part of this study. Future work should focus on the identification of more of these impurities to ascertain the effect that these have on the separation of these samples.

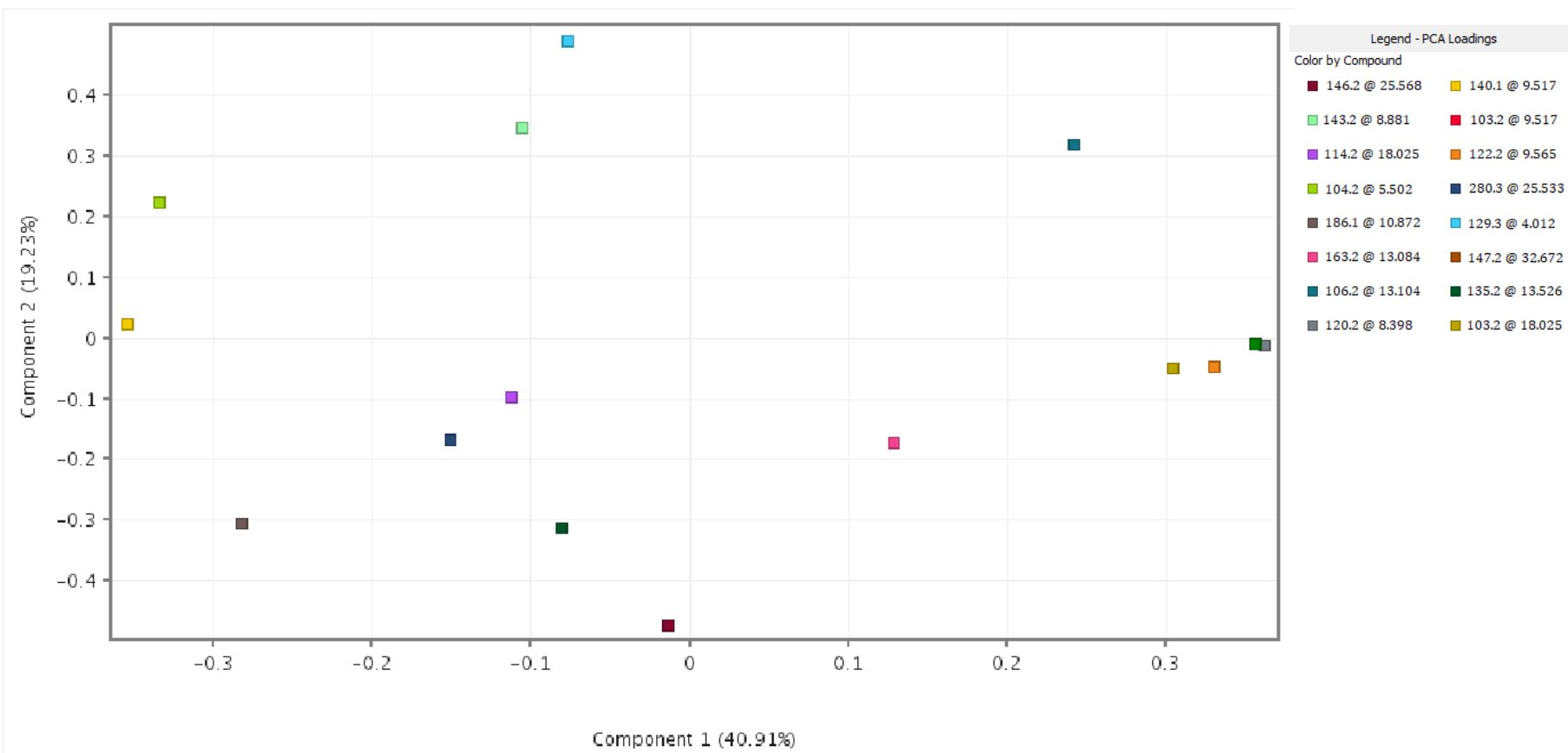


Figure 6-8: PCA loading plot showing the separation of the entities identified after GC-MS analysis

Reviewing the separation achieved from the LC-MS analysis (Figure 6-9), it can again be seen that there is some spread between the samples synthesised by the same method, however there is distinct separation between the different methods.

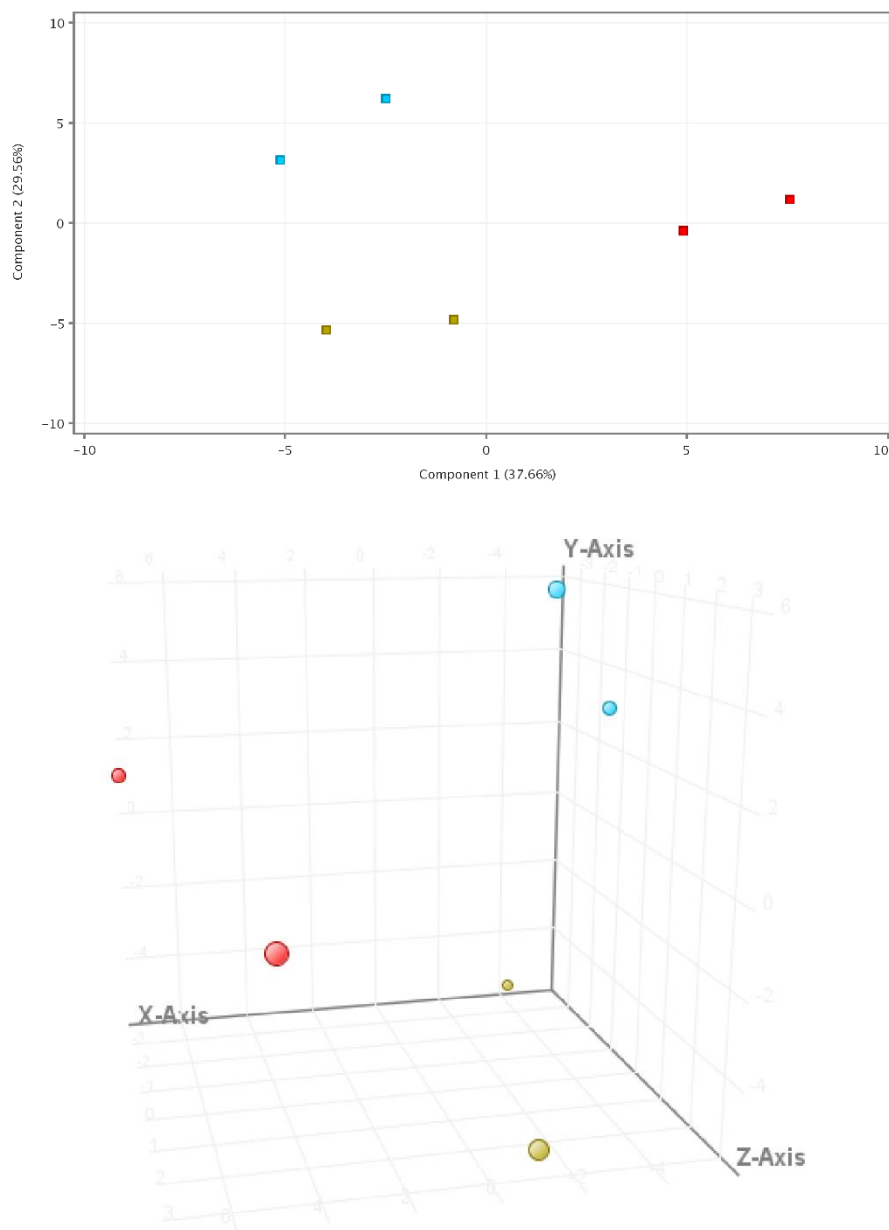


Figure 6-9: Principal component analysis showing the separation between different synthetic pathways based on LC-MS analysis for 2 components (**top**) and 3 components (**bottom**). Red: Valdez (Method 1); Blue: One-pot (Method 2); Yellow: Siegfried (Method 3). Each data point represents samples prepared in separate synthesis batches. The first 3 components accounted for 80.99% of the variance present.

After review of the PCAs, prediction models were constructed by the MPP software to provide the most accurate classification based on the inputted data. The prediction accuracies for the developed

models using the GC-MS data can be found in Table 6-1. It can be seen that the prediction accuracies of all trained models were quite high, with the least accurate model being Naïve Bayes, with an accuracy of 75%. The models trained using the LC-MS data all returned a prediction accuracy of 100%. Due to the small sample size used in this proof-of-concept study, these models were trained and evaluated using the same dataset. This is likely to cause an overestimation of the accuracy of the trained models, as evidenced by the 100% accuracy of the LC-MS models, however it may still be useful to get a rough idea of which model architectures may be most applicable to the data.

Table 6-1: Prediction accuracy for each trained model

Model Archetype	Correctly Predicted	Incorrectly Predicted	Overall Accuracy (%)
Partial Least Squares – Discriminant Analysis	7	1	87.5
Support Vector Machine	8	0	100
Naïve Bayes	6	2	75
Decision Tree	7	1	87.5
Artificial Neural Network	7	1	87.5

There were a number of limitations of this proof-of-concept study and, therefore, these results need to be reviewed in context. This proof-of-concept study had a very small sample size. In order to truly determine the accuracy of a prediction model at comparing synthetic pathways, a much larger sample size needs to be assessed. This would also allow for the inclusion of more appropriate model validation processes, such as k-fold cross-validation, as discussed in Chapter 4. This would allow for a clearer indication of the true accuracy of the trained models. Ideally, the models should also include authentic samples with a known origin, so that they can be validated under more realistic conditions. These samples could incorporate a number of different purification processes that are employed by clandestine laboratories in order to obtain a higher purity product. Conversely, authentic samples may also contain adulterants and diluents which would affect the impurity profile and, therefore, the statistical models that are created. The examination of these adulterants may also be a useful adjunct to the impurity profiling and may provide useful information for the geo-location and linking of seizures, as demonstrated in the Methamphetamine Profiling Program at NMI in Australia ^[255], however this is not relevant to the synthetic process itself. Practically, authentic samples can be difficult to source and, therefore, additional in-house synthesised samples should be included in future to further refined and optimise the models and evaluate the reproducibility of the impurity profile.

While this study has laid a foundation for synthetic route classification of acetyl fentanyl, there is still more work that could be done. The statistical analysis conducted in this proof-of-concept study shows promise for the differentiation of acetyl fentanyl samples synthesised by different methods. Future work, however, should focus on the expansion and optimisation of the developed models based on impurity profiles. The prediction models presented in Table 6-1 are trained and optimised by the MPP software with minimal user input. In order to refine these models, additional samples should be included, and model parameters should be investigated so that further optimisation can be completed to achieve the best possible prediction accuracy. Before final implementation of the models, blind trials should be completed and the models should be validated against authentic samples, where possible. As identified previously, there is a trend towards the increasing purity of seized fentanyl samples. This can limit the usefulness of the impurity profile and, therefore, future work should also investigate the effects that increasingly pure samples have on the developed prediction models.

6.5 Conclusion

The organic impurity profile of a sample can help provide information about the synthetic pathway that was used in the creation of that sample. This proof-of-concept study provided a preliminary investigation into acetyl fentanyl samples synthesised by three common synthetic pathways. Several major impurities were identified that were common to all pathways, as well as a number of impurities specific to the Valdez and One-pot synthesis methods. Initial statistical analysis showed promise, with clear separation being achieved between the different synthetic pathways from both GC-MS and LC-MS analysis. Future work should focus on further expansion and optimisation of the developed models by analysis of a larger sample set and investigation of the effects that increasing purity has on the impurity profile.

6.6 References

1. United Nations Office on Drugs and Crime. *World Drug Report 2019*. Vienna: United Nations Office on Drugs and Crime; June 2019.
43. Valdez, C.A., R.N. Leif, and B.P. Mayer, *An Efficient, Optimized Synthesis of Fentanyl and Related Analogs*. PLoS ONE, 2014. **9**(9): p. 1-8.
87. Gupta, P.K., et al., *A convenient one pot synthesis of fentanyl*. J. Chem. Res., 2005. **2005**(7).

110. Janssen, P.A.J. and J.F. Gardocki, *Method for producing analgesia*, in *Google Patents*, United States Patent and Trademark Office, Editor. 1964, N. V. Research Laboratorium, Dr. C. Janssen: United States. p. 1.
111. Siegfried. *Synthesis of Fentanyl*. 2004 [cited 2017 29th March]; Available from: <https://erowid.org/archive/rhodium/chemistry/fentanyl.html>.
112. Lurie, I.S., et al., *Profiling of illicit fentanyl using UHPLC–MS/MS*. *Forensic Sci. Int.*, 2012. **220**(1–3): p. 191-196.
191. Collins, M., et al., *Illicit drug profiling: the Australian experience*. *Aust. J. Forensic Sci.*, 2007. **39**(1): p. 25-32.
204. Mayer, B.P., et al., *Chemical Attribution of Fentanyl Using Multivariate Statistical Analysis of Orthogonal Mass Spectral Data*. *Anal. Chem.*, 2016. **88**(8): p. 4303-4310.
219. Cooreman, S., et al., *A comprehensive LC-MS-based quantitative analysis of fentanyl-like drugs in plasma and urine*. *J. Sep. Sci.*, 2010. **33**(17-18): p. 2654-2662.
255. Collins, M., *Illicit drug profiling: the Australian experience – revisited*. *Aust. J. Forensic Sci.*, 2017. **49**(6): p. 591-604.
256. Agilent Technologies. *Mass Profiler Professional Operation for Chemometric Analysis*. United States 2019. SW-MPP-3101c.
257. Mertens, B.J.A., *Transformation, Normalization, and Batch Effect in the Analysis of Mass Spectrometry Data for Omics Studies*, in *Statistical Analysis of Proteomics, Metabolomics, and Lipidomics Data Using Mass Spectrometry*, S. Datta and B.J.A. Mertens, Editors. 2017, Springer International Publishing: Cham. p. 1-21. 978-3-319-45809-0.
258. Smith, H.S., *Opioid metabolism*. *Mayo Clin Proc*, 2009. **84**(7): p. 613-624.
259. Aylward, G.H. and T.J.V. Findlay, *SI Chemical Data*. 2008: Wiley.
260. Kuhn, W., *Method for stabilizing phenylacetaldehyde*, in *Google Patents*, United States Patent and Trademark Office, Editor. 2003, Haarman & Reimer GmbH: United States. p. 3.
261. Drug Enforcement Administration. *2018 National Drug Threat Assessment*. Online: U.S. Department of Justice; 2018. DEA-DCT-DIR-032-18.

Chapter 7:

Overall Conclusions and

Recommendations for Future Work

Chapter 7: Overall Conclusions and Recommendations for Future Work

7.1 Overall Conclusions

The continual evolution of the illicit drug market and introduction of new NPS compounds poses significant challenges for forensic toxicologists. The development of innovative screening and analysis procedures is required to ensure the detection and timely identification of these novel analogues. This work found that specific subclasses of synthetic opioids exhibited class-specific cleavages when subjected to CID that give rise to diagnostic product ions. These diagnostic product ions can be exploited for the development of PIS techniques to detect unknown compounds within relevant biological matrices. These PIS techniques displayed suitable sensitivity for non-targeted screening, with an estimated screening cut-off of 0.05 ng/mL being established. In addition to the developed PIS method, data filtering using KMD values was evaluated. KMD analysis was found to be effective for a broad range of synthetic opioid compounds, with an estimated screening cut-off of 0.1 ng/mL being proposed. While these non-targeted techniques were not as sensitive as a traditional targeted screening extraction, they can be useful for the implementation of a non-targeted screening workflow. The use of an RFE algorithm was also assessed for inclusion into a non-targeted workflow. While this approach demonstrated excellent sensitivity, the extraction was non-specific for exogenous analytes, meaning that there were large amounts of data to review. From this, it was decided that this approach was not fit-for-purpose.

In addition to non-targeted detection of synthetic opioids, various machine learning approaches were investigated for the identification of unknown compounds. A classification model was optimised to predict the class of synthetic opioids belonging to the fentanyl analogues, AH series and U series. This Naïve Bayes model used MS² data generated for individual compounds, again exploiting the class-specific cleavages identified, to provide an overall prediction accuracy of 89.5%. A regression model was also trained and optimised for the prediction of experimental retention times from calculated molecular features of a given compound. It was found that a Gaussian Process Regression model provided the highest accuracy, with the optimised model displaying an RMSE of 0.084348. When these machine learning models are used as complementary techniques in a comprehensive compound identification workflow, they can provide a putative identification of an unknown compound with a high degree of confidence.

While the detection and identification of unknown compounds are important in biological matrices, the analysis of the drug samples themselves can provide useful intelligence to an analyst. To this end, this work presented a preliminary proof-of-concept study into the synthetic route profiling of acetyl fentanyl samples. From this study, it was found that a number of different impurities were identified that were common to all of the synthetic approaches used. Additionally, several impurities were identified that were specific to a particular synthetic route or found in 2 of the 3 methods. These findings were supported by previous studies conducted into the profiling of fentanyl itself. When the samples were considered from a statistical perspective, the use of PCA indicated separation in multivariate space between samples synthesised by the different methods although far greater within and between batch replication is required to allow assessment of the true within synthesis route variability of the impurity profiles. In addition, the development of classification models showed high accuracy for the synthetic route prediction of these samples. These results show promise for the use of an organic impurity profile to predict the synthetic route used in the production of seized drug samples, however, the current trends towards increasing purity of seized samples may compromise the usefulness of this information.

7.2 Recommendations for Future Work

The development of non-targeted screening strategies for illicit drugs is an on-going process. While the current work presents a comprehensive workflow for the detection and identification of a number of different synthetic opioid subclasses, these can be expanded to cover further drug classes. Whether this involves newly emerging opioids, such as the benzimidazole class recently reported, or entirely new NPS compound classes such as synthetic cannabinoids or cathinones, there is room for the expansion of the presented techniques. From an operational perspective, further automation of the developed methods can be investigated to reduce the time requirement for an analyst in the analysis of data for review.

The synthetic route profiling study conducted in this work represents a preliminary study. Therefore, the results produced by this study could be further validated by increasing the sample size to increase the confidence in the statistical models trained. By increasing the replication of the synthesis, it would also be possible to determine the variability in the impurity profile between different batches of samples. It may also be beneficial to investigate different sources of the various starting materials used in the synthetic procedure to evaluate the effect this might have on the profile obtained. This increased sample size would also allow for more rigorous validation of the statistical models and allow

for a better determination of their true accuracy. Additionally, the inclusion of purification steps in the synthetic procedure and production of salts of the final product should be investigated to determine the effects that these have on the organic impurity profile obtained as this would be more related to an operational context. Due to the increasing purity seen in many seizures of fentanyl analogues, additional analytical techniques, such as isotope ratio mass spectrometry, could be investigated to provide further information about the synthetic process without relying solely on the organic impurity profile.

Appendices

Appendix 1

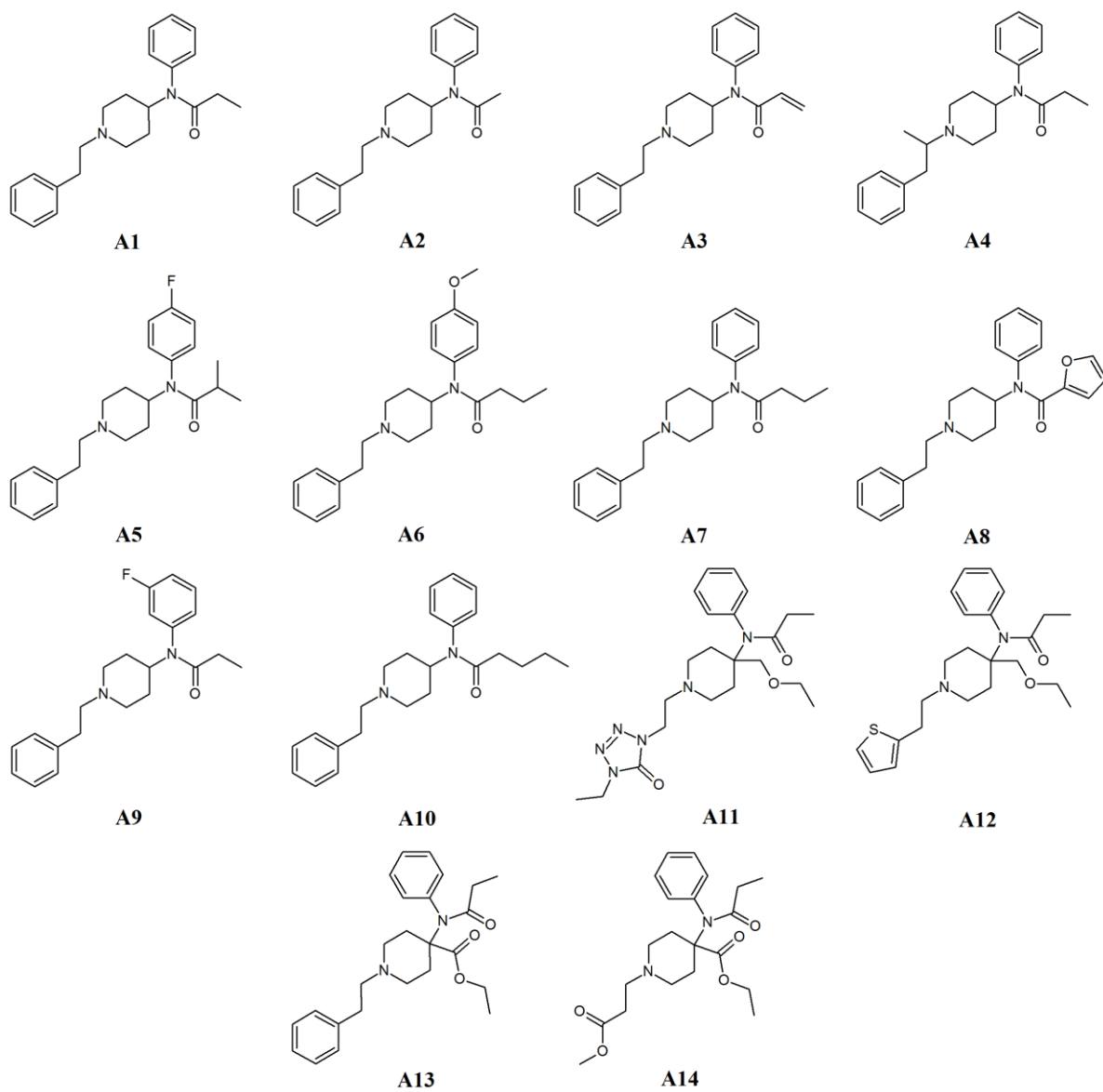


Figure A1-1: Structures of fentanyl derivatives included in the CID study

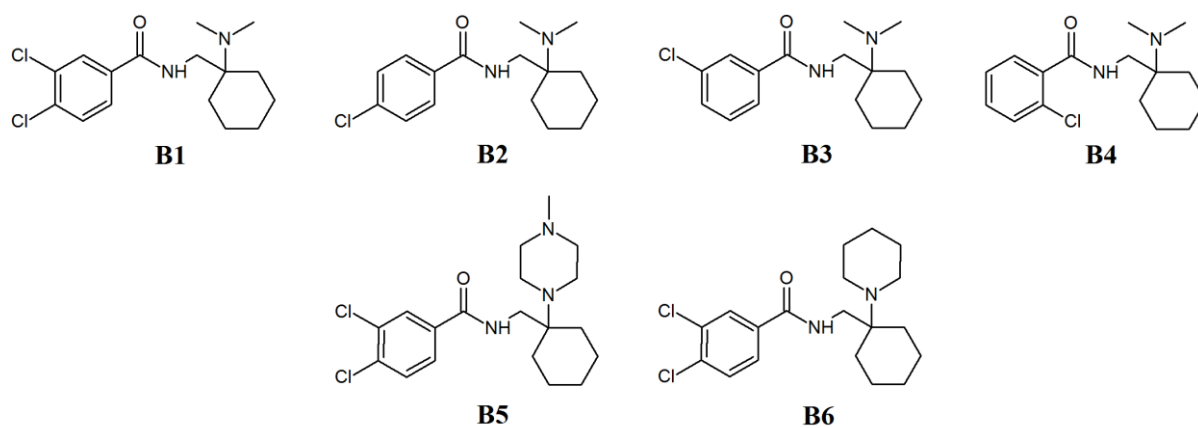


Figure A1-2: Structures of AH series opioids included in the CID study

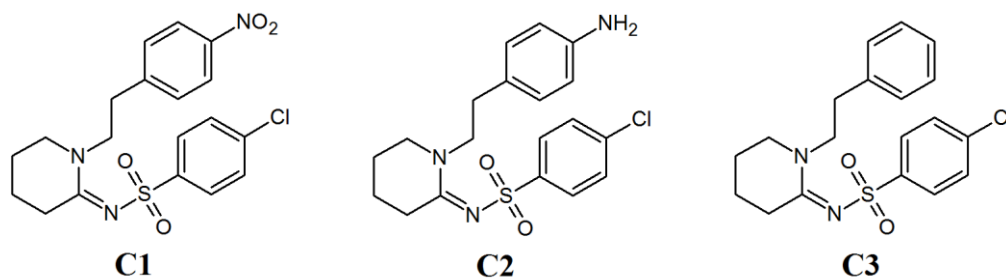


Figure A1-3: Structures of W series opioids included in the CID study

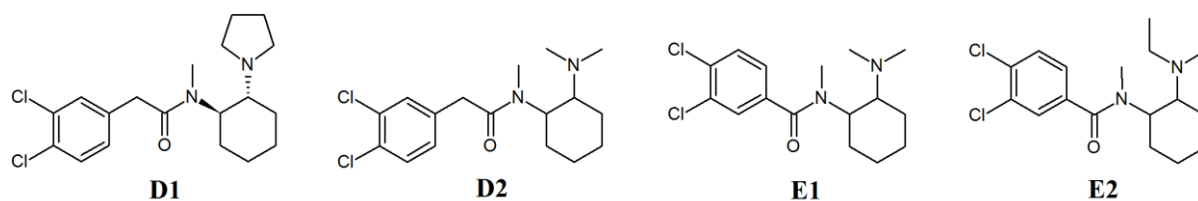


Figure A1-4: Structures of U series opioids with and without methylene spacers included in the CID study

Appendix 2

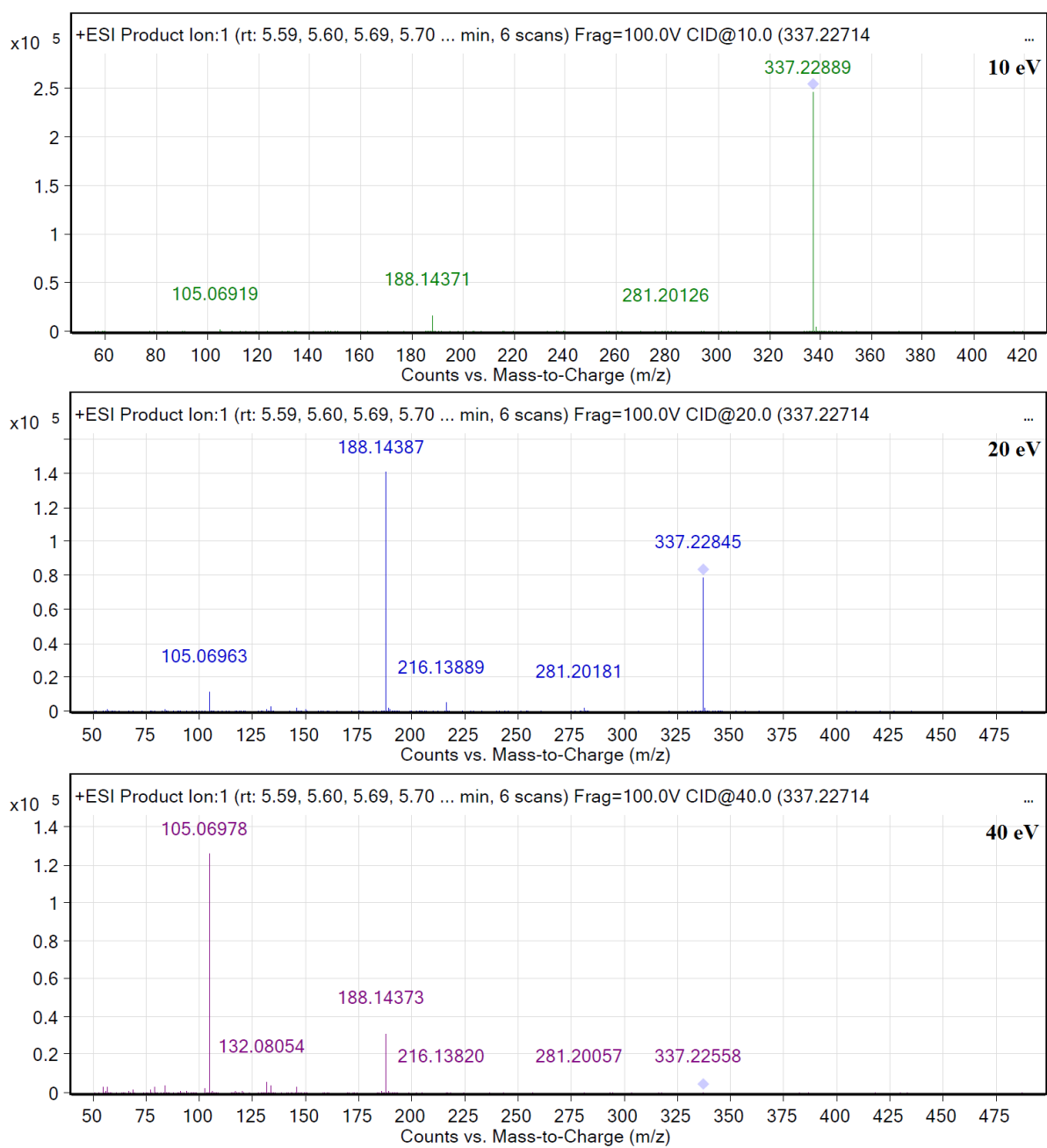


Figure A2-1: MS/MS spectra obtained from fentanyl

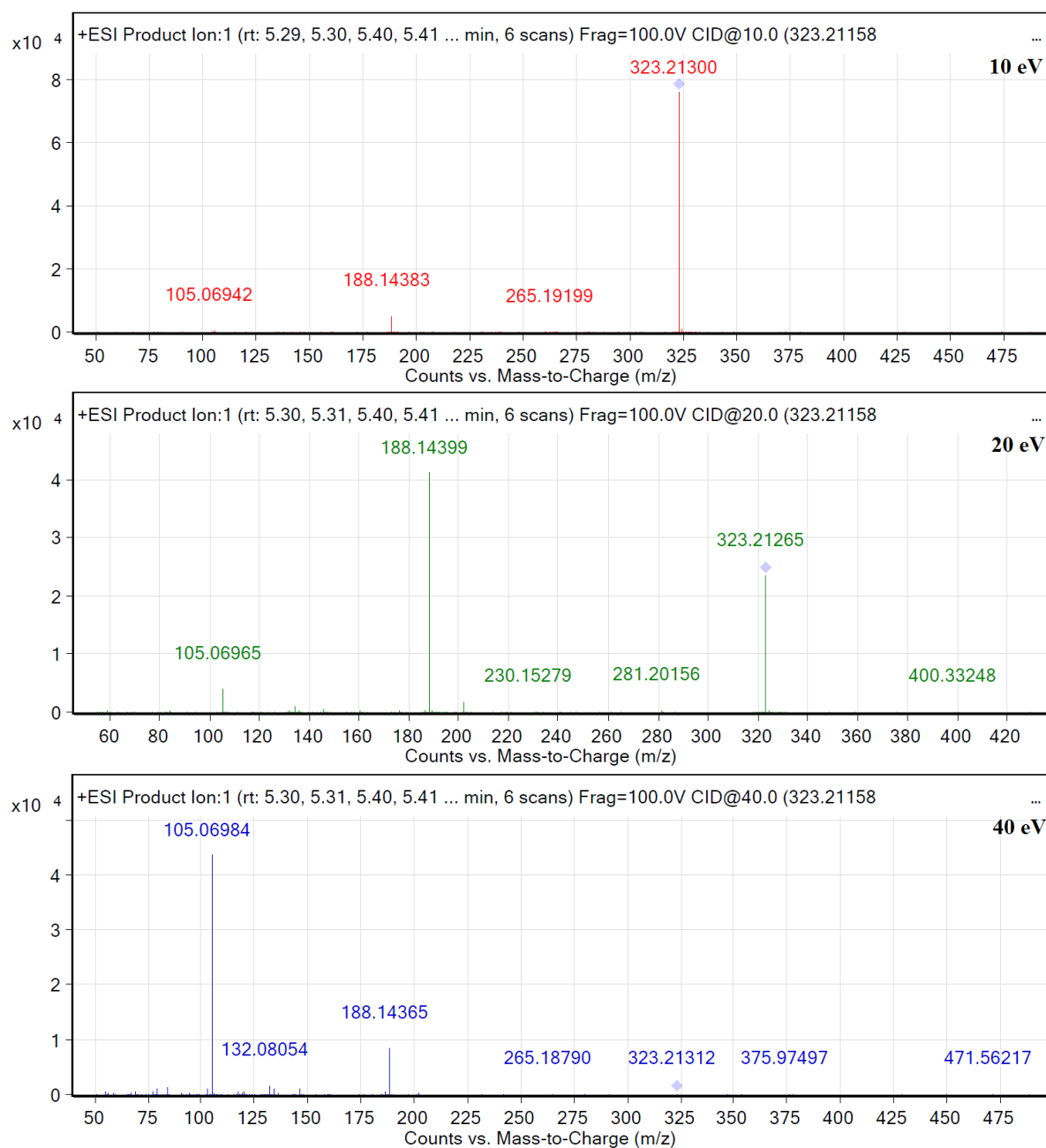


Figure A2-2: MS/MS spectra obtained from acetyl fentanyl

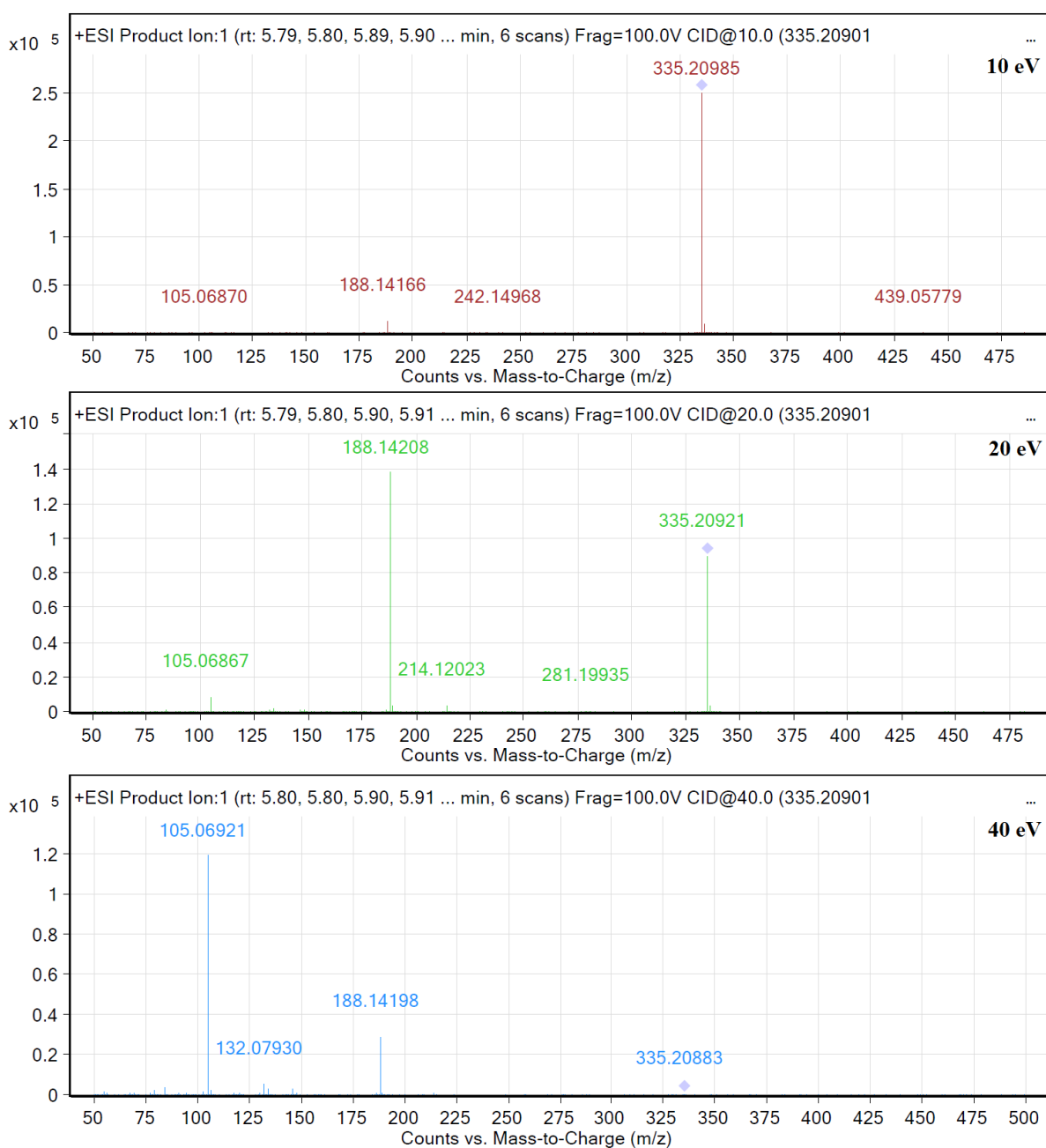


Figure A2-3: MS/MS spectra obtained from acryl fentanyl

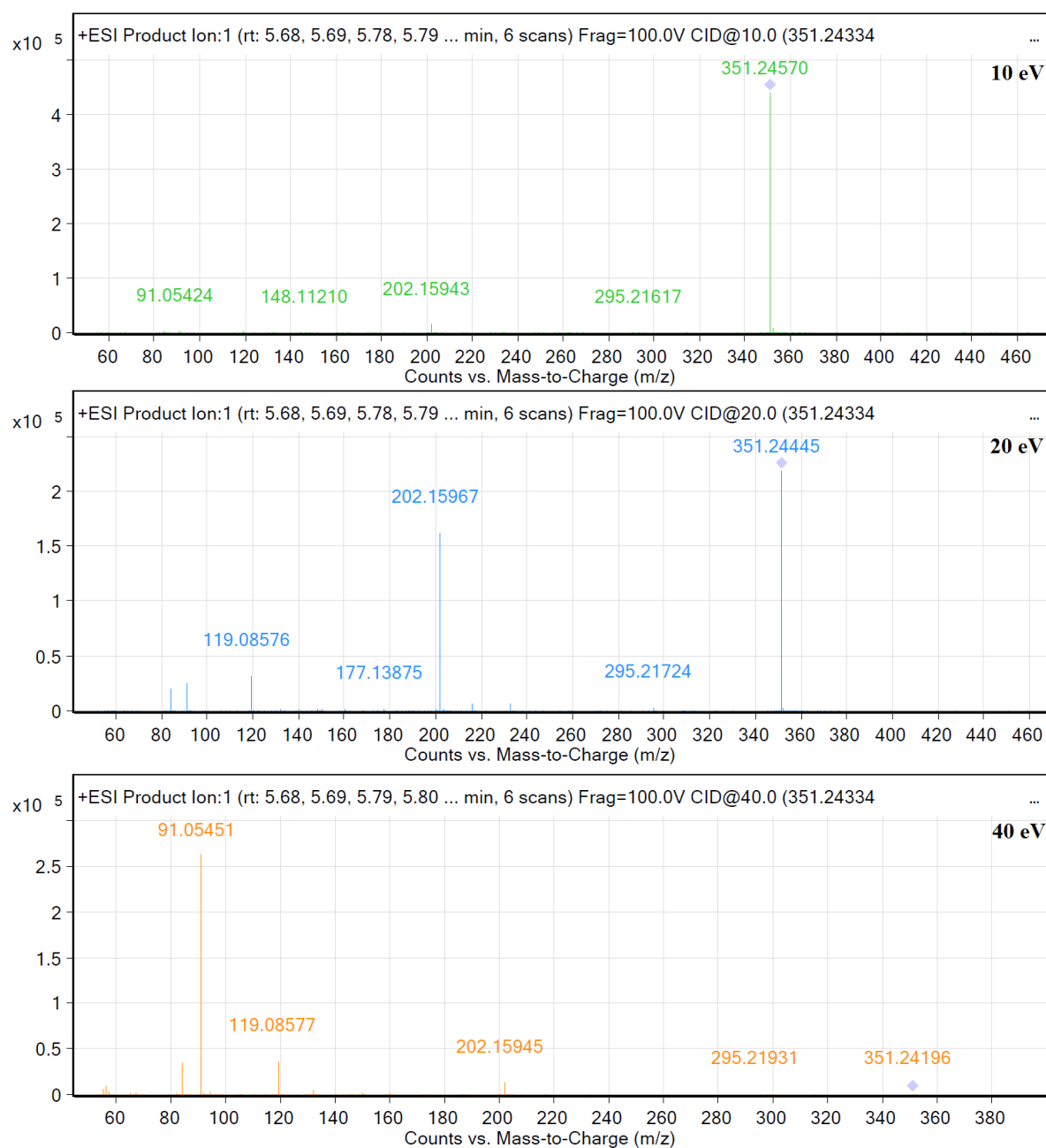


Figure A2-4: MS/MS spectra obtained from α -methyl fentanyl

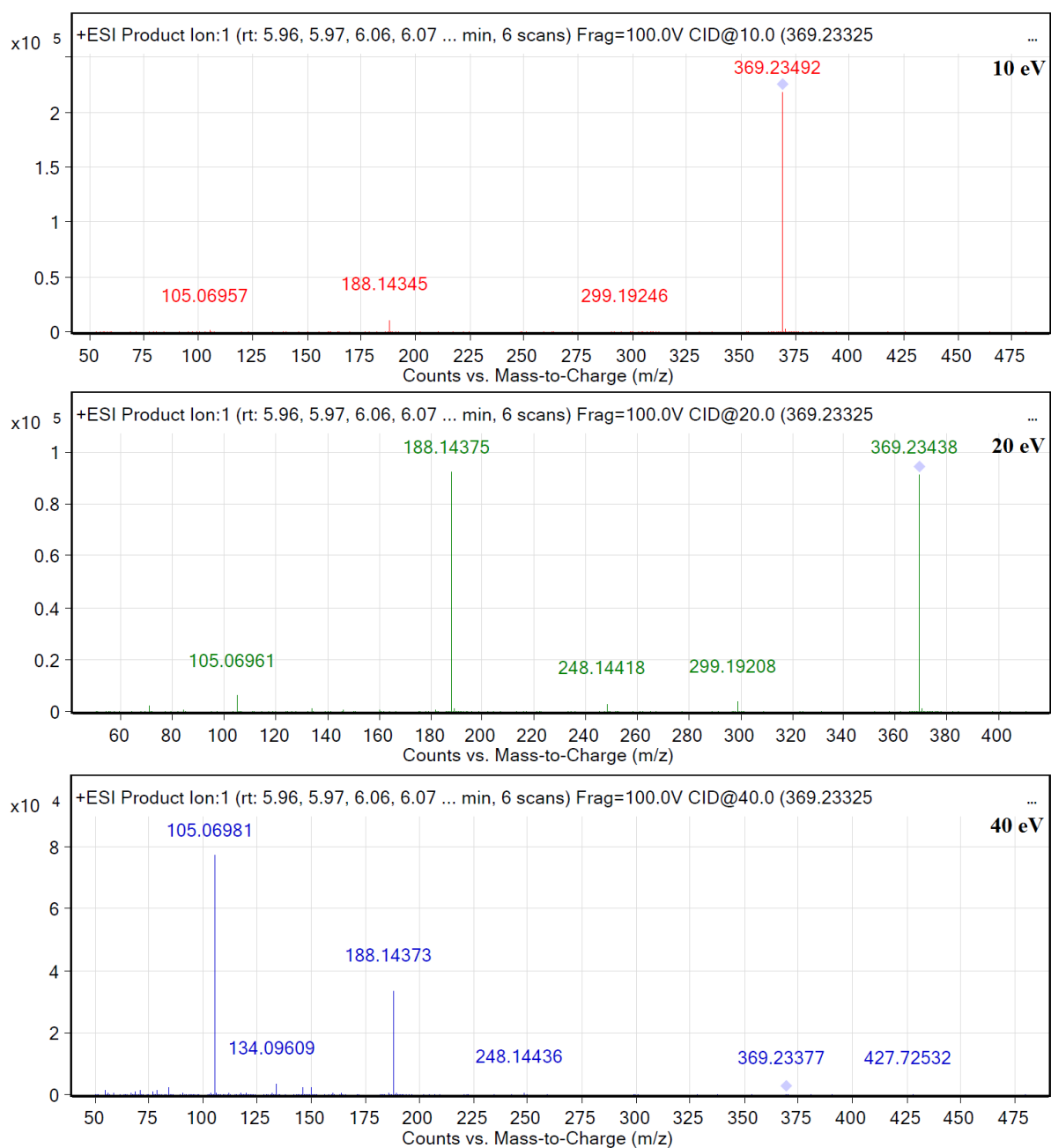


Figure A2-5: MS/MS spectra obtained from 4-fluoroisobutyryl fentanyl

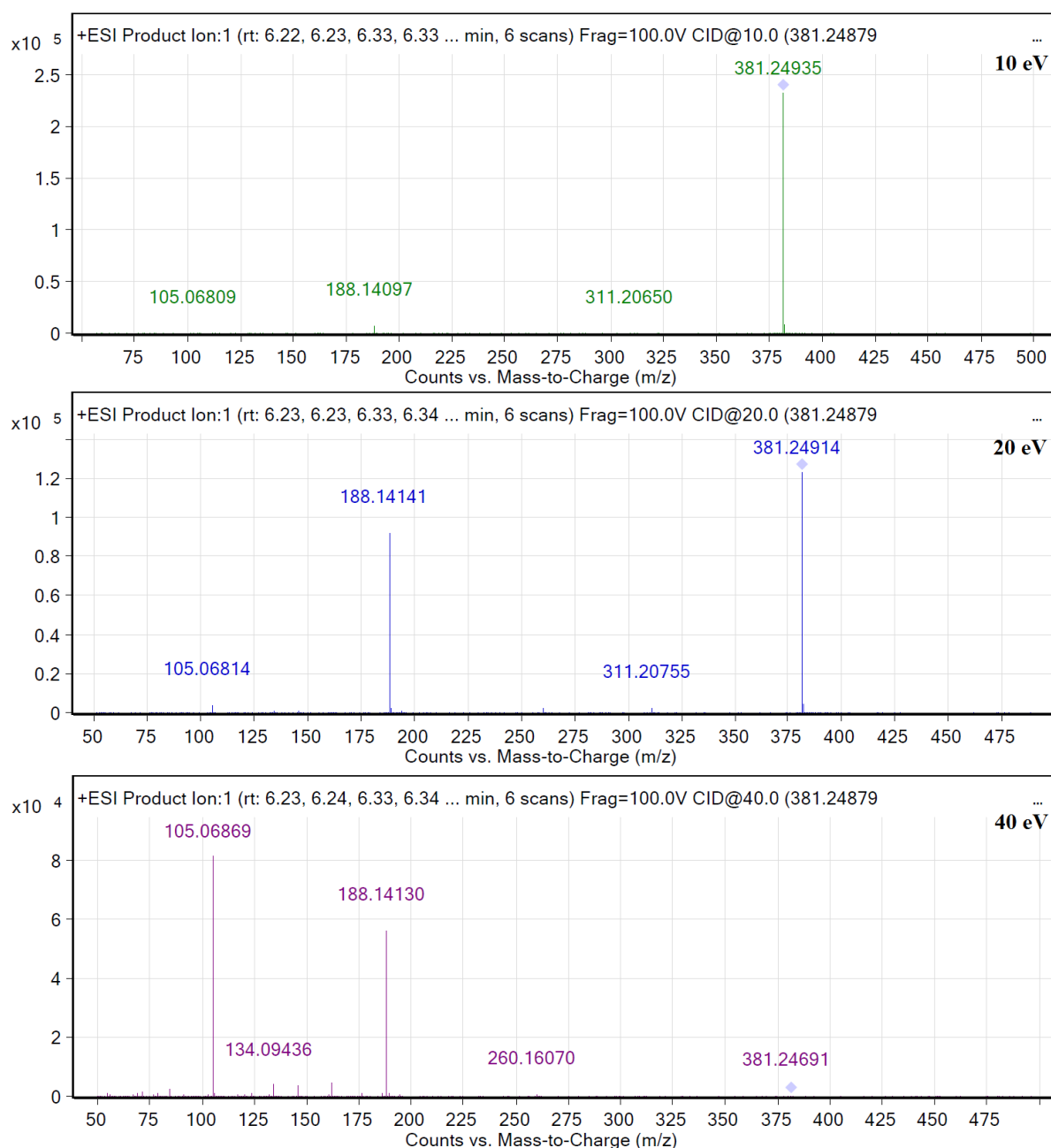


Figure A2-6: MS/MS spectra obtained from 4-methoxybutyryl fentanyl

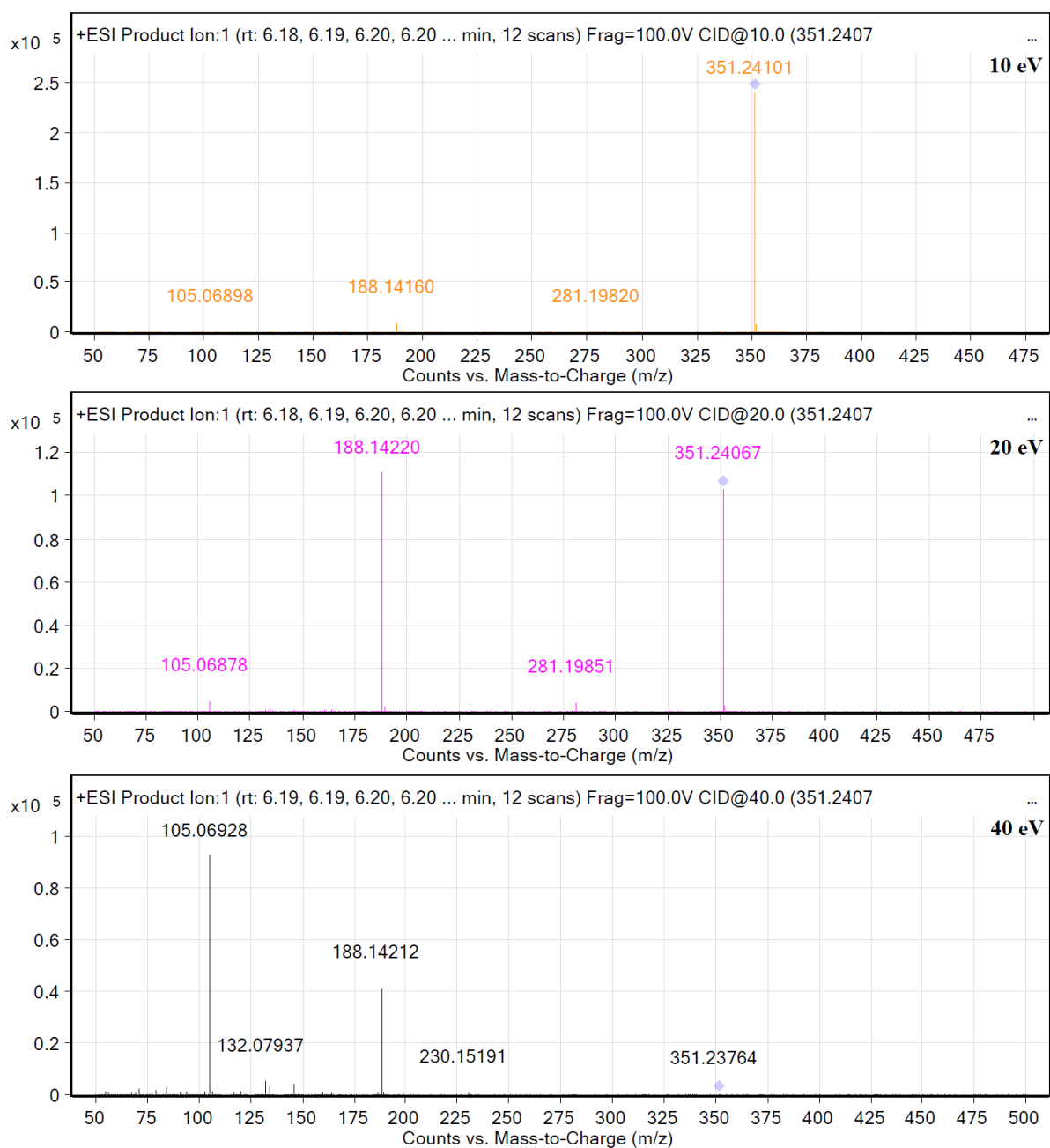


Figure A2-7: MS/MS spectra obtained from butyryl fentanyl

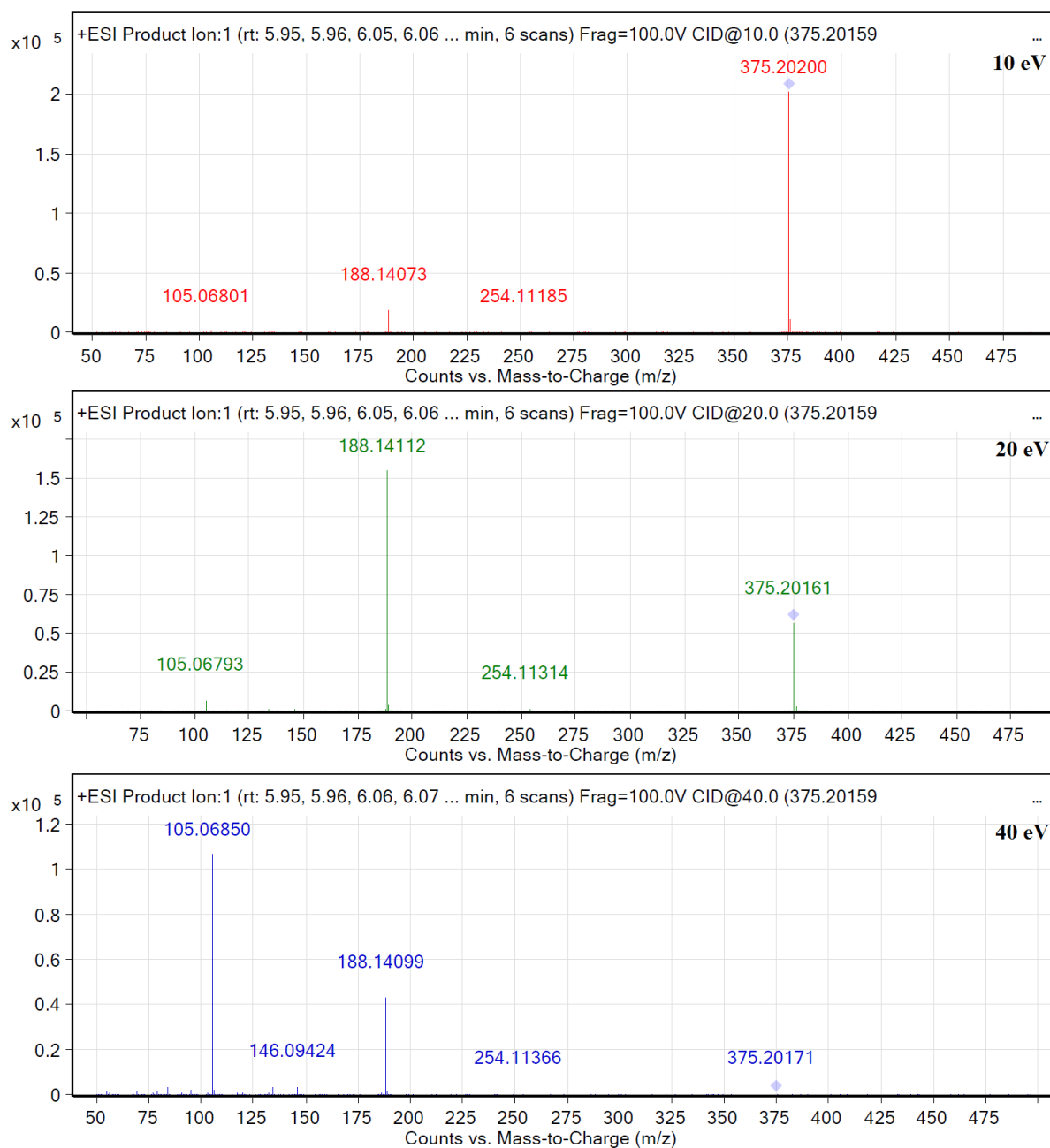


Figure A2-8: MS/MS spectra obtained from furanyl fentanyl

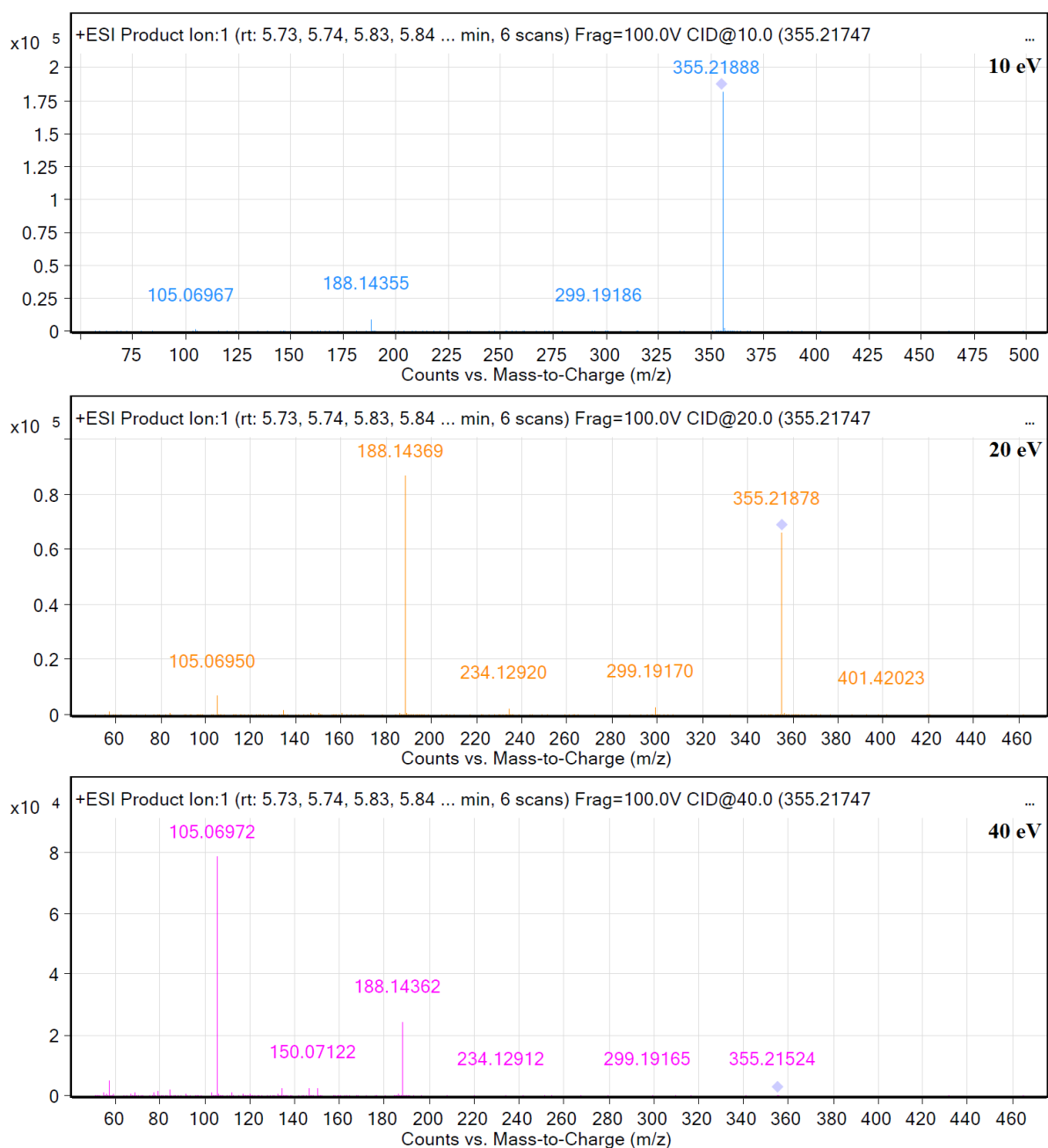


Figure A2-9: MS/MS spectra obtained from meta-fluoro fentanyl

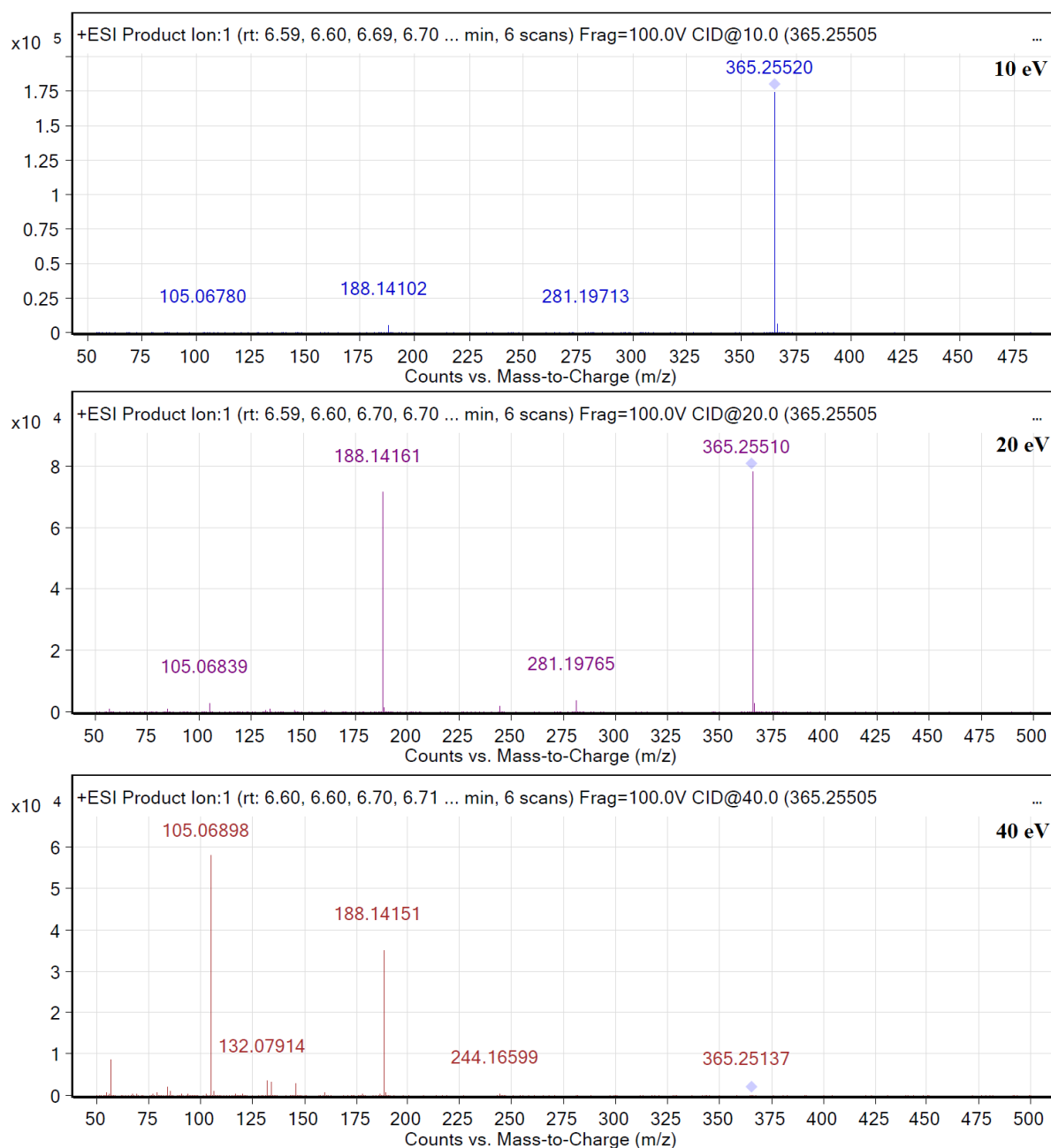


Figure A2-10: MS/MS spectra obtained from valeryl fentanyl

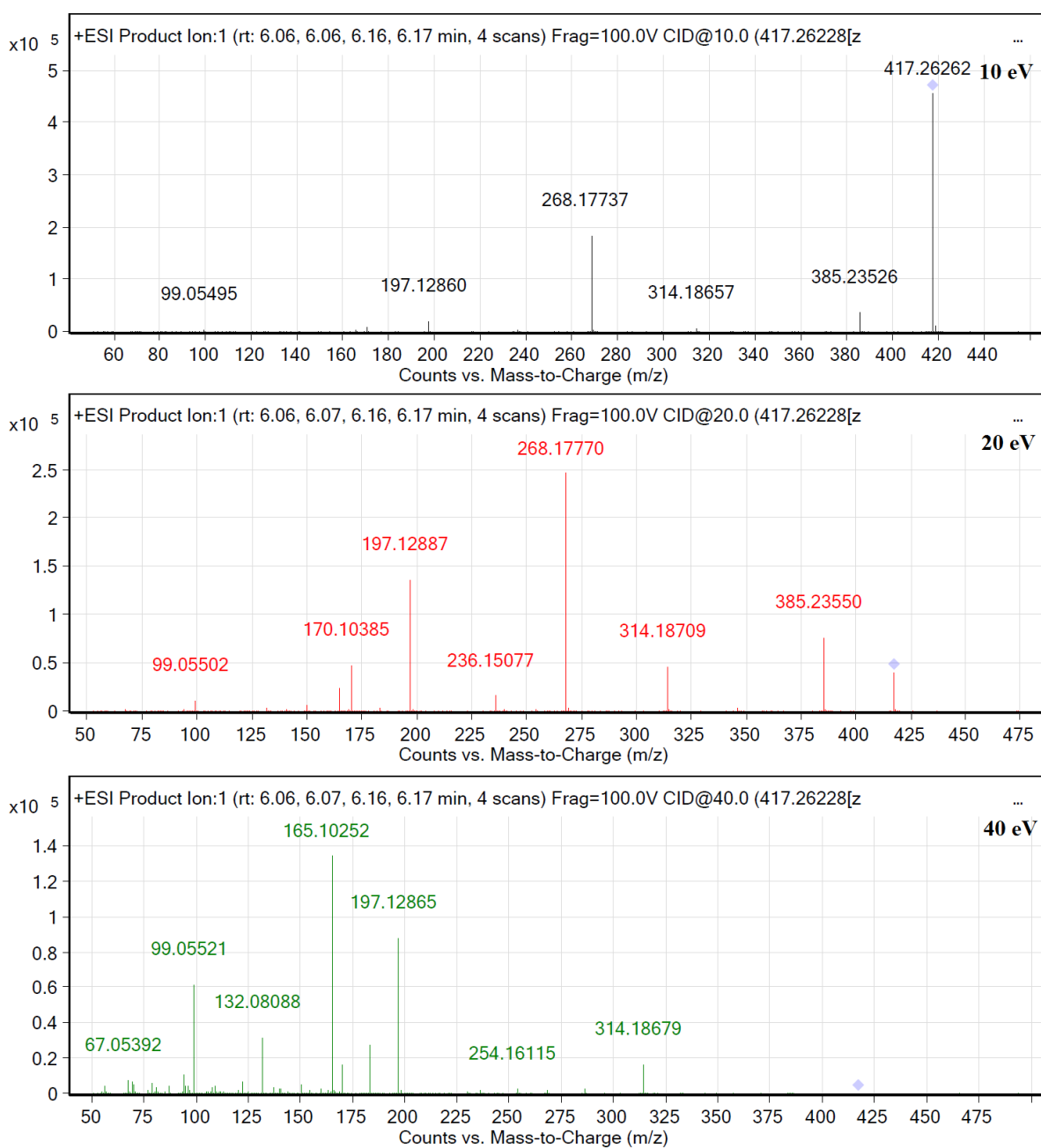


Figure A2-11: MS/MS spectra obtained from alfentanil

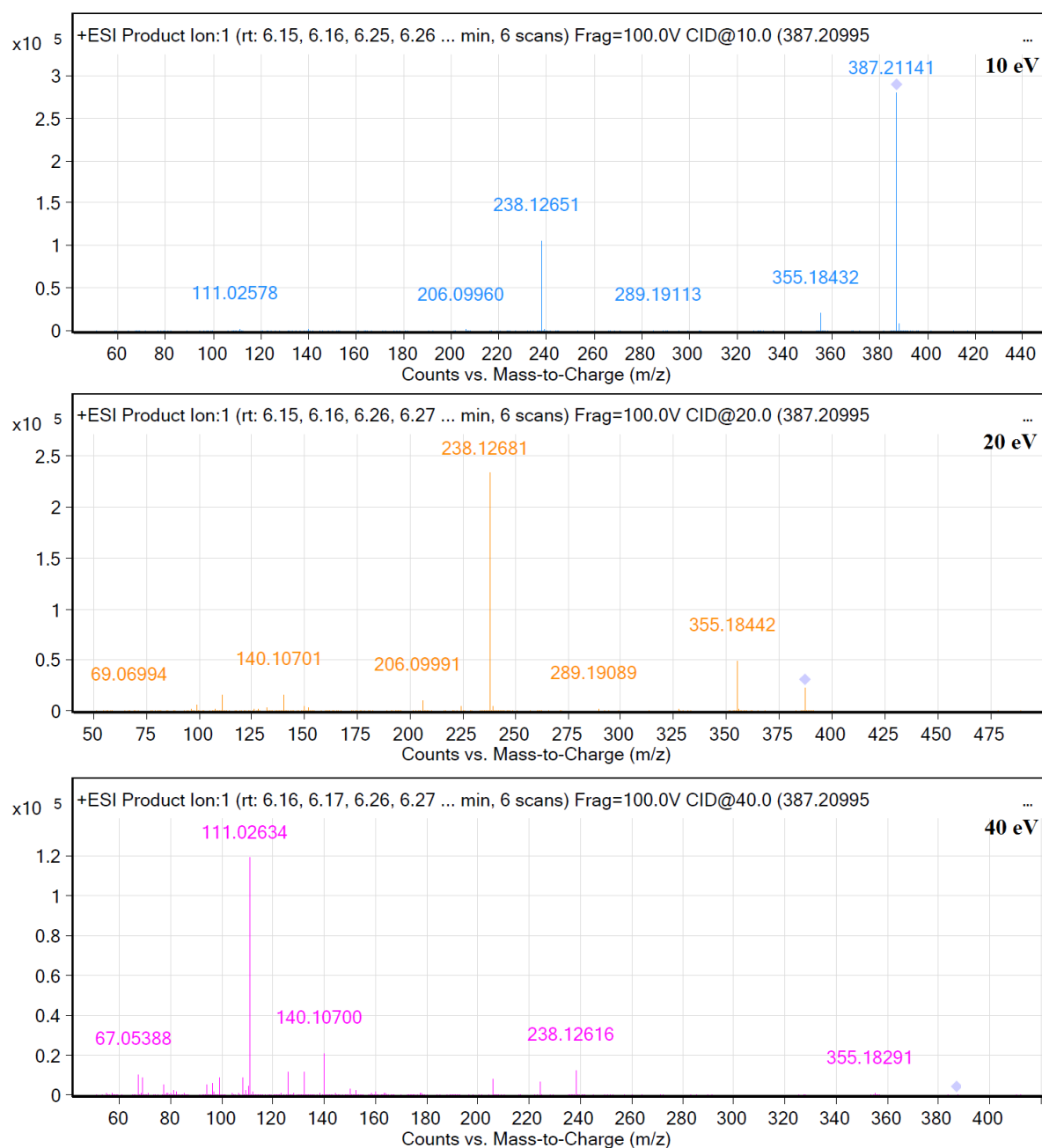


Figure A2-12: MS/MS spectra obtained from sufentanil

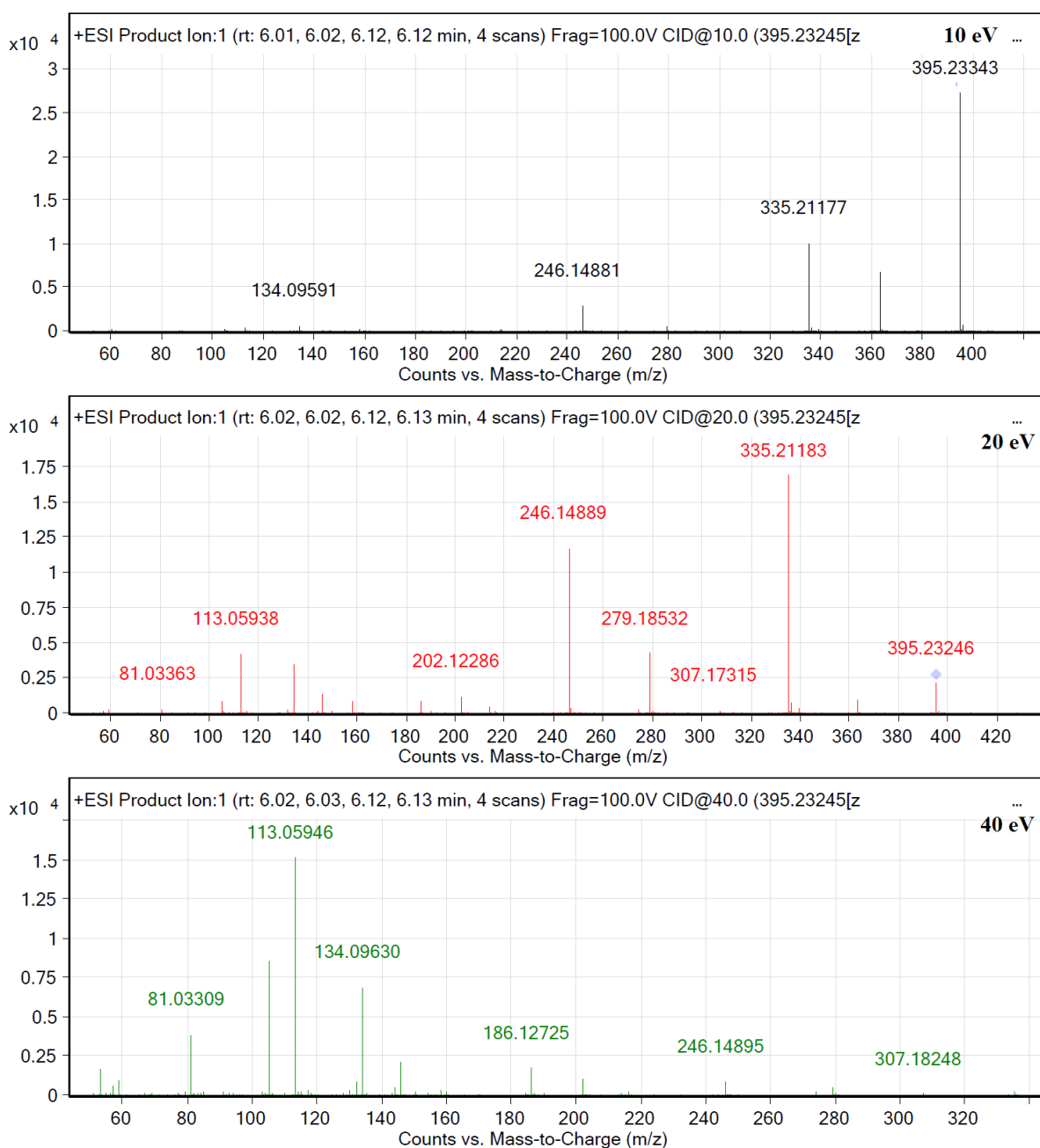


Figure A2-13: MS/MS spectra obtained from carfentanil

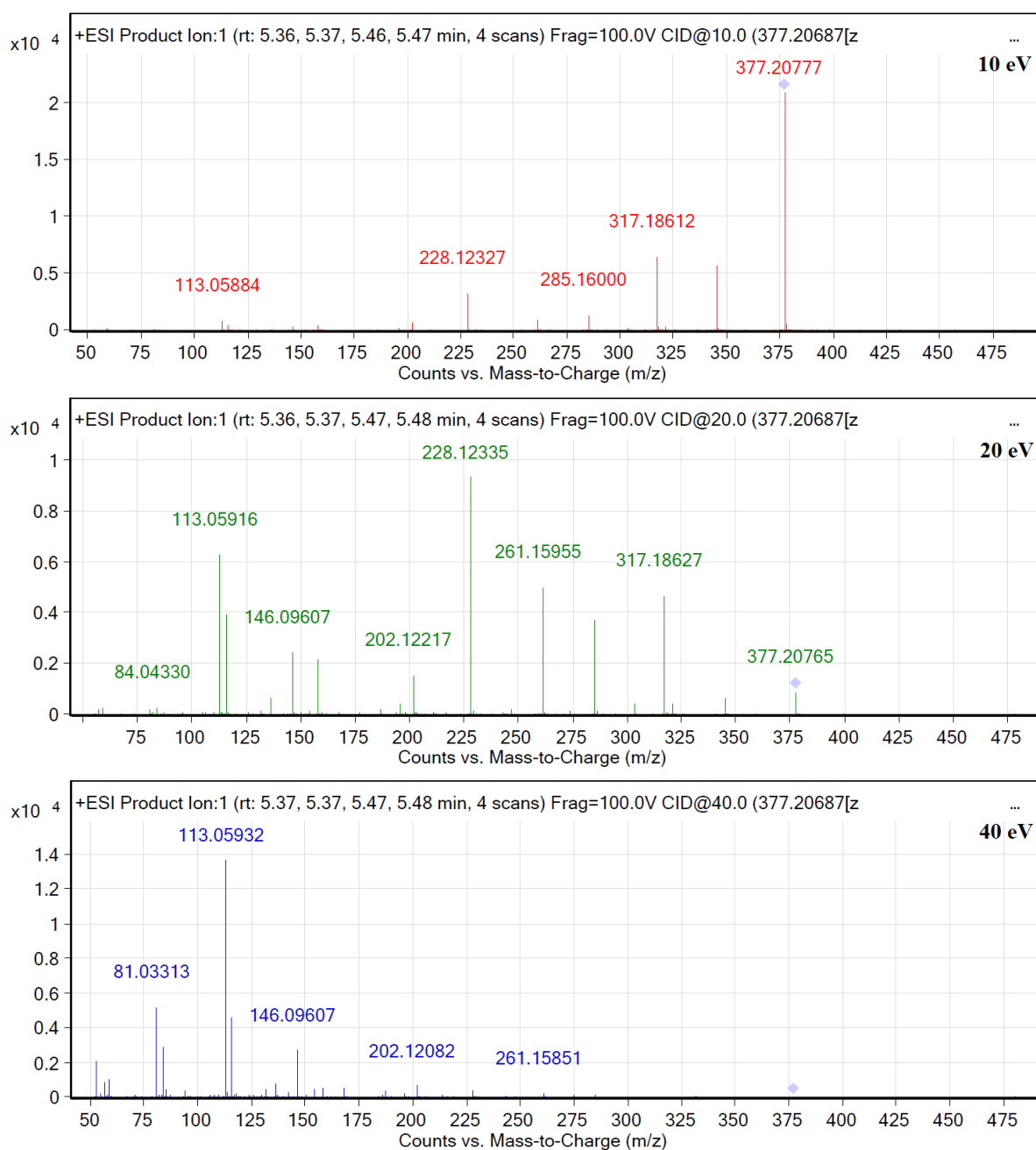


Figure A2-14: MS/MS spectra obtained from remifentanyl

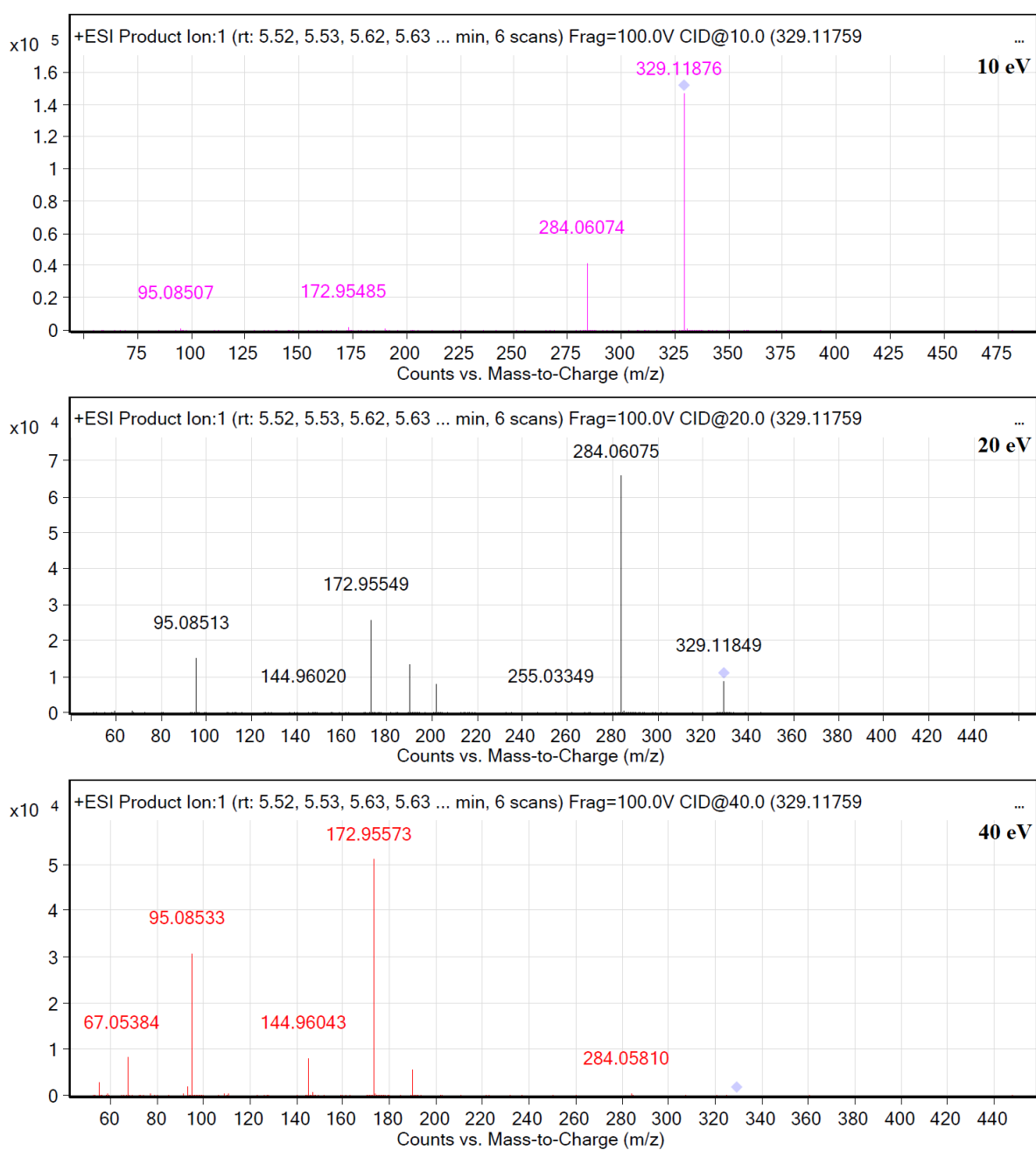


Figure A2-15: MS/MS spectra obtained from AH-7921

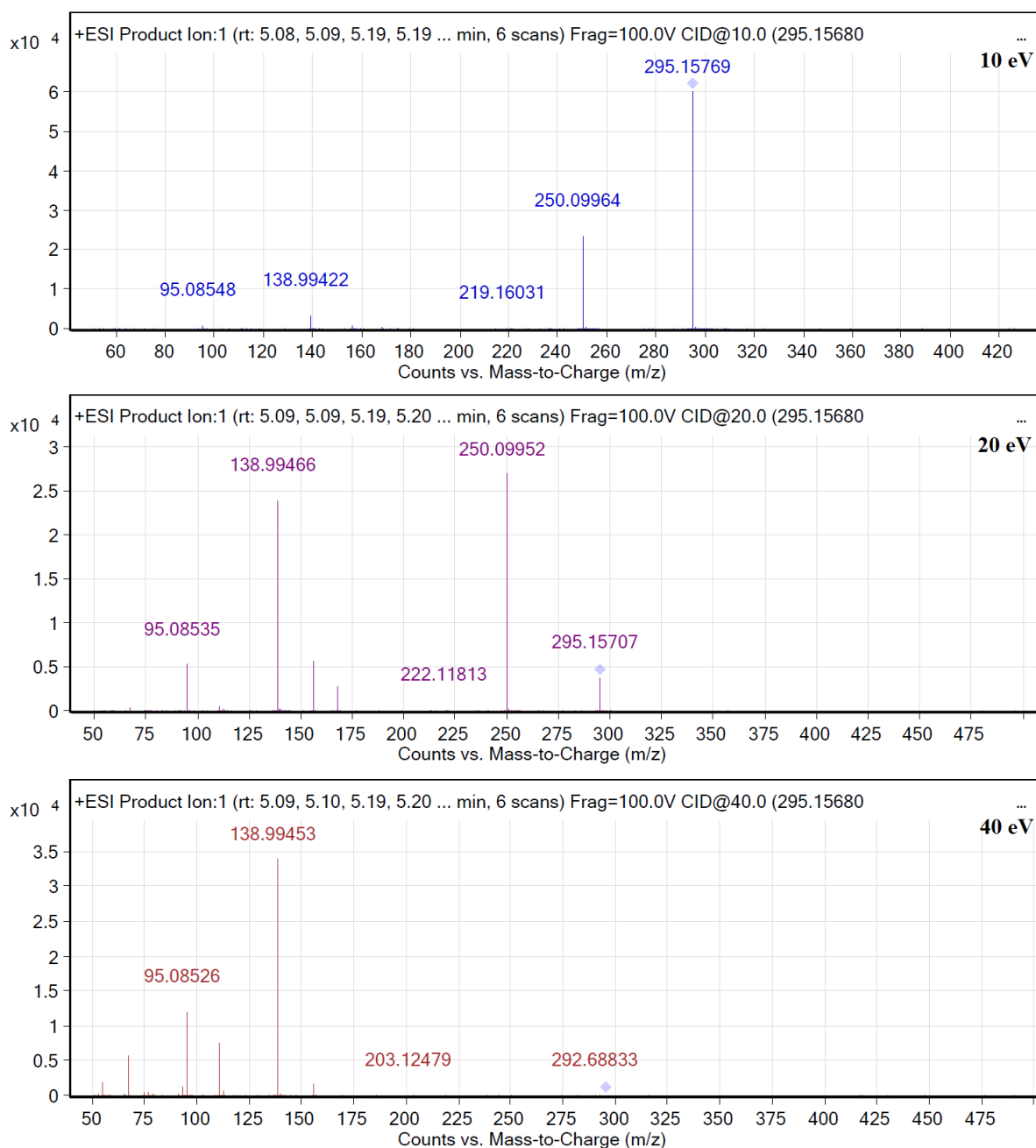


Figure A2-16: MS/MS spectra obtained from AH-8529

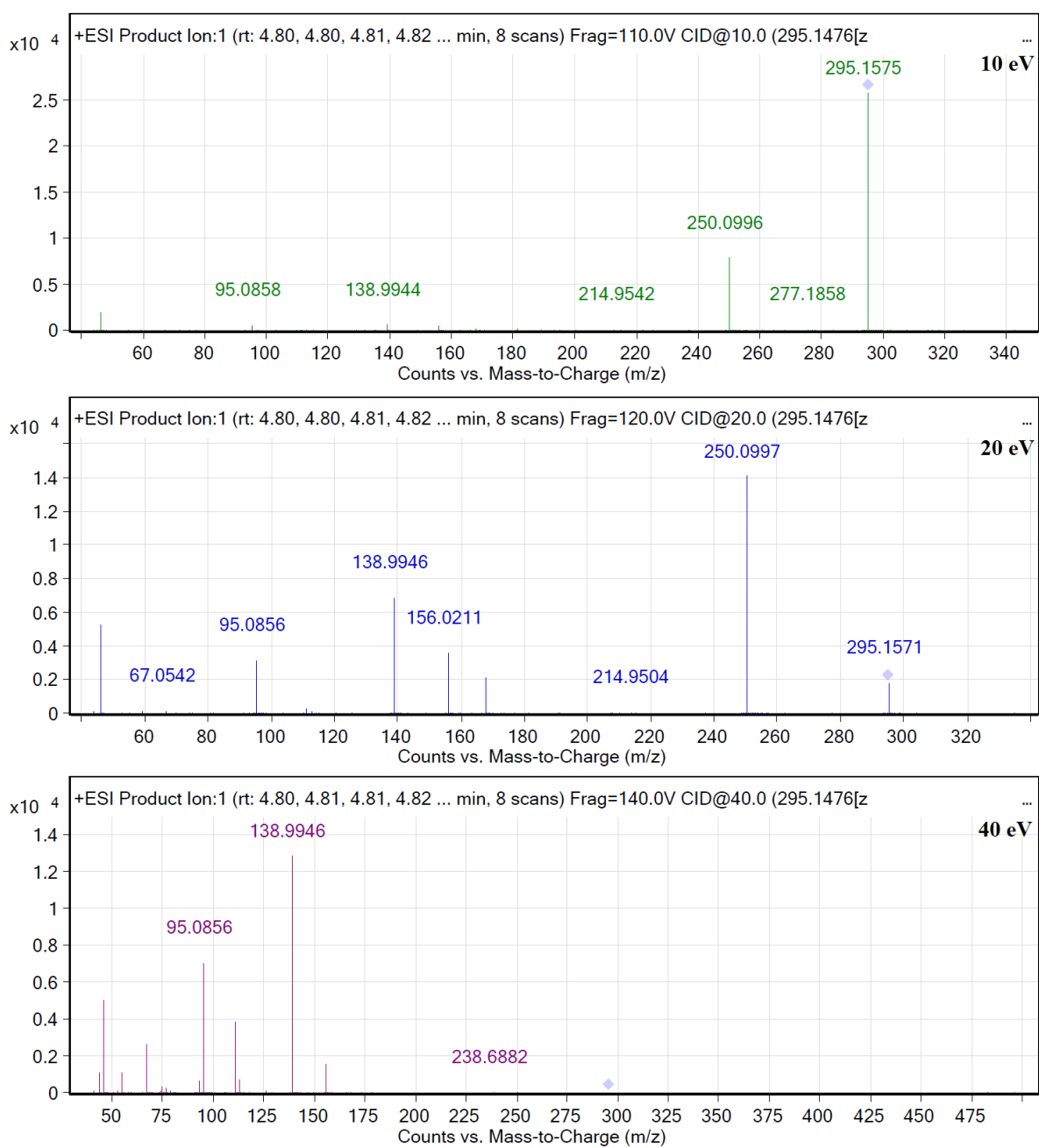


Figure A2-17: MS/MS spectra obtained from AH-8532

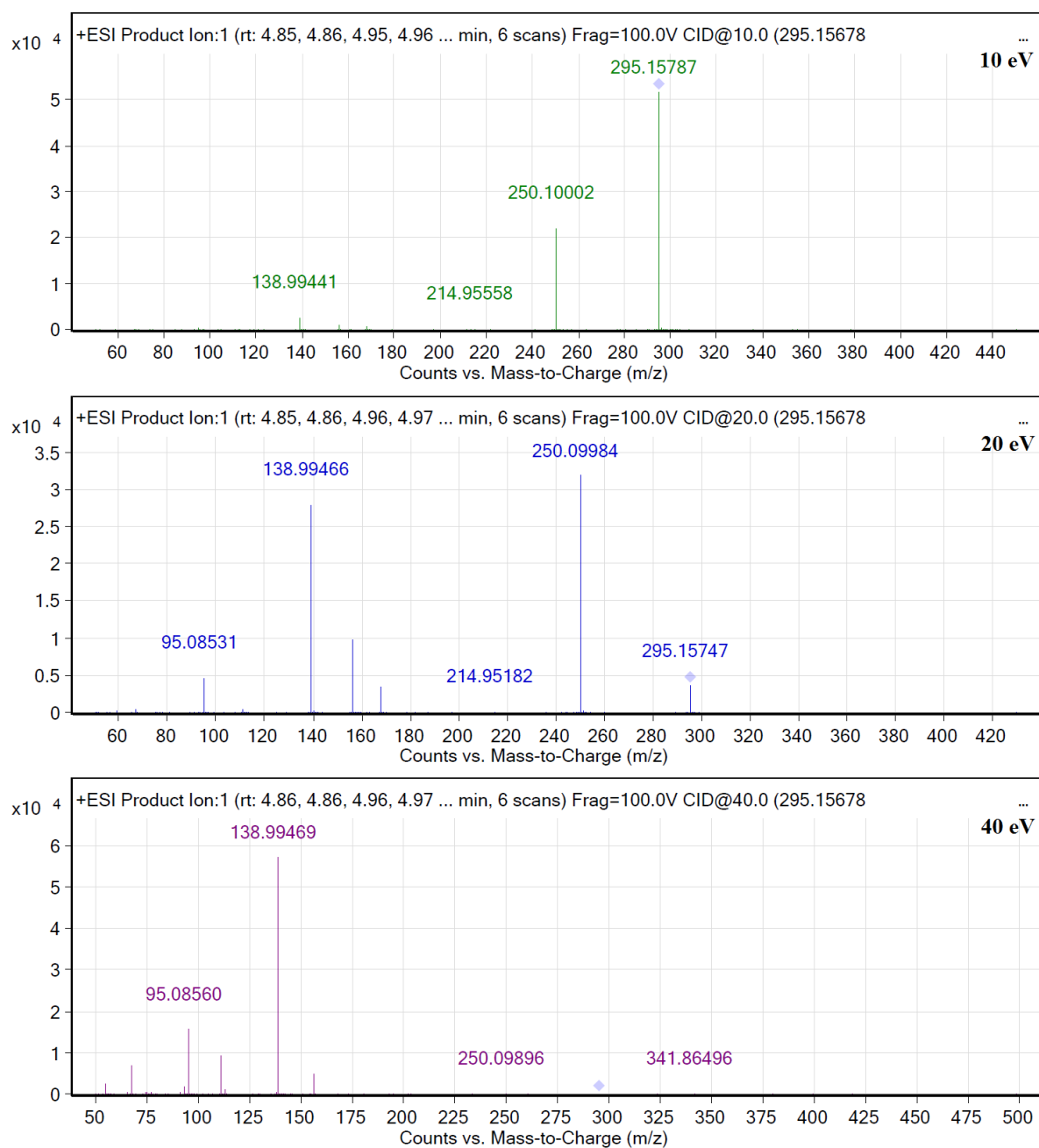


Figure A2-18: MS/MS spectra obtained from AH-8533

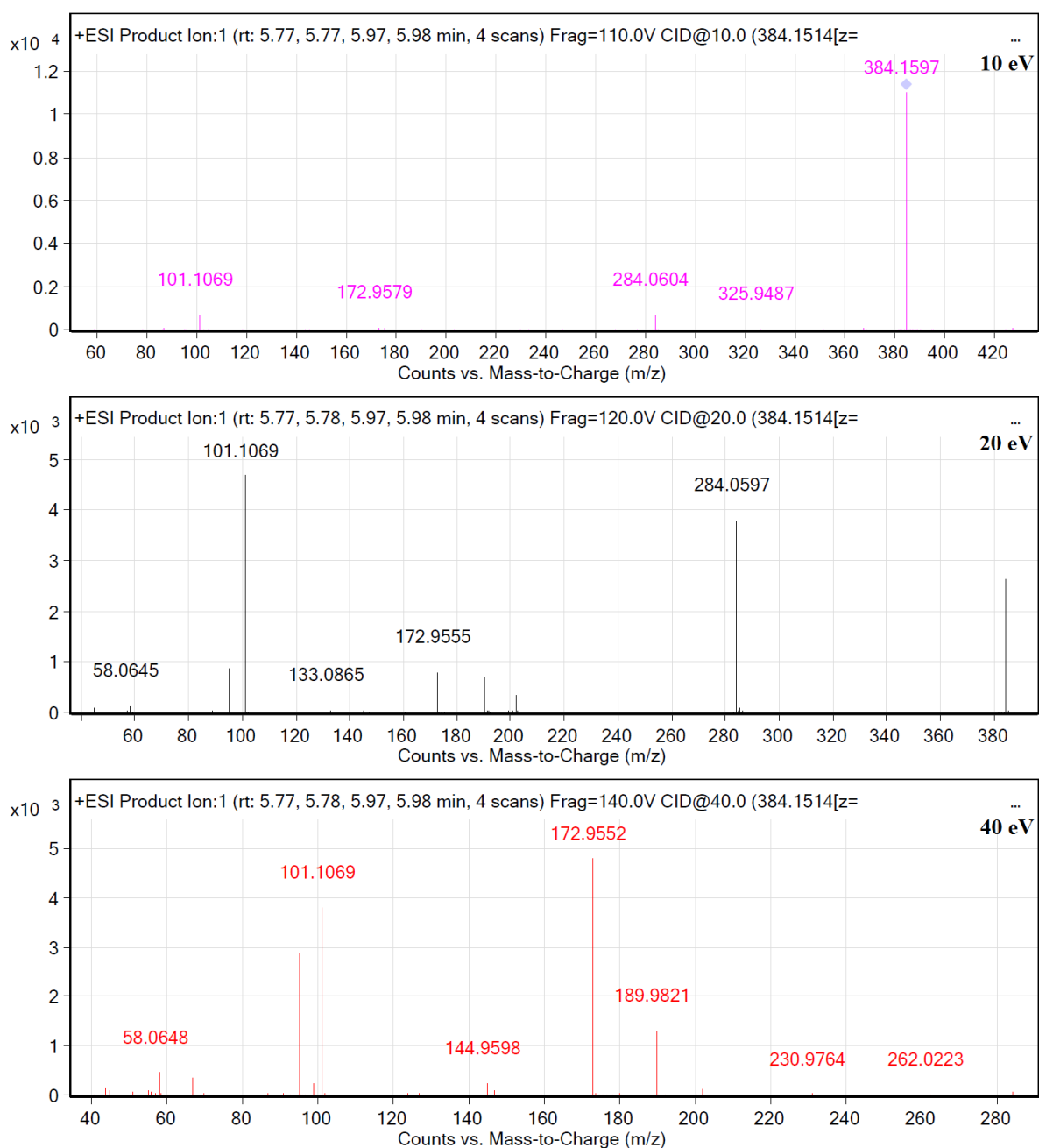


Figure A2-19: MS/MS spectra obtained from AH-8507



Figure A2-20: MS/MS spectra obtained from AH-7959

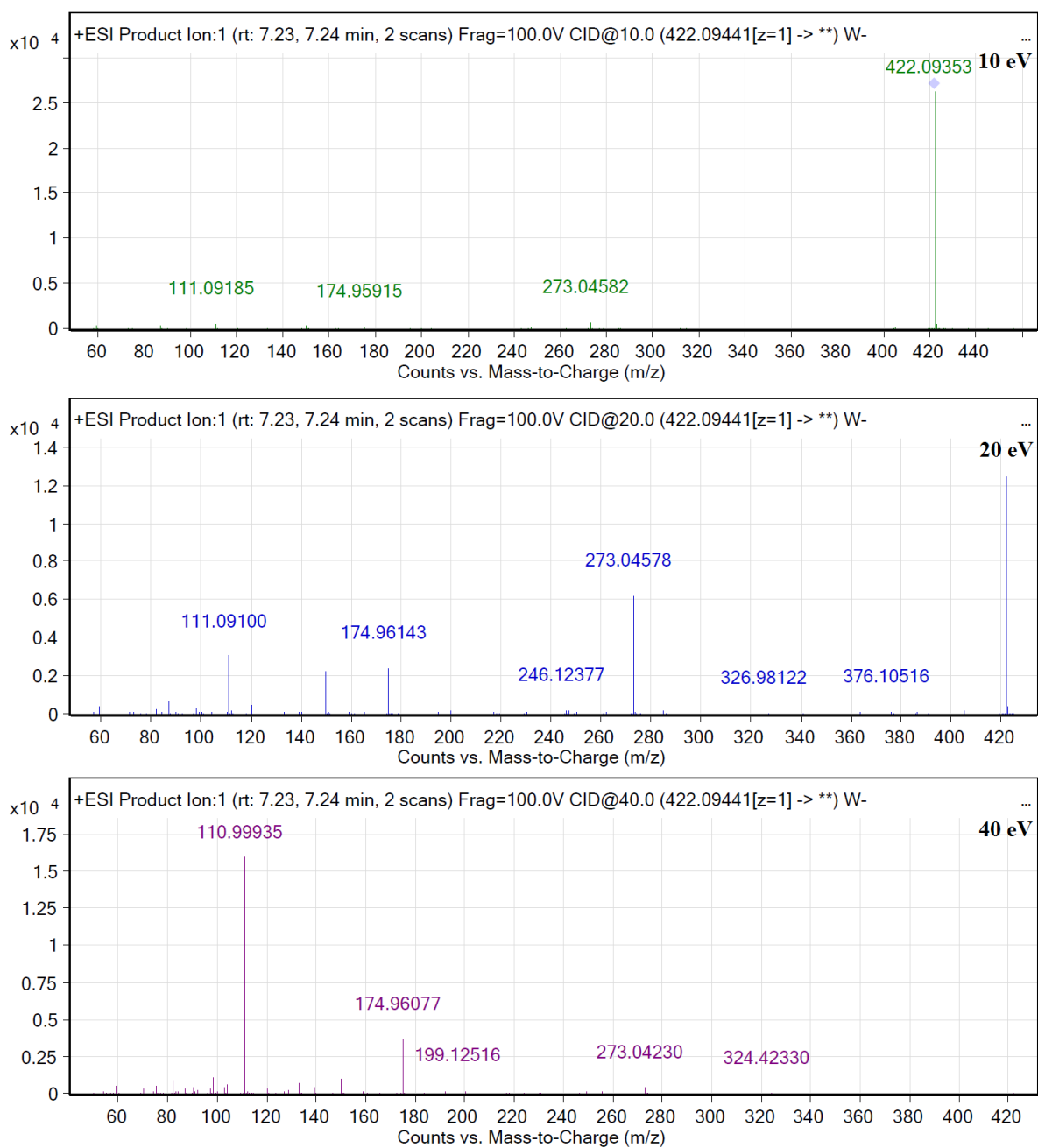


Figure A2-21: MS/MS spectra obtained from W-18

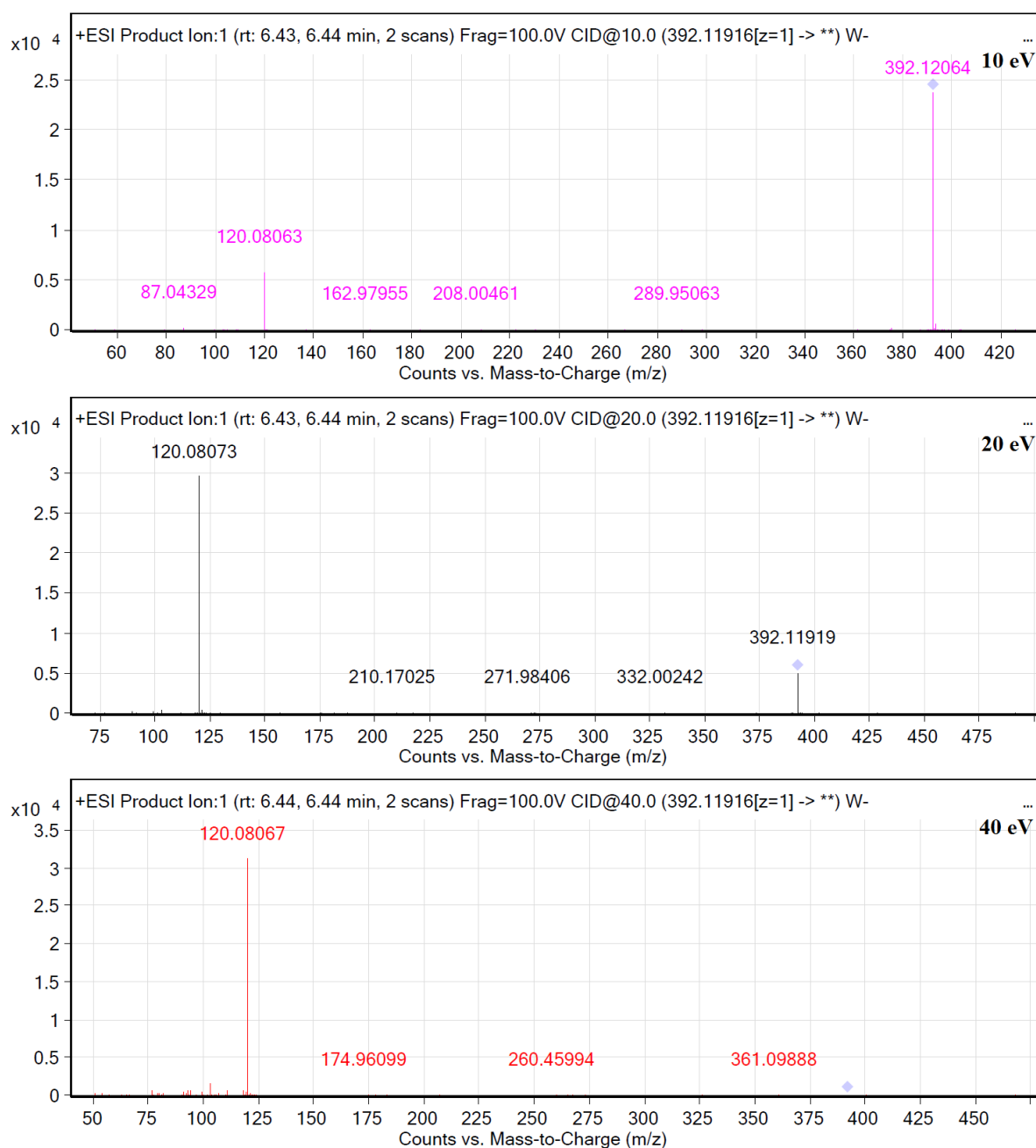


Figure A2-22: MS/MS spectra obtained from W-19

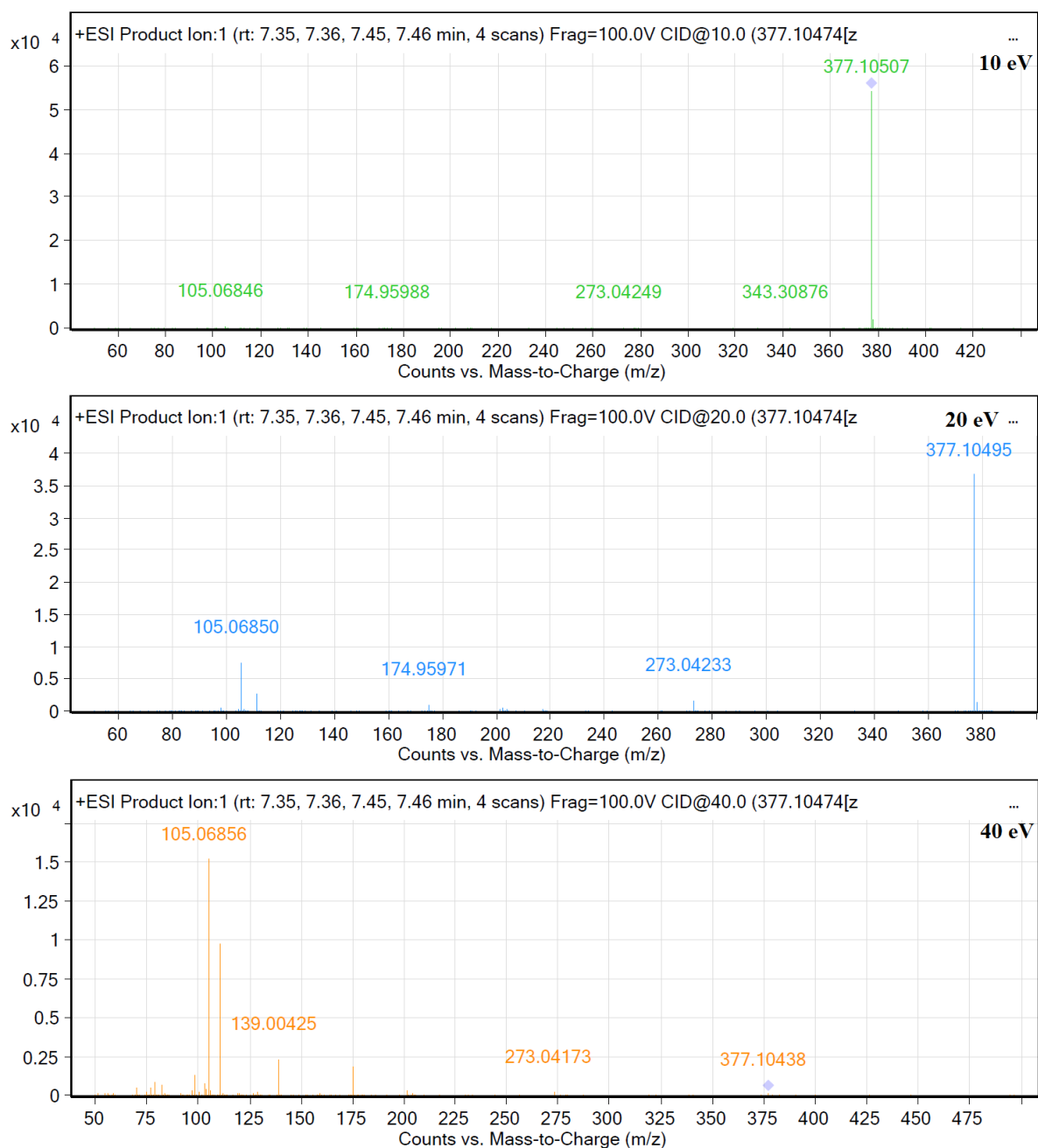


Figure A2-23: MS/MS spectra obtained from W-15

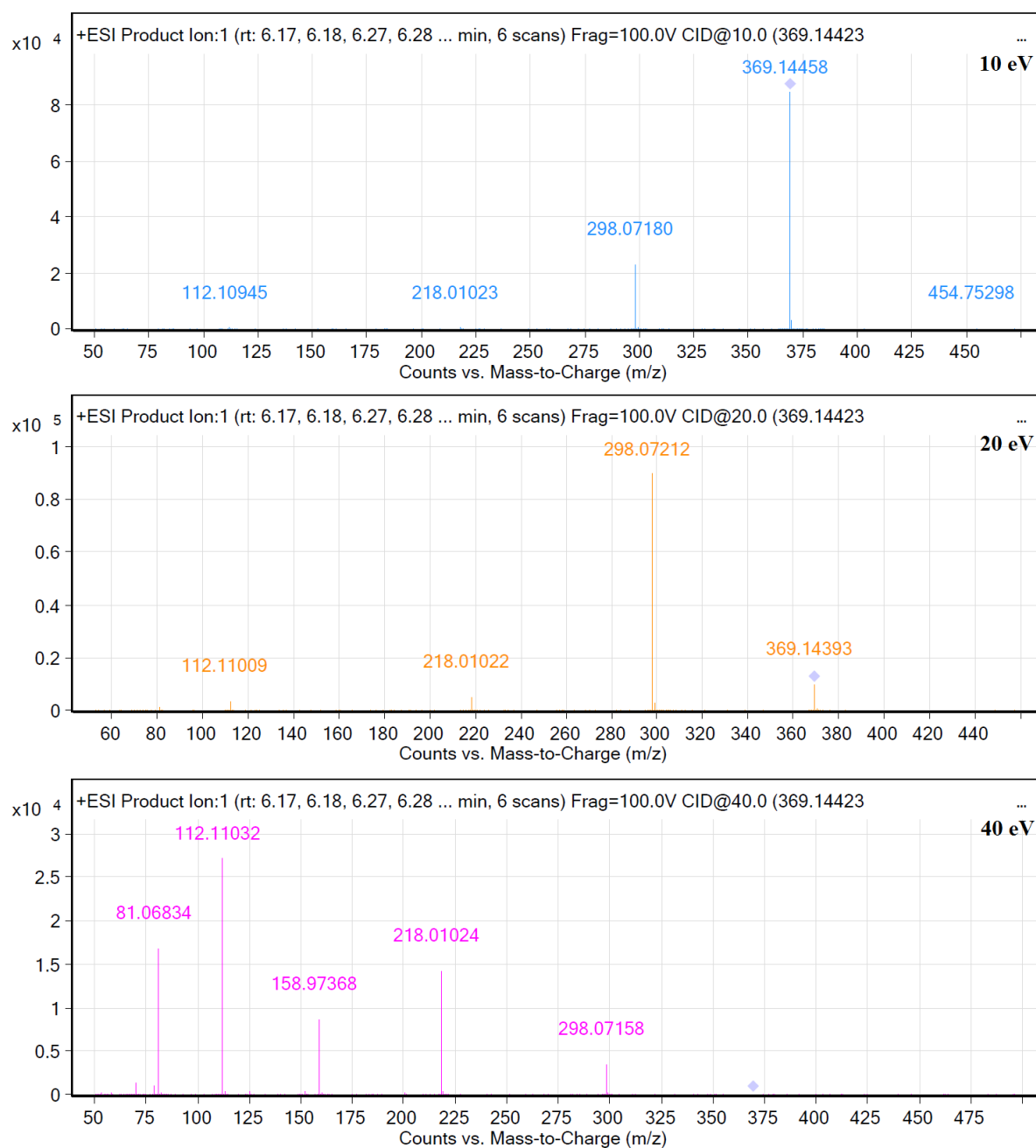


Figure A2-24: MS/MS spectra obtained from U-50488

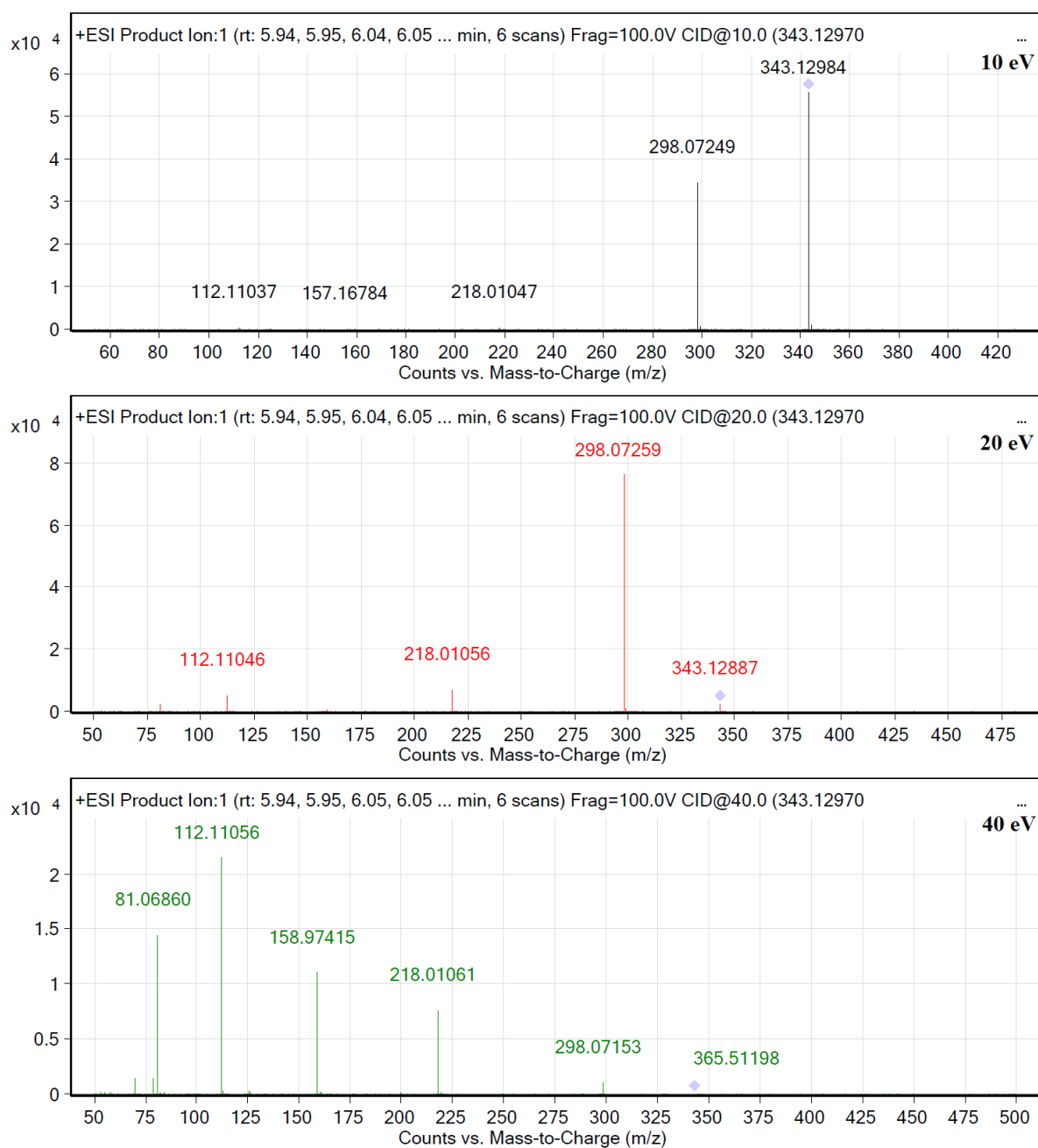


Figure A2-25: MS/MS spectra obtained from U-51754

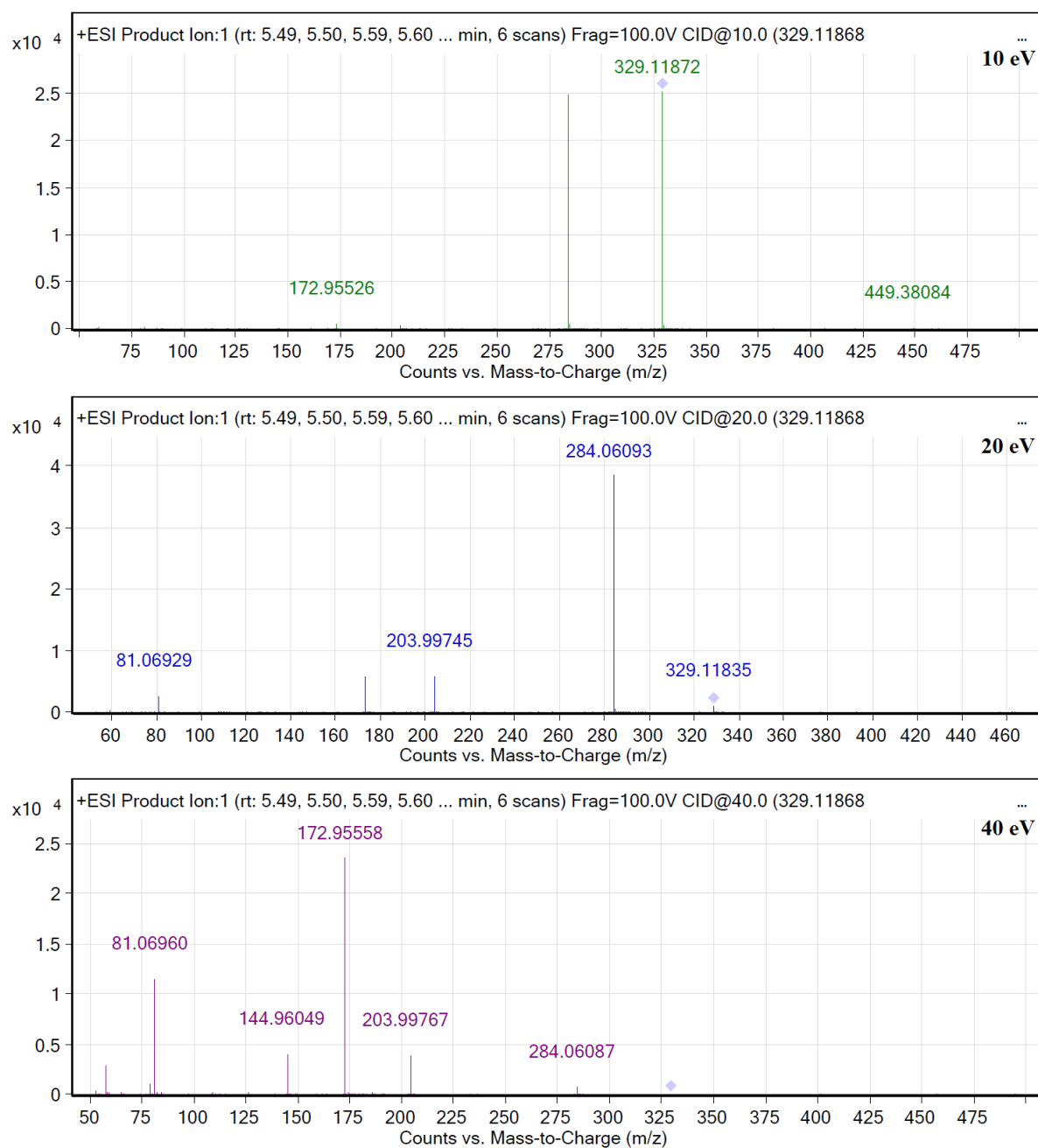


Figure A2-26: MS/MS spectra obtained from U-47700

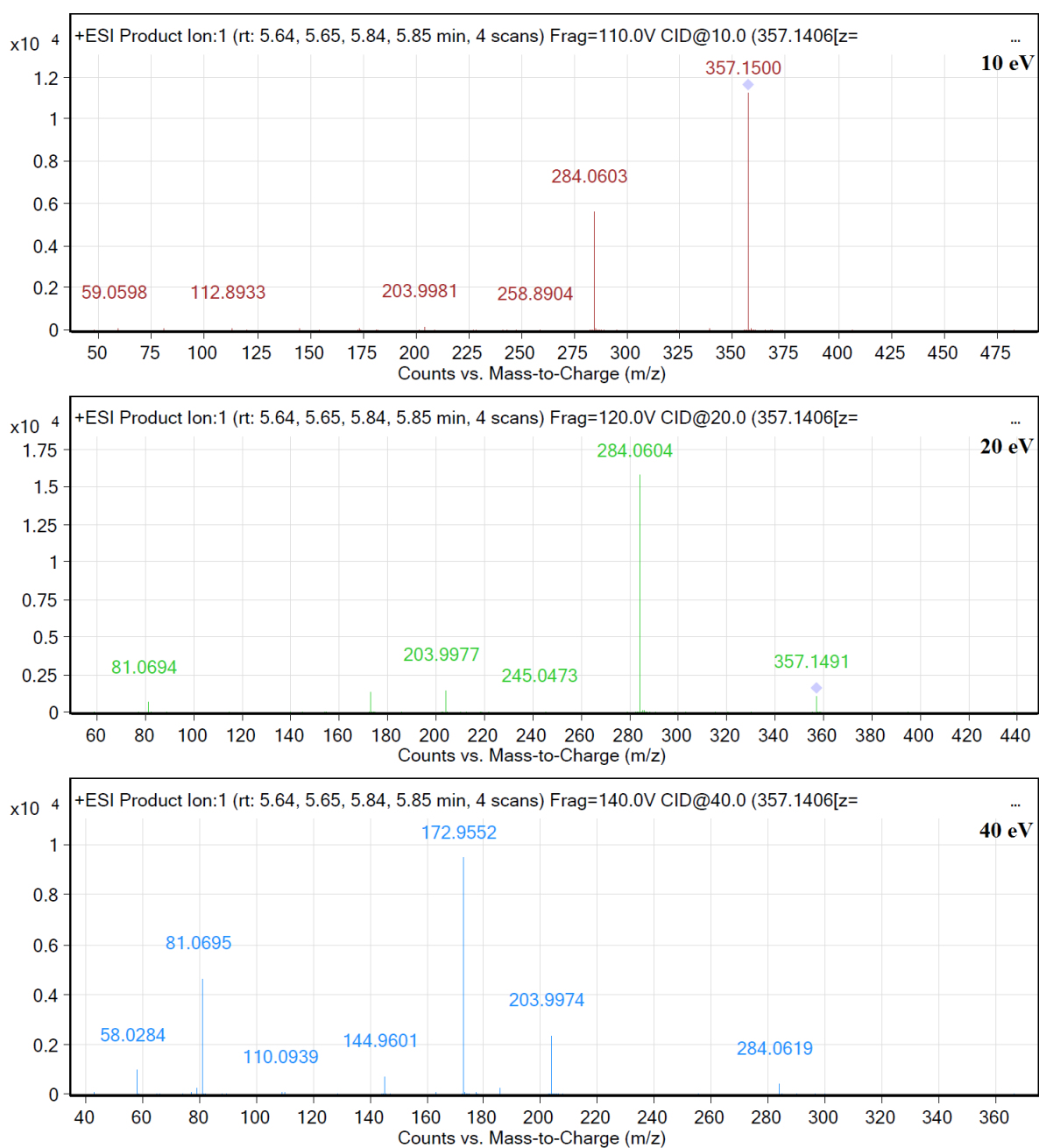
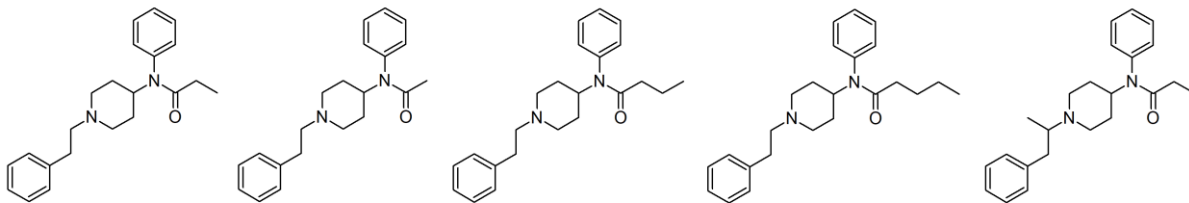


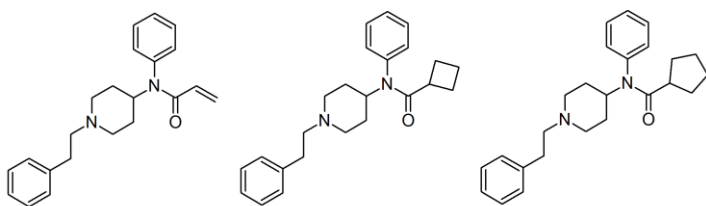
Figure A2-27: MS/MS spectra obtained from U-49900

Appendix 3

Amide Chain



Cyclic



F Analogs

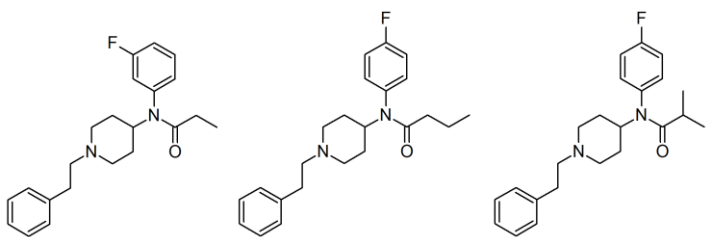


Figure A3-1: Structures of some compounds belonging to each fentanyl group included in the KMD filters

DefectDetect v1.0

Kendrick Mass Defect

Normalisation factor

☒ Methyl (-CH2)
 ☐ Oxygen (-O)
 ☐ Fluorine (-F)
 ☐ Chlorine (-Cl)

☐ Bromine (-Br)
 ☐ Iodine (-I)

☐ Custom
 Nominal Mass = Normalisation factor

Exact Mass

	Name	Kendrick Mass Defect	Minimum m/z	Strict Analysis?
Class 1	Fentanyl (Amide chain)	0.8510	323	<input checked="" type="checkbox"/>
Class 2	Fentanyl (Cyclic)	0.8380	335	<input checked="" type="checkbox"/>
Class 3	Fentanyl (F analogs)	0.8215	355	<input checked="" type="checkbox"/>
Class 4	Carfentanil	0.7916	395	<input checked="" type="checkbox"/>
Class 5	AH + U	0.7510	329	<input checked="" type="checkbox"/>
Class 6	AH	0.8276	295	<input checked="" type="checkbox"/>
Class 7				<input checked="" type="checkbox"/>
Class 8				<input checked="" type="checkbox"/>

Kendrick mass defect tolerance \pm Da

Filter

Mass range m/z -

Intensity \geq counts

Mass defect filter - Da

Even/Odd m/z? ☐ Even ☐ Odd ☒ Both

Internal Standard

Name m/z Tolerance (Da)

Sort

☒ Decreasing intensity (default)
 ☐ Decreasing m/z

☐ Increasing intensity
 ☐ Increasing m/z

Display

☒ only matched results
 ☐ all filtered results

Kendrick Mass Defect Calculator

m/z

☒ Methyl (-CH2)
 ☐ Custom

Nominal Mass

Exact Mass

Normalisation factor
 Kendrick Mass
 Kendrick Mass Defect

Cancel OK

Figure A3-2: Layout of the DefectDetect user interface to enter parameters for KMD analysis

Results - DefectDetect 1.2 (Josh Opioids)					
Parameters		Internal Standard			
Mass-to-charge (m/z) range	200 - 500	No internal standard monitored			
Intensity threshold (counts)	No filter selected				
Mass defect filter (Da)	0.1 - 0.27	Fentanyl (Amide chain)	0.846 - 0.856 (min. m/z = 323)	AH + U	0.746 - 0.756 (min. m/z = 329)
Even/Odd m/z	Both	Fentanyl (Cyclic)	0.833 - 0.843 (min. m/z = 335)	AH	0.8226 - 0.8326 (min. m/z = 295)
Kendrick tolerance (Da)	\pm 0.005	Fentanyl (F analogs)	0.8165 - 0.8265 (min. m/z = 355)	Class 7	No KMD value applied
Kendrick Normalisation	Methyl (-CH2)	Carfentanil	0.7866 - 0.7966 (min. m/z = 395)	Class 8	No KMD value applied
File Name	0.1 A-3				
Mass-to-charge ratio (m/z)	Intensity (counts)	Even m/z?	Mass defect (Da)	Kendrick Mass	Kendrick Mass Defect
323.1914	33	FALSE	0.1914	322.8293	0.8293
323.2095	96	FALSE	0.2095	322.8474	0.8474
329.1155	32	FALSE	0.1155	328.7468	0.7468
335.2181	19	FALSE	0.2181	334.8426	0.8426
337.2255	113	FALSE	0.2255	336.8478	0.8478
343.1329	4	FALSE	0.1329	342.7485	0.7485
349.2317	50	FALSE	0.2317	348.8405	0.8405
355.2146	260	FALSE	0.2146	354.8167	0.8167
357.1465	10	FALSE	0.1465	356.7465	0.7465
363.2445	32	FALSE	0.2445	362.8376	0.8376
365.2401	4	FALSE	0.2401	364.8309	0.8309
365.2649	7	FALSE	0.2649	364.8558	0.8558
369.2336	33	FALSE	0.2336	368.8200	0.8200
371.1658	26	FALSE	0.1658	370.7500	0.7500
377.2583	23	FALSE	0.2583	376.8357	0.8357
383.2547	20	FALSE	0.2547	382.8254	0.8254
385.1780	10	FALSE	0.1780	384.7465	0.7465
395.2346	51	FALSE	0.2346	394.7919	0.7919
427.2274	18	FALSE	0.2274	426.7488	0.7488

Figure A3-3: Output of the DefectDetect program showing a filtered mass list based on user-defined parameters

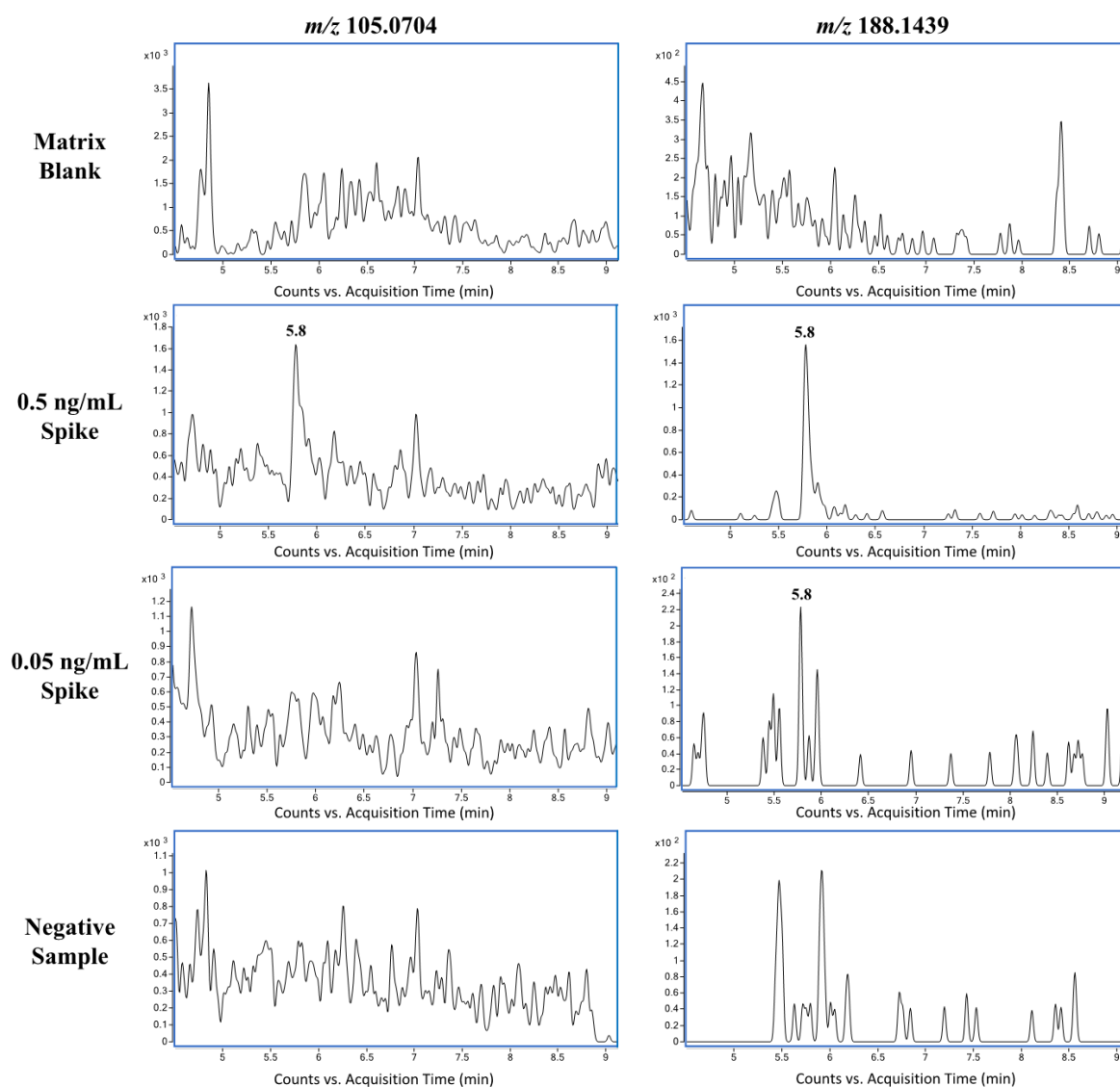


Figure A3-4: PIS results obtained from the blind trial fentanyl spike showing detection of both monitored ions in the 0.5 ng/mL spike and only detection of the m/z 188.1439 ion in the 0.05 ng/mL spike

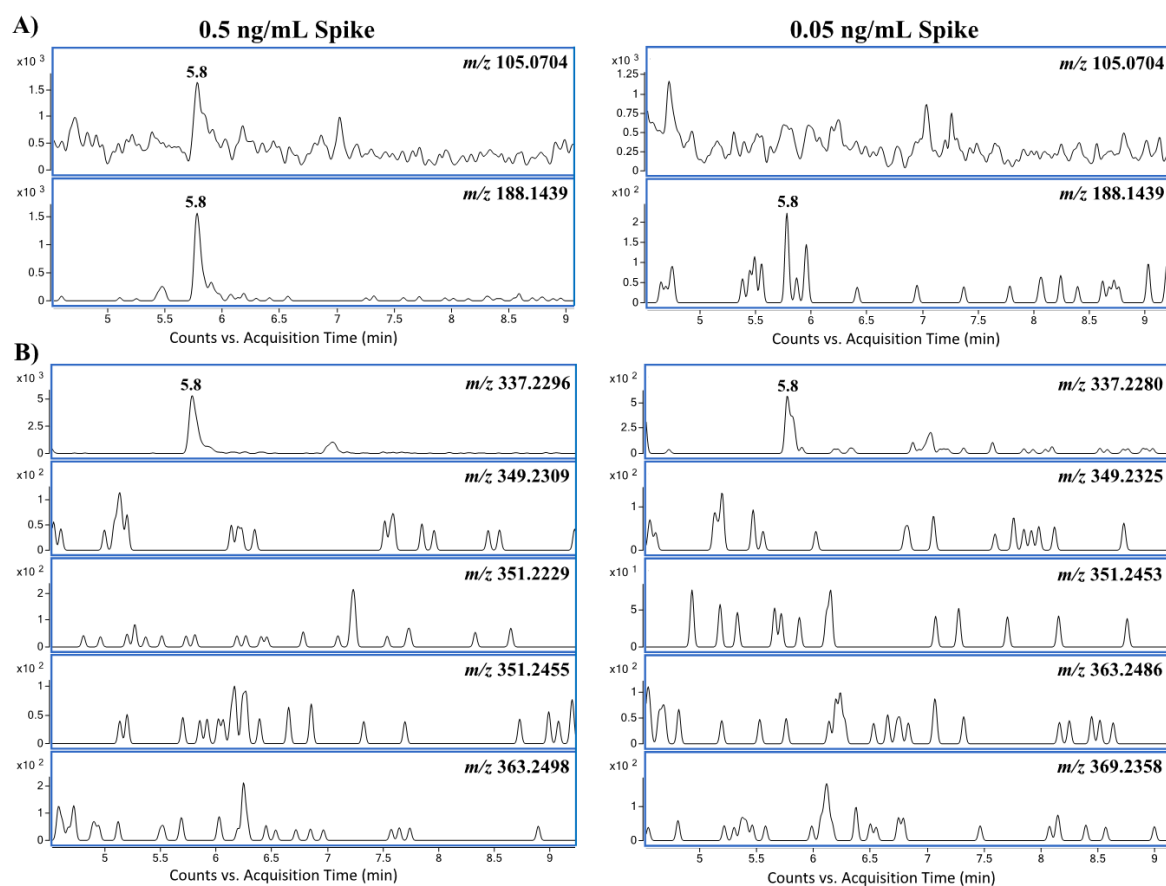


Figure A3-5: PIS results for the spiked fentanyl samples (A) compared with the associated KMD results for each sample (B)

Compound Groups (311 total)											
Automatically Show Columns											
Group	Found	Missed	Score (Tgt, max)	%RSD (Tgt)	Score (MFE, max)	%RSD (MFE)	Mass (Tgt)	Mass (avg)			
1	19	39	100	16.6	87	49.2	370.3089	370.3076			
2	51	7	99.9	83.3			234.1727	234.1726			
3	57	1	99.9	40.1			234.1732	234.1729			
4	20	38	86.9	19.1	87	65.9	392.2907	392.2921			
5	44	14	100	4.6	100	60.6	100.0523	100.0524			
6	58	0	100	0.4	100	0	122.0477	122.0477			
7	58	0	100	4.6	100	16.4	187.0633	187.0634			
8	58	0	99.9	7.9	100	8.4	472.2571	472.2586			
9	39	19	99.9	2.3	100	2.1	241.2773	241.2771			
10	58	0	99.9	5.3	100	8.5	165.0788	165.0795			
11	49	9	99.6	2	100	3.5	159.0683	159.0681			
12	43	15	100	85.7			292.2266	292.2255			
13	58	0	100	16.6			267.0968	267.0975			
14	19	39	87.6	14.7	82.9	122.5	146.0581	146.0578			
15	58	0	99	6			159.0684	159.0682			
16	20	38	87.4	10.4	100	4.6	87.1046	87.1046			
17	49	9	99.9	16.2	99.5	15	303.293	303.2924			
18	21	37	99.9	166.6			292.226	292.2255			

Compound Details											
Automatically Show Columns											
Diff (Tgt, mDa)	Diff (Tgt, ppm)	RT (Tgt)	RT Diff (Tgt)	Score (Tgt)	Flags (Tgt)	Flag Severity (Tgt)	Mass (Tgt)				
-0.6	-3.79	2.9	-0.019	97.02		Pass	159.0684				
0.14	0.85	2.9	-0.011	97.02		Pass	159.0684				
-0.49	-3.07	2.9	-0.032	97.73		Pass	159.0684				
-0.24	-1.52	2.9	-0.025	97.88		Pass	159.0684				
-0.83	-5.24	2.9	-0.026	93.12		Pass	159.0684				
-0.39	-2.44	2.9	-0.417	87.75		Pass	159.0684				
-0.57	-3.57	2.9	-0.418	96.2		Pass	159.0684				
-0.21	-1.35	2.9	-0.417	90.42		Pass	159.0684				
-0.73	-4.58	2.9	-0.416	80.94		Pass	159.0684				
-0.63	-3.93	2.9	-0.322	96.77		Pass	159.0684				
-0.32	-2	2.9	-0.417	80.16		Pass	159.0684				
-0.59	-3.69	2.9	-0.417	88.16		Pass	159.0684				
-0.35	-2.19	2.9	-0.417	97.77		Pass	159.0684				
-0.17	-1.05	2.9	-0.01	87		Pass	159.0684				
-0.12	-0.78	2.9	-0.009	94		Pass	159.0684				
-0.2	-1.23	2.9	-0.001	94.26		Pass	159.0684				
0.39	2.48	2.9	0.027	86.34		Pass	159.0684				
1.1	5.81	2.9	0.001	80.04		Pass	159.0684				

Figure A3-6: Compound group information provided by Profinder following recursive feature extraction

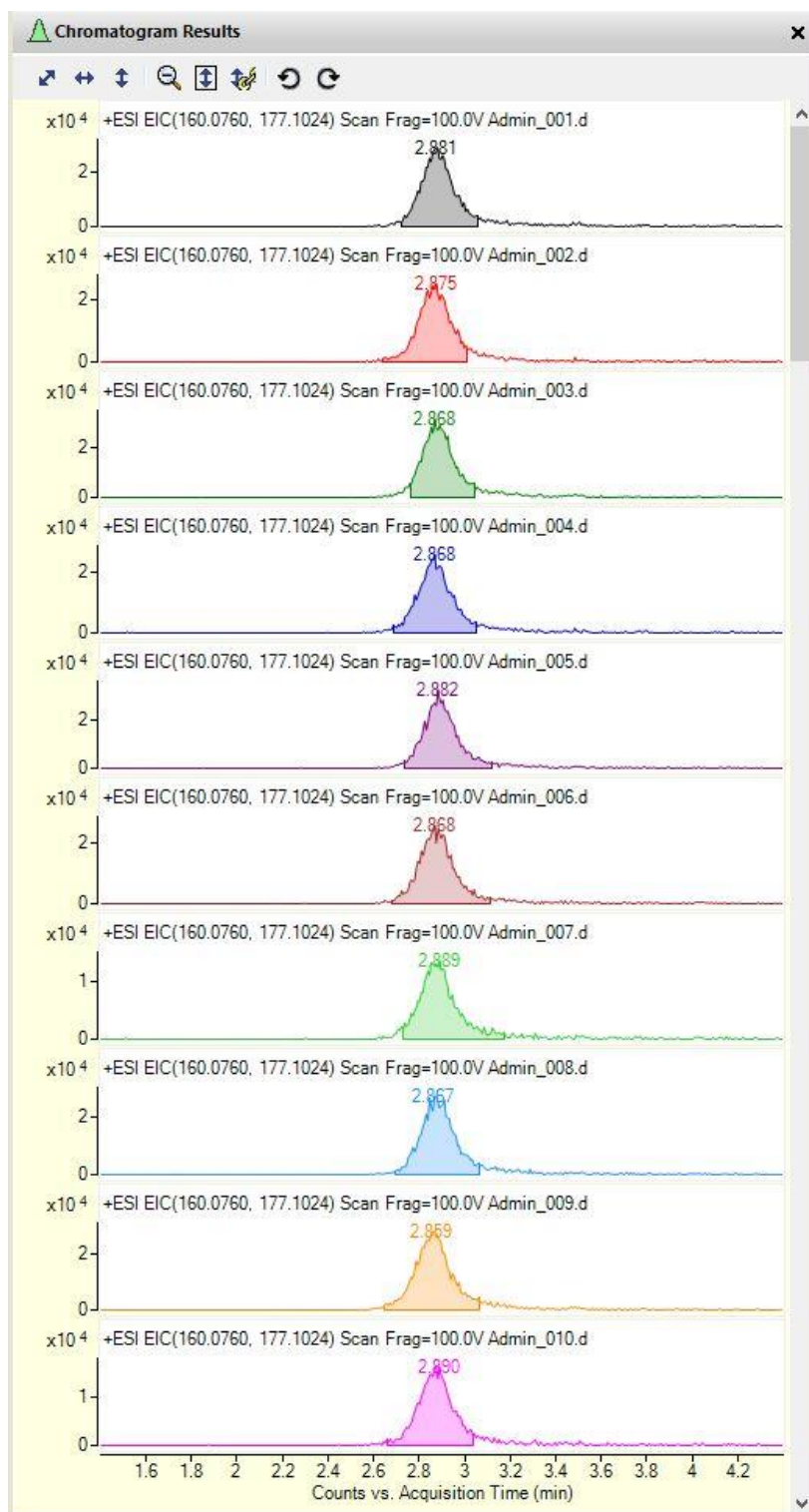


Figure A3-7: Extracted ion chromatograms for a selected compound group

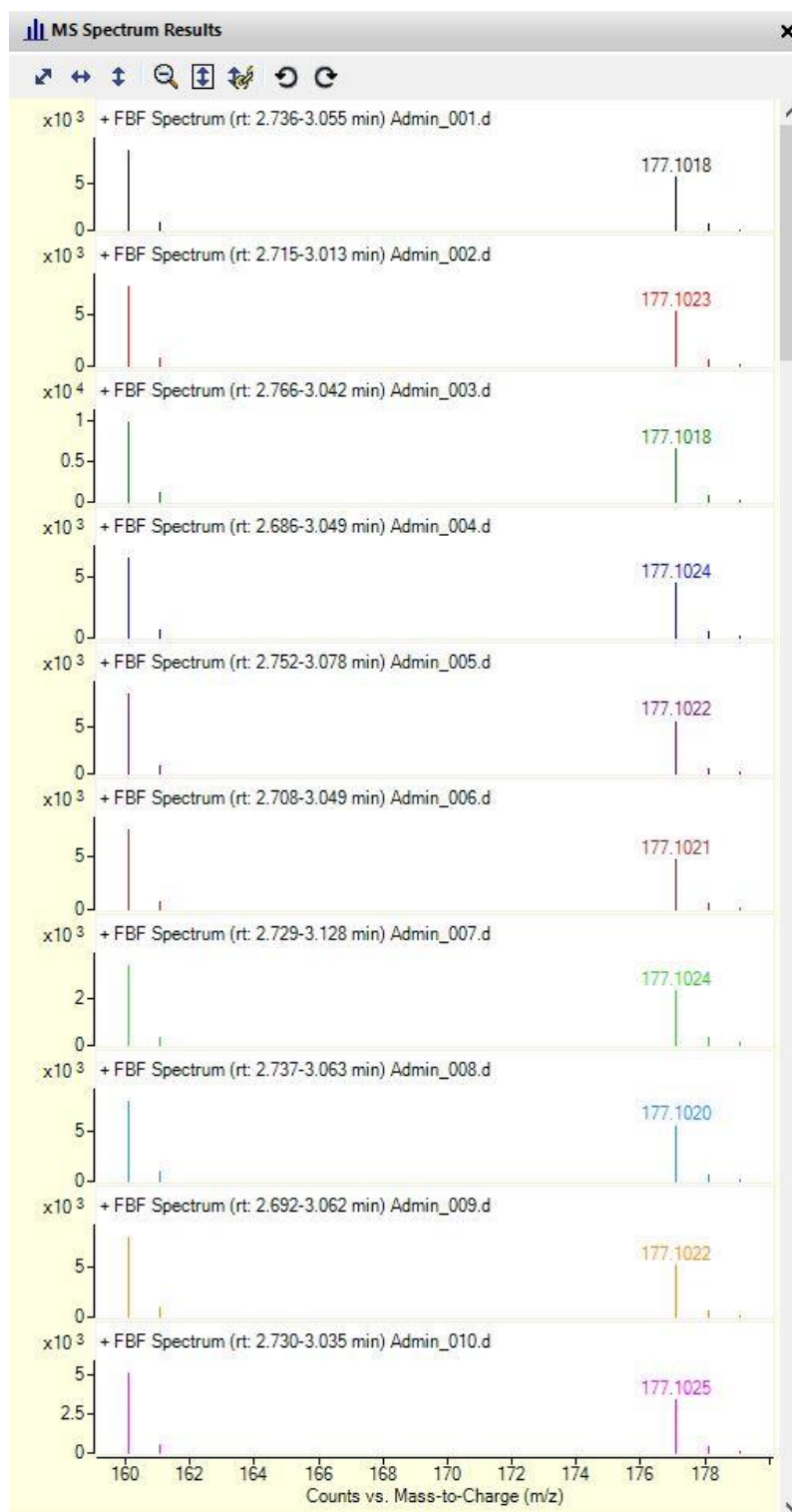


Figure A3-8: Find by Formula spectra for a selected compound group showing the molecular ion and identified adducts

Appendix 4

Table A4-1: Chemical information of the opioid standards used in the compound identification study

Name	IUPAC Name	Molecular Formula	Exact Mass
3,4-ethylenedioxy U-47700	N-[2-(dimethylamino)cyclohexyl]-N-methyl-2,3-dihydro-1,4-benzodioxine-6-carboxamide	C ₁₈ H ₂₆ N ₂ O ₃	318.1943
3,4-ethylenedioxy U-51754	2-(2,3-dihydro-1,4-benzodioxin-6-yl)-N-[2-(dimethylamino)cyclohexyl]-N-methylacetamide	C ₁₉ H ₂₈ N ₂ O ₃	332.2100
3,4-methylenedioxy U-47700	N-[2-(dimethylamino)cyclohexyl]-N-methyl-2H-1,3-benzodioxole-5-carboxamide	C ₁₇ H ₂₄ N ₂ O ₃	304.1787
4-chloroisobutyryl fentanyl	N-(4-chlorophenyl)-2-methyl-N-[1-(2-phenylethyl)-4-piperidinyl]-propanamide	C ₂₃ H ₂₉ ClN ₂ O	384.1968
4-fluorobutyryl fentanyl	N-(4-fluorophenyl)-N-[1-(2-phenylethyl)-4-piperidinyl]-butanamide	C ₂₂ H ₂₇ FN ₂ O	368.2264
4-fluoroisobutyryl fentanyl	N-(4-fluorophenyl)-2-methyl-N-[1-(2-phenylethyl)-4-piperidinyl]-propanamide	C ₂₂ H ₂₇ FN ₂ O	368.2264
4-methoxybutyryl fentanyl	N-(4-methoxyphenyl)-N-(1-phenethylpiperidin-4-yl)butyramide	C ₂₄ H ₃₂ N ₂ O ₂	380.2464
Acetyl fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidinyl]-acetamide	C ₂₁ H ₂₆ N ₂ O	322.2045
Acryl fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidinyl]-2-propenamide	C ₂₂ H ₂₆ N ₂ O	334.2045
AH-7563	N-[[1-(dimethylamino)cyclohexyl]methyl]-benzamide	C ₁₆ H ₂₄ N ₂ O	260.1889
AH-7921	3,4-dichloro-N-[[1-(dimethylamino)cyclohexyl]methyl]-benzamide	C ₁₆ H ₂₂ Cl ₂ N ₂ O	328.1109
AH-7959	3,4-dichloro-N-[[1-(1-piperidinyl)cyclohexyl]methyl]-benzamide	C ₁₉ H ₂₆ Cl ₂ N ₂ O	368.1422
AH-8507	3,4-dichloro-N-[[1-(4-methyl-1-piperazinyl)cyclohexyl]methyl]-benzamide	C ₁₉ H ₂₇ Cl ₂ N ₃ O	383.1531
AH-8529	4-chloro-N-[[1-(dimethylamino)cyclohexyl]methyl]-benzamide	C ₁₆ H ₂₃ ClN ₂ O	294.1499
AH-8532	3-chloro-N-[[1-(dimethylamino)cyclohexyl]methyl]-benzamide	C ₁₆ H ₂₃ ClN ₂ O	294.1499
AH-8533	2-chloro-N-[[1-(dimethylamino)cyclohexyl]methyl]-benzamide	C ₁₆ H ₂₃ ClN ₂ O	294.1499

Name	IUPAC Name	Molecular Formula	Exact Mass
Alfentanil	N-[1-[2-(4-ethyl-4,5-dihydro-5-oxo-1H-tetrazol-1-yl)ethyl]-4-(methoxymethyl)-4-piperidiny]-N-phenyl-propanamide	C ₂₁ H ₃₂ N ₆ O ₃	416.2536
Benzodioxole fentanyl	N-(1-phenethylpiperidin-4-yl)-N-phenylbenzo[d][1,3]dioxole-5-carboxamide	C ₂₇ H ₂₈ N ₂ O ₃	428.2100
Benzyl fentanyl	N-phenyl-N-[1-(phenylmethyl)-4-piperidiny]-propanamide	C ₂₁ H ₂₆ N ₂ O	322.2045
Butyryl fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidiny]-butanamide	C ₂₃ H ₃₀ N ₂ O	350.2358
Carfentanil	4-[(1-oxopropyl)phenylamino]-1-(2-phenylethyl)-4-piperidinecarboxylic acid, methyl ester	C ₂₄ H ₃₀ N ₂ O ₃	394.2256
Crotonyl fentanyl	(2E)-N-phenyl-N-[1-(2-phenylethyl)-4-piperidiny]-2-butenamide	C ₂₃ H ₂₈ N ₂ O	348.2202
Cyclopentyl fentanyl	N-(1-phenethylpiperidin-4-yl)-N-phenylcyclopentanecarboxamide	C ₂₅ H ₃₂ N ₂ O	376.2515
Cyclopropyl fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidiny]-cyclopropanecarboxamide	C ₂₃ H ₂₈ N ₂ O	348.2202
Despropionyl p-fluoro fentanyl	N-(4-fluorophenyl)-1-(2-phenylethyl)-4-piperidinamine	C ₁₉ H ₂₃ FN ₂	298.1845
Fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidiny]-propanamide	C ₂₂ H ₂₈ N ₂ O	336.2202
Furanyl fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidiny]-2-furancarboxamide	C ₂₄ H ₂₆ N ₂ O ₂	374.1994
Isopropyl U-47700	3,4-dichloro-N-[2-(dimethylamino)cyclohexyl]-N-(propan-2-yl)benzamide	C ₁₈ H ₂₆ Cl ₂ N ₂ O	356.1422
Methacryl fentanyl	N-(1-phenethylpiperidin-4-yl)-N-phenylmethacrylamide	C ₂₃ H ₂₈ N ₂ O	348.2202
m-fluoro fentanyl	N-(3-fluorophenyl)-N-[1-(2-phenylethyl)-4-piperidiny]-propanamide	C ₂₂ H ₂₇ FN ₂ O	354.2107
m-fluoro methoxyacetyl fentanyl	N-(3-fluorophenyl)-2-methoxy-N-(1-phenethylpiperidin-4-yl)acetamide	C ₂₂ H ₂₇ FN ₂ O ₂	370.2057
MT-45	1-cyclohexyl-4-(1,2-diphenylethyl)-piperazine	C ₂₄ H ₃₂ N ₂	348.2565
N-methyl carfentanil	1-methyl-4-[(1-oxopropyl)phenylamino]-4-piperidinecarboxylic acid, methyl ester	C ₁₇ H ₂₄ N ₂ O ₃	304.1787
N-methyl U-47931E	<i>trans</i> -4-bromo-N-[2-(dimethylamino)cyclohexyl]-N-methyl-benzamide	C ₁₆ H ₂₃ BrN ₂ O	338.0994
Norfentanyl	N-phenyl-N-4-piperidiny-propanamide	C ₁₄ H ₂₀ N ₂ O	232.1576
Ocfentanil	N-(2-fluorophenyl)-2-methoxy-N-[1-(2-phenylethyl)-4-piperidiny]-acetamide	C ₂₂ H ₂₇ FN ₂ O ₂	370.2057
o-fluoro fentanyl	N-(2-fluorophenyl)-N-[1-(2-phenylethyl)-4-piperidiny]-propanamide	C ₂₂ H ₂₇ FN ₂ O	354.2107
p-fluoro methoxyacetyl fentanyl	N-(4-fluorophenyl)-2-methoxy-N-(1-phenethylpiperidin-4-yl)acetamide	C ₂₂ H ₂₇ FN ₂ O ₂	370.2057

Name	IUPAC Name	Molecular Formula	Exact Mass
Phenyl fentanyl	N-(1-phenethylpiperidin-4-yl)-N-phenylbenzamide	C ₂₆ H ₂₈ N ₂ O	384.2202
Phenylpropionyl fentanyl	N,3-diphenyl-N-[1-(2-phenylethyl)piperidin-4-yl]propanamide	C ₂₈ H ₃₂ N ₂ O	412.2515
Propyl U-47700	3,4-dichloro-N-[2-(dimethylamino)cyclohexyl]-N-propylbenzamide	C ₁₈ H ₂₆ Cl ₂ N ₂ O	356.1422
Remifentanyl	4-(methoxycarbonyl)-4-[(1-oxopropyl)phenylamino]-1-piperidinepropanoic acid, methyl ester	C ₂₀ H ₂₈ N ₂ O ₅	376.1998
Senecioyl fentanyl	3-methyl-N-(1-phenethylpiperidin-4-yl)-N-phenylbut-2-enamide	C ₂₄ H ₃₀ N ₂ O	362.2358
Sufentanyl	N-[4-(methoxymethyl)-1-[2-(2-thienyl)ethyl]-4-piperidinyl]-N-phenyl-propanamide	C ₂₂ H ₃₀ N ₂ O ₂ S	386.2028
Tetrahydrofuran fentanyl	tetrahydro-N-phenyl-N-[1-(2-phenylethyl)-4-piperidinyl]-2-furancarboxamide	C ₂₄ H ₃₀ N ₂ O ₂	378.2307
Thiofentanyl	N-phenyl-N-[1-[2-(2-thienyl)ethyl]-4-piperidinyl]-propanamide	C ₂₀ H ₂₆ N ₂ OS	342.1766
U-47109	<i>trans</i> -3,4-dichloro-N-[2-(dimethylamino)cyclohexyl]-benzamide	C ₁₅ H ₂₀ Cl ₂ N ₂ O	314.0953
U-47700	<i>trans</i> -3,4-dichloro-N-[2-(dimethylamino)cyclohexyl]-N-methyl-benzamide	C ₁₆ H ₂₂ Cl ₂ N ₂ O	328.1109
U-47931E	<i>trans</i> -4-bromo-N-[2-(dimethylamino)cyclohexyl]-benzamide	C ₁₅ H ₂₁ BrN ₂ O	324.0837
U-48520	<i>trans</i> -4-chloro-N-[2-(dimethylamino)cyclohexyl]-N-methyl-benzamide	C ₁₆ H ₂₃ ClN ₂ O	249.1499
U-48800	2-(2,4-dichlorophenyl)-N-((1S,2S)-2-(dimethylamino)cyclohexyl)-N-methylacetamide	C ₁₇ H ₂₄ Cl ₂ N ₂ O	342.1266
U-49900	<i>trans</i> -3,4-dichloro-N-[2-(diethylamino)cyclohexyl]-N-methyl-benzamide	C ₁₈ H ₂₆ Cl ₂ N ₂ O	356.1422
U-50488	2-(3,4-dichlorophenyl)-N-methyl-N-[2-(pyrrolidin-1-yl)cyclohexyl]acetamide	C ₁₉ H ₂₆ Cl ₂ N ₂ O	368.1422
U-51754	<i>trans</i> -3,4-dichloro-N-[2-(dimethylamino)cyclohexyl]-N-methyl-benzeneacetamide	C ₁₇ H ₂₄ Cl ₂ N ₂ O	342.1266
U-62066	<i>rel</i> -3,4-dichloro-N-methyl-N-[(5S,7R,8R)-7-(1-pyrrolidinyl)-1-oxaspiro[4.5]dec-8-yl]-benzeneacetamide	C ₂₂ H ₃₀ Cl ₂ N ₂ O ₂	424.1684
UF-17	N-[2-(dimethylamino)cyclohexyl]-N-phenylpropanamide	C ₁₇ H ₂₆ N ₂ O	274.2045
Valeryl fentanyl	N-phenyl-N-[1-(2-phenylethyl)-4-piperidinyl]-pentanamide	C ₂₄ H ₃₂ N ₂ O	364.2515
α-methyl fentanyl	N-[1-(1-methyl-2-phenylethyl)-4-piperidinyl]-N-phenyl-propanamide	C ₂₃ H ₃₀ N ₂ O	350.2358
β-hydroxy thiofentanyl	N-[1-[2-hydroxy-2-(2-thienyl)ethyl]-4-piperidinyl]-N-phenyl-propanamide	C ₂₀ H ₂₆ N ₂ O ₂ S	358.1715

Table A4-2: Input data for the class prediction model containing the top 10 most abundant product ions in the 40 eV MS² spectra and known compound classes

Name	Ion 1	Ion 2	Ion 3	Ion 4	Ion 5	Ion 6	Ion 7	Ion 8	Ion 9	Ion 10	Class
AH-7563	105.03380	77.03843	95.08575	67.05442	79.05416	55.05414	122.06036	93.06967	51.02278	95.01221	AH Series
AH-7921	172.95549	95.08541	67.05411	144.96032	189.98163	55.05389	93.06987	146.97568	95.12127	173.00468	AH Series
AH-7959	172.95560	86.09640	95.08570	189.98220	67.05430	144.96070	44.01280	55.05390	284.06010	146.97620	AH Series
AH-8507	172.95547	101.10729	95.08554	58.06454	189.98151	144.95900	93.06938	99.09167	173.00373	174.95313	AH Series
AH-8529	138.99485	95.08575	110.99970	67.05446	156.02127	55.05396	139.03850	93.06960	113.01502	77.03819	AH Series
AH-8532	138.99466	95.08565	67.05412	110.99975	156.02099	55.05398	93.06944	113.01525	75.02218	77.03807	AH Series
AH-8533	138.99451	95.08540	110.99938	67.05419	156.02073	55.05391	139.03812	93.06951	113.01470	75.02243	AH Series
4-chloroisobutyril fentanyl	105.06982	188.14322	134.09610	166.04060	71.04869	146.09615	84.08050	105.10824	55.05414	59.05994	Fentanyl
4-fluorobutyril fentanyl	105.06991	188.14349	71.04953	55.05395	134.09465	84.07956	150.06986	105.10792	146.09398	117.06830	Fentanyl
4-fluoroisobutyril fentanyl	105.07002	188.14349	134.09652	150.07132	146.09672	105.10820	84.08066	71.04908	55.05383	69.06972	Fentanyl
4-methoxybutyril fentanyl	105.06992	188.14315	134.09624	146.09615	162.09148	84.08059	71.04886	79.05377	188.19288	176.10625	Fentanyl
Acetyl fentanyl	105.07010	188.14338	132.08068	79.05403	84.08060	146.09658	134.09608	55.05418	103.05421	77.03840	Fentanyl
Acryl fentanyl	105.06987	188.14353	132.08048	84.08070	79.05398	146.09659	134.09593	55.05379	103.05412	120.08074	Fentanyl
Alfentanil	165.10158	197.12783	99.05495	132.08048	183.11250	170.10313	314.18530	99.07936	94.06484	67.05394	Fentanyl
Benzodioxole fentanyl	149.02237	105.06912	188.14182	121.02774	134.09526	150.02564	65.03834	146.09490	84.08020	189.14516	Fentanyl
Benzyl fentanyl	91.05449	174.12813	82.06527	132.08085	65.03863	57.03371	94.06542	84.08077	91.08950	145.10108	Fentanyl
Butyril fentanyl	105.06998	188.14347	132.08057	146.09664	134.09616	71.04906	84.08069	55.05414	79.05425	69.06950	Fentanyl
Carfentanil	113.05987	105.06984	134.09636	81.03371	146.09651	186.12727	279.18588	53.03834	246.14768	202.12181	Fentanyl
Crotonyl fentanyl	105.07063	188.14445	69.03380	132.08115	134.09677	146.09701	84.08063	105.10875	55.05421	79.05440	Fentanyl
Cyclopentyl fentanyl	69.06980	105.06972	188.14280	134.09620	146.09613	132.08079	59.06009	69.10080	160.11205	97.06428	Fentanyl
Cyclopropyl fentanyl	105.07012	188.14340	69.03349	132.08090	146.09635	134.09602	84.08107	79.05428	69.06961	55.05363	Fentanyl
Despropionyl p-fluoro fentanyl	105.06980	79.05377	134.09598	188.14369	103.05356	105.10723	77.03875	164.08580	98.09649	55.05429	Fentanyl

Name	Ion 1	Ion 2	Ion 3	Ion 4	Ion 5	Ion 6	Ion 7	Ion 8	Ion 9	Ion 10	Class
Fentanyl	105.06971	188.14317	132.08053	84.08037	57.03331	146.09594	79.05399	134.09617	55.05416	103.05442	Fentanyl
Furanyl fentanyl	105.07018	188.14358	84.08107	146.09657	134.09664	95.01284	132.08064	55.05411	79.05422	106.07376	Fentanyl
m-fluoro fentanyl	105.06972	188.14335	57.03314	150.07113	84.08065	134.09608	146.09612	79.05364	55.05387	69.06993	Fentanyl
m-fluoro methoxyacetyl fentanyl	105.06991	188.14371	84.08010	134.09606	105.10960	79.05370	146.09581	77.03804	186.12669	55.05335	Fentanyl
Methacryl fentanyl	105.06967	188.14291	69.03305	134.09574	79.05345	84.08002	146.09603	103.05398	55.05367	105.10787	Fentanyl
N-methyl norcarfentanil	146.09636	113.05993	81.03287	53.03841	59.04922	91.05375	94.06325	57.03335	202.11867	158.09875	Fentanyl
Norfentanyl	55.05388	56.04938	84.08048	94.06454	57.03347	55.08136	67.05361	57.06977	53.03790	69.05758	Fentanyl
Ocfentanil	105.06971	188.14276	134.09576	146.09683	84.08046	105.10829	55.05403	69.06940	79.05338	132.08040	Fentanyl
o-fluoro fentanyl	105.06972	188.14313	150.07031	57.03344	134.09557	103.05385	84.08038	146.09600	55.05433	105.10970	Fentanyl
p-fluoro methoxyacetyl fentanyl	105.07023	188.14410	146.09671	134.09642	84.08033	79.05413	103.05362	55.05411	186.12951	120.08097	Fentanyl
Phenyl fentanyl	105.06953	105.03305	188.14302	77.03802	84.07999	134.09555	146.09546	105.10783	55.05385	79.05309	Fentanyl
Phenylpropionyl fentanyl	105.06968	188.14325	134.09603	132.08045	91.05427	146.09601	84.08041	105.10791	292.16720	189.14539	Fentanyl
Remifentanil	113.06010	116.07096	81.03374	146.09689	84.04449	53.03840	57.03343	136.07559	228.12259	113.09942	Fentanyl
Seneciyl fentanyl	83.04869	105.06939	55.05385	188.14310	134.09496	79.05366	132.07947	83.08358	146.09571	77.03995	Fentanyl
Sufentanil	111.02636	140.10706	132.08075	238.12572	126.09135	99.08049	67.05406	69.06982	206.09974	108.08076	Fentanyl
Tetrahydrofuran fentanyl	105.06992	188.14356	71.04896	134.09688	146.09659	132.08046	84.07995	55.05414	120.08096	79.05395	Fentanyl
Thiofentanyl	111.02597	84.08041	194.09919	146.09597	132.08079	77.03772	67.05377	56.04899	57.03278	96.07998	Fentanyl
Valeryl fentanyl	105.06997	188.14317	57.06987	134.09630	146.09671	132.08055	84.08122	55.05421	160.11213	59.06030	Fentanyl
α -methyl fentanyl	91.05472	119.08577	84.08101	202.15924	56.04938	55.05433	132.08103	57.03355	94.06519	67.05402	Fentanyl
β -hydroxy thiofentanyl	111.02585	97.01043	192.08408	132.08046	55.05414	57.03328	146.09611	70.99465	151.04627	138.03640	Fentanyl
3,4-ethylenedioxy U-47700	163.03908	137.05966	81.06990	194.08068	135.04377	274.14249	58.02881	107.01260	79.05372	176.06930	U series
3,4-ethylenedioxy U-51754	149.05952	112.11213	81.06985	288.15863	208.09618	70.06478	79.05372	112.15046	121.06506	126.12741	U series
3,4-methylenedioxy U-47700	149.02345	81.06987	121.02847	58.02881	123.04408	180.06532	95.04917	93.03346	260.12721	79.05428	U series

Name	Ion 1	Ion 2	Ion 3	Ion 4	Ion 5	Ion 6	Ion 7	Ion 8	Ion 9	Ion 10	Class
Isopropyl U-47700	172.95531	81.06984	189.98193	144.96064	270.04412	79.05389	232.02814	312.08916	171.97140	81.10352	U series
N-methyl U-47931E	182.94357	81.06962	213.98530	154.94848	58.02879	79.05406	294.04729	182.99397	195.97479	126.12673	U series
Propyl U-47700	172.95510	81.06952	232.02839	144.95950	189.98093	79.05334	312.09063	171.97131	81.10345	173.00424	U series
U-47109	172.95547	81.06990	144.96019	189.98149	79.05438	126.12841	84.08035	81.10371	77.03739	110.99985	U series
U-47700	172.95601	81.07014	203.99770	144.96067	58.02896	79.05447	284.06087	173.95820	81.10419	53.03829	U series
U-47931E	182.94344	81.06990	154.94838	199.96974	79.05376	126.12752	156.96452	182.99315	280.02987	81.10299	U series
U-48520	138.99491	81.07008	110.99965	170.03666	58.02884	79.05445	139.03677	250.09854	53.03900	81.10279	U series
U-48800	112.11203	81.07003	158.97603	218.01251	70.06515	79.05399	298.07411	125.01451	81.10374	53.03837	U series
U-49900	172.95570	81.06978	203.99752	58.02853	144.95982	284.05975	79.05373	81.10266	185.98584	53.03853	U series
U-50488	112.11239	81.06994	218.01363	158.97640	298.07666	70.06548	79.05384	112.15204	113.11541	152.14477	U series
U-51754	112.11190	81.06975	158.97604	218.01322	79.05380	70.06496	298.07499	112.15090	125.01511	81.10287	U series
U-62066	168.13810	137.09599	71.04913	354.10143	91.05404	67.05434	69.03350	158.97601	218.01282	93.06988	U series

Table A4-3: Input spreadsheet for the retention time prediction model containing the calculated molecular features for each compound, as well as the average relative retention time (RRT)

Name	pKa	nC	nO	nDB	logP	AlogP	MlogP	nR06	logD (pH 9)	nBNZ	logS (pH 9)	DBE	Hy	RRT
3,4-ethylenedioxy U-47700	9.20	18	3	4	2.217	-1.8525	2.89	3	1.8064	1	-2.8147	7	-0.7721	0.312
3,4-ethylenedioxy U-51754	9.20	19	3	4	2.198	-1.8177	3.00	3	1.7863	1	-2.3492	7	-0.7816	0.346
3,4-methylenedioxy U-47700	9.20	17	3	4	2.328	-1.8702	2.78	2	1.9165	1	-2.7856	7	-0.7618	0.291
4-chloroisobutyryl fentanyl	8.74	23	1	7	4.963	0.5210	3.55	3	4.7714	2	-4.9942	10	-0.8426	1.387
4-fluorobutyryl fentanyl	8.74	23	1	7	4.403	-0.6798	3.55	3	4.2129	2	-4.6316	10	-0.8426	1.113
4-fluoroisobutyryl fentanyl	8.74	23	1	7	4.501	0.2564	3.55	3	4.3113	2	-4.5695	10	-0.8426	1.060
4-methoxybutyryl fentanyl	8.76	24	2	7	4.102	-1.2113	3.66	3	3.9042	2	-4.3254	10	-0.8482	1.051
Acetyl fentanyl	8.77	21	1	7	3.115	-0.9477	3.44	3	2.9148	2	-3.5271	10	-0.8639	0.556
Acryl fentanyl	8.80	22	1	8	3.871	-0.3568	3.55	3	3.6607	2	-4.1903	11	-0.8694	0.719
AH-7563	9.48	16	1	4	2.789	-1.1593	3.11	2	2.1827	1	-2.7305	6	-0.3528	0.243
AH-7921	9.48	16	1	4	3.997	0.4455	2.67	2	3.3908	1	-4.1698	6	-0.2913	0.731
AH-7959	9.40	19	1	4	4.848	-0.5999	3.00	3	4.3052	1	-4.7352	7	-0.3420	1.229
AH-8507	8.10	19	1	4	3.844	0.4028	2.89	3	3.7925	1	-4.0974	7	-0.3168	1.237
AH-8529	9.48	16	1	4	3.393	-0.2988	2.78	2	2.7867	1	-3.4595	6	-0.3205	0.470
AH-8532	9.48	16	1	4	3.393	-0.2988	2.78	2	2.7867	1	-3.4595	6	-0.3205	0.472
AH-8533	9.48	16	1	4	3.393	-0.2988	2.78	2	2.7867	1	-3.4595	6	-0.3205	0.292
Alfentanil	7.50	21	3	6	2.807	-0.8236	2.78	2	2.7937	1	-2.1832	9	-0.6933	1.131
α -methyl fentanyl	9.00	23	1	7	4.232	-1.0754	3.66	3	3.9287	2	-3.9243	10	-0.8745	0.807
Benzodioxole fentanyl	8.75	27	3	10	4.592	-1.2603	3.88	4	4.3996	3	-5.7148	15	-0.8363	1.122
Benzyl fentanyl	8.38	21	1	7	3.527	-1.1271	3.44	3	3.4328	2	-3.9618	10	-0.8639	0.770
β -hydroxy thiofentanyl	8.16	20	2	6	2.809	-1.0472	3.11	2	2.7506	1	-3.7899	9	-0.3564	0.585
Butyryl fentanyl	8.77	23	1	7	4.260	-1.1595	3.66	3	4.0600	2	-4.3671	10	-0.8745	0.991
Carfentanil	8.05	24	3	8	3.668	-1.0948	3.55	3	3.6218	2	-4.3472	11	-0.8193	1.055

Name	pKa	nC	nO	nDB	logP	AlogP	MlogP	nR06	logD (pH 9)	nBNZ	logS (pH 9)	DBE	Hy	RRT
Crotonyl fentanyl	8.80	23	1	8	4.258	-0.1570	3.66	3	4.0474	2	-4.5310	11	-0.8745	0.876
Cyclopentyl fentanyl	8.77	25	1	7	4.784	-1.1671	3.88	3	4.5838	2	-5.1412	11	-0.8836	1.406
Cyclopropyl fentanyl	8.77	23	1	7	3.895	-0.5911	3.66	3	3.6947	2	-4.1692	11	-0.8745	0.858
Despropionyl p-fluoro fentanyl	9.12	19	0	6	3.629	-0.2635	3.22	3	3.2627	2	-3.3914	9	-0.3992	0.878
Fentanyl	8.77	22	1	7	3.815	-0.8715	3.55	3	3.6154	2	-3.8770	10	-0.8694	0.731
Furanyl fentanyl	8.66	24	2	9	4.029	-1.1117	3.66	3	3.8657	2	-4.8436	13	-0.8482	0.847
Isopropyl U-47700	9.20	18	1	4	4.686	-0.1290	2.89	2	4.2754	1	-4.2550	6	-0.7721	1.214
Methacryl fentanyl	8.70	22	1	7	3.958	-0.3918	3.44	3	3.7806	2	-4.1578	10	-0.8365	0.843
m-fluoro fentanyl	8.80	23	1	8	4.267	0.1963	3.66	3	4.0563	2	-4.1393	11	-0.8745	0.822
m-fluoro methoxyacetyl fentanyl	8.70	22	2	7	3.084	-0.7315	3.33	3	2.9064	2	-3.6809	10	-0.8059	0.590
MT-45	9.00	24	0	6	5.634	-1.4056	3.88	4	5.3353	2	-3.5935	10	-0.9125	1.529
N-methyl norcarfentanil	7.56	17	3	5	1.654	-1.2597	2.78	2	1.6389	1	-2.6819	7	-0.7618	0.304
N-methyl U-47931E	9.20	16	1	4	3.473	-0.6787	2.78	2	3.0621	1	-3.3669	6	-0.7872	0.545
Norfentanyl	10.03	14	1	4	1.419	-1.5723	2.67	2	0.3552	1	-1.2397	6	-0.3135	0.204
Ocfentanil	8.62	22	2	7	3.084	-0.7315	3.33	3	2.9316	2	-3.7062	10	-0.8059	0.593
o-fluoro fentanyl	8.62	22	1	7	3.958	-0.3918	3.44	3	3.8060	2	-4.1832	10	-0.8365	0.860
p-fluoro methoxyacetyl fentanyl	8.74	22	2	7	3.084	-0.7315	3.33	3	2.8941	2	-3.6687	10	-0.8095	0.583
Phenyl fentanyl	8.75	26	1	10	4.969	-0.9576	3.99	4	4.7764	3	-5.3648	14	-0.8876	1.128
Phenylpropionyl fentanyl	8.77	28	1	10	5.394	-0.8604	4.21	4	5.1938	3	-5.1255	14	-0.8950	1.642
Propyl U-47700	9.20	18	1	4	4.792	-0.4063	2.89	2	4.3810	1	-4.4697	6	-0.7721	1.288
Remifentanil	7.51	20	5	6	1.515	-1.3044	2.89	2	1.5015	1	-2.9284	8	-0.7327	0.578
Seneciyl fentanyl	8.80	24	1	8	4.501	0.4919	3.77	3	4.2907	2	-4.7055	11	-0.8792	1.116
Sufentanil	8.86	22	2	6	3.605	-1.0992	3.33	2	3.3667	1	-3.8479	9	-0.8059	1.233
Tetrahydrofuran fentanyl	8.77	24	2	7	3.247	-1.8332	3.66	3	3.0473	2	-3.8090	11	-0.8484	0.619

Name	pKa	nC	nO	nDB	logP	AlogP	MlogP	nR06	logD (pH 9)	nBNZ	logS (pH 9)	DBE	Hy	RRT
Thiofentanyl	8.76	20	1	6	3.728	-0.8245	3.22	2	3.5326	1	-3.8742	9	-0.8228	0.789
U-47109	9.18	15	1	4	3.689	-0.1812	2.56	2	3.2876	1	-3.9257	6	-0.2713	0.643
U-47700	9.20	16	1	4	3.912	-0.0789	2.67	2	3.5014	1	-3.7892	6	-0.7504	0.693
U-47931E	9.18	15	1	4	3.249	-0.7810	2.67	2	2.8482	1	-3.5055	6	-0.3010	0.454
U-48520	9.20	16	1	4	3.308	-0.8232	2.78	2	2.8974	1	-3.0789	6	-0.7872	0.481
U-48800	9.20	17	1	4	3.893	-0.0441	2.78	2	3.4813	1	-3.3213	6	-0.7618	0.795
U-49900	9.76	18	1	4	4.626	0.4219	2.89	2	3.8014	1	-4.0039	6	-0.7721	1.106
U-50488	9.65	19	1	4	4.298	-0.8015	3.00	2	3.5576	1	-3.2295	7	-0.7816	0.985
U-51754	9.20	17	1	4	3.893	-0.0441	2.78	2	3.4813	1	-3.3213	6	-0.7618	0.857
U-62066	9.76	22	2	4	3.812	0.1139	3.22	2	2.9863	1	-3.2521	8	-0.7775	0.972
UF-17	8.96	17	1	4	3.209	-1.5837	3.00	2	2.9276	1	-2.8721	6	-0.8364	0.467
Valeryl fentanyl	8.77	24	1	7	4.705	-1.4475	3.77	3	4.5045	2	-4.8542	10	-0.8792	1.329

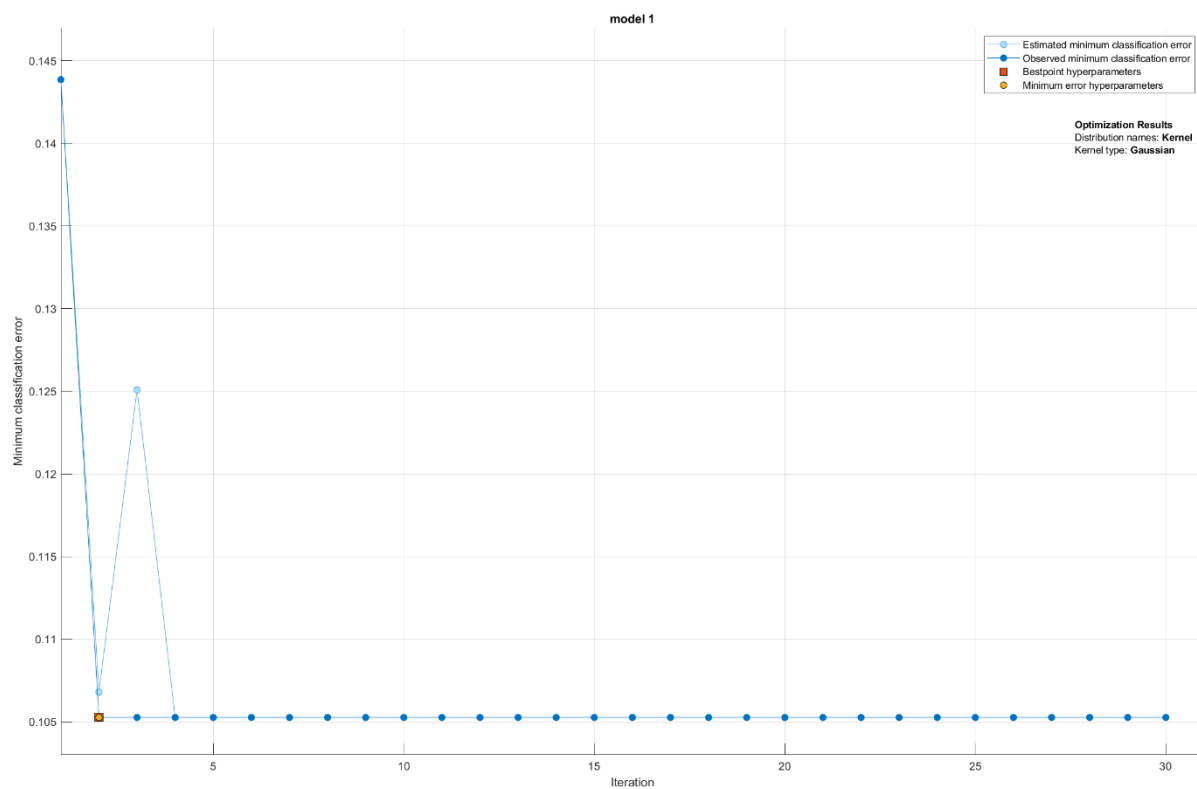


Figure A4-1: Plot showing the minimum classification error for each iteration of the Naïve Bayes classification model

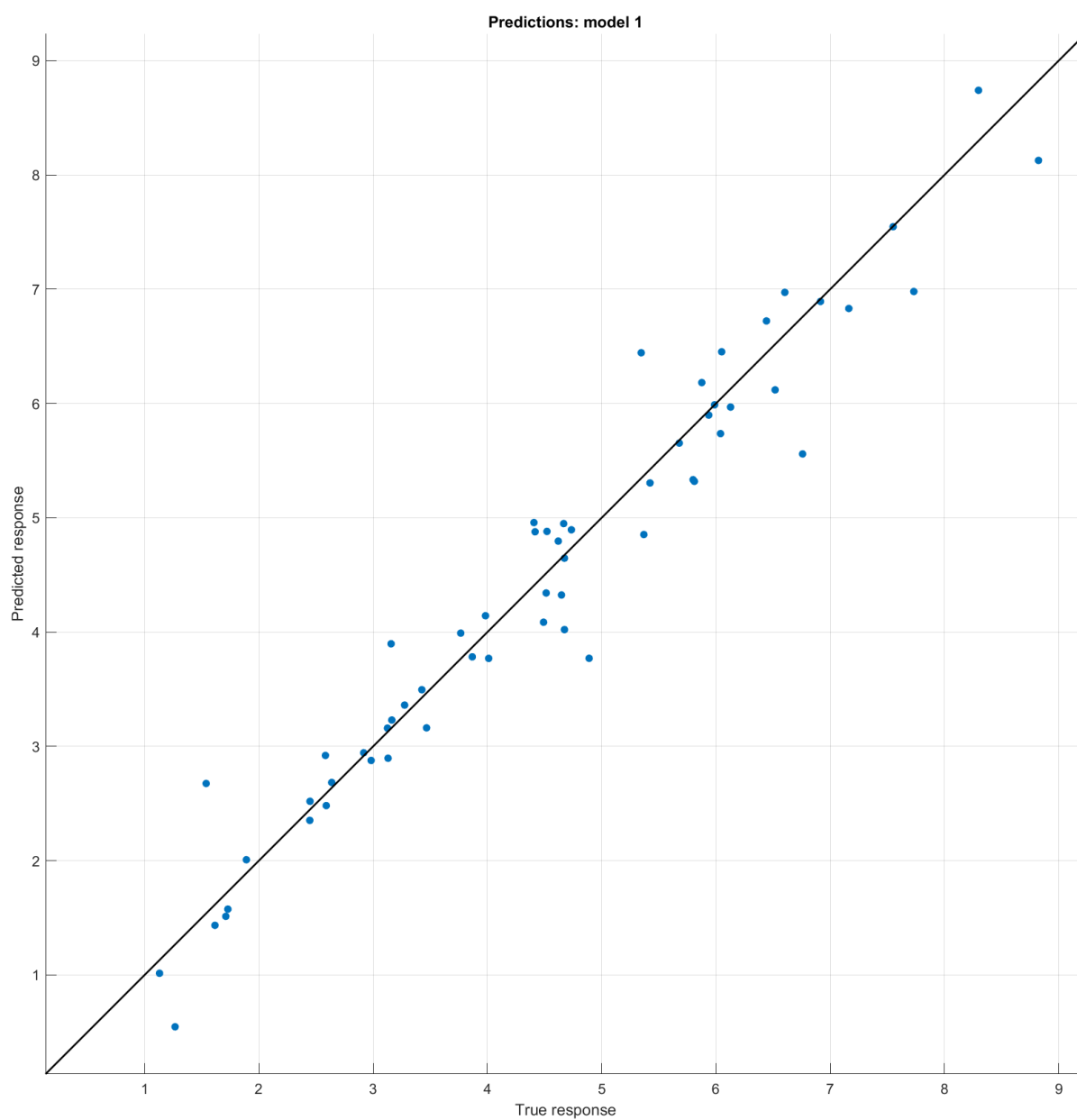


Figure A4-2: Predicted (RT_p) vs. experimental (RT_e) for the developed Gaussian Process Regression model using absolute retention time values.

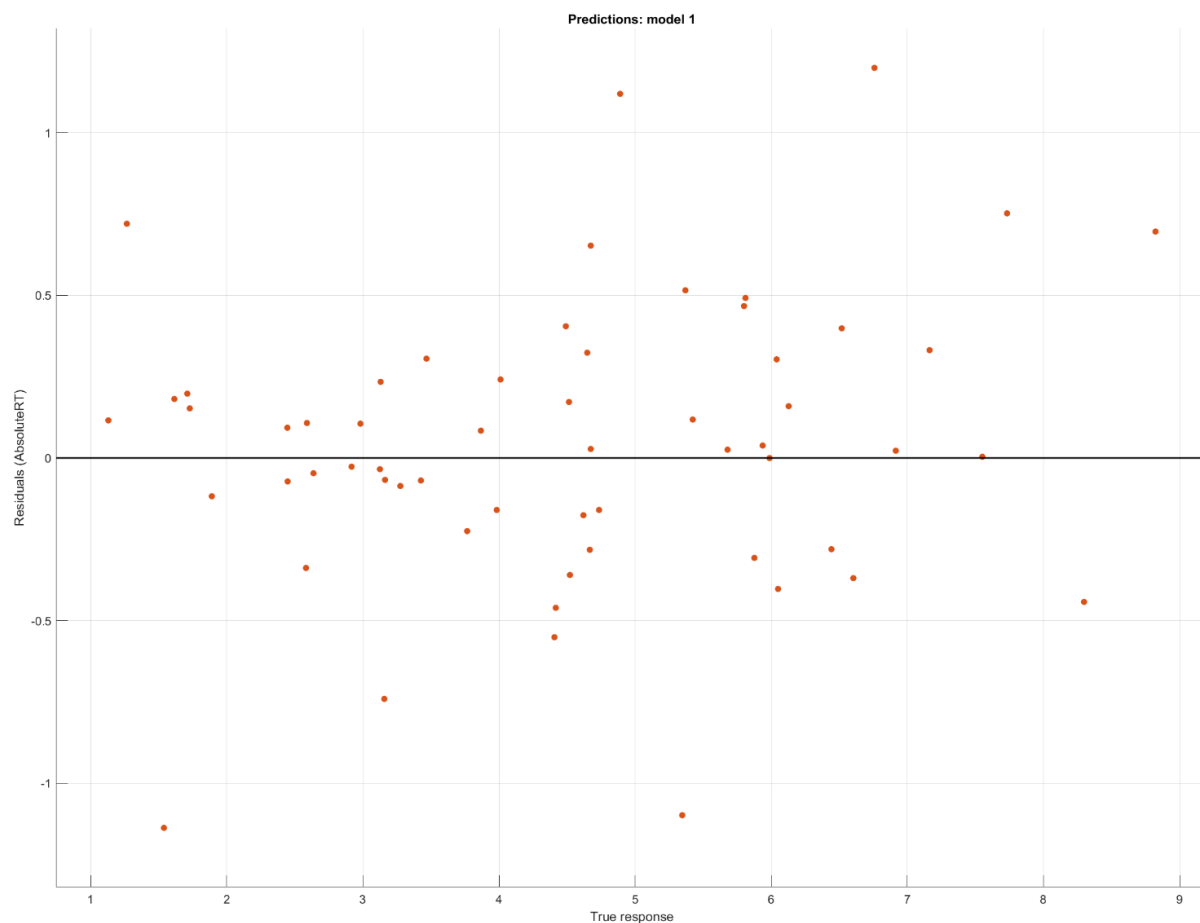


Figure A4-3: Residuals produced from the retention time prediction model using absolute retention time values.

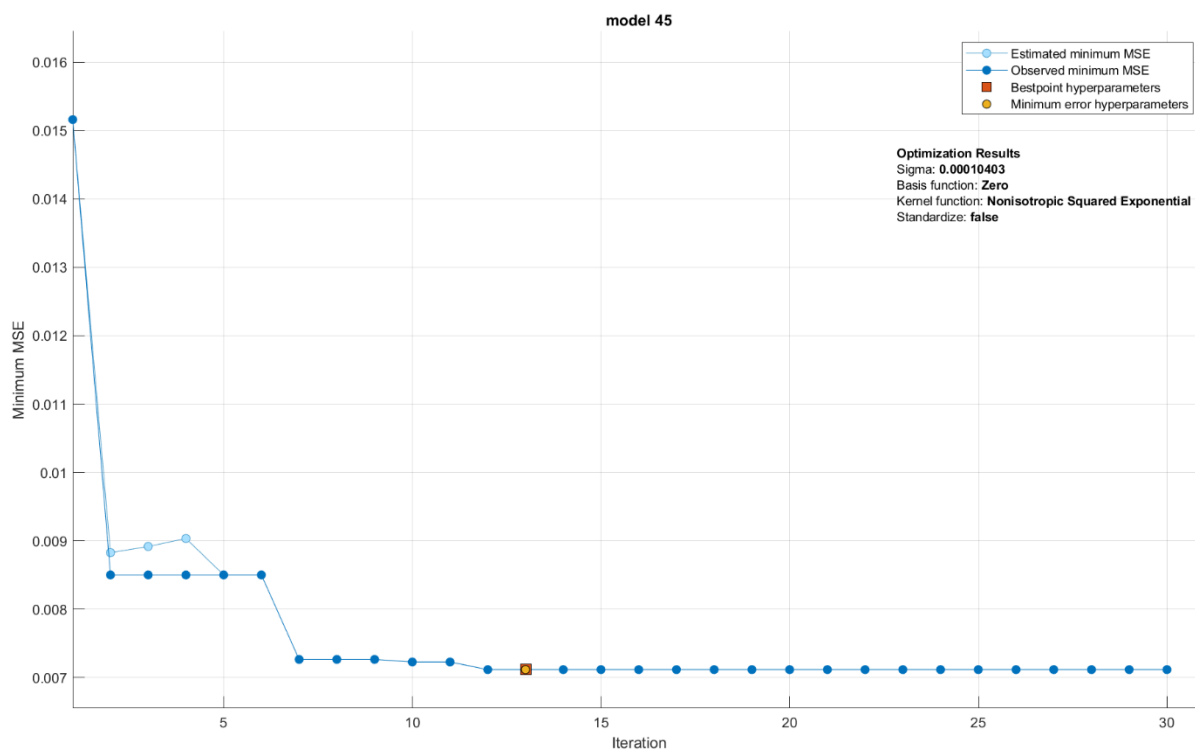


Figure A4-4: Plot showing the minimum mean square error (MSE) for each iteration of the Gaussian process regression model

Appendix 5

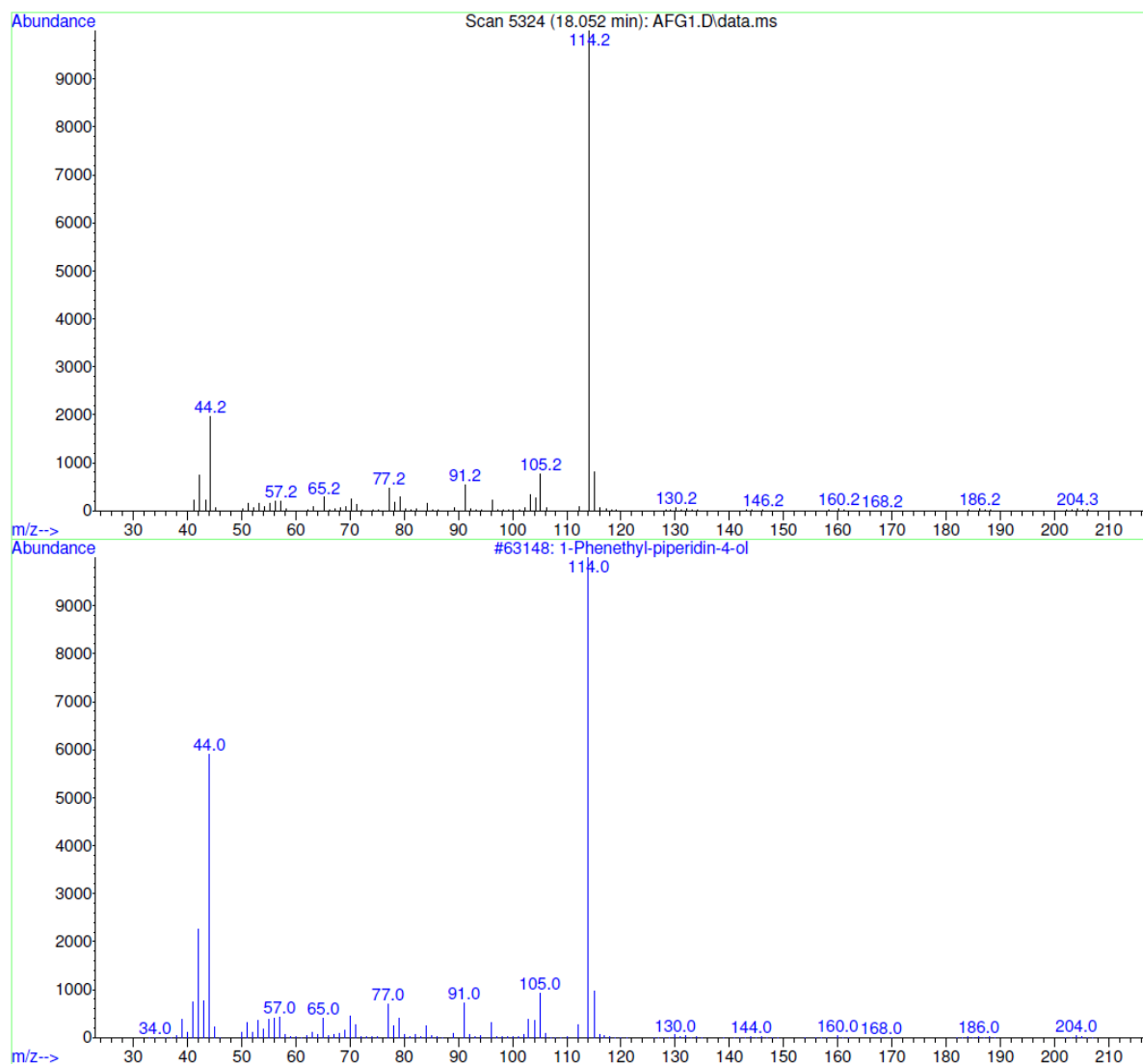


Figure A5-1: Library match for 1-phenethyl-piperidine-4-ol (Similarity 85%)

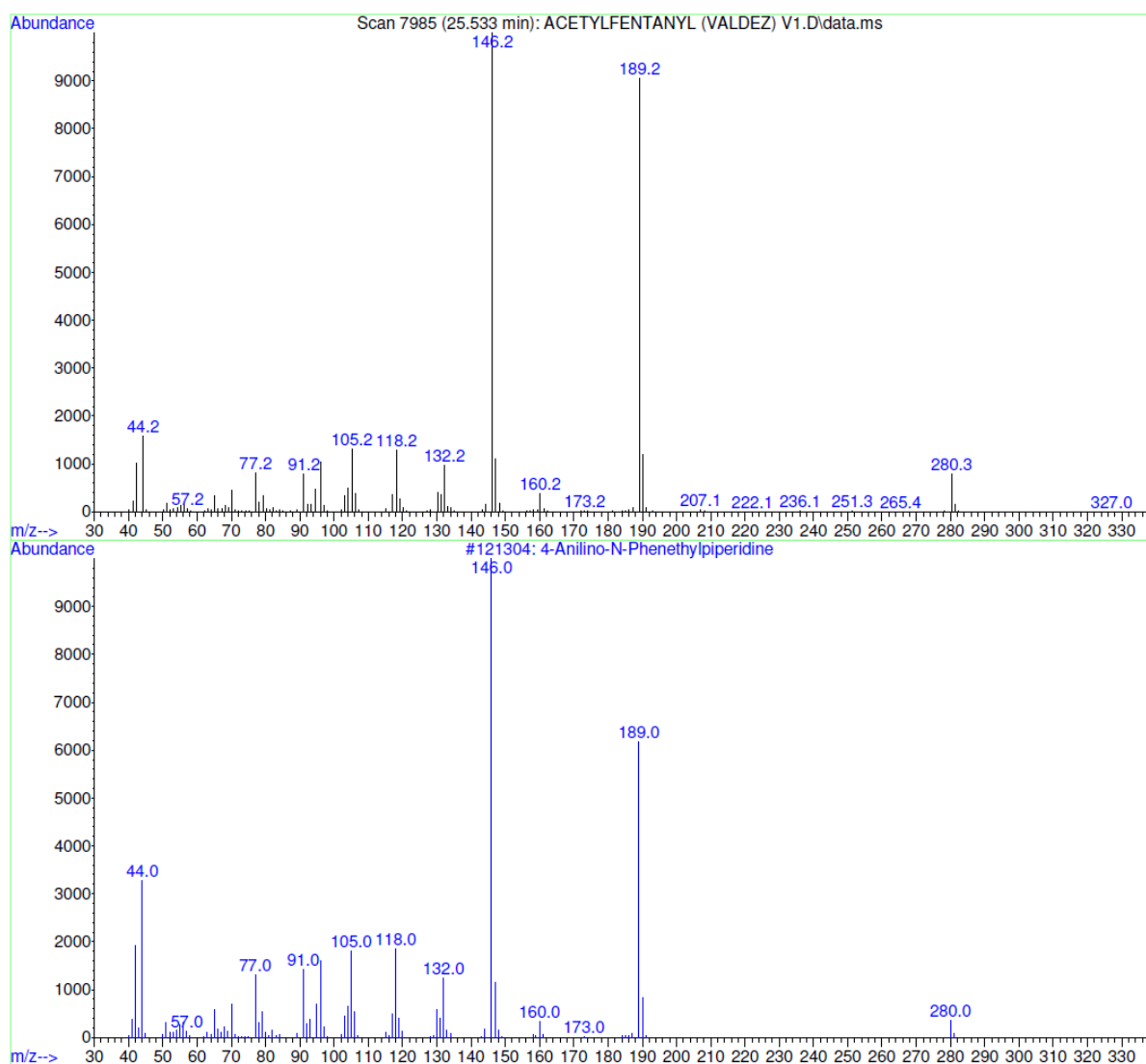


Figure A5-2: Library match for 4-anilino-N-phenethylpiperidine (Similarity 99%)

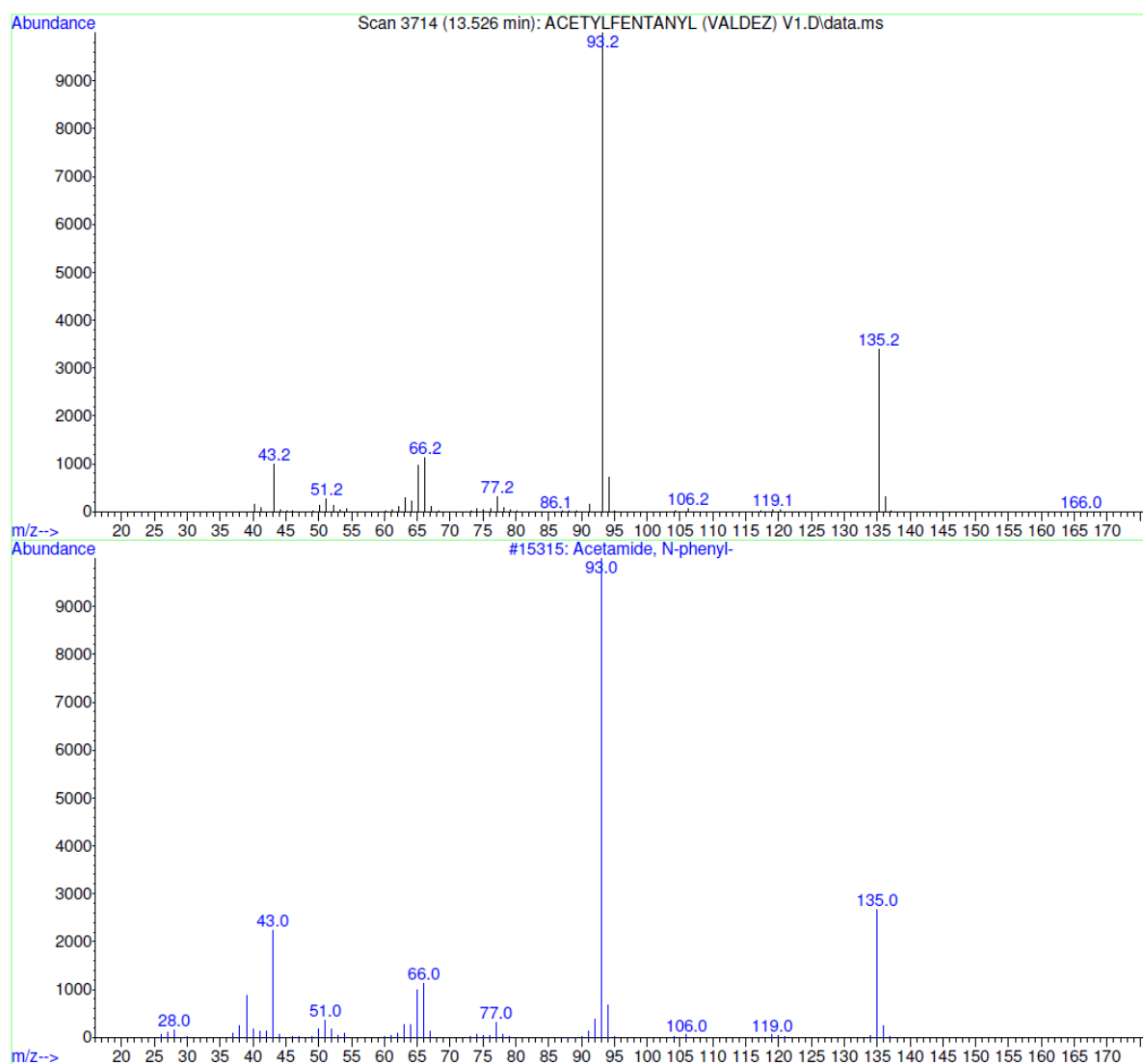


Figure A5-3: Library match for acetanilide (Similarity 91%)

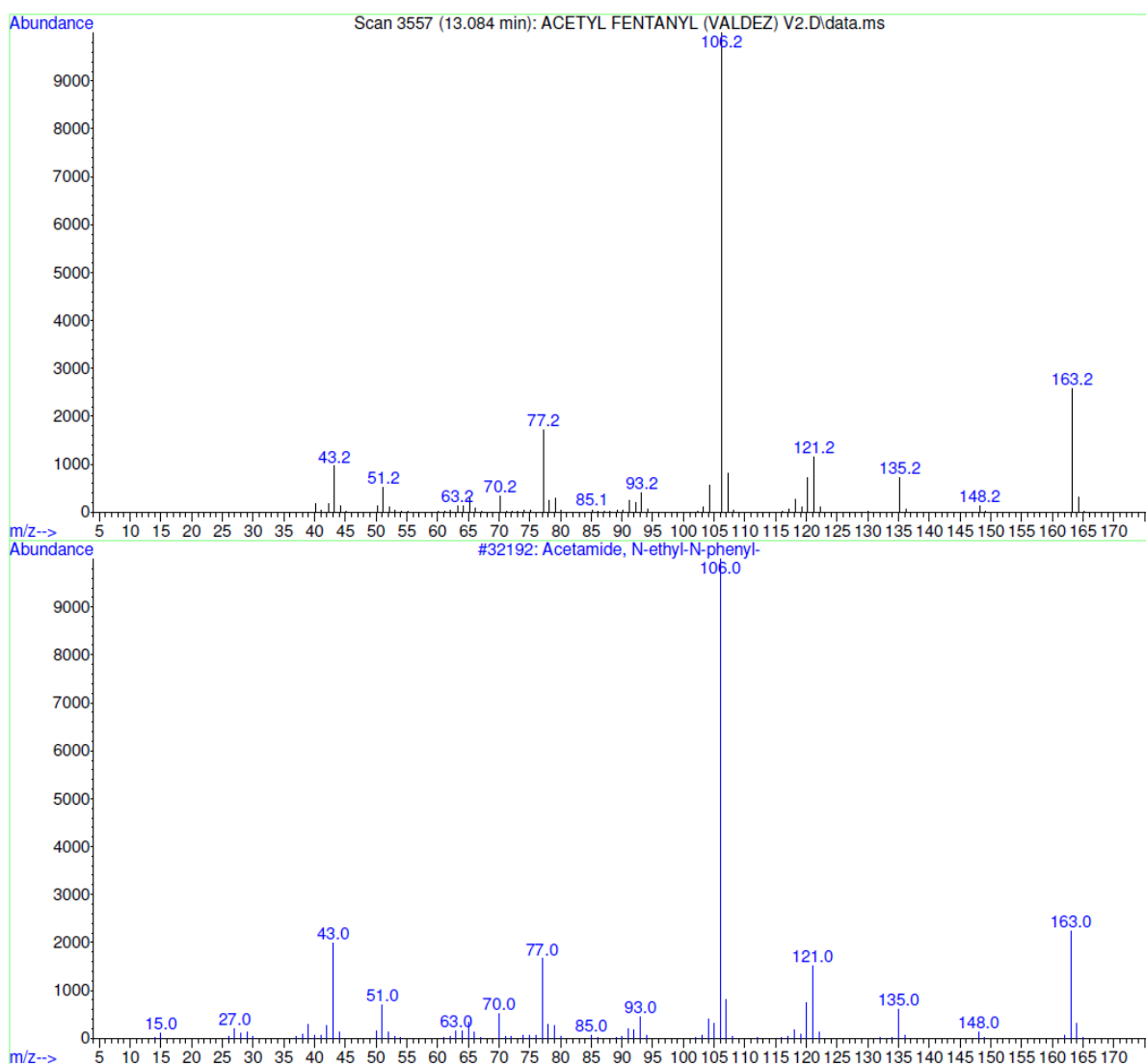


Figure A5-4: Library match for N-ethylacetanilide (Similarity 90%)

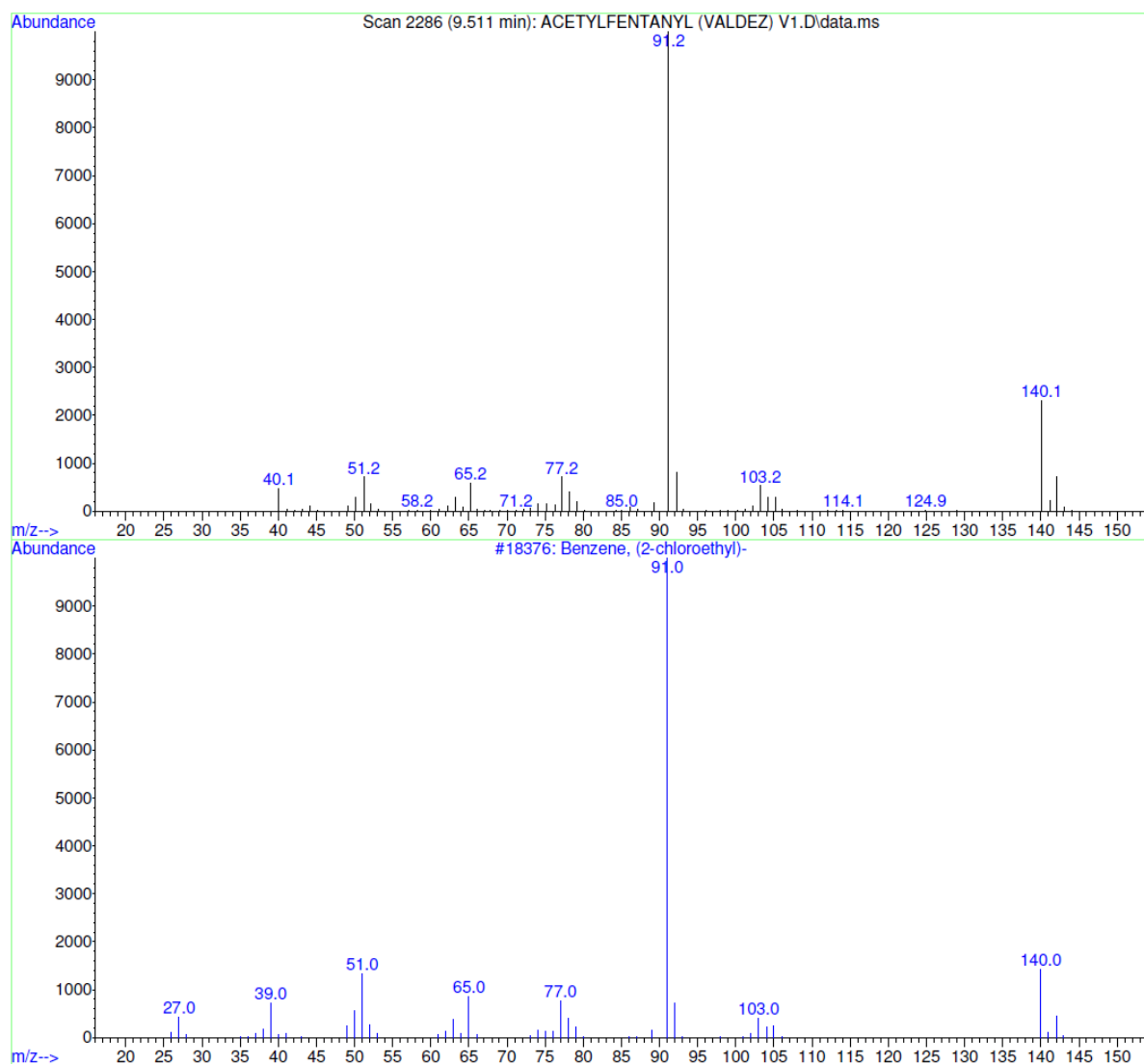


Figure A5-5: Library match for (2-chloroethyl)benzene (Similarity 87%)

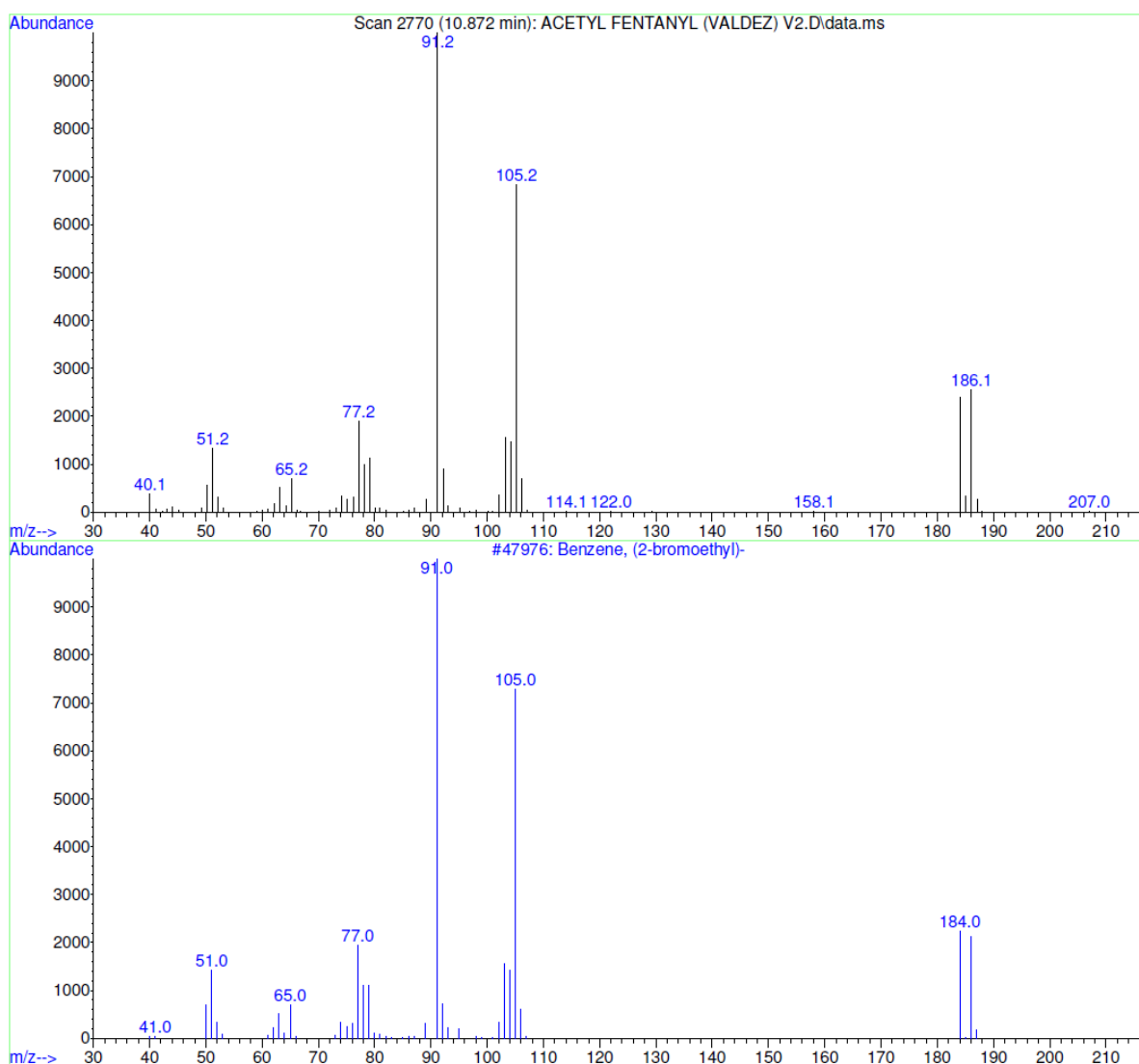


Figure A5-6: Library match for (2-bromoethyl)benzene (Similarity 98%)

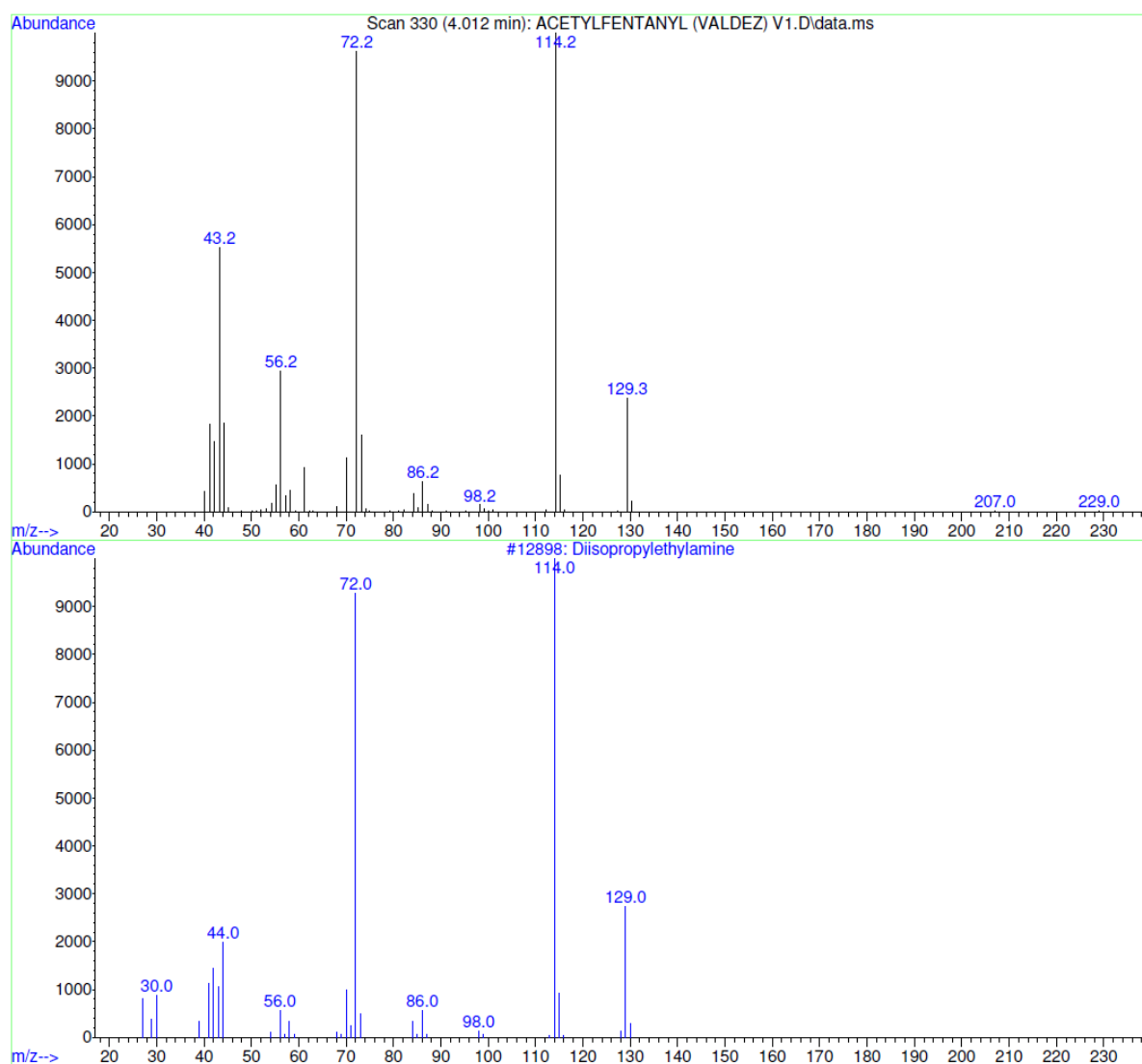


Figure A5-7: Library match for diisopropylethylamine (Similarity 87%)

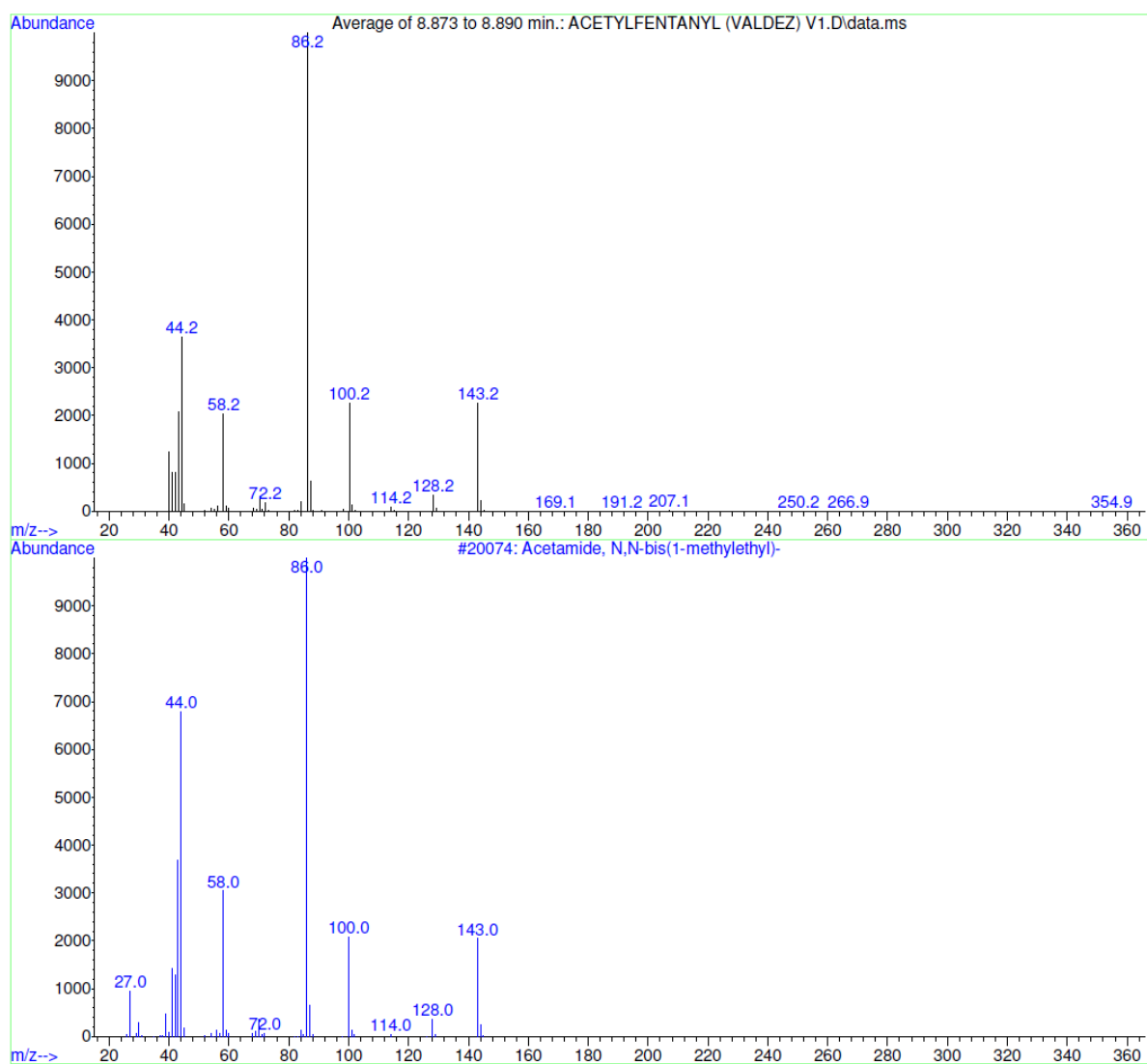


Figure A5-8: Library match for *N,N*-diisopropylacetamide (Similarity 91%)

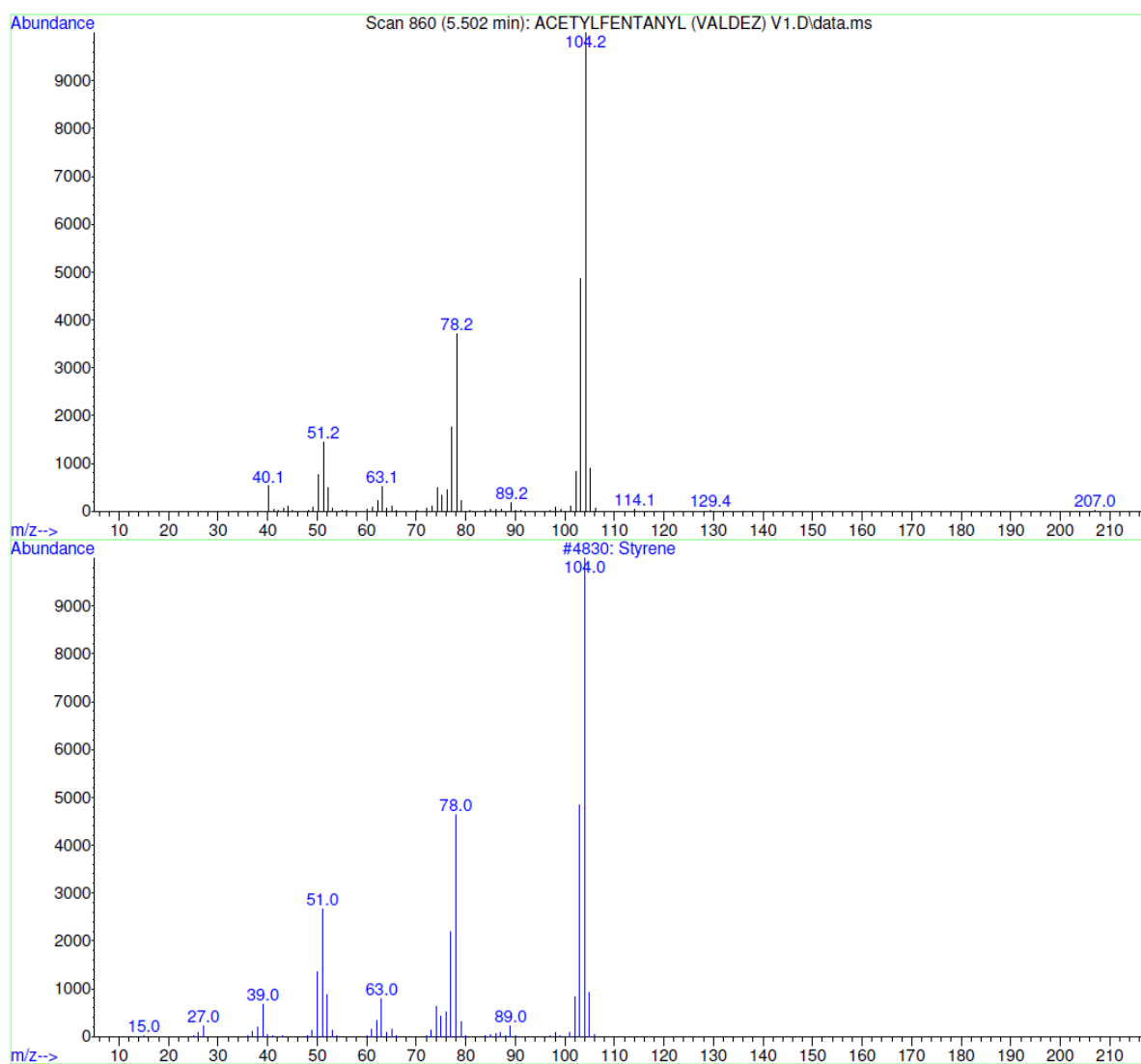


Figure A5-9: Library match for styrene (Similarity 94%)

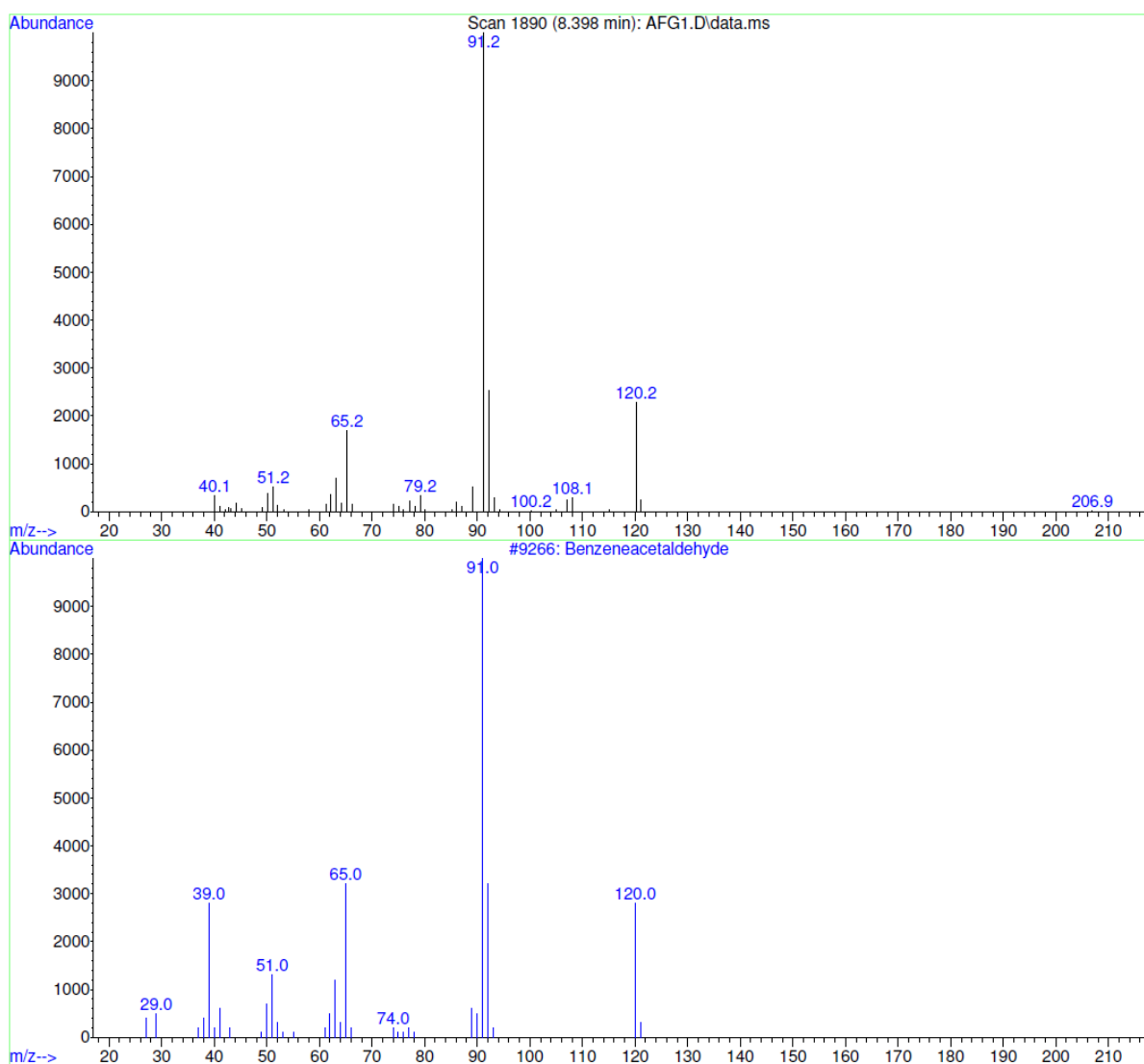


Figure A5-10: Library match for phenacetaldehyde (Similarity 90%)

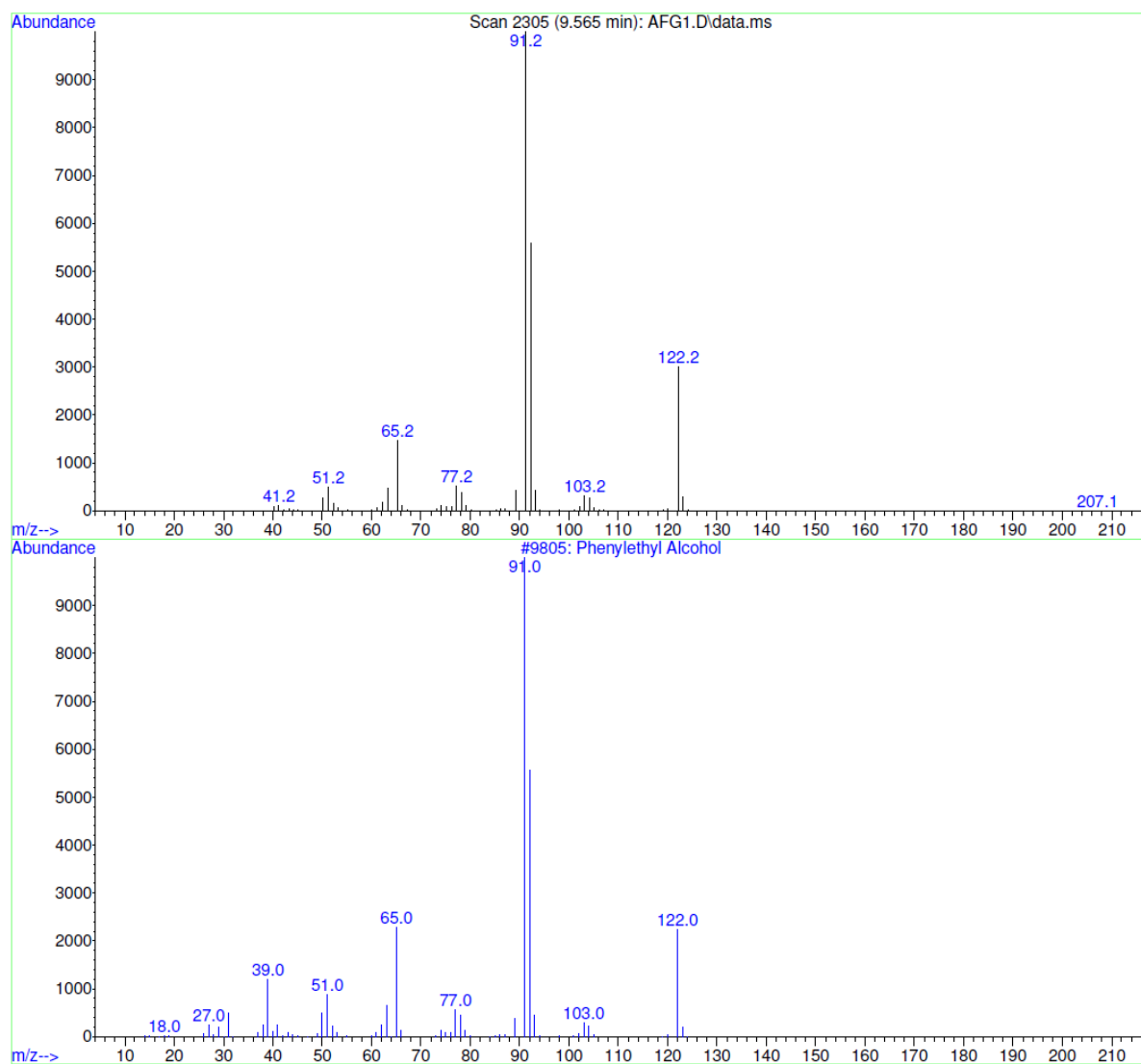


Figure A5-11: Library match for phenethyl alcohol (Similarity 94%)

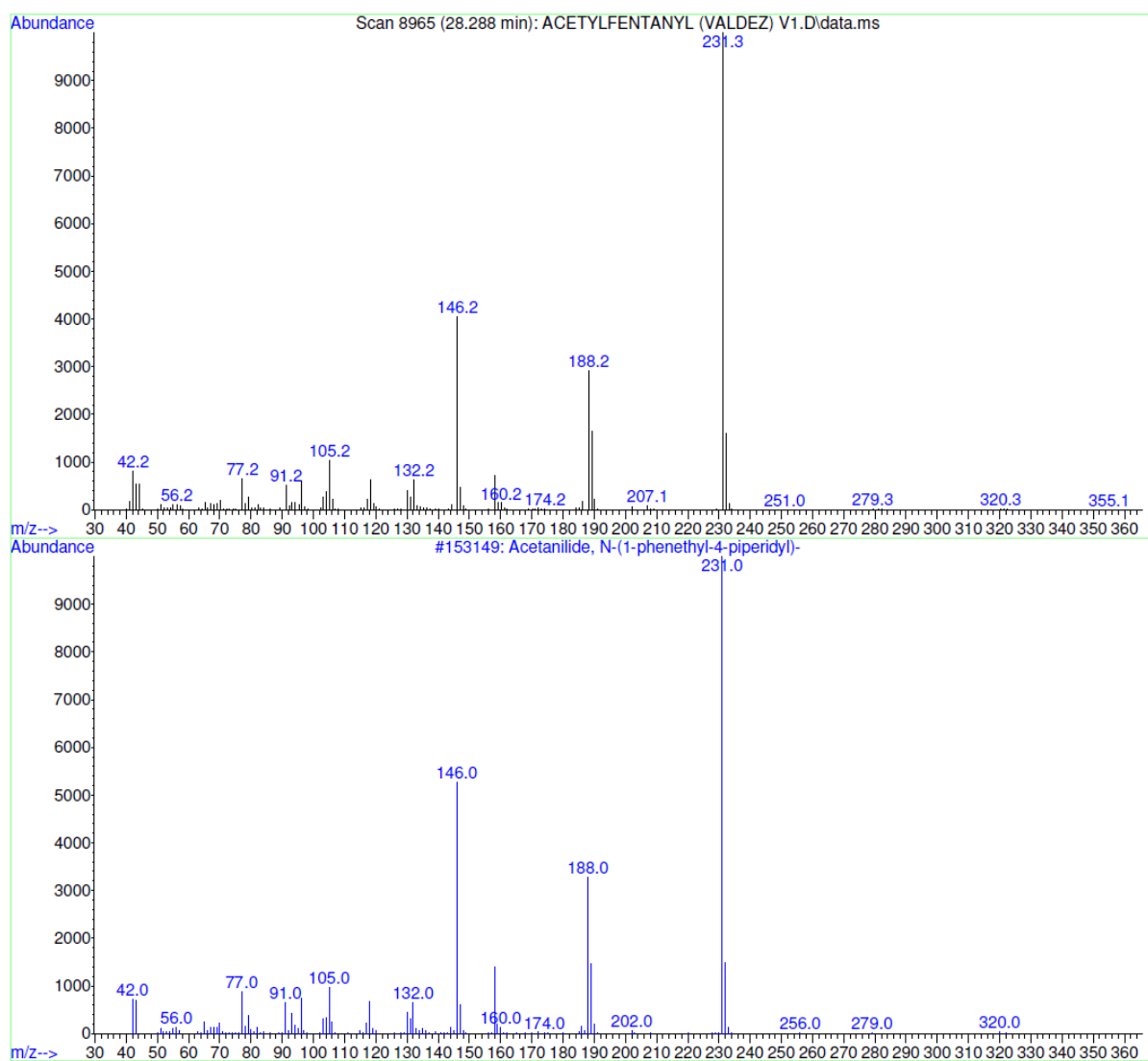


Figure A5-12: Library match for acetyl fentanyl (Similarity 94%)

Table A5-1: Entity list produced from the GC-MS data used for statistical analysis

Entity Number	Mass (u)	Retention Time (min)
1	146.2	25.568
2	143.2	8.881
3	114.2	18.025
4	104.2	5.502
5	186.1	10.872
6	163.2	13.084
7	106.2	13.104
8	120.2	8.398
9	140.1	9.517
10	103.2	9.517
11	122.2	9.565
12	280.3	25.533
13	129.3	4.012
14	147.2	32.672
15	135.2	13.526
16	103.2	18.025

Table A5-2: Entity list produced from the LC-MS data used for statistical analysis

Entity Number	Mass (u)	Retention Time (min)	Entity Number	Mass (u)	Retention Time (min)
1	94.0590	0.672	33	160.0720	0.686
2	322.2014	0.675	34	150.1027	0.678
3	322.2024	1.627	35	212.1697	0.670
4	322.2010	2.154	36	126.0654	0.682
5	135.0668	1.856	37	158.0551	0.682
6	363.2635	0.672	38	268.1987	0.675
7	190.0824	0.695	39	70.0416	0.678
8	247.1515	0.675	40	328.0960	0.678
9	190.0830	1.967	41	280.2372	0.675
10	188.0680	0.686	42	414.2368	0.662
11	256.0561	2.282	43	400.1103	0.678
12	294.2149	0.678	44	168.0731	0.678
13	100.0503	0.678	45	379.1798	0.682
14	384.1358	0.682	46	180.0781	0.678
15	255.2514	0.670	47	140.9508	0.549
16	208.1055	0.686	48	458.2648	0.662
17	334.2065	0.678	49	142.0604	0.678
18	228.0587	0.692	50	121.9564	0.549
19	168.0151	0.686	51	122.0711	0.678
20	148.0153	0.678	52	80.9297	0.549
21	84.0550	0.682	53	502.2893	0.666
22	152.0431	0.678	54	318.0240	2.278
23	148.0159	1.540	55	98.0709	0.678
24	126.0999	0.675	56	267.0646	0.675
25	188.0676	2.284	57	273.2627	0.668
26	122.1083	0.678	58	256.2381	0.675
27	228.0588	1.584	59	154.0989	0.682
28	123.9831	0.736	60	82.0413	0.682
29	114.0668	0.682	61	106.0765	0.675
30	296.2315	0.675	62	371.0776	0.678
31	138.1012	0.678	63	432.1249	0.672
32	202.0489	0.678	64	280.1918	2.801

Entity Number	Mass (u)	Retention Time (min)	Entity Number	Mass (u)	Retention Time (min)
65	288.2186	0.661	68	363.2629	1.424
66	405.2727	0.662	69	166.1407	0.662
67	205.1439	1.391			

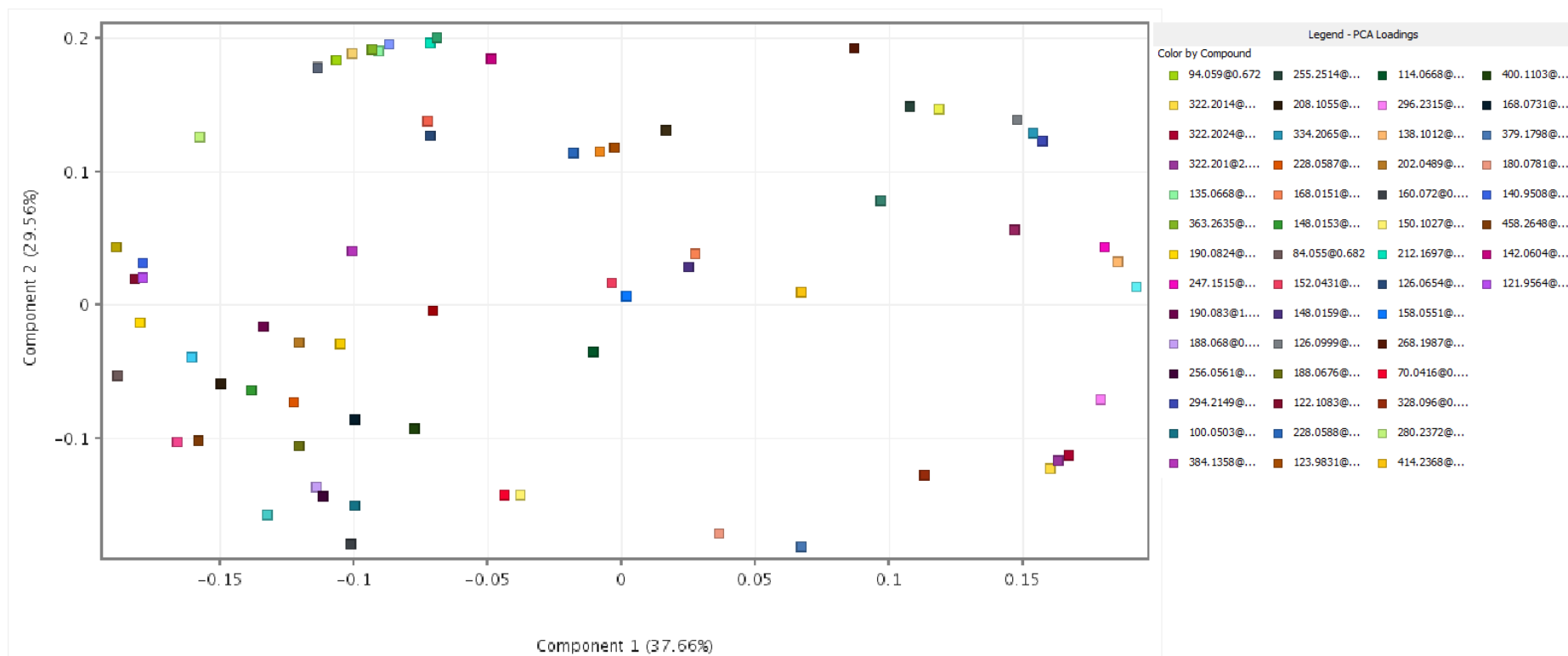


Figure A5-13: PCA loading plot for acetyl fentanyl samples analysed by LC-MS