CLASSIFICATION OF CANNABINOIDS USING MASS SPECTRAL DATA TO ASSIST IN THE IDENTIFICATION OF NOVEL SYNTHETIC CANNABINOIDS

A thesis presented to the faculty of the Graduate School of Western Carolina University in partial fulfillment of the requirements for the degree of Master of Science in Chemistry

By
Kristopher Charles Evans-Newman

Advisor: Dr. Nuwan Perera
Assistant Professor of Forensic Science
Department of Chemistry and Physics
Western Carolina University

Committee Members: Dr. Scott Huffman, Department of Chemistry and Physics
Dr. Rangika Koralege, Department of Chemistry and Physics

April 2024

Acknowledgements

A special thanks to my committee members for helping me with advice, assistance, and their time in not only every class but every committee meeting. Dr. Perera you helped drive me to do more and better yet challenging tasks throughout my time on this research. As well as Garion Schneider who helped me learn PLS-Toolbox and MATLAB.

My parents for always pushing me for knowledge and success especially my mother who passed away before she could fulfill her dream to see me graduate at Western Carolina with my bachelor's degree and missed the opportunity to see me continue higher education.

My loving wife Lynn, who helped me better explain complicated topics in this research for general understanding. She has and always will be a special reason to call Cullowhee home as she is a piece of the foundation for my motivation.

TABLE OF CONTENTS

LIST OF TABLES

\

LIST OF FIGURES

List of Abbreviations

BUT: Butyl
CB1: Cannabinoid receptor 1
CB2: Cannabinoid receptor 2
CCC: Cayman Chemical Company
CDB: Cannabidiol
CV: Cross Validation
DEA: Drug Enforcement Agency
EI: Electron Ionization
FNR; False Negative Rate
FPR: False Positive Rate
FUB: 4-Fluorobenzyl
GA: Genetic Algorithm
GC: Gas Chromatography
GCMS: Gas Chromatography Mass Spectroscopy
HCA: Hierarchical Clustering
HRMS: High Resolution-Mass Spectroscopy
LV: Loading Vector
MATLAB: MATrix LABoratory
MDMB: MethylDiMethylButononate
MIA: Multi Image Analysis
MS: Mass Spectroscopy
NMR: Nuclear Magnetic Resonance
NPSs: New Psychoactive Substances
NSCs: New Synthetic Cannabinoids
PCA: Principal Component Analysis
PLS-DA: Partial Least Squares-Discriminate Analysis
SCRA: Synthetic Cannabinoid Receptor Agonist
SWGDrug: Scientific Working Group for The Analysis of Seized Drugs
THC: Tetrahydrocannabinol
THCa: Tetrahydrocannabinolic acid
TNR: True Negative Rate
TPR: True Positive Rate
UNODC: United Nations Office of Drugs and Crime

ABSTRACT

Detection and characterization of newly synthesized cannabinoids (NSCs) is challenging due to the lack of availability of reference standards and chemical data. Identifying these substances usually involves the use of intelligence gathering or prior knowledge on new psychoactive substances (NPSs). The current methods for the structural elucidation of NPSs are using expensive and time-consuming analysis methods such as high-resolution mass spectrometry (HRMS), and nuclear magnetic resonance spectroscopy (NMR). The focus of this study is to develop a solution to identify NSCs. A classification system was developed using existing mass spectral data of synthetic and classical cannabinoids to determine the presence of previously unknown cannabinoid-related substances encountered in laboratories. Research used computer learning software from Eigenvector, called PLS -Toolbox, which is an addon in a MATLAB programming language. Principle component analysis (PCA) and partial least square discriminant analysis (PLS-DA) were used to develop a binary classification system using mass spectra obtained from a freely available spectral database maintained by the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG). Genetic algorithm (GA) was used to select the most discriminatory mass to charge (m/z) ratios of the mass spectral data. First, a binary classification model was developed to discriminate cannabinoids and cannabinoid-related compounds from other drug classes such as opioids, tryptamines, fentanyl and fentanyl derivatives. Then, a classification model was developed to discriminate classical and synthetic cannabinoids. Finally, sub-models were developed to discriminate the presence and absence of functional groups found on commonly encountered synthetic cannabinoids. Hierarchical cluster analysis (HCA) in conjunction with PCA was used to determine the possible drug classes and HCA along with chemical structure similarities resulted in the different unique groups seen with synthetic

cannabinoids that were classified. These groups include benzopyrrole (indole), isoindazole (indazole), naphthalene (Naphthyl), 4-Fluorobenzyl **(**FUB**),** and 1-amino-3,3-dimethyl-1-oxobutan-2-yl (BUT). Classification models were developed for the determination of the presence or absence of functional groups in an unknown cannabinoid. Current results show that these models are highly accurate (>95%) and applicable in determining the presence of cannabinoid-related substances and different functional groups. Limitations encountered include issues with mass spectral fragmentation patterns for similar structural compounds such as indazoles and azaindoles. Future research motivation includes building on to the binary classification system. Some of these groups include but are not limited to 3-dimethylbutanoate (MDMB) and N-[1-(Aminocarbonyl)-2,2-dimethylpropyl]-1-butyl (ADB).

CHAPTER ONE: INTRODUCTION

Background

New Psychoactive substances (NPSs) are newly discovered compounds that mimic known illicit drugs such as opioids, tryptamines, and cannabinoids. [1] Between 2008 and 2016 240 New synthesized cannabinoids have been reported to the United Nations Office of Drugs and Crimes (UNODC). NPSs also show outbreaks of acute toxicity for example in 2016 there were 33 cases of a new synthetic cannabinoid of 33 cases of exposure to AMB-FUBINACA.[7] These drugs are hard to detect and identify because they have no reference or standards as a base. This is important because forensic analysis relies on accurate standards and references to identify and quantify drugs and other illicit substances in seized evidence. Forensic labs use chromatography and spectrometry techniques to analyze samples to submit to court hearings. The most relied on methods are gas chromatography mass spectrometry (GCMS) and liquid chromatography mass spectrometry (LCMS). The method of mass spectroscopy relies on a library to match the data to a reference molecule in the library. NPSs however, are not in reference libraries for these techniques and must be updated if a new molecule is identified. To identify NPSs, forensic labs must rely on intelligence data on NPSs or expensive and time-consuming structural elucidation methods such as nuclear resonance spectroscopy (NMR) or high-resolution mass spectrometry. Most laboratories do not have these means of analysis, which means these samples are stored until a later date or are sent to other labs that could be behind on sample processing.[5]

Cannabinoids: Natural and Synthetic Cannabinoids

Cannabinoids are natural compounds found in *Cannabis sativa* plant. There are multiple isomers and derivatives of cannabinoids that are naturally occurring in the plant. Cannabinoids that are primary targeted from the plant is Δ^9-Tetrahydrocannabinol (Δ^9-THC) and cannabidiol (CDB). In the body, Δ^9-THC causes psychedelic effects as well as in some cases pain and anti-anxiety affects. This occurs when Δ^9-THC binds to the cannabinoid receptors in the body. The compounds that bind to these receptors are sometimes referred to as cannabinoid receptor agonists. CBD, a cannabinoid presents in hemp in abundance, does not cause psychedelic effects but helps in reducing pain and show relaxation effects. The chemical structure of most cannabinoids found in Cannabis plant share a tricyclic core that contains a hydrocarbon tail (Figure 1.). [5,6]

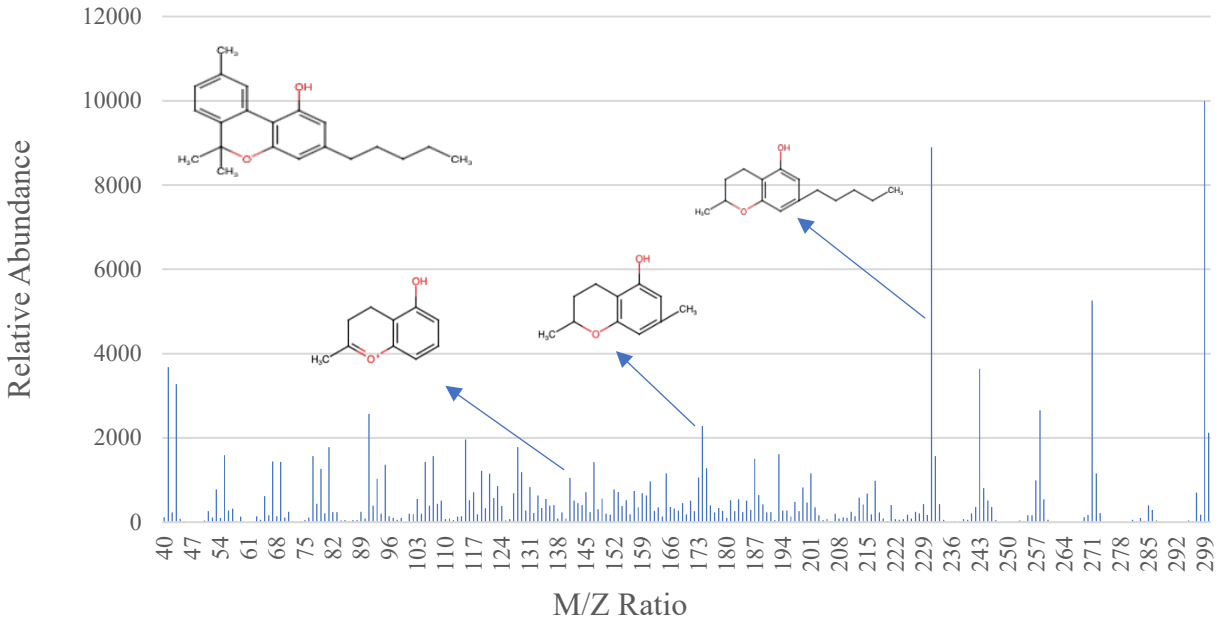


Figure 1: Mass spectrum of Delta-9-THC and common fragments.

Synthetic cannabinoids are designed compounds that mimic the effects of naturally occurring cannabinoids in the body. These are mainly manufactured in Asia where quality control and precursor laws are low or nonexistent. These compounds are mixed with plant matter or incense to smoke or a liquid that can be vaporized. [5] Both natural and synthetic cannabinoids have an affinity to the CB1 and CB2 receptors, commonly referred to as cannabinoid receptors, primarily in the brain but found throughout the body as a part of the endocannabinoid system that is responsible for homeostasis and neurotransmitter communication. Thus, these synthetic compounds are sometimes referred to as synthetic cannabinoid receptor agonists (SCRAs). These receptors are a part of a unique system called the endocrine system where our body produces natural occurring endocannabinoids.[9] Synthetic cannabinoids made their street debut in 2004 when JWH-018 was encountered in Europe. JWH is short for Dr. John W. Huffman who researched the synthesis and the effects of over 400 synthetic cannabinoids. It was found that synthetic cannabinoids not only could mimic the effects of THC but also be up to 100 times stronger. There have been cases of toxicity and health issues related to the renal system due to the tacticity of these cannabinoids. Synthetic cannabinoid nomenclature starts either with the initials of the synthesizer or common names which are abbreviations that arise from the International Union of Pure and Applied Chemist (IUPAC) naming system. The IUPAC names of synthetic cannabinoids are impractical due to their length as well as the complex structures of synthetic cannabinoids. The abbreviation should breakdown all key functional groups of the synthetic cannabinoid.[5,9] Structurally these compounds have four unique groups known as the head, linker, core, and tail groups (see Figure 2).[5]

Group	Color
Head	Blue
Linker	Green
Core	Red
Tail	Yellow

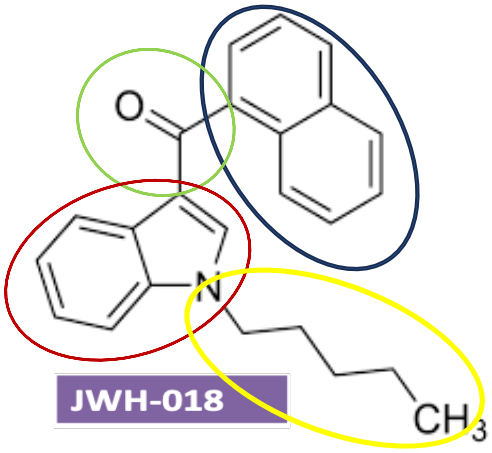


Figure 2. Structure of JWH-018 demonstrating the groups of a typical synthetic cannabinoid.

Each of these structural groups provide a specific mass spectral pattern as the tail group and head group are the primary point of cleavage during the ionization process of the mass spectrometer. The core structures of these compounds are multiple cyclic groups that are protected from the election impact due to the surrounding groups. Synthetic cannabinoids primarily have less fragmentation when ionized due to these bulky groups surround the core of the molecule. In contrast, in natural cannabinoids the cyclic core is more exposed, thus more fragments can be formed during the ionization process. The core structure of synthetic cannabinoids is less electronegative as well which can provide general protection against ionization. The difficulty comes to pinpointing the linker group due to the point of cleavage at the carbonyl as it can cleave with either at the head or core groups.

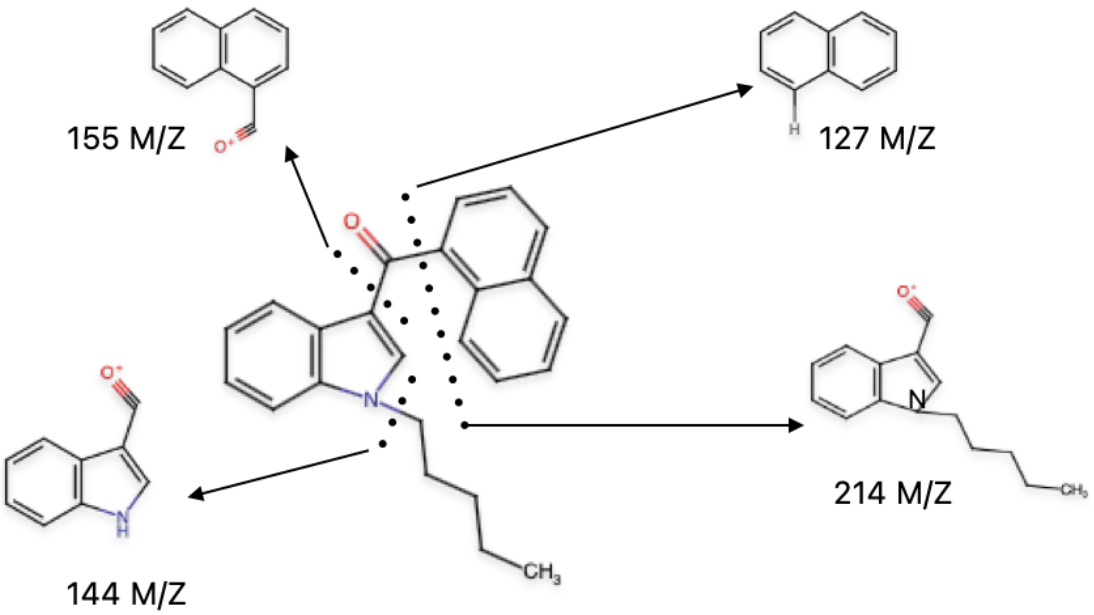


Figure 3. Fragmentation of synthetic cannabinoid JWH-018

In the reaction with the electron ionization (EI), it is shown that the core group tend to cleave with the linker group as well as the tail group. This shows how the core group trend to stay intact during the EI process. This also shows how the tail and head groups ten to cleave from the core structure to produce smaller fragments. These structures are most common with all synthetic cannabinoids. This trend and their structures can be seen on the mass spectra in figure 4 for JWH-018. The smaller fragments are produced when the tail and head structure cleave into smaller ionized fragments which can produce issues when it comes to discriminating these compounds as these smaller fragments with lower abundance can be seen in multiple different tail and head structures.

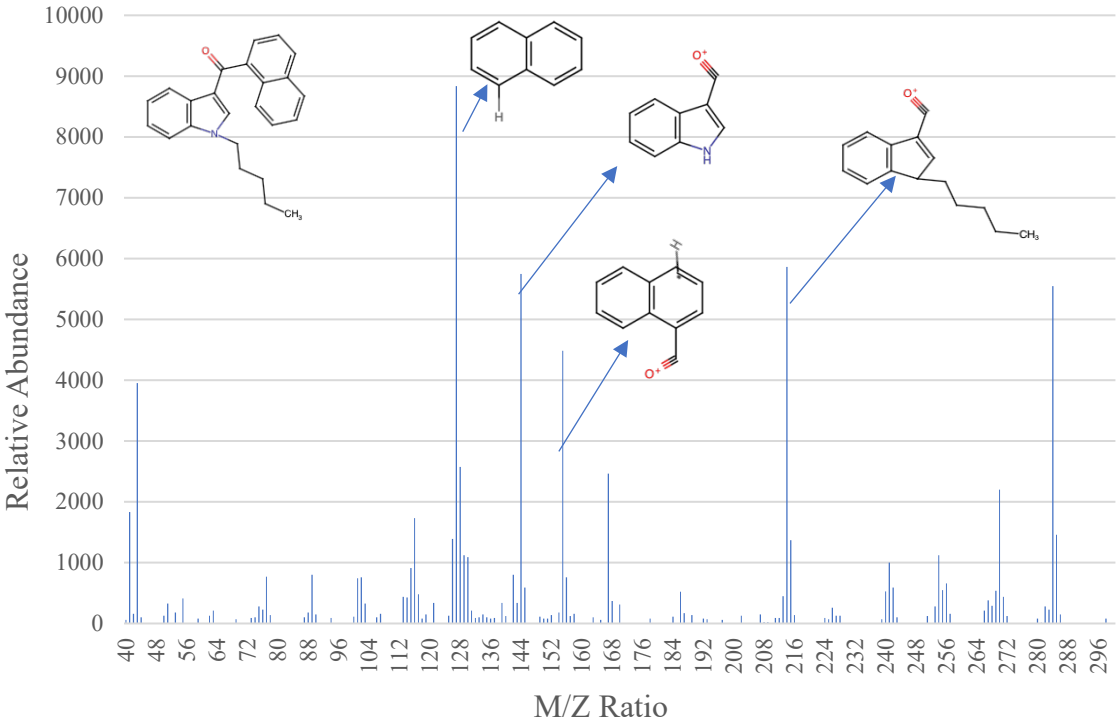
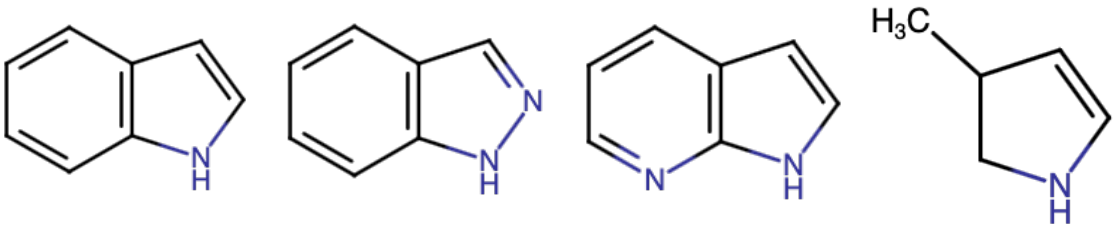


Figure 4. Fragmentation pattern with fragment structures of synthetic cannabinoid JWH-018 as an example of common fragmentations for synthetic cannabinoids

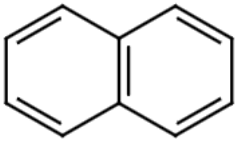
Figure 5 below shows common structural groups found in synthetic cannabinoids. These groups are the target groups of this research because of their commonality. These structures have also shown to be not only stable but the backbone of other possible synthetic cannabinoid structures that can be encountered. These modifications are relatively small such as adding a methyl group or changed though halogenating the structure. These structural groups because of this stability also play a part in their ability to be agonist and antagonist to the cannabinoid receptors throughout the body. There are also many other groups that need to be researched and modeled using this research method for a more complete analysis of the fragmentation patterns to use for future predictions.

Core Groups



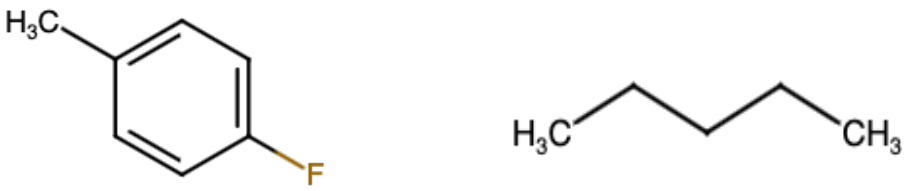
Indole Indazole Azaindole Nathtrolpyrole

Head Groups



Naphthyl

Tail Groups



FUB BUT

Figure 5. Common structures of primary interest of synthetic cannabinoids

Gas Chromatography

Forensic science laboratories widely use gas chromatography (GC) as a confirmatory technique for the presence of a compounds as well to separate mixtures for analysis.[1] These analyses include drug, alcohol, and other toxicology methods. GC is excellent at the separation of

complex matrixes like blood, other bodily fluids, and unknown powder and liquid mixes that are commonly encountered. When the sample is introduced to the instrument the sample is heated and volatilized and pushed through a capillary column by the mobile phase which is normally an inert gas such has helium, nitrogen, and argon called the carrier gas. In the capillary column there is a powder or gel known the stationary phase that can delay the elution of compounds based on the interactions. The GC and ramp the temperature so it slowly rises and can separate compounds based off boiling point as well as interaction such as polarity in the column.. GC has multiple different types of detectors such as thermal conductivity, flame ionization detector, and ultraviolet detectors. GC systems can also be in tandem with other instruments such as infrared spectrometers or mass spectrometers.[2,7,1]

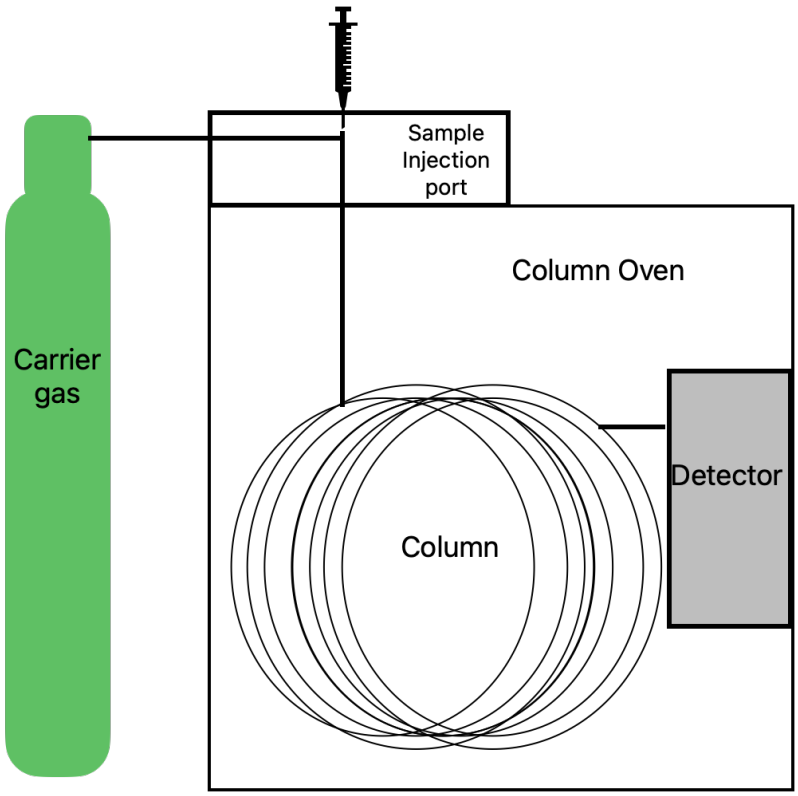


Figure 6. Gas Chromatography schematic

Mass Spectrometry

Mass spectrometry is an instrument that analyzes the mass to charge ratio of fragments of compounds. When the sample is injected into the instrument it encounters a three-step process. The first is ionization of the sample from an electron energy source. These sources can be soft or hard and this affects the size of the fragments to the abundance of the ionization. A soft technique is electrospray ionization which is done in tandem with liquid chromatograph. A hard electron impact source is electron impact ionization (EI) that has a high energy electron beam. When the high energy electrons collide with the analyte molecules in the ionization chamber, the bonds in the analyte compound can break and form fragments. The bonds breaking will produce a positively charged ion fragment and neutral ion fragments. Larger molecule as well as complex molecules can be protected causing larger or heavier fragments. After ionization the fragments are sent into the mass analyzer that has a quadrupole system. Based off the charges and size of the fragments a quadrupole mass analyzer can filter ion fragments with a particular mass to charge ration (m/z). The last step in the spectrometer is the detector which generates an electrical signal based on the abundance of the ions. A mass spectrum is then produced by the computer attached to the mass spectrometer to show the abundance and the mass to charge ratios of fragments reached to the detector (see Figure 1) . In a typical laboratory analysis, the mass spectrum generated for a sample is compared with a mass spectral library to find the identity of an unknown compound. [7,17]

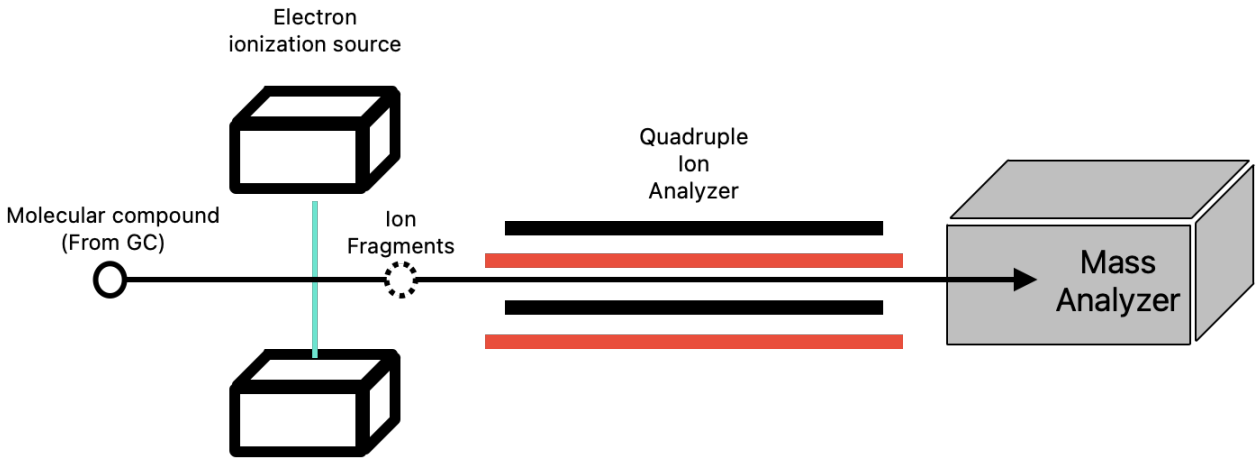


Figure 7. Schematic of Mass Spectrometer

Spectral Libraries

Two mass spectral libraries were used in this study. The first library was used for the development of the model and this set is commonly known as the training set. The library that was used for the training and development of the models was obtained from the Scientific Working Group for the Analysis of Seized Drugs (SWGDrug).[3] This library has over 5000 spectra of illicit drugs, metabolites and precursors. To test the models a prediction set was used. This prediction set consists of cannabinoids that were not included in the training set. The mass spectral database published by Cayman Chemical Company (CCC) was used for this purpose. The CCC database is compiled using their inhouse GCMS with over 4000 compounds.[4] Both databases used electron impact ionization (EI) with an electron voltage of 70. This is a universal electron voltage that is used to get the best ionization by matching the de Broglie wavelength of bonds within organic molecules as well as producing a stable mass spectrum.[5,12]

PCA

Principal component analysis (PCA) is an unsupervised technique that is used for data prepressing that is used with machine learning algorithms. PCA is used to analyze the covariance of large complex data for the reduction of dimensionality without loss of information. It is an unsupervised meaning that the algorithm does not need to consider the class or categories in each data set. PCA reduces the data drawing a line through the data with the greatest variance this is known as the individual principal components (PC) which are linear components with the original data to help reduce dimensionality and variables. The first PC shows the most covariance while the second PC shows the second most and further on. [16,17,18]

$$cov_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{N-1} \quad (1)$$

Where, X = random variable 1, Y= Random variable 2, X_i= sum of all x variables, Y_i= sum of all Y variables, N=Number of variables.

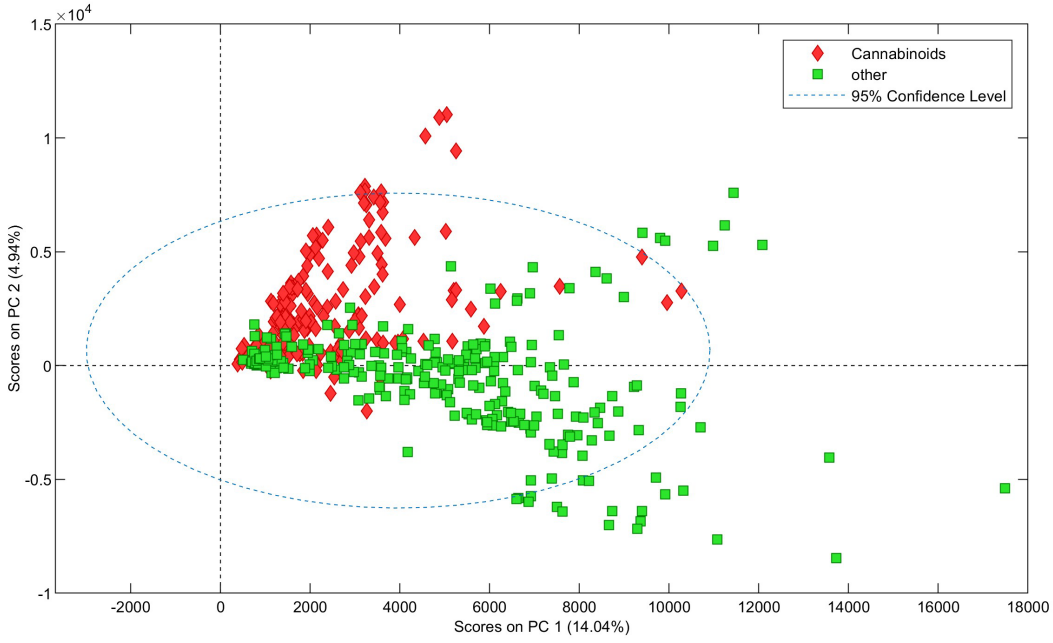
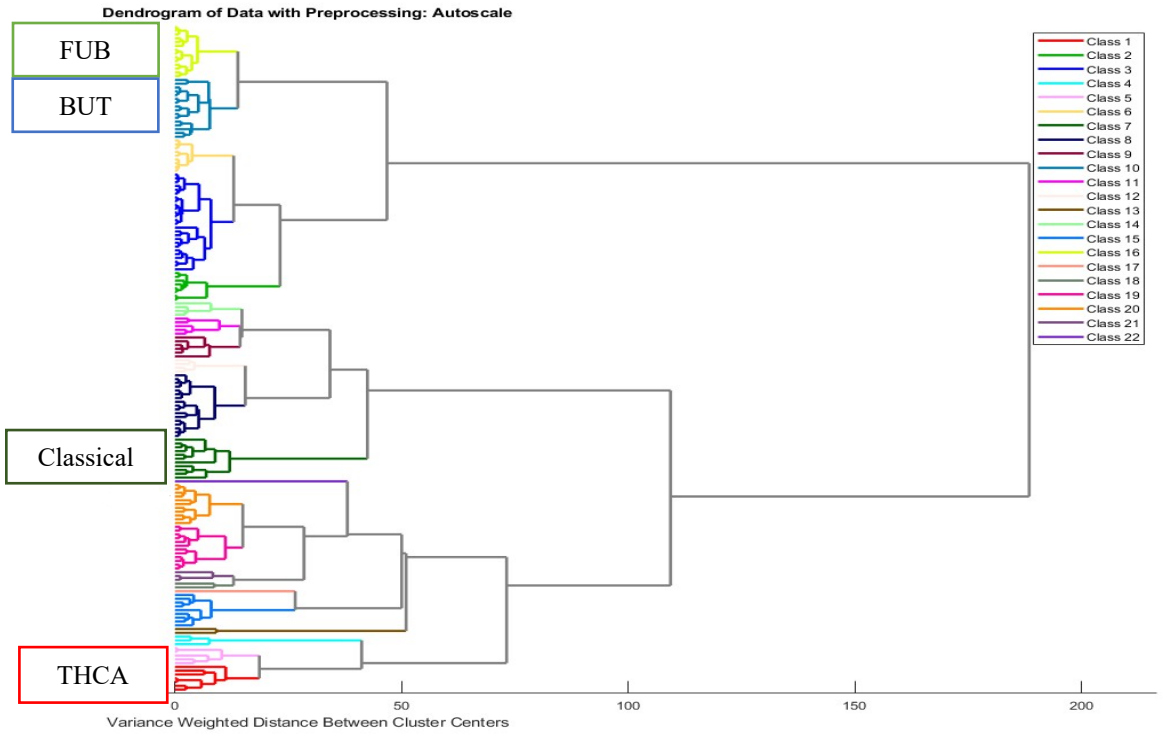


Figure 8. PCA of Cannabinoids (Red) and other drug classes (Green) with maximum variance

HCA

Hierarchal clustering analysis (HCA) can be used to identify classes and subclasses in large data sets. These clusters show the variance in the form of a dengogram that defines each class based off cluster characteristics. The Variance in this case comes from the PCA scores of four principal components. The classes or groups suggested by HCA may or may not be used in classification problems. The careful study of the structural details is done before deciding the drug classes in the current study. For example, class 7 (Dark Green) was THC and other natural and classical cannabinoids while the class 1 was THCA which is an acidic derivative. Class 16 (lime-yellow) is where the FUB group was located, and Class 15 (dark blue) is where the BUT tail group is located. These were used as two models because they showed great variance in HCA as well as PCA and great discrimination in PLS-DA. The biggest issue in the synthetic cannabinoid core groups could not be fully separated using HCA due the selectivity. The HCA analysis selectively targets the readily cleaving groups in the fragmentation patterns so manly head groups and tail groups that give high abundance on certain mass-to-charge ratios. The FUB group is unique due to the fragment that has m/z of 109 while BUT has a unique fragment that has a m/z of 201 and the selectivity of the HCA can use this to designate classes with variance whereas the core groups do not have much variance.



Dendrogram of cannabinoids showing the precision of the computer to find and assign classes based off variables.

PLS-DA

Partial least squared discriminate analysis (PLS-DA) is a supervised machine learning method. PLS-DA is used for the reduce of dimensionality like PCA, but the computer uses known classes and discriminate between them. Since the base is partial least squares the computer must use an independent and a dependent variable and these being the mass spectral information as the independent variable classification of compounds as the dependent variable.

PLS-DA gives a variety of ways of accuracy. true positive rates (TPR) and true negative rates (TNR). TPR is the rate that the model predicts a sample as a positive identification while TNR is the exact opposite in which the computer can predict that a sample does not belong in that class. TPR shows the sensitivity of the model while TNR shows the specificity of the model.[18]

$$TPR = \frac{TP}{(TP+FN)} \quad (2)$$

$$TNR = \frac{TN}{(TN+FP)} \quad (3)$$

TN= Number of true negatives

TP= Number of true positives

PLS-DA also gives false positive rates (FPR) and false negative rates (FNR) that work in the same context as TPR and TNR. FPR is the proportion that were negative but incorrectly identified as a positive. While FNR is the proportion that were positive but identified as negative.

$$FPR = \frac{FP}{(TN+FP)} \quad (4)$$

$$FNR = \frac{FN}{(FN+FP)} \quad (5)$$

FN= Number of false negatives

FP= Number of false positives

PLS-DA also has a calculation for accuracy of the model. The accuracy is correlated with the error of the model. The error is the proportion of the samples that were incorrectly identified. The accuracy tells the proportion of the sample Identified correctly in the population.[18]

$$1 - Error = Accuracy \quad (6)$$

$$1 - Accuracy = Error \quad (7)$$

Cross Validation

Cross validation (CV) is used to test the robustness of the models. Here, a set of samples are taken out during the model development, and they are used to determine if the model can predict the unknowns accurately. Random subset of ten percent of samples in the training set were

used in this research with 10 iterations. Random subsets are when the program takes random samples out of the model and tries to fit them as an unknown as a way of validation of accuracy and precision of the model and takes the average of all iterations. The benefits are this is a versatile and gives all samples an equal change for validation.

Feature Selection using GA

Genetic algorithm is an evolutionary based algorithm in which the program uses generations with a population and uses the concept of natural selection. This can be used to select the variable that carry the most important information for a given problem. Mass spectra are collected over a wide range of m/z not all these m/z values are not significant in each chemometric problem. Therefore, using GA the number of features or variables (m/z) used in calculating model can be minimized. Furthermore, this reduces the mass spectral fragments with lower importance that could cause discrimination issues. A typical GA analysis is performed for multiple generations. Each generation uses the best variables in the next until the best fit is achieved. When the best variables or mass spectral features are calculated for the best fit the algorithm then can input those variables in the model. In this research for example it was found after 50 generations the algorithm starts to overfit. To solve this issue, the process was done at 50 generations with a population of 64 models, 6 loading vectors, a 0.005 mutation rate, three replicate runs for an average, and a window width of one. The use of GA will also help identifying highlight important m/z ratios that can be used to discriminate each class from one another.

Research Objective

This research focuses on development of a system of binary classification models to aid in the identification of newly synthetized cannabinoids chemical structures to conserve resources and time as well as open possible avenues for laboratories that do not have the instrumentation available to identify newly developed cannabinoids.

Binary Classification System

The modeling system is built from multiple different binary models that can help aid in the identification of novel cannabinoids. The model system was developed this way because of the specific mass spectral fragmentation patterns if there were three or more classes the analysis would start to overfit data from each class into each other causing the accuracy to decline. It was found that the modeling went best with the discrimination of two classes at a time and to use each output as a building block for the whole molecule. Notice that there is no linking group model, and this is because the linking group is likely to cleave as it is impacted by the electron source in the mass spectrometer. This causes concern as an important building block of a cannabinoid is the linking group from the core to the head group. The system takes and unknown mass spectrum and then can help aid in the identification of the unknown cannabinoid.

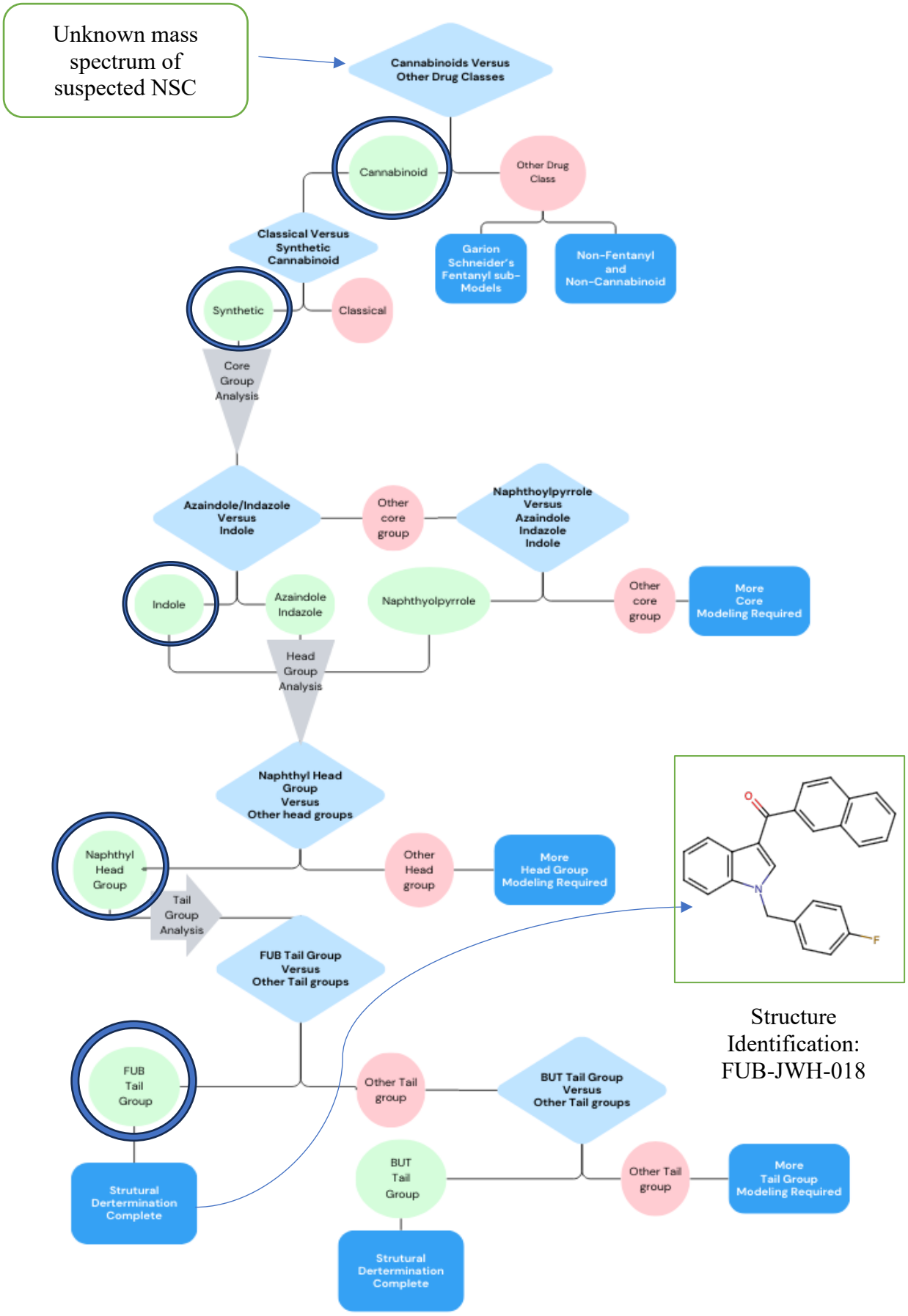


Figure 9. Flow map of Binary modeling system with example of identification

CHAPTER TWO: EXPERIMENTAL

Mass Spectral Data

All mass spectra were obtained from the SWGDrug database (Versions 3.9, 3.10, 3.11, and 3.12).[3] This database was only used as the training set during the model development. The two largest groups of interest are cannabinoids and fentanyl as they are the most common encountered NPSs. A total of 185 cannabinoids were used in this study along with other drugs of abuse.

Table 1. Total number of spectra taken from both the SWGDrug and Cayman Chemical Company database

Drug Type	SWGDRUG (Training set)	Cayman Chemical Company (Prediction set)
Cannabinoids	185	25
Fentanyl	174	10
Opioids	27	5
Cathinones	48	5
Tryptamines	7	5
Phenylethylamines	14	5
Total	454	55

Preprocessing of Mass Spectral Data

The SWGDrug mass spectral library need various preprocessing with the first being uniformity. The area of the spectra was also focused between 40-300 mass-to-charge ratios (m/z). If the spectra were shorter the missing data was replacing with zeros until the desired length. The reason behind this is because this research focuses on the fragmentation pattern for the aid of classification and identification of newly synthesized compounds. Most cannabinoids have a molecular ion above 300 m/z even though the SWGdrug spectra were collected from 40-662 m/z. The prediction sets for the models were done using Cayman Chemical Company mass spectral database.[4] All preprocessing was done in the same way for SWGDrug with one exception the

abundance for SWGdrug went to 9999 while the CCC database abundance only went to 999 and this had to be adjusted using an equation for accurate model predictions.

$$(\boldsymbol{CCC\ Abundance}/\mathbf{999}) * \mathbf{9999} \qquad \textbf{(8)}$$

Software

The programming language used as the base for the model development was MATLAB also known as MATrix LABoratory. The MATLAB program can perform large number of functions containing large amount of information and variables. The program can do basic functions to advanced arithmetic functions. While MATLAB can perform these functions there are toolboxes available for easy use and access. In this method of model development, a toolbox engineered by Eigenvector research incorporated was used. This toolbox is called PLS-Toolbox and is a chemometric tool for machine learning containing over 300 algorithms and tools that might be needed for analysis and development with the primary function being partial least squares. This toolbox can be used in MATLAB as well as a stand-alone program known as solo. There are other programs that were not used in this research but could be used in conjunction with PLS-Toolbox such as the Multivariant Image Analysis (MIA).

Methods

The machine learning techniques that were used were Principal component analysis and partial least squares discriminate analysis. Both methods are reduction tools that were used in the large variable input. The goal of this research is to form a predive model using classification for the identification of unknown cannabinoids. Using PCA shows the different variance of the data to use to classify the cannabinoids. Using PLS-DA, shows discrimination between two classes

using a plot with the best fit as the discrimination threshold. PLS-DA was accompanied by cross validation and genetic algorithm as a tool for identification and selection of important variables to aid in the reduction of dimensionality. Cross validation was used for the verification of accuracy, TPR and NPR. GA was used to not only provide better discrimination but to supply the best fit of loading vectors. These loading vectors were used to provide an understanding of the mass fragmentation patterns of cannabinoids and helped develop a system to verify and classify the type of group on the cannabinoid.

CHAPTER THREE: RESULTS AND DISCUSSION

3.1 Classification of Cannabinoids Versus Other Drug Classes

The first classification model was developed to discriminate cannabinoids from the other drugs. First, a PLS-DA model was developed using all mass to charge ratios in mass spectral data. Feature selection algorithms were not used to reduce the number of variable (mass charge ratios). The PLS-DA plot is shown in figure 11. Here, the discriminant line shown in red was calculated by the computer to show where the classes are separated from each other. The discrimination of the classes is evident with high accuracy. However, there are some misclassified samples.

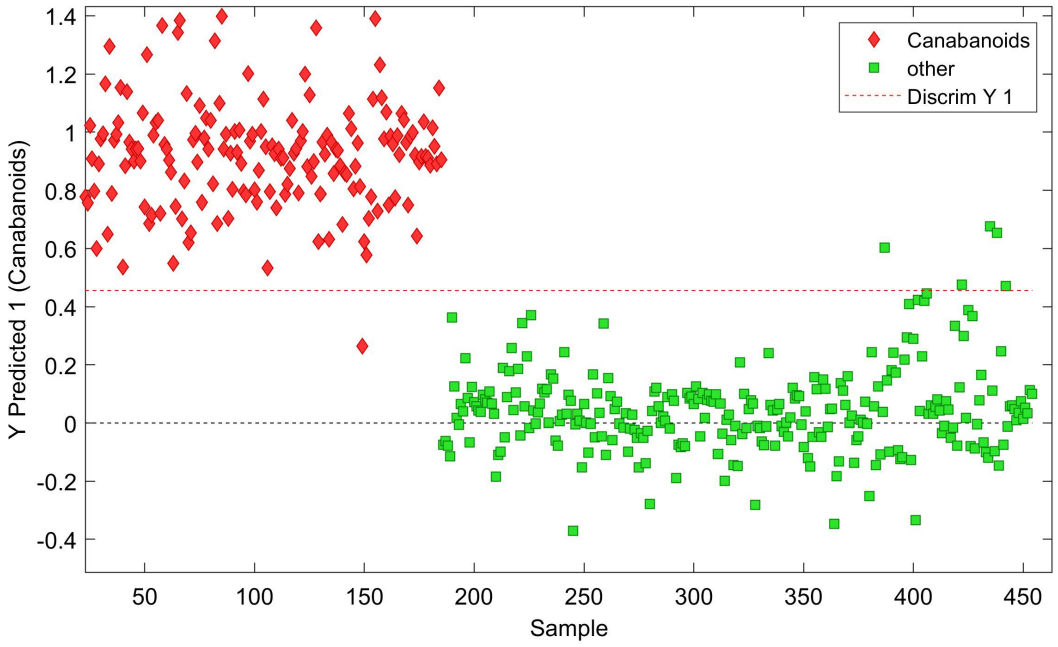


Figure 10. PLS-DA Plot of cannabinoids versus other drug classes without GA

Figure 12 shows the first latent variable (LV) with the loadings of each mass to charge ratio used in the model development. Here no variable selection was performed and hence all m/z values were included in model development.

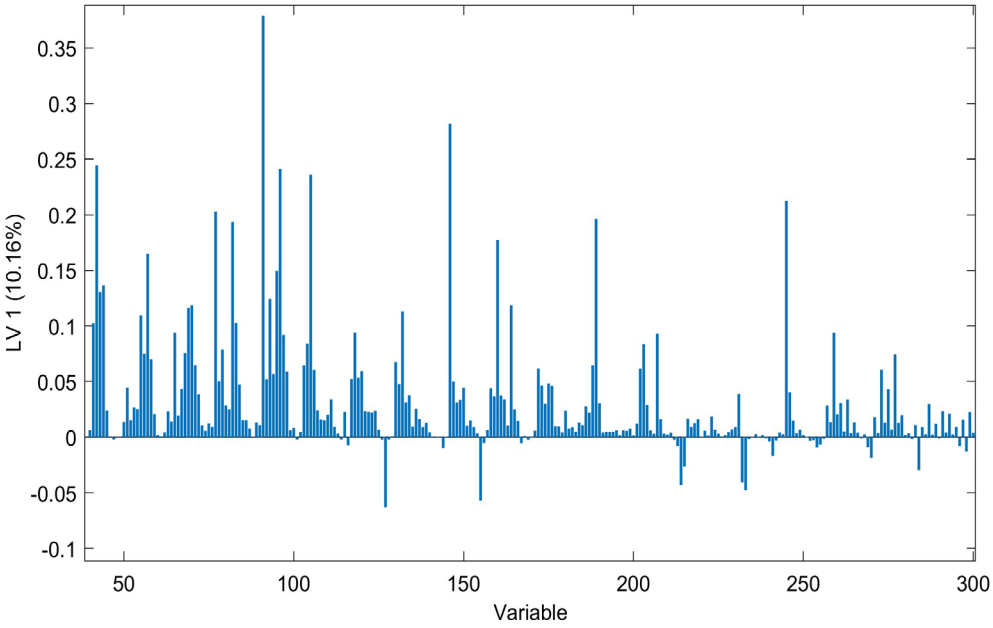


Figure 11. LV1 without the variable selection process. The x-axis is the independent variable mass-to-charge ratio.

Next, the GA was performed to reduce the variables used in the model development (see Figures 13, and 14). The first LV with the loadings is shown in Figure 15. that indicates some variables were removed from the model development. There could be some variables that do not have importance for the discrimination of a particular classification problem however can arise from predominantly present mass fragments of some samples. This can cause misclassifications as the loadings of each sample was calculated using all the variables if GA was not in use. The PLS-DA plots developed for cannabinoid versus other drug classes using GA as the feature selection method showed great accuracy of 97%. In the latent variables (LV) it shows the important m/z peaks for both cannabinoids and other drug classes. The highlighted variables are 105, 245, 144,145 and 215. The peak at m/z 105 is the fragment of fentanyl as well as some other groups on other drug classes. The peak at m/z of 245 shows the molecular ion for fentanyl and fentanyl derivatives. Fragment peaks are seen at m/z of 144 and 145 and are the base peaks for the

cannabinoid that contain indazole and indole. The fragment at m/z of 215 is important as well as 214 m/z fragment as it correlates with the core with some of the tail groups and linking group of cannabinoids. The peak at 233 is the and important fragment of natural THC and derivatizes. This shows the importance of the use of GA; thus it was used in all the models discusses below.

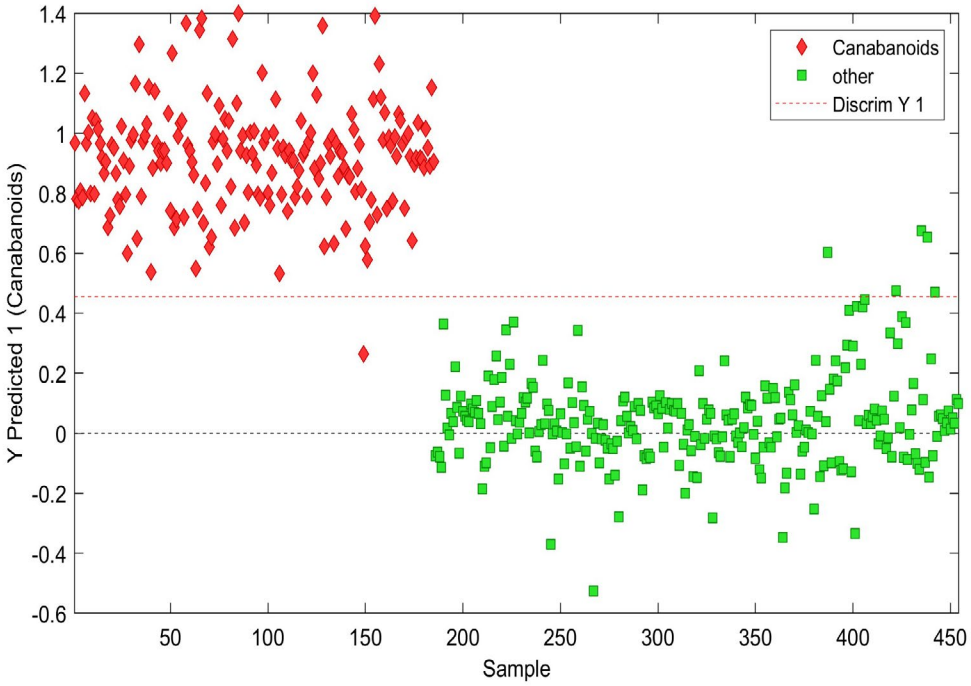


Figure 12. PLS-DA plot of cannabinoids versus other without cross validation.

However, Figure 14 shows the PLS-DA plot with cross validation. The cross-validation plots show the average of each sample placement. This means that the program runs the sample multiple times through cross validation and then averages the scores of the sample to produce the PLS-DA plot. This explains why the validation process accuracy and TPR and TNR are lower than the prediction most probable. GA was used as well to reduce the number of loadings the model has, and this can contribute to some samples being correctly identified as well as incorrectly identified if only a few samples have that loading then it can be dropped from the model overall causing misidentification. The loadings in Figure 15. further shows the core structures respective

m/z for cannabinoids that are useful in the development and discrimination of the model. This shows that the algorithm can pick the most useful variables to use during development. This suggest that the spectra have fragments that are not necessarily important for the discrimination of two structural groups.

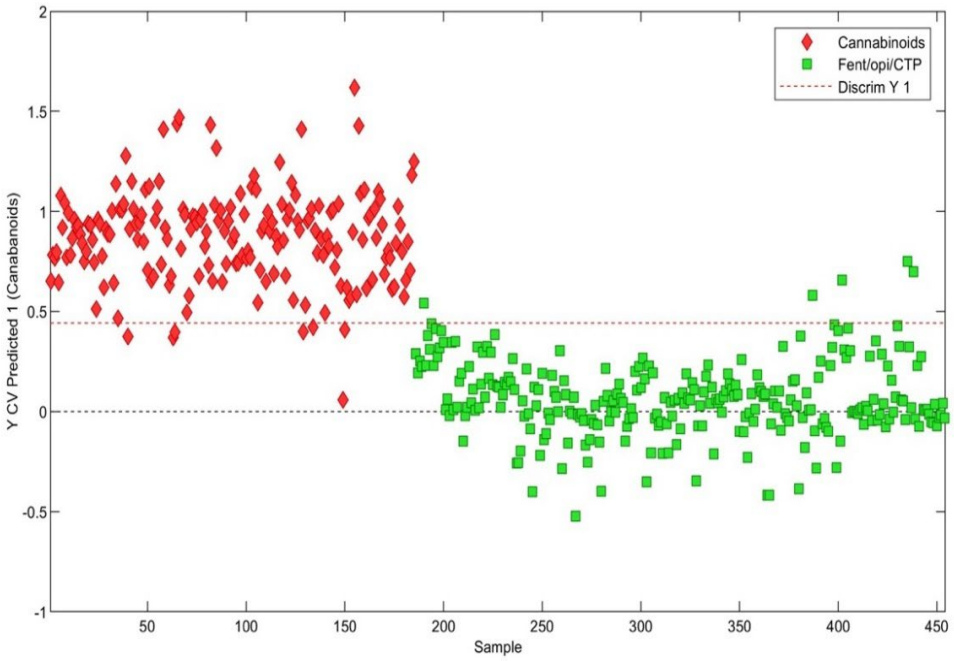


Figure 13. PLS-DA plot of Cannabinoids (Red) and other drug groups (Green) with cross validation

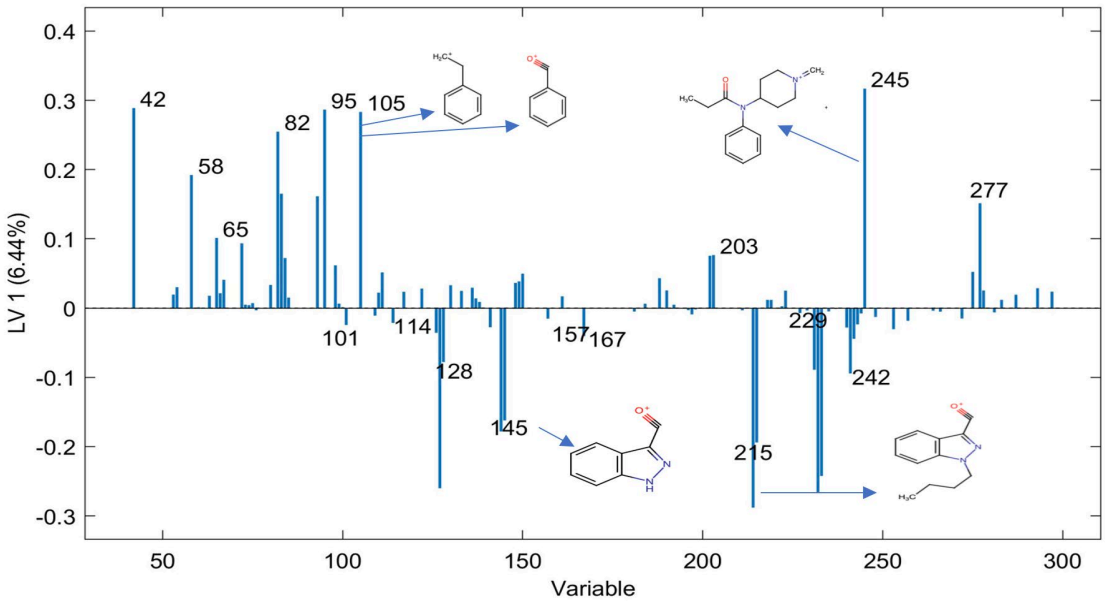


Figure 14. LV1 with the uses of GA. The areas with no bars show zeroes and these are variables not seen as important to the loading of the model. The x-axis is the independent variable mass-to-charge ratio.

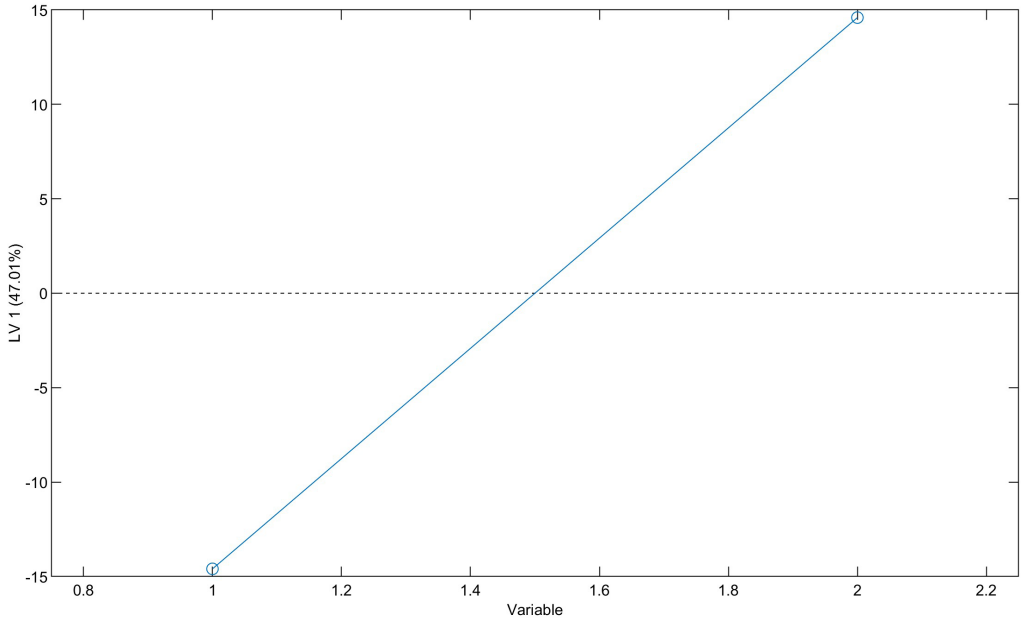


Figure 15. Y loading vectors for cannabinoids versus other drug classes. The x-axis is the dependent variable of the assigned class. 1 relates to the cannabinoid class while 2 relates to the other drug classes.

Figure 16 shows the Y-loadings, and these demonstrate why there are positive and negative loading vectors. If the loading vector weighs as a negative it is a variable one which correlates with cannabinoids while any vector that weighs positive, it is variable two which correlates with the other drug classes. The correlation is inverse in the PLS-DA plot as the cannabinoids are positive while other drug classes are negative.

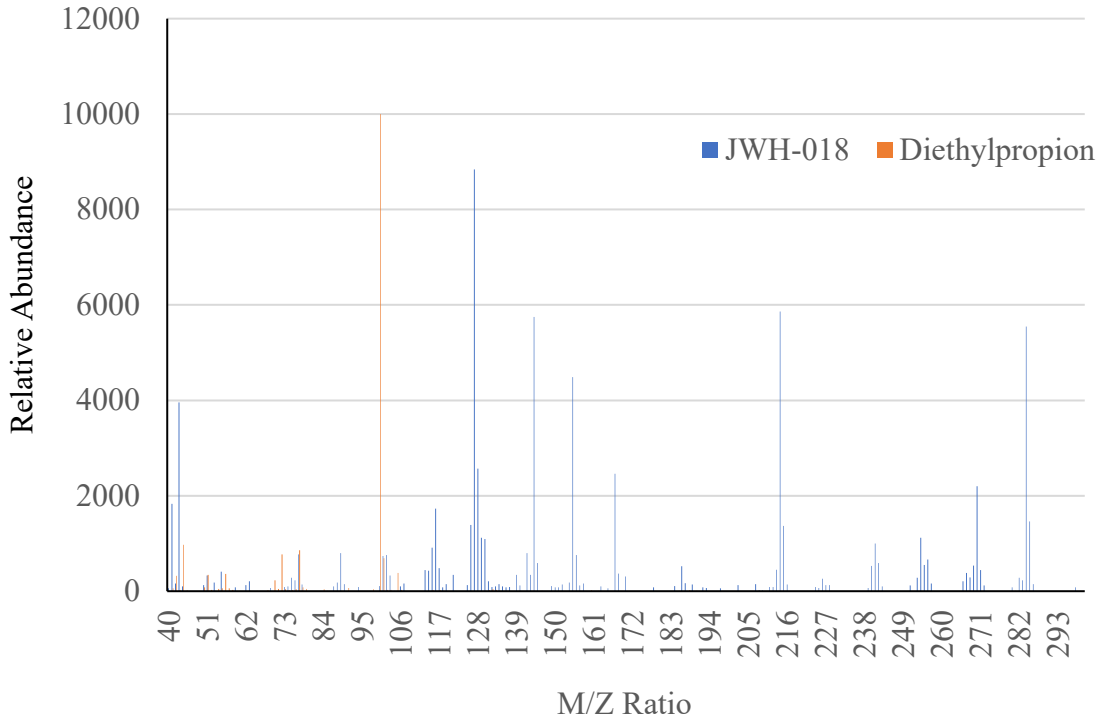


Figure 16. Mass Spectra comparison of synthetic cannabinoid JWH-018 (Blue) and cathinone diethylpropion (orange)

Misclassified samples in these models were thoroughly examined by comparing the mass spectral data and LVs. Some predictive error was associated with cathinones which are likely caused by the peak for mass fragments at m/z of 101. This fragment was found to be common in

some cathinones such as diethylpropion. The mass spectral comparison of diethylpropion (a cathinone) and JWH-018 (a synthetic cannabinoid that has the typical fragmentation pattern of synthetic cannabinoids) in Figure 17. shows the presence of a strong peak at m/z of 101. Additionally, according to Figure 15. the fragment at m/z of 101 is associated with cannabinoids. Having a very strong 101 peak in diethylpropion indicates that the scores calculated by the model for diethylpropion would be closer to that of cannabinoids; hence this compound is predicted as a cannabinoid.

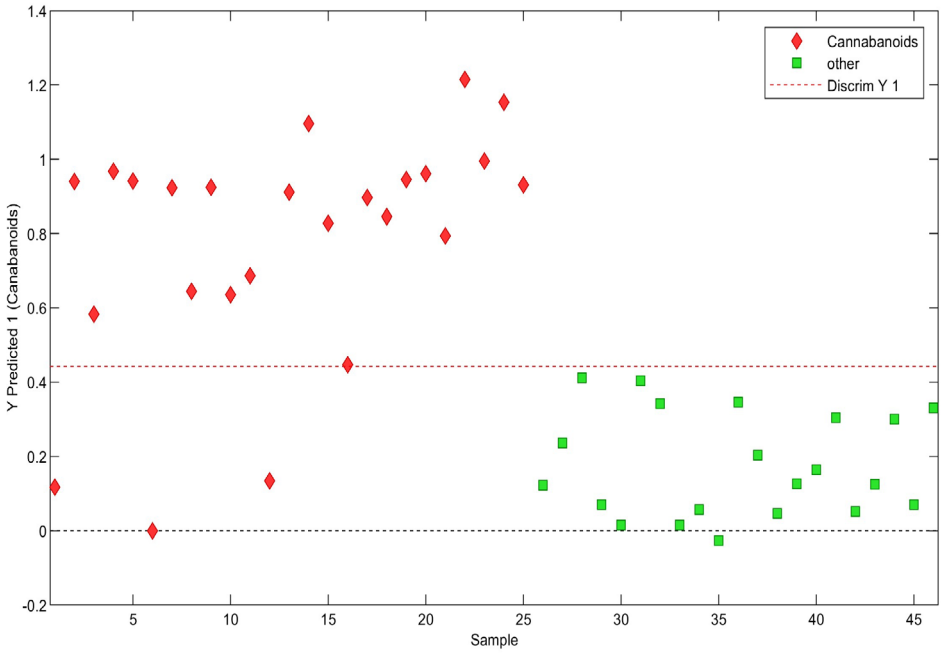


Figure 17. Prediction plot of the cannabinoid versus other drug classes.

The results table (Table 2) summarizes all the numerical results of the models has three different sections; the first section shows the prediction results of samples used in the training set. The model was calculated, and all the samples were predicted against the developed model. model results that do not use cross validation averaging of the samples. This process only uses the

prediction most probable method as self-validation. The second section of this table consists of the results obtained with cross validation. The last part of the table has the information for the prediction (or test) sets were completed using the Cayman Chemical Company (CCC) mass spectral database. The cohort of 55 compounds chosen as the test set were not used in model development. The accuracy of the test set shows the predictive ability of this model. The misclassification of the prediction set is likely caused by the mass spectral range of 40 m/z to 115 m/z as there are multiple fragments in synthetic and classical cannabinoids that are present in this region. However, the LVs show that these peaks are associated with other drugs compared to cannabinoids, and this could be solved possibly if those variables were removed and use only the m/z range of 100 to 300 of the mass spectrum. However, this was not tested in this present work.

Table 2. PLS-DA Results for cannabinoids versus other drug classes. The self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Cannabinoids	0.9755	0.9675	0.9814	0.0185	0.0324
Other Drugs		0.9814	0.9675	0.0324	0.0185

Cross Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Cannabinoids	0.9735	0.9621	0.9814	0.0185	0.0378
Other Drugs		0.9814	0.9621	0.0378	0.0378

Test Set Results

	Accuracy	TPR	TNR	FPR	FNR
Cannabinoids	0.9398	0.8800	1.000	0.000	1.200
Other Drugs		1.000	0.8800	1.200	0.000

3.2 Classification of Tricyclic Based Cannabinoids versus Other Synthetic Cannabinoids

The next step was to develop a model to discriminate tricyclic based cannabinoids (classical) vs other synthetic cannabinoids. These two groups show significant structural differences. The PLS-DA score plot is shown in Figure 19. The analysis of loadings (Figure 20.) shows that the key features that are important for this discrimination were mass-to-charge fragments 144, 155, 214. Furthermore, the m/z fragments of 231. 144 and 214 in the synthetic cannabinoids are assigned to the common core structures present in these molecules including indole and indazole while the m/z of 155 fits with the naphthyl head group that is a common moiety

found on synthetic cannabinoids. The m/z of 231 is the large tricyclic group associated with the classical cannabinoid structure.

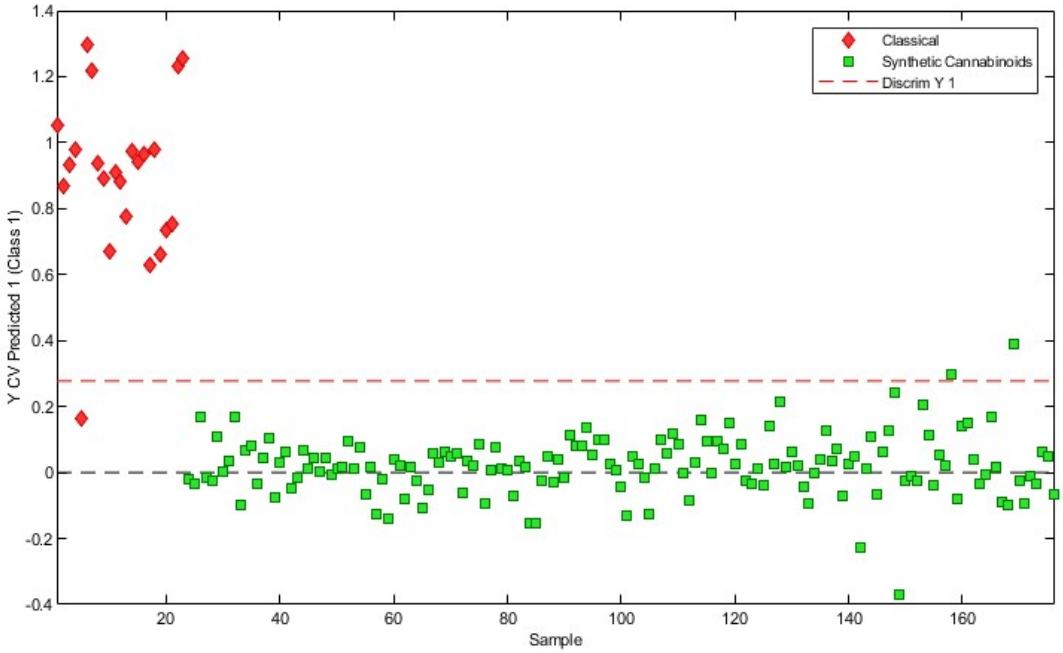


Figure 18. PLS-DA plot with cross validation average of classical cannabinoids versus synthetic cannabinoids

One of the misclassified samples of most concern is cannabidiol (CBD) especially with the current legality surrounding it. In figure 21. an overlay with CBD and JWH-018 has shown that CBD had a fragmentation pattern with several fragments and some of these fragments are around the m/z range of 94-97 as well as around m/z of 144 which correlates with the selected loading vectors as a synthetic cannabinoid. This model shows 98 percent accuracy (see Table 3). There are some major fragments that overlay with each group such as the m/z of 231 fragment for both the synthetic cannabinoid core group as well as the core group on the classical cannabinoid.

Table 3. PLS-DA Results with cross validation results for classical cannabinoids versus Synthetic cannabinoids the self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Classical	0.9886	0.9565	0.9935	0.0065	0.0435
Synthetic		0.9935	0.9565	0.0435	0.0065

Cross-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Classical	0.9830	0.9565	0.9869	0.0131	0.0435
Synthetic		0.9869	0.9565	0.0435	0.0131

Test Set Results

	Accuracy	TPR	TNR	FPR	FNR
Classical	1.000	1.000	1.000	0.000	0.000
Synthetic		1.000	1.000	0.000	0.000

The mass spectra of natural tricyclic cannabinoids have a lot of fragments that can be less valuable for the discrimination process throughout the spectrum unlike the synthetic cannabinoids. The LV in Figure 20 shows that a large number of smaller mass fragments are associated with discriminating synthetic cannabinoids from classical cannabinoids. Most of these smaller peaks are associated with the classical cannabinoids (see the negative loadings in Figure 20).

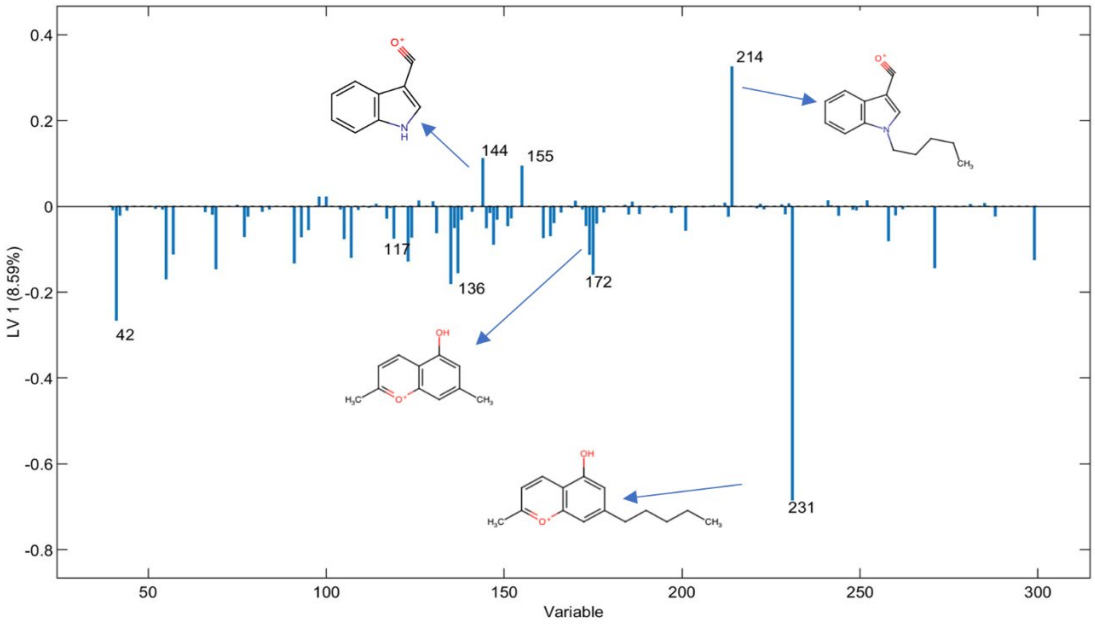


Figure 19. Loading vectors for classical cannabinoids (bottom). And synthetic cannabinoids (top) The x-axis is the independent variable mass-to-charge fragments with the most weight.

The comparison of the mass spectra of cannabidiol (the misclassified sample in red in Figure 19.) and JWH-018 (synthetic cannabinoid) shows that there are intense peaks at m/z of 127 and 231 for cannabidiol compared to JWH-018. Additionally, the first LV shown in Figure 19 has stronger peaks that are correlated to synthetic cannabinoids can cause cannabidiol to be misclassified as a synthetic cannabinoid.

The prediction results that were obtained for this model using CCC library data showed no misclassifying of any of the compounds. This shows the high accuracy of the model as well as the distinction of structural difference of classical and synthetic cannabinoids. A topic to be furthered investigated is with smaller fragmentation involving the cyclic groups such as naphthyl and their effect on the discrimination and prediction process. This group is common in the samples used to develop the model.

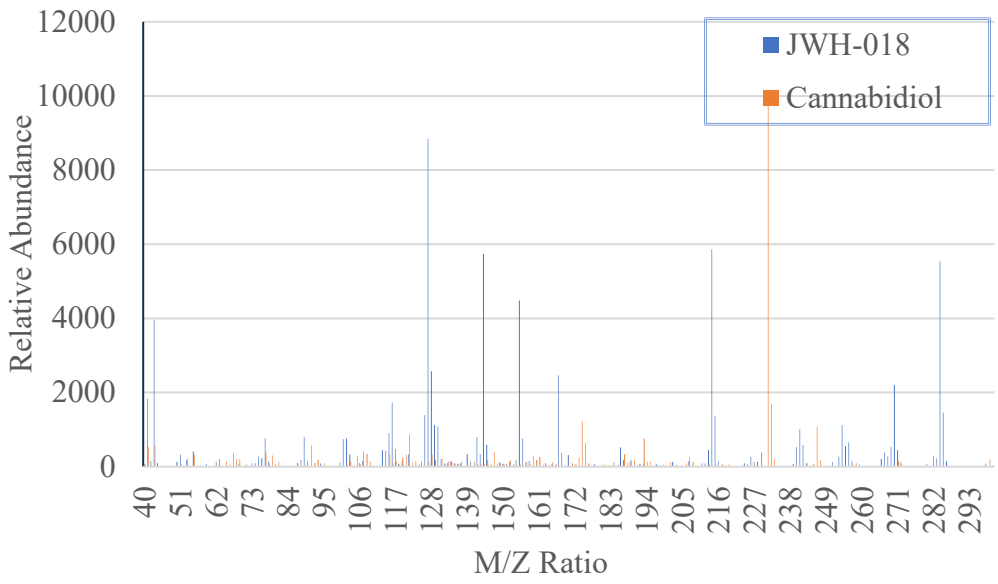


Figure 20. Mass spectra Overlay of JWH-018 and cannabidiol

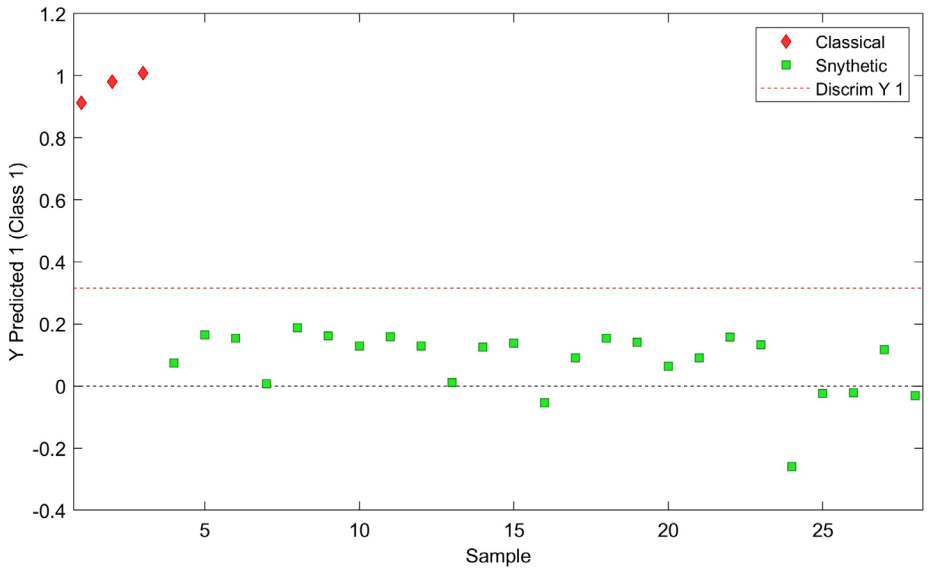


Figure 21. Prediction PLS-DA plot for Cannabinoids versus Synthetic cannabinoids using CCC Database of unknown compounds to model.

3.3 Classification of Napththoylpyrroles vs other synthetic cannabinoid core groups

The next step was to discriminate the synthetic cannabinoids using core structures. Three main core structures that were observed are indazole, indole, azaindole and napththyolpyrrole. Napththyolpyrrole containing synthetic cannabinoids were easier to discriminate from indoles, indazoles and other core structures. The results of the model shows a successful discrimination with only one misclassification resulting in 0.99 accuracy. However, there are only a small number of napththyolpyrrole containing synthetic cannabinoids compared to other cannabinoids used to develop this model. This can add unwanted bias for not only prediction but discrimination and should be investigated further as more napththyolpyrrole containing cannabinoids are identified.

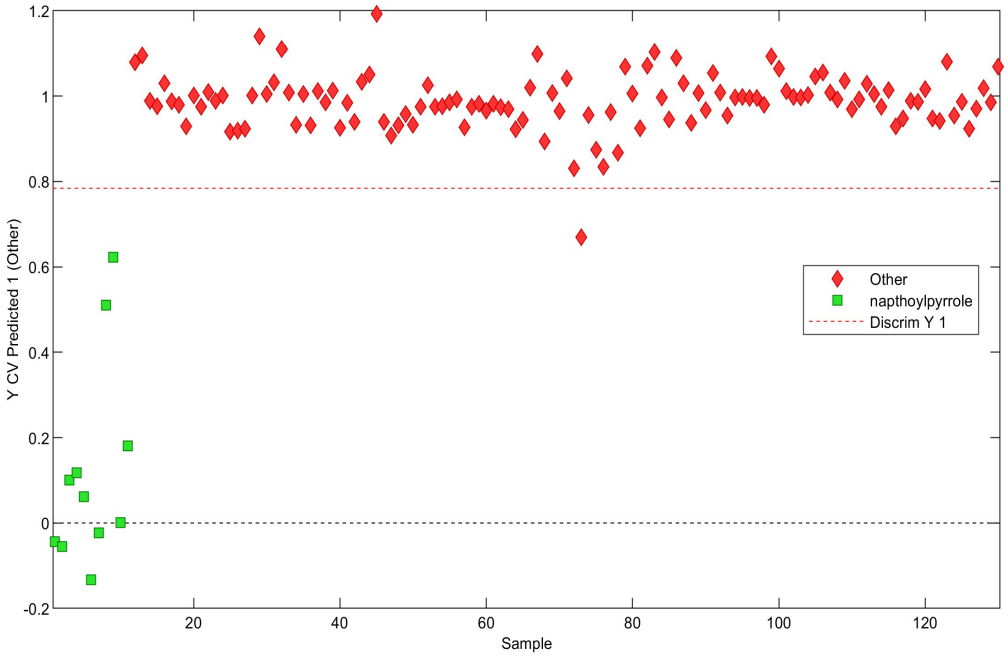


Figure 23. PLSDA plot of napththyolpyrrole (green) versus other synthetic cannabinoids (red)

The only misclassification in Figure likely caused by a strong m/z of 155 base peak found in the spectra. The misclassified compound was identified as ADB-BINACA that has a cyclohexane tail group. This can correlate with common functional groups associated with indole, azaindole and indazole core groups can cause the misclassification due to the formation of mass fragments with similar mass to charge ratios.

Table 4 PLS-DA results with CV results for naphthoylpyrrole vs other core structures (Azaindole indazole, indole) The self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Napththyolpyrrole	1.000	1.000	1.000	0.000	0.000
Other core groups		1.000	1.000	0.000	0.000

Cross-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Napththyolpyrrole	0.9923	0.9744	1.000	0.000	0.0256
Other core groups		1.000	0.9744	0.0256	0.000

Test Set Results

	Accuracy	TPR	TNR	FPR	FNR
Napththyolpyrrole	!.000	1.000	1.000	0.000	0.000
Other Core Groups		1.000	1.000	0.000	0.000

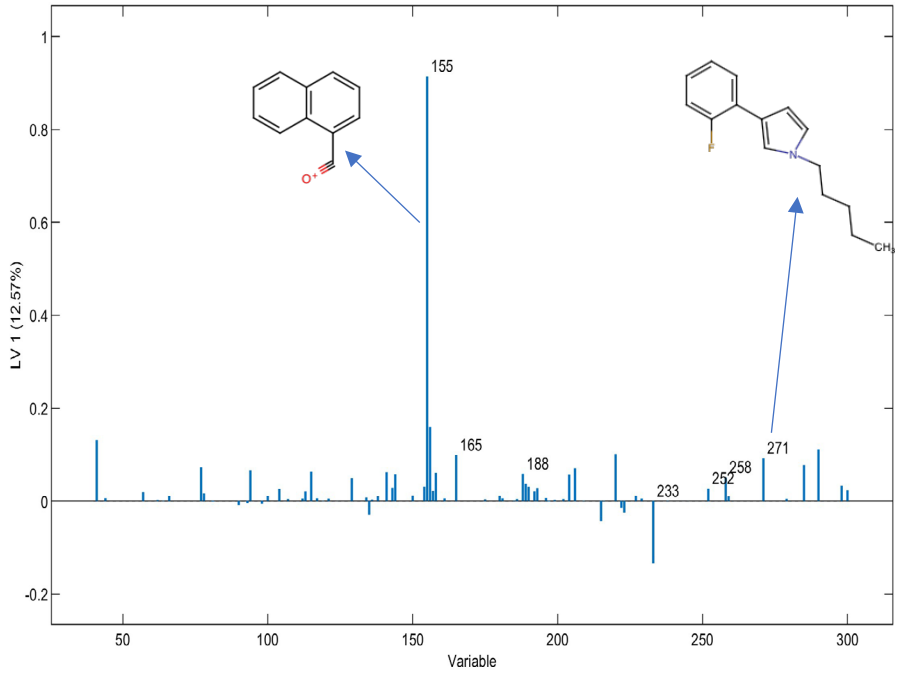


Figure 24. LV1 plot of Naphthoylpyrrole (top) versus Other Tail Groups (bottom) with structures. The x-axis is the independent variable mass-to-charge fragments with the most weight.

The first loading vector in figure 24 shows that the napththyolpyrrole class agrees with the fragmentation pattern for the napththyolpyrrole containing cannabinoids. The major peak at 155 m/z shows that this core group is linked to the naphthyl head group most often but there is a peak at 271 m/z that shows the fragment where the core is connected to its tail as another major fragment with this core group. This means that the GA picked out important fragments for the napththyolpyrrole class, but it also picked other fragments with smaller weights that can cause issues or discrimination, The prediction results in future 25 show a correct prediction of all the samples used from CCC database. The model statistics are shown in Table 4.

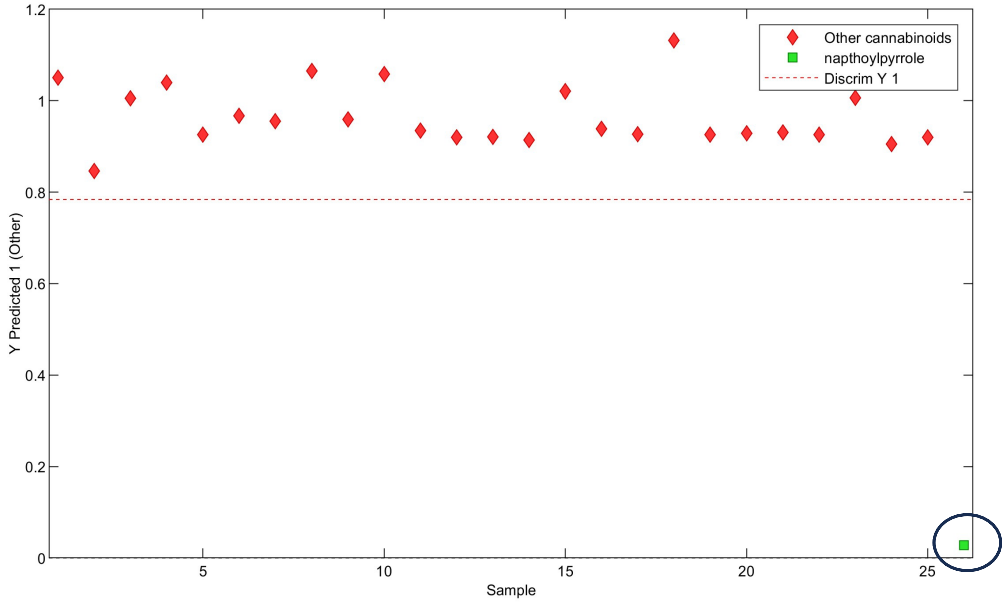


Figure 25 PLS-DA Prediction plot with CCC database and previously unknow cannabinoids to the model

3.4 Classification of azaindole/indazoles Core structure Versus indole core structure

After separating the napththoylpyrroles, the next goal was to separate indoles from azaindoles and indazoles. Azaindoles and indazoles were group together as they both have similar structures with varying the position of nitrogen atoms. Therefore, these molecules would fragment in a similar fashion producing ions with the same m/z ratios. Successful discrimination is evident in Figure 14. Interestingly, the PLS-DA was able to pick the correct variables for this model although they are off by only one mass to charge ratio. The 144 and 214 are a part of the indole core group fragments while the 145 and 215 are one higher due to the extra nitrogen in the azaindole and indazole core group.

The PLS-DA plot in Figure 26. shows an accurate discrimination but there is one compound, 5F-PCN, of concern that cross validation averaged just below the discriminate line for

the azaindole/indazole class. This compound is an azaindole that has a naphthyl head group. This head group in the loading vectors (Figure 27) is seen as an indole at 155 M/Z.

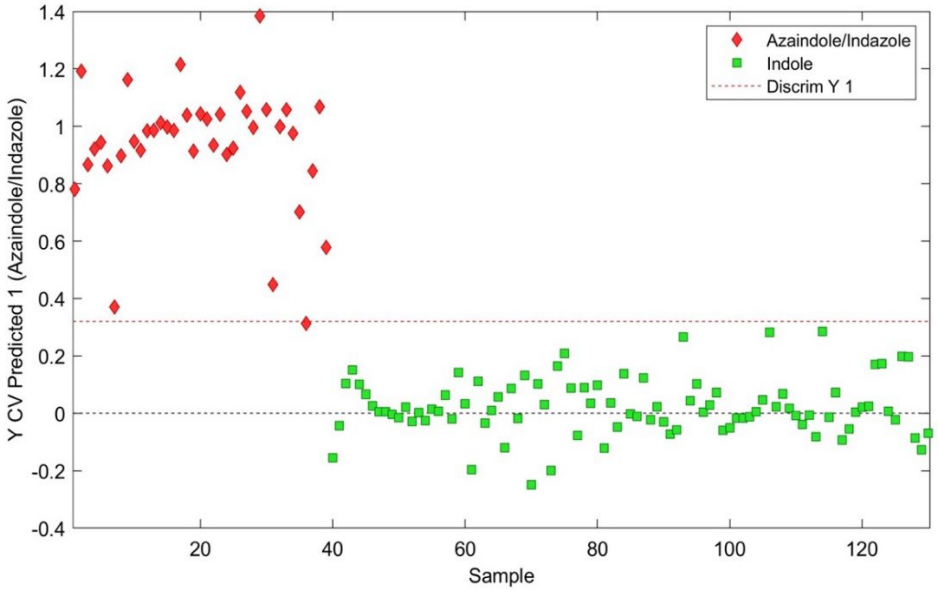


Figure 26. PLS-DA of azaindoles/indazoles (red) versus indoles (green).

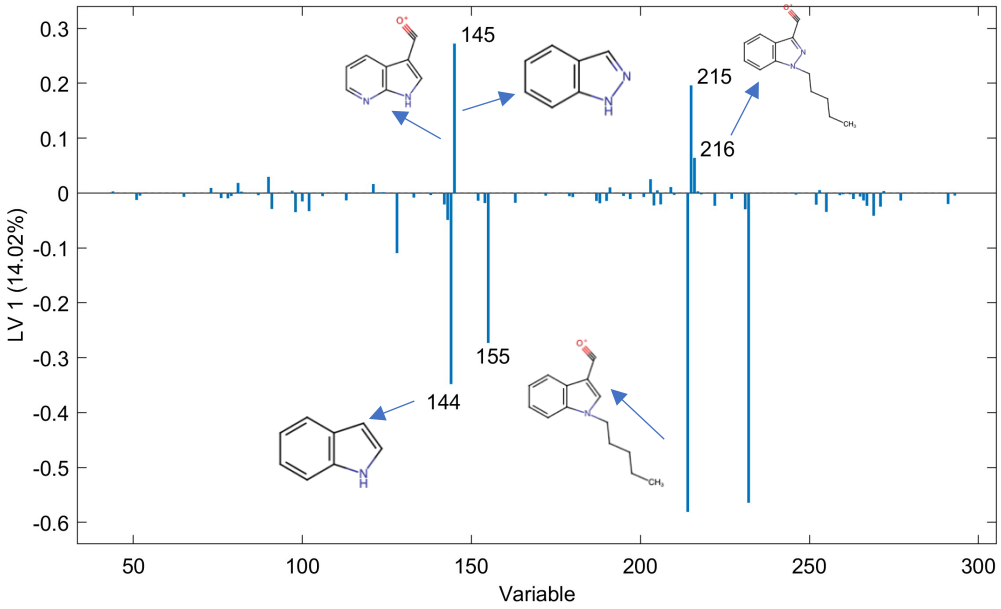


Figure 27. LV1 plot of azaindoles/indazoles (top) versus indoles (bottom).

Additionally, the loading fragment at m/z of 232 (see Figure 27.) and in the presence of the a peak that has the same m/z in mass spectrum of 5F-PCN (see Figure 28.) explains the misclassification of that sample. This loading vector is shown on the LV plot below as a feature of the indole group.

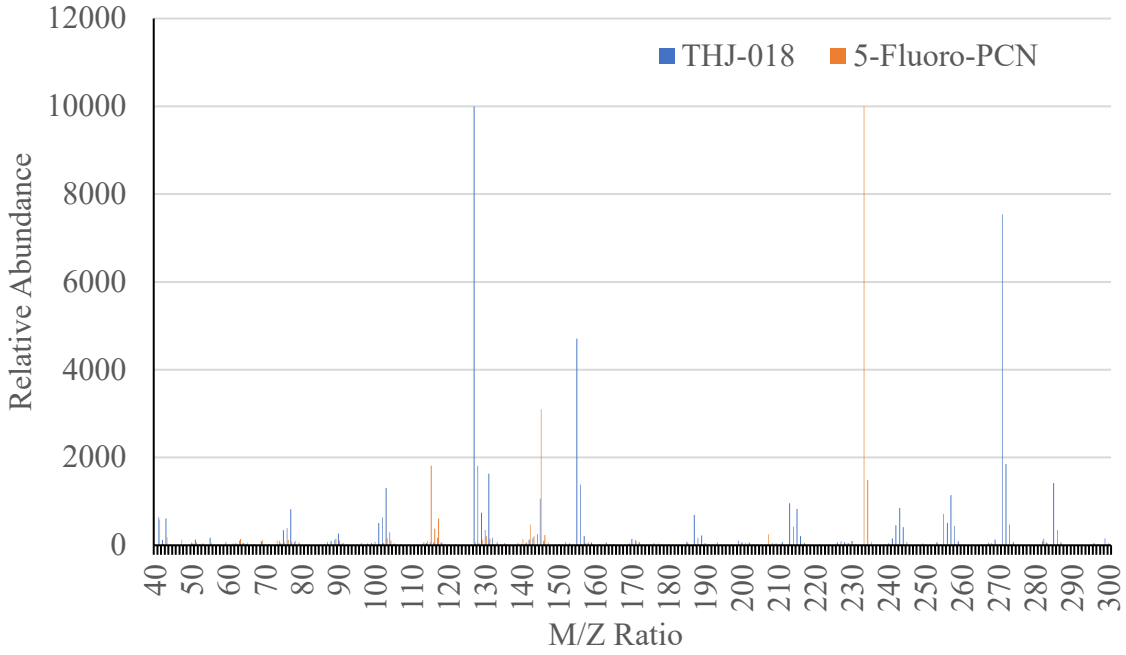


Figure 28. Mass spectra comparison to show issue of chemical structure isomer of THJ-018 (indazole) (blue) and 5F-PCN (azaindole) (orange). Notice around 145 and 127 of similarities that the model looks for in Figure 15.

Further discrimination of azaindoles from indazoles was not possible due to the similar mass spectra forms by these core groups. Azaindoles and indazoles core structure have a core peak at m/z of 145 which is a large loading on the loading vectors.

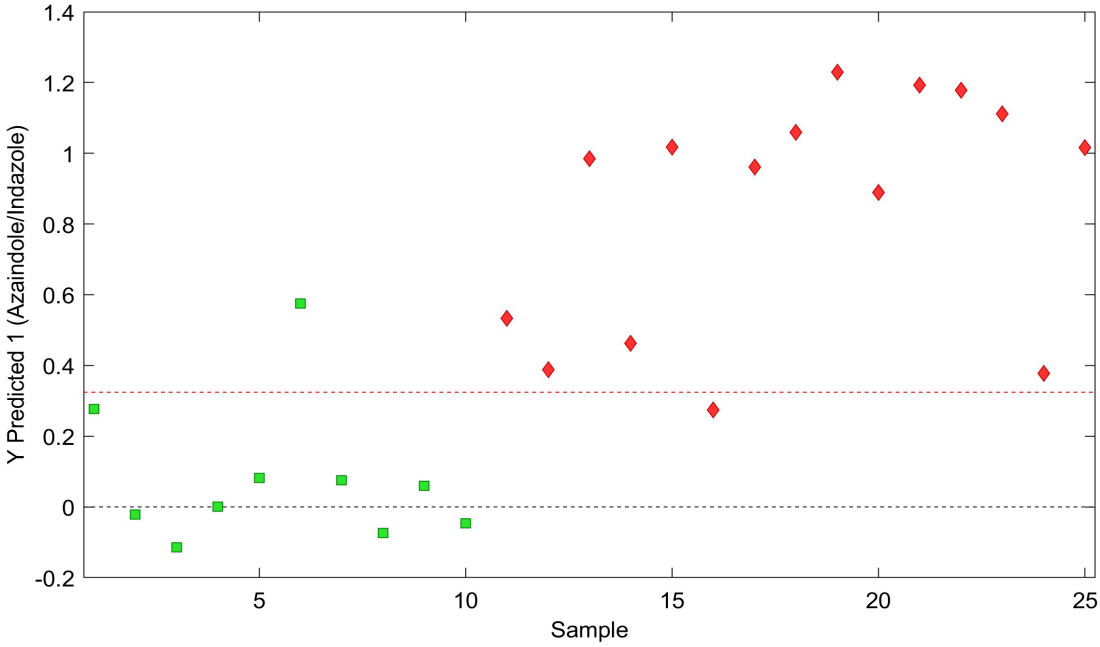


Figure 29. Prediction PLS-DA plot of azaindole/indazole versus indoles using CCC database

The prediction using the spectra obtained from CCC database shows two misclassified compounds. The first compound was an azaindole that has a naphthyl head group along with a peak at 231 m/z which corresponds with the indoles in the loading vectors. The other misidentified sample was FUB-JWH. This compound in general is unique due having a FUB tail group and a naphthyl group. This is a rare combination of these two groups in one compound and the mass spectral data shows that this compound fragments to give a very strong peak at m/z of 109 along with many smaller fragments.

Table 5. Model results with CV for the PLS-DA of azaindoles/indazole versus indoles. The self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Azaindole/Indazole	1.000	1.000	1.000	0.000	0.000
Indole		1.000	1.000	0.000	0.000

Cross-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Azaindole/Indazole	0.9923	0.9744	1.000	0.000	0.0256
Indole		1.000	0.9744	0.0256	0.000

Test Set Results

	Accuracy	TPR	TNR	FPR	FNR
Azaindole/Indazole	0.9200	0.9333	0.9000	0.1000	0.0667
Indole		0.9000	0.9333	0.0667	0.1000

3.5 Classification of Naphthyl containing versus other head containing cannabinoids

Next phase of this research was focused on classifying synthetic cannabinoids based on their head groups. The naphthyl containing cannabinoids were grouped against all the other cannabinoids. The PLS-DA score plot shows successful discrimination. However, model showed some difficulty in separation when it came to different tail groups such as FUB. While the accuracy is 0.97, this is the lowest TPR and TNR in all the models. Data manipulation could solve this problem by removing the prominent m/z of 109 peak in the other group as well as m/z of 194 in the data entry. While this method was not tested, it important to note that the naphthyl head group

is most common with cannabinoids that were developed by John W. Huffman, hence the name starts with JWH.

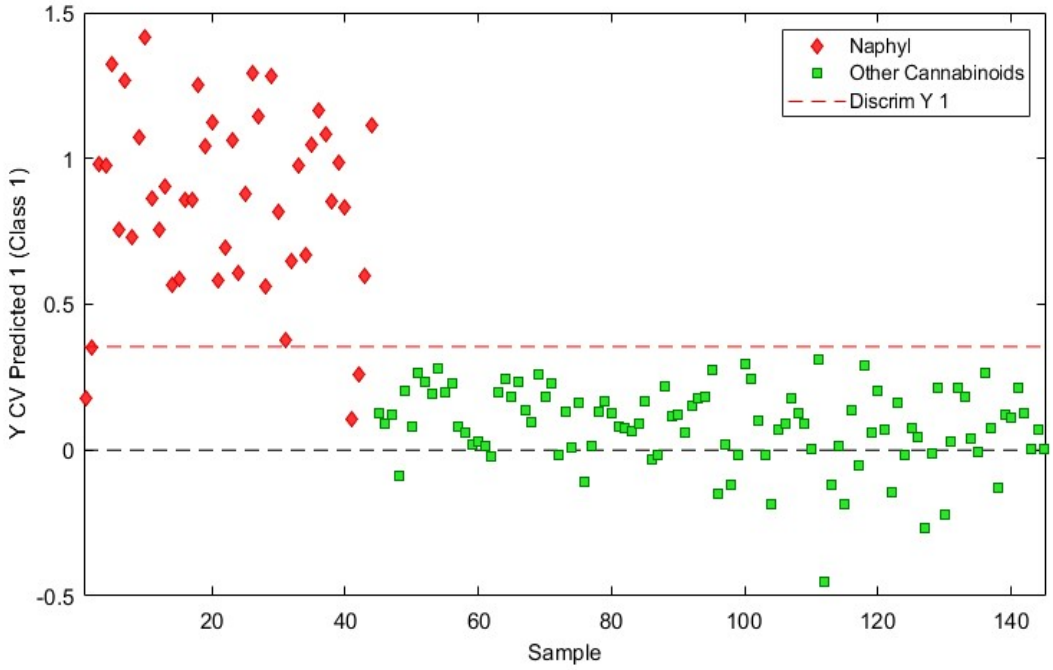


Figure 30. PLS-DA results of Naphthyl head group cannabinoids (red) versus other head group cannabinoids (green)

The LVs and loadings in Figure 31. shows the weight of m/z of 127 and 155 in the discrimination. Both these fragments show the influence of the naphthyl groups. Additionally, the in naphthyl head group containing compounds are usually combined with multiple different core and tail groups.

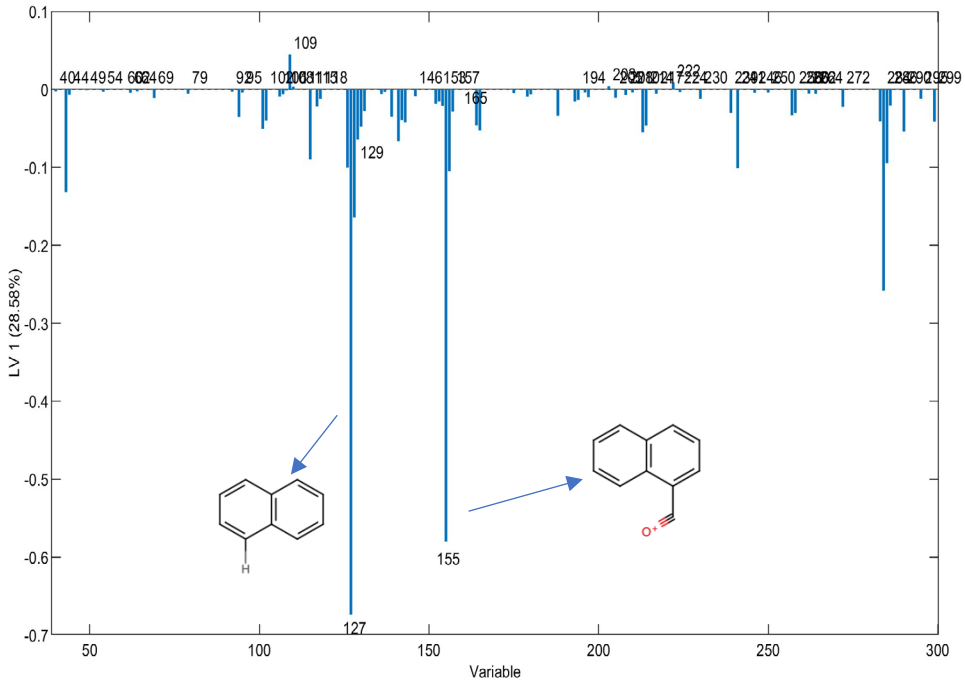


Figure 31. LV1 plot of Naphthyl (bottom) versus other head groups cannabinoids (bottom) with structures. The x-axis is the independent variable mass-to-charge fragments with the most weight.

Figure 32 shows the comparison of a mass spectra of JWH-018 which contains the naphthyl head group and AB-FUBINACA, a misclassified sample, that has the AB headgroup. It is evident that on AB-FUBINACA the tail group FUB forms mass fragments with a large abundance, and the rest of the spectra has lower abundance. This is important because if the model is discriminating the head group, and loadings related to tail groups should not play an important role in discrimination. However, a peak for m/z a 109 can be seen as an important parameter in this discrimination (see Figure 31.) and this suggest that the tail group FUB is being used to discriminate the classes. This can be caused by having many samples that contain the FUB group used in the training set.

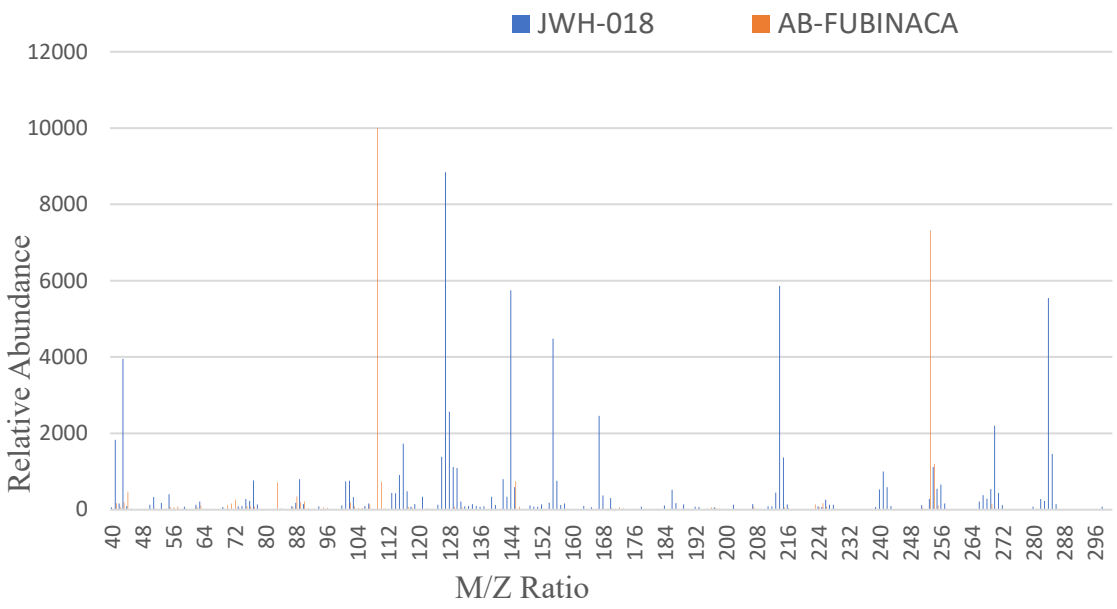


Figure 32. Mass spectra comparison of JWH-018 (Blue) and AB-Fubinaca (Orange)

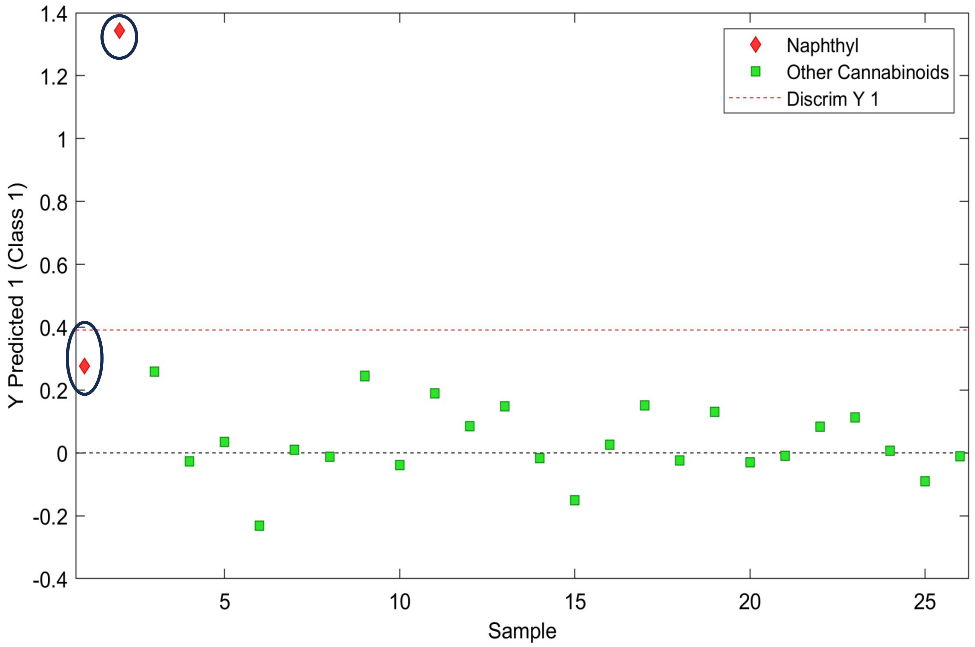


Figure 33. Prediction PLS-DA plot of the Naphthyl head group versus other head group containing cannabinoids.

The prediction set results are shown in Figure 33. shows one misidentified sample FUB-JWH-018. The chemical structure and mass spectral patterns of this compound was discussed in section 3.4 of this thesis.

Table 6. PLS-DA results of Naphthyl head group cannabinoids versus other head group cannabinoids with CV The self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Naphthyl	0.9731	0.9318	1.000	0.000	0.0682
Non-Naphthyl		1.000	0.9318	0.0682	0.000

Cross-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
Naphthyl	0.9724	0.9091	1.000	0.000	0.0909
Non-Naphthyl		1.000	0.9091	0.0909	0.000

Test Set Result

	Accuracy	TPR	TNR	FPR	FNR
Naphthyl	0.9200	0.9333	1.000	0.000	0.0256
Non-Naphthyl		1.000	0.9333	0.0256	0.000

3.6 Classification of FUB containing cannabinoids versus other tail containing cannabinoids

The investigation of head groups revealed that there are several head groups that can generate unique fragments that can be used to discriminate them. The idea is to use these models to predict if an unknown compound had a particular head group. First, FUB group was considered as a classification problem. The PLS-DA score plot shows a clear discrimination. No misclassifications were found and in cross validation. However, this group highlights a unique problem with the number of samples used. Given the low number of samples for the FUB model

the program demonstrates how accurate it can be with a major variable in a particular group. FUB is a unique example of how the base peak plays a role in the computer interpretation and discrimination.

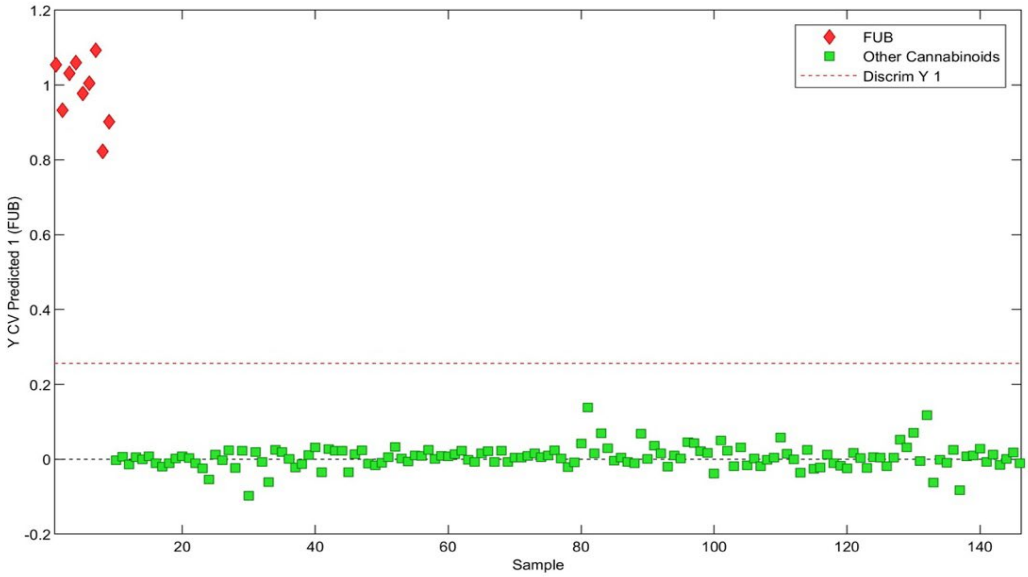


Figure 34. PLS-DA plot of FUB tail group versus other tail group containing cannabinoids

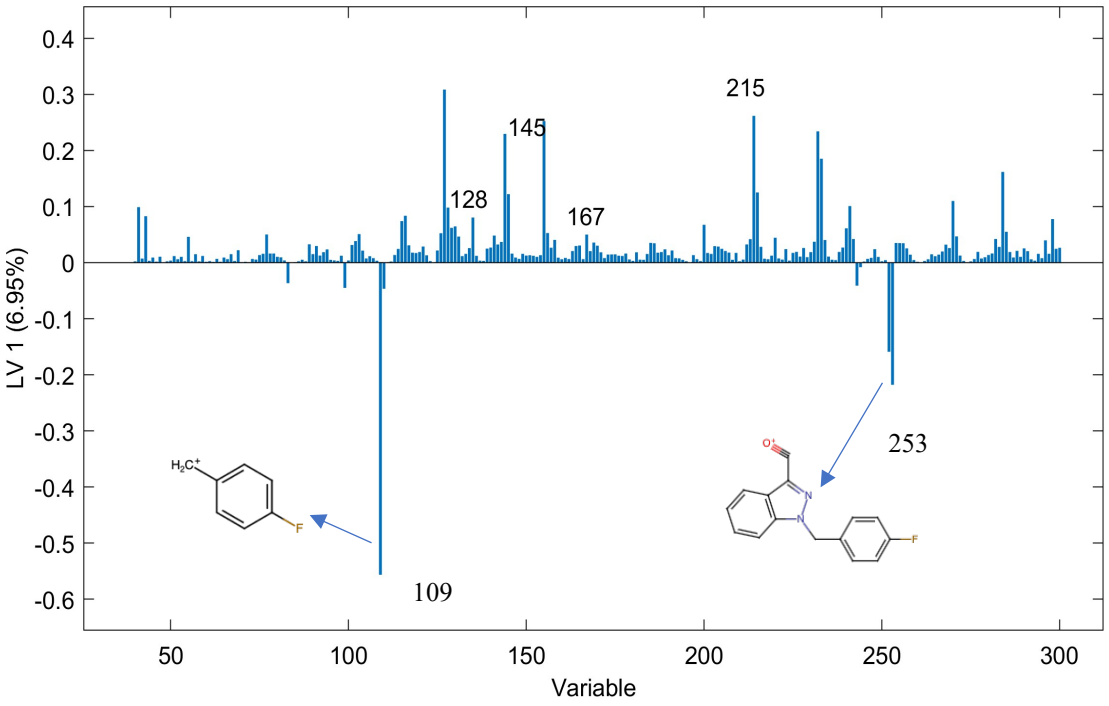


Figure 35. LV1 of FUB (Bottom) and Other Tail Groups (top) The x-axis is the independent variable mass-to-charge fragments with the most weight.

The loading variables show that the FUB group has a major loading at m/z of 109 which is unique to this class of cannabinoids and can be seen in the overlay mass spectra with JWH-018 below in Figure 35. While there is a core fragment at m/z of 145 mixed with the other tail groups the m/z of 109 fragment is the most dominant peak, and the model shows it is very selective. This peak was causing problems with highest loading vectors especially for other models in the beginning stages of method development because of the strong abundance of the fragment as well as the unique group it was representing. The overlay mass spectra in figure 35 shows the overlay of JWH-018 and AB-FUBINACA. This demonstrates what the loadings show that there are two unique peaks one at m/z of 109 and the other at m/z at 253 for the FUB group in cannabinoids.

There is also less overlay of AB-FUBINACA on unique fragments for the JWH-018 which drives further the discrimination of this tail group and its important fragments.

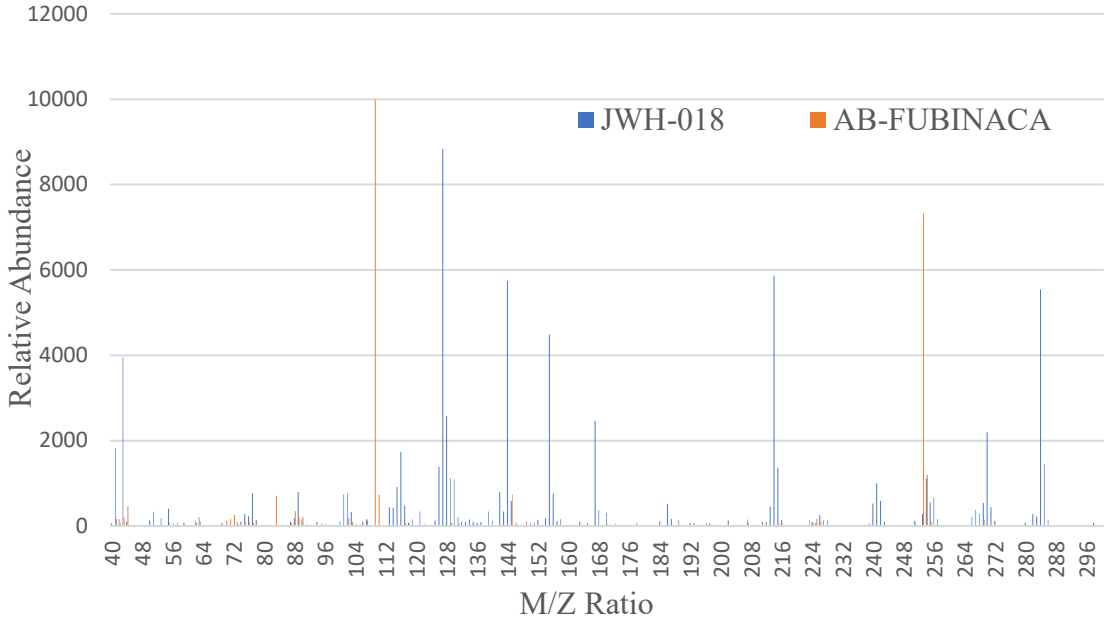


Figure 36. Mass spectra comparison of JWH-018 (Blue) and AB-Fubinaca (Orange)

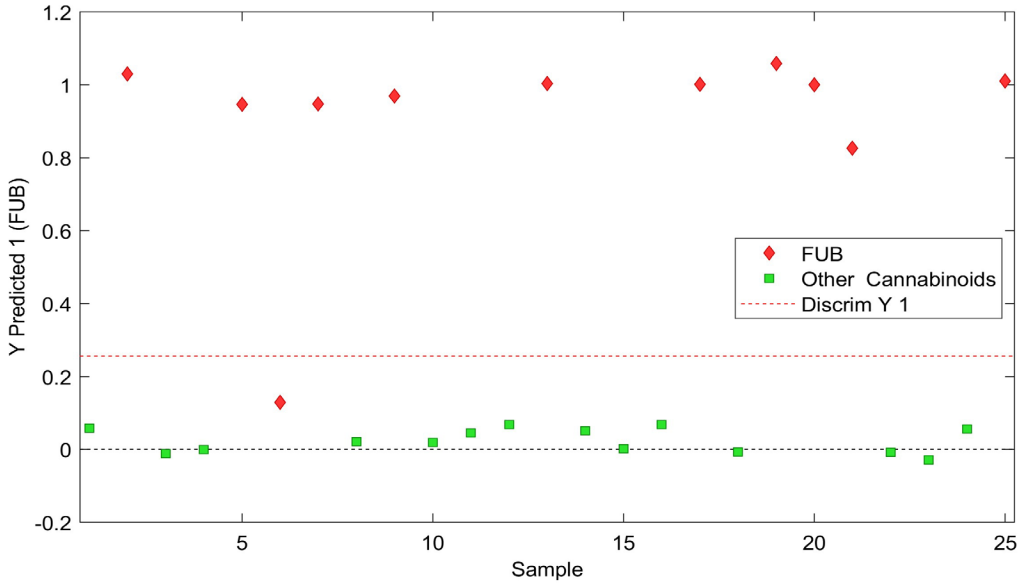


Figure 37. Prediction PLS-DA plot of FUB versus other tail group containing cannabinoids.

The prediction shows one sample incorrectly, FUB-JWH-018 and this situation is similar to the previously state in section 3.4 of this thesis relating to the naphthyl head group and that this type of structural features is rarely seen with one other, and this has previously been unseen in model train and was a real challenge for prediction. The next course of action for model development with this type of compound is to either remove the head group data from the training set such as the m/z of 155 for the naphthyl head group and then try prediction to see if this improves the unknown data.

Table 7. PLS-DA results of FUB tail group cannabinoids versus other tail group cannabinoids with CV The self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
FUB	1.000	1.000	1.000	0.000	0.000
Other Tail Groups		1.000	1.000	0.000	0.000

Cross-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
FUB	1.000	1.000	1.000	0.000	0.000
Other Tail Groups		1.000	1.000	0.000	0.000

Test Set Results

	Accuracy	TPR	TNR	FPR	FNR
FUB	0.960	0.909	1.000	0.000	0.091
Other Tail Groups		1.000	0.909	0.091	0.000

3.7 Classification of BUT containing versus other tail containing cannabinoids

The PLS-DA model developed for BUT group containing cannabinoids versus other synthetic cannabinoids shows accuracy of more than 0.98. The major concern with this group is that the hydrocarbon chain (BUT) attached to the core groups can have different number of carbon atoms and when longer hydrocarbon chains are attached, they have a high probability to cleave and the resulting fragment can have similar fragments that are generated by other cannabinoids that are not included in this class. This can lead to misclassification due to having the fragments with the same m/z ratios. The comparison of mass spectra between JWH-018 which has a pentyl chain and MDMB-BUTINACA in the mass spectra overlay in Figure 40. Shows this behavior in low m/s region.

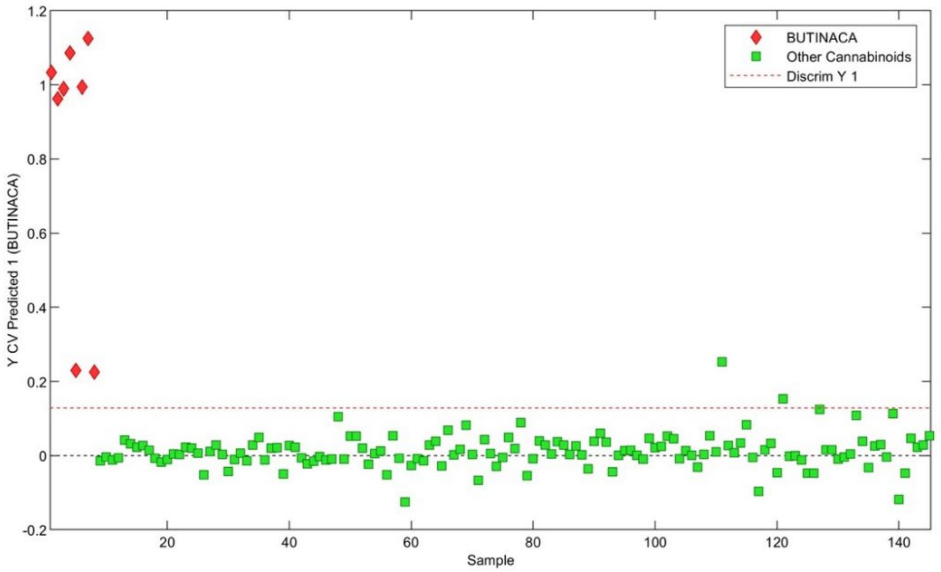


Figure 38.PLS-DA plot of BUT tail group cannabinoids (red) versus other tail group cannabinoids (green)

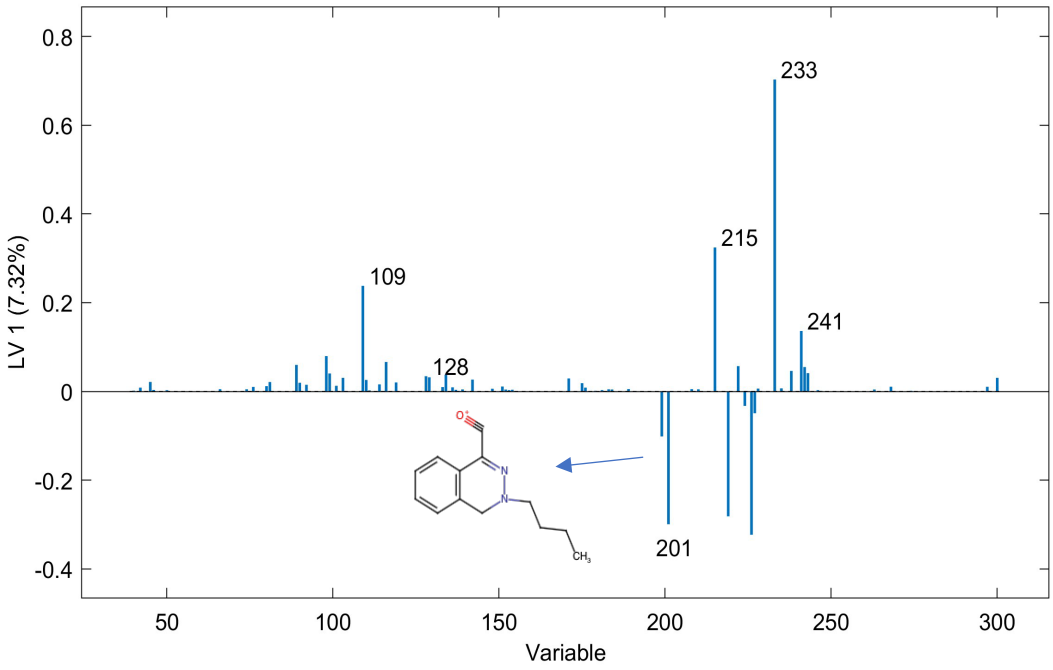


Figure 39. LV1 plot of BUT tail cannabinoids (bottom) versus other tail group cannabinoids (top) The x-axis is the independent variable mass-to-charge fragments with the most weight.

The first loading vector for the BUT versus other tail containing cannabinoids shows that the butyl group cannot be identified on its own that it can only be identified as a useful loading vector with its respective core group around the m/z of 200. While this is a good way to discriminate the BUT tail group from the other tail groups it can cause an area for concern if new cannabinoids have an unknown core structure with a BUT group. More information needs to be added and examined for the possibility of isolating this fragment but since this is a linear hydrocarbon chain and thus, there is a higher chance of fragment being similar to the fragments generated by other functional groups cannabinoids.

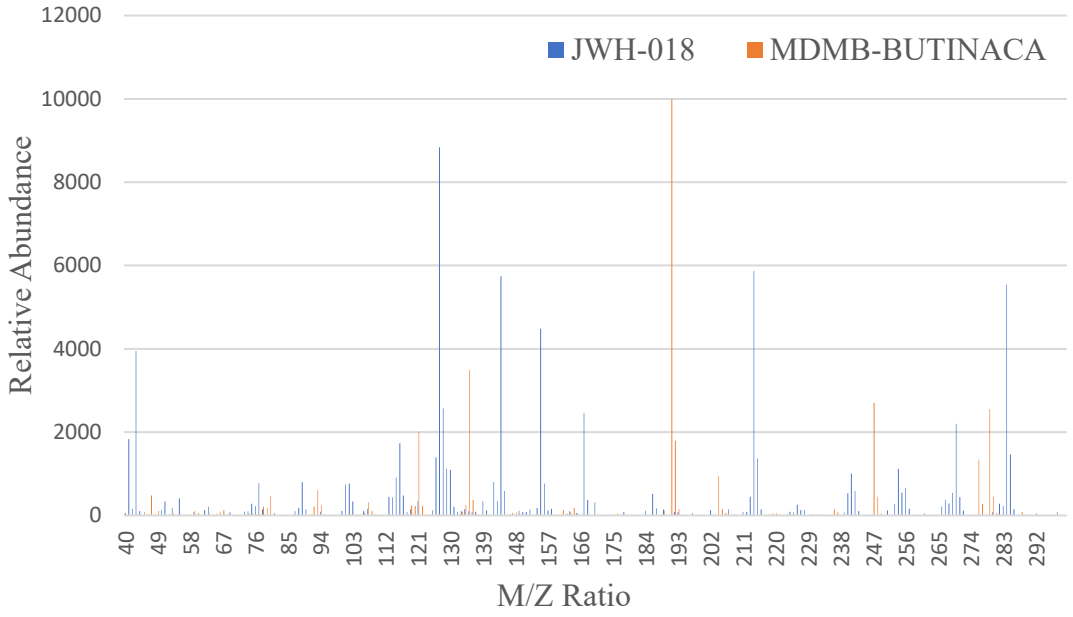


Figure 40. Mass spectra comparison of JWH-018 (Blue) and MDMB-BUTINACA(Orange)

Figure 40 uses a mass spectra overlay to demonstrate that the BUT tail group containing cannabinoid MDMB-BUTINACA has that base fragment around 200 which is the area of interest

in the LV plot for the BUT group. This also demonstrates the difference in tail groups as JWH-018 has a pentyl tail group and this cleaves differently than the BUT group.

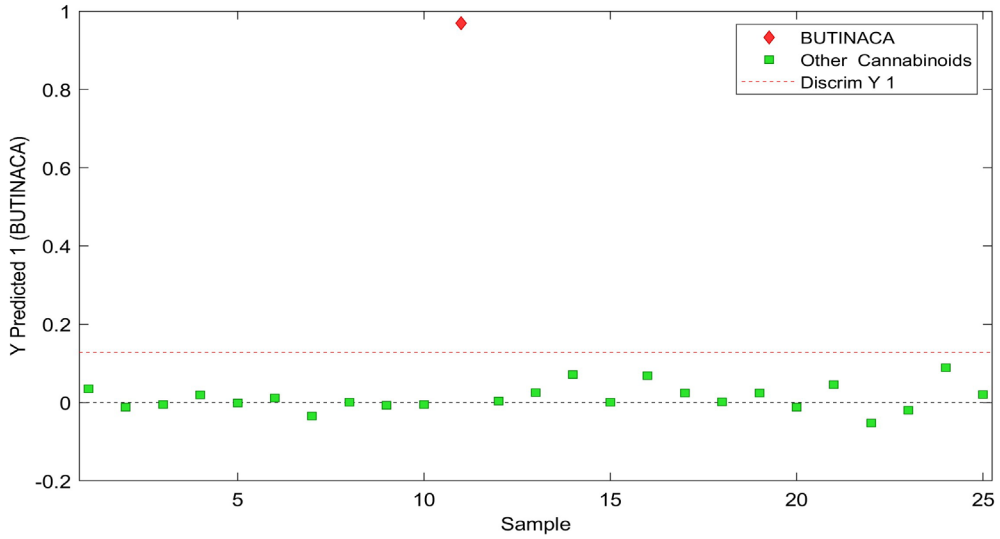


Figure 41. Prediction PLS-DA plot of BUT containing cannabinoids versus other tail containing cannabinoids.

Table 8. PLS-DA results of BUT tail group cannabinoids versus other tail group cannabinoids with cross validation. The self-validation is completed using the prediction most probable method with no cross validation. The cross-validation results are an average of the random subset methods using 10% of samples over 10 iterations. And the test sets are done using cannabinoids not previously in the training set from the CCC database.

Self-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
BUT	0.9927	1.000	0.9927	0.0073	0.000
Other Tail Groups		0.9927	1.000	0.000	0.0073

Cross-Validation Results

	Accuracy	TPR	TNR	FPR	FNR
BUT	0.9854	1.000	0.9854	0.0146	0.0000
Other Tail Groups		0.9854	1.000	0.0000	0.0146

Test Set Results

	Accuracy	TPR	TNR	FPR	FNR
BUT	1.000	1.000	1.000	0.000	0.0000
Other Tail Groups		1.000	1.000	.0000	0.000

CHAPTER FOUR: CONCLUSIONS AND FUTURE DIRECTIONS

This research has shown the ability to use mass spectral data of known compounds to aid in the classification and identification of possible new cannabinoids. Classification models developed using readily available mass spectral data showed successful discrimination of cannabinoid related compounds from other drug classes. Additionally, cannabinoids were classified into several subgroups based on their functional groups including benzopyrrole (indole), isoindazole (indazole), naphthalene (Naphthyl), 4-Fluorobenzyl (FUB), and 1-amino-3,3-dimethyl-1-oxobutan-2-yl (BUT). The cross was used to validate the developed models and all the models developed show more than 95% accuracy. The prediction set used to further test the robustness of these models is compiled using a different database and the cannabinoids that were not used to build the models. The results of the test set demonstrate the applicability of the developed models. Furthermore, the analysis of LVs and loadings shows that the basis of discrimination of each model. Furthermore, the information provided by the LVs and loadings can be used to find the unique and most abundant mass to charge ratios (m/z) of each drug class. The detailed analysis of misclassified samples demonstrates the limitations of the use of mass spectral data. The models also show how the abundance of fragmentation can affect predicting the newly synthesized cannabinoids. The mass spectral data has a unique problem when used in building discriminant models. The discrete m/z values and the formation of fragments that have similar m/z ratios from different functional groups can lead to misclassifications. To obviate this problem, mass spectral data can be combined with chemical data obtained using other methods such as gas chromatography infrared spectroscopy (GC-IR). However, GC-IR spectral libraries are not readily available for use. This is an interest in the future research as the use of GC-IR and related

techniques are becoming more common in the forensic science field. Although these methods would not eliminate the use of structural determination techniques such as NMR and HRMS, the detection of novel drugs in variety of matrices and the identification of functional groups using such methods on them would greatly benefit the forensic analysis process.

The focus of the future research is on other main drug classes as well as a deeper look on structural isomers. As more compounds are identified their spectra can be added to increase the precision and accuracy of the model development.

REFERENCES

(1) Shipman, R.; Conti, T.; Tighe, T.; Buel, E. *Forensic Drug Identification by Gas Chromatography - Infrared Spectroscopy*; Vermont Forensic Laboratory, 2013.

(2) Bovens, M.; Bissig, C.; Staeheli, S. N.; Poetzsch, M.; Pfeiffer, B.; Kraemer, T. Structural Characterization of the New Synthetic Cannabinoids CUMYL-PINACA, 5F-CUMYL-PINACA, CUMYL-4CN-BINACA, 5F-CUMYL-P7AICA and CUMYL-4CN-B7AICA. *Forensic Science International* **2017**, *281*, 98–105. https://doi.org/10.1016/j.forsciint.2017.10.020.

(3) SWGDRUG https://swgdrug.org/ms.htm (Accessed April 25, 2024)

(4) Cayman Chemical Company https://www.caymanchem.com (Accessed April 25, 2024)

(5) Bell, S. *Forensic Chemistry*, Third edition.; CRC Press: Boca Raton London New York, 2022. https://doi.org/10.4324/9780429440915.

(6) Flavin, A.; Azizi, P.; Murataeva, N.; Yust, K.; Du, W.; Ross, R.; Greig, I.; Nguyen, T.; Zhang, Y.; Mackie, K.; Straiker, A. CB1 Receptor Negative Allosteric Modulators as a Potential Tool to Reverse Cannabinoid Toxicity. *Molecules* **2024**, *29* (8). https://doi.org/10.3390/molecules29081881.

(7) Adams, A. J.; Banister, S. D.; Irizarry, L.; Trecki, J.; Schwartz, M.; Gerona, R. "Zombie" Outbreak Caused by the Synthetic Cannabinoid AMB-FUBINACA in New York. *New England journal of medicine* **2017**, *376* (3), 235–242.

(8) UNODC https://www.unodc.org (Accessed April 25, 2024)

(9) A. J. Potts, S. H. L. T., C. Cano; Hill, S. L. Synthetic Cannabinoid Receptor Agonists: Classification and Nomenclature. *Clinical Toxicology* **2020**, *58* (2), 82–98. https://doi.org/10.1080/15563650.2019.1661425.

(10) Namera, A.; Kawamura, M.; Nakamoto, A.; Saito, T.; Nagao, M. Comprehensive Review of the Detection Methods for Synthetic Cannabinoids and Cathinones. *Forensic Toxicology* **2015**, *33* (2), 175–194. https://doi.org/10.1007/s11419-015-0270-0.

(11) Mohr, A. L. A.; Logan, B. K.; Fogarty, M. F.; Krotulski, A. J.; Papsun, D. M.; Kacinko, S. L.; Huestis, M. A.; Ropero-Miller, J. D. Reports of Adverse Events Associated with Use of Novel Psychoactive Substances, 2017-2020: A Review. *J Anal Toxicol* **2022**, *46* (6), e116–e185. https://doi.org/10.1093/jat/bkac023.

(12) Abbate, V.; Schwenk, M.; Presley, B. C.; Uchiyama, N. The Ongoing Challenge of Novel Psychoactive Drugs of Abuse. Part I. Synthetic Cannabinoids (IUPAC Technical Report). *Pure and Applied Chemistry* **2018**, *90* (8), 1255–1282. https://doi.org/doi:10.1515/pac-2017-0605.

(13) CFSRE https://www.cfsre.org (Accessed April 25, 2024)

(14) Huffman, S. Chem455 course. https://doi.org/10.15139/S3/6J9ZAU (Accessed April 25, 2024)).

(15) Gromski, P. S.; Xu, Y.; Correa, E.; Ellis, D. I.; Turner, M. L.; Goodacre, R. A Comparative Investigation of Modern Feature Selection and Classification Approaches for the Analysis of Mass Spectrometry Data. *Analytica Chimica Acta* **2014**, *829*, 1–8. https://doi.org/10.1016/j.aca.2014.03.039.

(16) Kranenburg, R. F.; Peroni, D.; Affourtit, S.; Westerhuis, J. A.; Smilde, A. K.; Asten, A. C. van. Revealing Hidden Information in GC–MS Spectra from Isomeric Drugs: Chemometrics Based Identification from 15 eV and 70 eV EI Mass Spectra. *Forensic Chemistry* **2020**, *18*, 100225. https://doi.org/10.1016/j.forc.2020.100225.
(17) Tharwat, A. Classification Assessment Methods. *ACI* **2021**, *17* (1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003.
(18) Kyzar, E. J.; Nichols, C. D.; Gainetdinov, R. R.; Nichols, D. E.; Kalueff, A. V. Psychedelic Drugs in Biomedicine. *Trends in Pharmacological Sciences* **2017**, *38* (11), 992–1005. https://doi.org/10.1016/j.tips.2017.08.003.
(19) Norwood, A. P. Criminal Law—When Apples Tatse Like Oranges, You Cannot Judge a Book by Its Cover: How to Fight Emerging Synthetic" Designer" Drugs of Abuse. *University of Arkansas at Little Rock Law Review* **2017**, *39* (2), 323.

APPINDEX A

Table 9.. List of compounds with respective classes for Cannabinoids versus other Drug classes model

Label	Class
Cannabigerol	Cannabinoids
Cannabicyclol	Cannabinoids
Cannabichromene	Cannabinoids
Cannabidiol	Cannabinoids
Cannabinol	Cannabinoids
delta-9-Tetrahydrocannabinol	Cannabinoids
delta-8-Tetrahydrocannabinol	Cannabinoids
5-Fluoro-MN-18	Cannabinoids
Cannabidiolic Acid	Cannabinoids
MMB2201	Cannabinoids
AB-PINACA N-(3-fluoropentyl) isomer	Cannabinoids
Tetrahydrocannabivarin	Cannabinoids
5-Fluoro-ADBICA	Cannabinoids
5-Fluoro-ADB	Cannabinoids
5-Fluoro-ADB-PINACA isomer 2	Cannabinoids
5-Fluoro-PCN	Cannabinoids
5-Fluoro-AEB	Cannabinoids
Cannabipiperidiethanone	Cannabinoids
Cannabidivarin	Cannabinoids
4-Fluoro-AMB	Cannabinoids
5-Fluoro-CUMYL-PICA	Cannabinoids
Cannabicitran	Cannabinoids
Cannabichromevarin	Cannabinoids
Cannabichromeorcin	Cannabinoids
5-Fluoro-AB-001	Cannabinoids
Cannabidiol-2TMS	Cannabinoids
Cannabidiolic Acid-3TMS	Cannabinoids
Tetrahydrocannabinolic Acid-2TMS	Cannabinoids
delta-9-Tetrahydrocannabinol-TMS	Cannabinoids
Cannabigerorcinic Acid	Cannabinoids
Cannabigerorcin	Cannabinoids
5-Fluoro-7-APAICA	Cannabinoids
Cannabigerovarin	Cannabinoids
5-Fluoro MDMB-PICA metabolite 7	Cannabinoids
4-Fluoro MDMB-BUTINACA	Cannabinoids
5F-NNEI 2'-naphthyl isomer	Cannabinoids
5-Fluoro-PB-22	Cannabinoids

5-Fluoro-SDB-006	Cannabinoids
5-Fluoro-THJ	Cannabinoids
A-796,260	Cannabinoids
A-834,735	Cannabinoids
A-836,339	Cannabinoids
AB-CHMINACA	Cannabinoids
AB-FUBINACA	Cannabinoids
AB-PINACA	Cannabinoids
ADB-FUBINACA	Cannabinoids
ADBICA	Cannabinoids
ADB-PINACA	Cannabinoids
AKB48	Cannabinoids
AM1220	Cannabinoids
AM1235	Cannabinoids
AM1241	Cannabinoids
AM1248	Cannabinoids
AM2201	Cannabinoids
AM2201 8-quinolinyl carboxamide	Cannabinoids
AM2232	Cannabinoids
AM2233	Cannabinoids
AM679	Cannabinoids
AM694	Cannabinoids
AMB	Cannabinoids
BB-22	Cannabinoids
CB-13	Cannabinoids
CB-25	Cannabinoids
CB-52	Cannabinoids
CP 47,497	Cannabinoids
CP 47,497-C8 Homolog	Cannabinoids
CP 55,940	Cannabinoids
CUMYL-PICA	Cannabinoids
delta-9-Tetrahydrocannabinol	Cannabinoids
EG-018	Cannabinoids
EG-2201	Cannabinoids
FDU-PB-22	Cannabinoids
XLR11	Cannabinoids
FUB-144	Cannabinoids
FUB-AMB	Cannabinoids
AM2201 benzimidazole analog	Cannabinoids
FUB-PB-22	Cannabinoids
HU-210	Cannabinoids
HU-211	Cannabinoids
JWH-081-N-(cyclohexylmethyl) analog	Cannabinoids
JWH-007	Cannabinoids
JWH-011	Cannabinoids

JWH-015	Cannabinoids
JWH-016	Cannabinoids
JWH-018	Cannabinoids
AB-001	Cannabinoids
APICA	Cannabinoids
JWH-018 benzimidazole analog	Cannabinoids
THJ-018	Cannabinoids
JWH-019	Cannabinoids
JWH-020	Cannabinoids
JWH-022	Cannabinoids
JWH-030	Cannabinoids
JWH-031	Cannabinoids
JWH-072	Cannabinoids
JWH-073	Cannabinoids
JWH-073	Cannabinoids
JWH-073 2-methylnaphthyl analog	Cannabinoids
JWH-073 4-methylnaphthyl analog	Cannabinoids
JWH-081	Cannabinoids
JWH-098	Cannabinoids
JWH-122	Cannabinoids
JWH-122 N-(4-pentenyl) analog	Cannabinoids
JWH-145	Cannabinoids
JWH-147	Cannabinoids
JWH-175	Cannabinoids
JWH-180	Cannabinoids
JWH-182	Cannabinoids
JWH-203	Cannabinoids
JWH-210	Cannabinoids
JWH-249	Cannabinoids
JWH-167	Cannabinoids
JWH-251	Cannabinoids
JWH-307	Cannabinoids
JWH-309	Cannabinoids
JWH-368	Cannabinoids
JWH-369	Cannabinoids
JWH-370	Cannabinoids
JWH-398	Cannabinoids
JWH-424	Cannabinoids
ADB-CHMINACA	Cannabinoids
MA-CHMINACA	Cannabinoids
MAM2201	Cannabinoids
MDMB-CHMICA	Cannabinoids
MDMB-CHMINACA	Cannabinoids
MDMB-FUBINACA	Cannabinoids
MN-18	Cannabinoids

MO-CHMINACA	Cannabinoids
JWH-200	Cannabinoids
Nabilone	Cannabinoids
NM2201	Cannabinoids
NNEI	Cannabinoids
UR-144	Cannabinoids
Pravadoline	Cannabinoids
PX 1	Cannabinoids
PX 2	Cannabinoids
PB-22	Cannabinoids
RCS-4	Cannabinoids
RCS-4-C4 homolog	Cannabinoids
RCS-8	Cannabinoids
SDB-005	Cannabinoids
SDB-006	Cannabinoids
STS-135	Cannabinoids
THJ	Cannabinoids
THJ2201	Cannabinoids
UR-144 N-heptyl analog	Cannabinoids
URB597	Cannabinoids
URB602	Cannabinoids
URB754	Cannabinoids
WIN 55,212-2	Cannabinoids
XLR11 N-(4-pentenyl) analog	Cannabinoids
XLR12	Cannabinoids
4-Chloro MDMB-BUTICA	Cannabinoids
MDMB-FUB7AICA	Cannabinoids
MDMB-CHM7AICA	Cannabinoids
MPP-PICA	Cannabinoids
MMB-CHM7AICA	Cannabinoids
5-Fluoro MMB-P7AICA	Cannabinoids
5-Fluoro EDMB-PICA	Cannabinoids
5-Fluoro EMB-PICA	Cannabinoids
5-Fluoro AB-7-PAICA	Cannabinoids
CUMYL-CBMICA	Cannabinoids
5-Chloro MDMB-PICA	Cannabinoids
Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Cannabinoids
5-Chloro AMB-PICA	Cannabinoids
5-Bromo AMB-PICA	Cannabinoids
CUMYL-CBMINACA	Cannabinoids
MDMB-BUTINACA	Cannabinoids
ADB-4en-PINACA	Cannabinoids
ADB-PHETINACA	Cannabinoids
THQ-PINACA	Cannabinoids
iPDMB-FUBINACA	Cannabinoids

4-Cyano ADB-BUTINACA	Cannabinoids
Benzyl-4-cyano BUTINACA	Cannabinoids
4-Fluoro ABUTINACA	Cannabinoids
5F-EDMB-PINACA	Cannabinoids
ADB-BUTINACA	Cannabinoids
MMB-4en-PINACA	Cannabinoids
4-Cyano MDMB-BUTINACA	Cannabinoids
MDMB-3en-BUTINACA	Cannabinoids
4-Fluoro MDMB-BUTINACA	Cannabinoids
CUMYL-PINACA	Cannabinoids
ADB-BINACA	Cannabinoids
CUMYL-THPINACA	Cannabinoids
Adamantyl-THPINACA	Cannabinoids
"25C-NBOMe"	Phenylethylamines
"25D-NBOMe"	Phenylethylamines
"25E-NBOMe"	Phenylethylamines
"25I-NBOMe"	Phenylethylamines
"2C-B-fly"	Phenylethylamines
"2C-C"	Phenylethylamines
"2C-D"	Phenylethylamines
"2C-E"	Phenylethylamines
"2C-H"	Phenylethylamines
"2C-N"	Phenylethylamines
"2C-P"	Phenylethylamines
"2C-T-2"	Phenylethylamines
"2C-T-4"	Phenylethylamines
"30C-NBOMe"	Phenylethylamines
"4-AcO-DIPT"	Tryptamines
"Bufotenine"	Tryptamines
"Diethyltryptamine"	Tryptamines
"Dimethyltryptamine"	Tryptamines
"5-Methoxy-N,N-Diisopropyltryptamine"	Tryptamines
"5-Methoxy-N,N-Dimethyltryptamine"	Tryptamines
"5-Methoxy-.alpha.-methyltryptamine"	Tryptamines
.alpha.'-methyl Butyryl fentanyl	Fentanyl Derivatives
.alpha.-Methyl Butyryl fentanyl	Fentanyl Derivatives
.alpha.-Methyl Thiofentanyl	Fentanyl Derivatives
.alpha.-Methylfentanyl	Fentanyl Derivatives 2
.beta.-Hydroxyfentanyl	Fentanyl Derivatives 2
.beta.-Hydroxythiofentanyl	Fentanyl Derivatives
.beta.-methyl Acetyl fentanyl	Fentanyl Derivatives 2
.beta.-Methyl Fentanyl	Fentanyl Derivatives
2,2,3,3-Tetramethyl-Cyclopropyl fentanyl	Fentanyl Derivatives
2,3-Seco-Fentanyl	Fentanyl Derivatives
2'-Fluoro ortho-Fluorofentanyl	Fentanyl Derivatives

2'-Fluoro, ortho-fluoro-cis-3-methyl Fentanyl	Fentanyl Derivatives
2'-Fluoro, ortho-fluoro-trans-3-methyl Fentanyl	Fentanyl Derivatives 2
2'-Fluorofentanyl	Fentanyl Derivatives
2-Furanyl fentanyl	Fentanyl Derivatives
2'-Methyl Acetyl fentanyl	Fentanyl Derivatives 2
2'-Methyl Fentanyl	Fentanyl Derivatives
3'-Fluoro ortho-Fluorofentanyl	Fentanyl Derivatives
3-Furanyl fentanyl	Fentanyl Derivatives
3'-Methyl Acetyl fentanyl	Fentanyl Derivatives
3-Methyl butyryl fentanyl	Fentanyl Derivatives 2
3'-Methyl Fentanyl	Fentanyl Derivatives
3-Methylfentanyl	Fentanyl Derivatives
4'-Fluoro, ortho-fluoro-cis-3-methyl Fentanyl	Fentanyl Derivatives
4'-Fluoro, ortho-fluoro-trans-3-methyl Fentanyl	Fentanyl Derivatives
4'-Fluoro, para-fluoro-cis-3-methyl Fentanyl	Fentanyl Derivatives
4'-Fluoro, para-fluoro-trans-3-methyl Fentanyl	Fentanyl Derivatives
4-Fluorobutyrylfentanyl	Fentanyl Derivatives 2
4'-Fluorofentanyl	Fentanyl Derivatives
4'-Methyl acetyl fentanyl	Fentanyl Derivatives
4-Phenyl fentanyl	Fentanyl Derivatives
Acetyl fentanyl	Fentanyl Derivatives
Acetyl norfentanyl	Fentanyl Derivatives 2
Acetyl-.alpha.-methyl fentanyl	Fentanyl Derivatives
Acrylfentanyl	Fentanyl Derivatives
Benzodioxole fentanyl	Fentanyl Derivatives
Benzyl Acrylfentanyl	Fentanyl Derivatives
Benzylfentanyl	Fentanyl Derivatives
beta'-Phenyl fentanyl	Fentanyl Derivatives
Butyryl fentanyl	Fentanyl Derivatives
Butyryl norfentanyl	Fentanyl Derivatives
cis-3-Methyl Butyryl fentanyl	Fentanyl Derivatives
cis-3-methyl Norfentanyl	Fentanyl Derivatives
cis-3-Methyl Thiofentanyl	Fentanyl Derivatives
cis-Isofentanyl	Fentanyl Derivatives
Crotonyl fentanyl	Fentanyl Derivatives
Cyclobutyl fentanyl	Fentanyl Derivatives
Cyclohexyl fentanyl	Fentanyl Derivatives
Cyclopentenyl fentanyl	Fentanyl Derivatives
Cyclopentyl fentanyl	Fentanyl Derivatives
Cyclopropyl fentanyl	Fentanyl Derivatives
Cyclopropyl norfentanyl	Fentanyl Derivatives
Despropionyl 2'-fluoro ortho-Fluorofentanyl	Fentanyl Derivatives
Despropionyl meta-Fluorofentanyl	Fentanyl Derivatives
Despropionyl meta-Methylfentanyl	Fentanyl Derivatives
Despropionyl ortho-Methylfentanyl	Fentanyl Derivatives

Despropionyl para-Fluorofentanyl	Fentanyl Derivatives
Despropionyl para-Methylfentanyl	Fentanyl Derivatives
Despropionyl-2-fluorofentanyl	Fentanyl Derivatives
Ethoxyacetyl fentanyl	Fentanyl Derivatives
Fentanyl	Fentanyl Derivatives
Fentanyl Carbamate	Fentanyl Derivatives
Fentanyl meta methylphenyl acetyl analog	Fentanyl Derivatives 2
Fentanyl meta methylphenyl analog	Fentanyl Derivatives
Fentanyl meta tolyl analog	Fentanyl Derivatives
Fentanyl methyl acetyl analog	Fentanyl Derivatives
Fentanyl Methyl Carbamate	Fentanyl Derivatives
Fentanyl ortho methylphenyl acetyl analog	Fentanyl Derivatives 2
Fentanyl ortho tolyl acetyl analog	Fentanyl Derivatives
Fentanyl ortho tolyl analog	Fentanyl Derivatives 2
Fentanyl para methylphenyl analog	Fentanyl Derivatives 2
Fentanyl para tolyl acetyl analog	Fentanyl Derivatives
Fentanyl propyl acetyl analog	Fentanyl Derivatives 2
Fentanyl propyl analog	Fentanyl Derivatives
Furanyl norfentanyl	Fentanyl Derivatives 2
Furanylethyl fentanyl	Fentanyl Derivatives
Heptanoyl fentanyl	Fentanyl Derivatives
Hexanoyl fentanyl	Fentanyl Derivatives
Isobutyryl fentanyl	Fentanyl Derivatives
Isobutyryl norfentanyl	Fentanyl Derivatives
Isovaleryl fentanyl	Fentanyl Derivatives
meta-Fluoro Acrylfentanyl	Fentanyl Derivatives
meta-Fluoro Furanyl fentanyl	Fentanyl Derivatives
meta-Fluoro Methoxyacetyl fentanyl	Fentanyl Derivatives
meta-Fluoro Valeryl fentanyl	Fentanyl Derivatives
meta-Fluorobutyryl fentanyl	Fentanyl Derivatives
meta-Fluorofentanyl	Fentanyl Derivatives
meta-Fluoroisobutyryl fentanyl	Fentanyl Derivatives
meta-Methoxy Furanyl fentanyl	Fentanyl Derivatives 2
meta-Methyl Acetyl fentanyl	Fentanyl Derivatives
meta-Methyl Cyclopropyl fentanyl	Fentanyl Derivatives
meta-Methyl Furanyl fentanyl	Fentanyl Derivatives
meta-Methyl Methoxyacetyl fentanyl	Fentanyl Derivatives
meta-Methylfentanyl	Fentanyl Derivatives
Methacrylfentanyl	Fentanyl Derivatives
Methoxyacetyl fentanyl	Fentanyl Derivatives
Methoxyacetyl norfentanyl	Fentanyl Derivatives 2
N-(3-ethylindole) Norfentanyl	Fentanyl Derivatives
N,N-Dimethylamido-despropionyl fentanyl	Fentanyl Derivatives
N-benzyl Furanyl norfentanyl	Fentanyl Derivatives
N-Benzyl meta-fluoro Norfentanyl	Fentanyl Derivatives

N-benzyl para-fluoro Cyclopropyl norfentanyl	Fentanyl Derivatives
N-benzyl para-fluoro norfentanyl	Fentanyl Derivatives
N-Benzyl para-fluoro Norfentanyl	Fentanyl Derivatives
N-Benzyl phenyl norfentanyl	Fentanyl Derivatives
N-methyl Cyclopropyl norfentanyl	Fentanyl Derivatives
N-methyl Norfentanyl	Fentanyl Derivatives
Norfentanyl	Fentanyl Derivatives
o-Fluorofentanyl	Fentanyl Derivatives
ortho-Fluoro Acrylfentanyl	Fentanyl Derivatives
ortho-Fluoro Furanyl fentanyl	Fentanyl Derivatives
ortho-Fluorobutyryl fentanyl	Fentanyl Derivatives
ortho-Fluoroisobutyryl fentanyl	Fentanyl Derivatives
ortho-Methoxy Furanyl fentanyl	Fentanyl Derivatives
ortho-Methoxy-Butyryl fentanyl	Fentanyl Derivatives 2
ortho-Methyl Acetyl fentanyl	Fentanyl Derivatives
ortho-Methyl Acrylfentanyl	Fentanyl Derivatives
ortho-Methyl Cyclopropyl fentanyl	Fentanyl Derivatives
ortho-Methyl Furanyl fentanyl	Fentanyl Derivatives
ortho-Methyl Methoxyacetyl fentanyl	Fentanyl Derivatives
ortho-Methylfentanyl	Fentanyl Derivatives
para-Chloro Acrylfentanyl	Fentanyl Derivatives
para-Chloro Cyclobutyl fentanyl	Fentanyl Derivatives
para-Chloro Cyclopentyl fentanyl	Fentanyl Derivatives
para-Chloro Cyclopropyl fentanyl	Fentanyl Derivatives
para-Chloro Furanyl fentanyl	Fentanyl Derivatives
para-Chloro Furanyl fentanyl 3-furancarboxamide	Fentanyl Derivatives
para-Chloro Methoxyacetyl fentanyl	Fentanyl Derivatives
para-Chloro Valeryl fentanyl	Fentanyl Derivatives
para-Chlorobutyryl fentanyl	Fentanyl Derivatives
para-Chlorofentanyl	Fentanyl Derivatives
para-Chloroisobutyryl fentanyl	Fentanyl Derivatives
para-Fluoro Acrylfentanyl	Fentanyl Derivatives
para-Fluoro Crotonyl fentanyl	Fentanyl Derivatives
para-Fluoro Cyclopentyl fentanyl	Fentanyl Derivatives
para-Fluoro Cyclopropyl fentanyl	Fentanyl Derivatives
para-Fluoro Furanyl fentanyl	Fentanyl Derivatives
para-Fluoro Furanyl fentanyl 3-furancarboxamide isomer	Fentanyl Derivatives
para-Fluoro Methoxyacetyl fentanyl	Fentanyl Derivatives
para-Fluoro Tetrahydrofuran fentanyl	Fentanyl Derivatives
para-Fluoro Valeryl fentanyl	Fentanyl Derivatives
para-Fluoroacetyl fentanyl	Fentanyl Derivatives
para-Fluorobutyryl fentanyl	Fentanyl Derivatives
para-Fluoroisobutyryl fentanyl	Fentanyl Derivatives
para-Hydroxy Butyryl fentanyl	Fentanyl Derivatives
para-Methoxy Acrylfentanyl	Fentanyl Derivatives

para-Methoxy Furanyl fentanyl	Fentanyl Derivatives
para-Methoxy Methoxyacetyl fentanyl	Fentanyl Derivatives
para-Methoxy Valeryl fentanyl	Fentanyl Derivatives
para-Methoxy-Butyrylfentanyl	Fentanyl Derivatives
para-Methoxyfentanyl	Fentanyl Derivatives 2
para-Methyl Acetyl fentanyl	Fentanyl Derivatives
para-Methyl Acrylfentanyl	Fentanyl Derivatives
para-Methyl Cyclopentyl fentanyl	Fentanyl Derivatives
para-Methyl Cyclopropyl fentanyl	Fentanyl Derivatives
para-Methyl Furanyl fentanyl	Fentanyl Derivatives
para-Methylfentanyl	Fentanyl Derivatives
p-Fluorofentanyl	Fentanyl Derivatives
Phenoxyacetyl fentanyl	Fentanyl Derivatives
Phenyl fentanyl	Fentanyl Derivatives
Pivaloyl fentanyl	Fentanyl Derivatives
Senecioylfentanyl	Fentanyl Derivatives
Tetrahydrofuran fentanyl	Fentanyl Derivatives
Tetrahydrofuran fentanyl 3-tetrahydrofurancarboxamide isomer	Fentanyl Derivatives
Tetrahydrothiophene fentanyl	Fentanyl Derivatives
Thienyl fentanyl	Fentanyl Derivatives 2
Thiofentanyl	Fentanyl Derivatives
Thiofuranyl fentanyl	Fentanyl Derivatives
Thiophene fentanyl 3-thiophenecarboxamide	Fentanyl Derivatives
Tigloyl fentanyl	Fentanyl Derivatives
Trans-3-methyl Norfentanyl	Fentanyl Derivatives
trans-3-Methyl Thiofentanyl	Fentanyl Derivatives
Valeryl fentanyl	Fentanyl Derivatives
Morphine	Opioids
Desomorphine	Opioids
Dihydromorphine	Opioids
Meperidine	Opioids
Methadone	Opioids
Trimeperidine	Opioids
Normethadone	Opioids
Buprenorphine	Opioids
Heroin	Opioids
Codeine	Opioids
Acetylcodeine	Opioids
Ethylmorphine	Opioids
Hydrocodone	Opioids
Hydromorphone	Opioids
Oxycodone	Opioids
Oxymorphone	Opioids
Dihydrocodeine	Opioids
6.alpha.-Oxycodol	Opioids

Nalbuphine	Opioids
O-6-Monoacetylmorphine	Opioids
Butorphanol	Opioids
Levorphan	Opioids
Naloxone	Opioids
Nalorphine	Opioids
Naltrexone	Opioids
Loperamide	Opioids
Lefetamine	Opioids
Buphedrone	Cathinones
Butylone	Cathinones
3,4-Dimethylethcathinone	Cathinones
Dimethylone	Cathinones
4-Ethyl-N,N-dimethylcathinone	Cathinones
4-Ethylethcathinone	Cathinones
Ethylone	Cathinones
4-Fluoroethcathinone	Cathinones
4-Methylbuphedrone	Cathinones
Cathinone	Cathinones
Dibutylone	Cathinones
Eutylone	Cathinones
4-Fluoroisocathinone	Cathinones
Pentedrone	Cathinones
Hexedrone	Cathinones
Diethylpropion	Cathinones
Benzedrone	Cathinones
Mexedrone	Cathinones
Diethylone	Cathinones
Pentylone	Cathinones
Dipentylone	Cathinones
Pyrovalerone	Cathinones
Naphyrone	Cathinones
3,4-Methylenedioxy-PV8	Cathinones
3,4-Methylenedioxy-.alpha.-Pyrrolidinohexanophenone	Cathinones
3,4-Methylenedioxypyrovalerone	Cathinones
4-MeO-.alpha.-PVP	Cathinones
4-Fluoro-.alpha.-pyrrolidinopentiophenone	Cathinones
MePPP	Cathinones
.alpha.-PBP	Cathinones
.alpha.-PVP	Cathinones
.alpha.-Pyrrolidinopropiophenone	Cathinones
.alpha.-Piperidinobutiophenone	Cathinones
.alpha.-Phthalimidopropiophenone	Cathinones
4-Methylpentedrone	Cathinones
N,N-Dimethylpentylone	Cathinones

4-Methyl-.alpha.-ethylaminobutiophenone	Cathinones
3,4-Dimethylmethcathinone	Cathinones
4-Ethylmethcathinone	Cathinones
Flephedrone	Cathinones
Methcathinone	Cathinones
Mephedrone	Cathinones
Methedrone	Cathinones
Clephedrone	Cathinones
Brephedrone	Cathinones
Methylone	Cathinones
3,4-EDMC	Cathinones
3,4-Dichloromethcathinone	Cathinones

Table 10. List of Compounds and respective classes for Classical Tricyclic Cannabinoids versus synthetic cannabinoids

Name	Class
Cannabigerol	Classical Cannabinoids
Cannabicyclol	Classical Cannabinoids
Cannabichromene	Classical Cannabinoids
Cannabidiol	Classical Cannabinoids
Cannabinol	Classical Cannabinoids
delta-9-Tetrahydrocannabinol	Classical Cannabinoids
delta-8-Tetrahydrocannabinol	Classical Cannabinoids
Cannabidiolic Acid	Classical Cannabinoids
Tetrahydrocannabivarin	Classical Cannabinoids
Cannabidivarin	Classical Cannabinoids
Cannabicitran	Classical Cannabinoids
Cannabichromevarin	Classical Cannabinoids
Cannabichromeorcin	Classical Cannabinoids
Cannabigerorcinic Acid	Classical Cannabinoids
Cannabigerorcin	Classical Cannabinoids
Cannabigerovarin	Classical Cannabinoids
CB-25	Classical Cannabinoids
CB-52	Classical Cannabinoids
CP 47,497	Classical Cannabinoids
CP 47,497-C8 Homolog	Classical Cannabinoids
CP 55,940	Classical Cannabinoids
HU-210	Classical Cannabinoids
HU-211	Classical Cannabinoids
Nabilone	Classical Cannabinoids
Cannabipiperidiethanone	Synthetic Cannabinoids
AM1220	Synthetic Cannabinoids
AM1235	Synthetic Cannabinoids

AM1241	Synthetic Cannabinoids
AM1248	Synthetic Cannabinoids
AM2201	Synthetic Cannabinoids
AM2232	Synthetic Cannabinoids
AM2233	Synthetic Cannabinoids
AM679	Synthetic Cannabinoids
BB-22	Synthetic Cannabinoids
CB-13	Synthetic Cannabinoids
EG-018	Synthetic Cannabinoids
EG-2201	Synthetic Cannabinoids
AM2201 benzimidazole analog	Synthetic Cannabinoids
JWH-081-N-(cyclohexylmethyl) analog	Synthetic Cannabinoids
JWH-007	Synthetic Cannabinoids
JWH-011	Synthetic Cannabinoids
JWH-015	Synthetic Cannabinoids
JWH-016	Synthetic Cannabinoids
JWH-018	Synthetic Cannabinoids
AB-001	Synthetic Cannabinoids
APICA	Synthetic Cannabinoids
JWH-018 benzimidazole analog	Synthetic Cannabinoids
THJ-018	Synthetic Cannabinoids
JWH-019	Synthetic Cannabinoids
JWH-020	Synthetic Cannabinoids
JWH-022	Synthetic Cannabinoids
JWH-030	Synthetic Cannabinoids
JWH-031	Synthetic Cannabinoids
JWH-072	Synthetic Cannabinoids
JWH-073	Synthetic Cannabinoids
JWH-073	Synthetic Cannabinoids
JWH-073 2-methylnaphthyl analog	Synthetic Cannabinoids
JWH-073 4-methylnaphthyl analog	Synthetic Cannabinoids
JWH-081	Synthetic Cannabinoids
JWH-098	Synthetic Cannabinoids
JWH-122	Synthetic Cannabinoids
JWH-122 N-(4-pentenyl) analog	Synthetic Cannabinoids
JWH-145	Synthetic Cannabinoids
JWH-147	Synthetic Cannabinoids
JWH-175	Synthetic Cannabinoids
JWH-180	Synthetic Cannabinoids
JWH-182	Synthetic Cannabinoids
JWH-203	Synthetic Cannabinoids
JWH-210	Synthetic Cannabinoids
JWH-249	Synthetic Cannabinoids
JWH-167	Synthetic Cannabinoids
JWH-251	Synthetic Cannabinoids

JWH-307	Synthetic Cannabinoids
JWH-309	Synthetic Cannabinoids
JWH-368	Synthetic Cannabinoids
JWH-369	Synthetic Cannabinoids
JWH-370	Synthetic Cannabinoids
JWH-398	Synthetic Cannabinoids
JWH-424	Synthetic Cannabinoids
MAM2201	Synthetic Cannabinoids
JWH-200	Synthetic Cannabinoids
NM2201	Synthetic Cannabinoids
NNE1	Synthetic Cannabinoids
RCS-4	Synthetic Cannabinoids
RCS-4-C4 homolog	Synthetic Cannabinoids
RCS-8	Synthetic Cannabinoids
THJ2201	Synthetic Cannabinoids
XLR12	Synthetic Cannabinoids
5-Fluoro-MN-18	Synthetic Cannabinoids
MMB2201	Synthetic Cannabinoids
AB-PINACA N-(3-fluoropentyl) isomer	Synthetic Cannabinoids
5-Fluoro-ADBICA	Synthetic Cannabinoids
5-Fluoro-ADB	Synthetic Cannabinoids
5-Fluoro-ADB-PINACA isomer 2	Synthetic Cannabinoids
5-Fluoro-PCN	Synthetic Cannabinoids
5-Fluoro-AEB	Synthetic Cannabinoids
4-Fluoro-AMB	Synthetic Cannabinoids
5-Fluoro-CUMYL-PICA	Synthetic Cannabinoids
5-Fluoro-AB-001	Synthetic Cannabinoids
5-Fluoro-7-APAICA	Synthetic Cannabinoids
5-Fluoro MDMB-PICA metabolite 7	Synthetic Cannabinoids
5F-NNEI 2'-naphthyl isomer	Synthetic Cannabinoids
5-Fluoro-PB-22	Synthetic Cannabinoids
5-Fluoro-SDB-006	Synthetic Cannabinoids
5-Fluoro-THJ	Synthetic Cannabinoids
AB-CHMINACA	Synthetic Cannabinoids
AB-FUBINACA	Synthetic Cannabinoids
AB-PINACA	Synthetic Cannabinoids
ADB-FUBINACA	Synthetic Cannabinoids
ADBICA	Synthetic Cannabinoids
ADB-PINACA	Synthetic Cannabinoids
AKB48	Synthetic Cannabinoids
AM2201 8-quinolinyl carboxamide	Synthetic Cannabinoids
AM694	Synthetic Cannabinoids
AMB	Synthetic Cannabinoids
CUMYL-PICA	Synthetic Cannabinoids
FDU-PB-22	Synthetic Cannabinoids

XLR11	Synthetic Cannabinoids
FUB-144	Synthetic Cannabinoids
FUB-AMB	Synthetic Cannabinoids
FUB-PB-22	Synthetic Cannabinoids
ADB-CHMINACA	Synthetic Cannabinoids
MA-CHMINACA	Synthetic Cannabinoids
MDMB-CHMICA	Synthetic Cannabinoids
MDMB-CHMINACA	Synthetic Cannabinoids
MDMB-FUBINACA	Synthetic Cannabinoids
MN-18	Synthetic Cannabinoids
MO-CHMINACA	Synthetic Cannabinoids
PX 1	Synthetic Cannabinoids
PX 2	Synthetic Cannabinoids
SDB-005	Synthetic Cannabinoids
SDB-006	Synthetic Cannabinoids
THJ	Synthetic Cannabinoids
XLR11 N-(4-pentenyl) analog	Synthetic Cannabinoids
A-796,260	Synthetic Cannabinoids
A-834,735	Synthetic Cannabinoids
A-836,339	Synthetic Cannabinoids
UR-144	Synthetic Cannabinoids
Pravadoline	Synthetic Cannabinoids
PB-22	Synthetic Cannabinoids
STS-135	Synthetic Cannabinoids
UR-144 N-heptyl analog	Synthetic Cannabinoids
WIN 55,212-2	Synthetic Cannabinoids
4-Chloro MDMB-BUTICA	Synthetic Cannabinoids
MDMB-FUB7AICA	Synthetic Cannabinoids
MDMB-CHM7AICA	Synthetic Cannabinoids
MPP-PICA	Synthetic Cannabinoids
MMB-CHM7AICA	Synthetic Cannabinoids
5-Fluoro MMB-P7AICA	Synthetic Cannabinoids
5-Fluoro EDMB-PICA	Synthetic Cannabinoids
5-Fluoro EMB-PICA	Synthetic Cannabinoids
5-Fluoro AB-7-PAICA	Synthetic Cannabinoids
CUMYL-CBMICA	Synthetic Cannabinoids
5-Chloro MDMB-PICA	Synthetic Cannabinoids
Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Synthetic Cannabinoids
5-Chloro AMB-PICA	Synthetic Cannabinoids
5-Bromo AMB-PICA	Synthetic Cannabinoids
CUMYL-CBMINACA	Synthetic Cannabinoids
MDMB-BUTINACA	Synthetic Cannabinoids
ADB-4en-PINACA	Synthetic Cannabinoids
ADB-PHETINACA	Synthetic Cannabinoids
THQ-PINACA	Synthetic Cannabinoids

iPDMB-FUBINACA	Synthetic Cannabinoids
4-Cyano ADB-BUTINACA	Synthetic Cannabinoids
Benzyl-4-cyano BUTINACA	Synthetic Cannabinoids
4-Fluoro ABUTINACA	Synthetic Cannabinoids
5F-EDMB-PINACA	Synthetic Cannabinoids
ADB-BUTINACA	Synthetic Cannabinoids
MMB-4en-PINACA	Synthetic Cannabinoids
4-Cyano MDMB-BUTINACA	Synthetic Cannabinoids
MDMB-3en-BUTINACA	Synthetic Cannabinoids
4-Fluoro MDMB-BUTINACA	Synthetic Cannabinoids
CUMYL-PINACA	Synthetic Cannabinoids
ADB-BINACA	Synthetic Cannabinoids
CUMYL-THPINACA	Synthetic Cannabinoids
Adamantyl-THPINACA	Synthetic Cannabinoids

Table 11 . List of compounds with respective class in Azaindole/Indazole Versus Indole model

Name	Class
MDMB-FUB7AICA	Azaindole/Indazole
MDMB-CHM7AICA	Azaindole/Indazole
MMB-CHM7AICA	Azaindole/Indazole
5-Fluoro MMB-P7AICA	Azaindole/Indazole
5-Fluoro AB-7-PAICA	Azaindole/Indazole
5-Fluoro-PCN	Azaindole/Indazole
CUMYL-CBMINACA	Azaindole/Indazole
ADB-FUBINACA	Azaindole/Indazole
5-Fluoro-THJ	Azaindole/Indazole
MN-18	Azaindole/Indazole
5-Fluoro-AEB	Azaindole/Indazole
AB-PINACA	Azaindole/Indazole
SDB-005	Azaindole/Indazole
AB-CHMINACA	Azaindole/Indazole
AB-FUBINACA	Azaindole/Indazole
MA-CHMINACA	Azaindole/Indazole
5-Fluoro-7-APAICA	Azaindole/Indazole
MO-CHMINACA	Azaindole/Indazole
4-Fluoro-AMB	Azaindole/Indazole
5-Fluoro-ADB	Azaindole/Indazole
5-Fluoro-ADB-PINACA	Azaindole/Indazole
5-Fluoro-ADB-PINACA isomer 2	Azaindole/Indazole
5-Fluoro-MN-18	Azaindole/Indazole
AB-PINACA N-(3-fluoropentyl) isomer	Azaindole/Indazole
ADB-CHMINACA	Azaindole/Indazole
ADB-PINACA	Azaindole/Indazole

AKB48	Azaindole/Indazole
AMB	Azaindole/Indazole
FUB-AMB	Azaindole/Indazole
MDMB-CHMINACA	Azaindole/Indazole
MDMB-FUBINACA	Azaindole/Indazole
PX 2	Azaindole/Indazole
THJ	Azaindole/Indazole
THJ-018	Azaindole/Indazole
THJ2201	Azaindole/Indazole
ADB-BINACA	Azaindole/Indazole
CUMYL-THPINACA	Azaindole/Indazole
Adamantyl-THPINACA	Azaindole/Indazole
4-Fluoro MDMB-BUTINACA	Azaindole/Indazole
5-Fluoro MDMB-PICA metabolite 7	Indole
5-Fluoro-PB-22	Indole
5-Fluoro-SDB-006	Indole
5F-NNEI 2'-naphthyl isomer	Indole
AB-001	Indole
ADBICA	Indole
AM1220	Indole
AM1241	Indole
AM1248	Indole
AM2201	Indole
AM2201 8-quinolinyl carboxamide	Indole
AM2232	Indole
AM2233	Indole
AM679	Indole
AM694	Indole
APICA	Indole
BB-22	Indole
CUMYL-PICA	Indole
FDU-PB-22	Indole
FUB-144	Indole
FUB-PB-22	Indole
JWH-007	Indole
JWH-011	Indole
JWH-015	Indole
JWH-016	Indole
JWH-018	Indole
JWH-019	Indole
JWH-020	Indole
JWH-072	Indole
JWH-073	Indole
JWH-073	Indole
JWH-073 2-methylnaphthyl analog	Indole

JWH-073 4-methylnaphthyl analog	Indole
JWH-081	Indole
JWH-081-N-(cyclohexylmethyl) analog	Indole
JWH-098	Indole
JWH-122	Indole
JWH-122 N-(4-pentenyl) analog	Indole
JWH-167	Indole
JWH-175	Indole
JWH-180	Indole
JWH-182	Indole
JWH-200	Indole
JWH-203	Indole
JWH-210	Indole
JWH-249	Indole
JWH-251	Indole
JWH-398	Indole
JWH-424	Indole
MAM2201	Indole
MDMB-CHMICA	Indole
MMB2201	Indole
NM2201	Indole
NNE1	Indole
PB-22	Indole
Pravadoline	Indole
PX 1	Indole
RCS-4	Indole
RCS-4-C4 homolog	Indole
RCS-8	Indole
SDB-006	Indole
STS-135	Indole
UR-144	Indole
UR-144 N-heptyl analog	Indole
XLR11	Indole
XLR11 N-(4-pentenyl) analog	Indole
XLR12	Indole
5F-AMBICA	Indole
5-Fluoro-ADBICA	Indole
5-Fluoro-AB-001	Indole
5-Fluoro-CUMYL-PICA	Indole
4-Chloro MDMB-BUTICA	Indole
MPP-PICA	Indole
5-Fluoro EDMB-PICA	Indole
5-Fluoro EMB-PICA	Indole
CUMYL-CBMICA	Indole
5-Chloro MDMB-PICA	Indole

Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Indole
5-Chloro AMB-PICA	Indole
5-Bromo AMB-PICA	Indole
JWH-145	Indole
JWH-147	Indole
JWH-307	Indole
JWH-309	Indole
JWH-368	Indole
JWH-369	Indole
JWH-370	Indole
AM1235	Indole
JWH-022	Indole
JWH-030	Indole
JWH-031	Indole

Table 12. List of compounds with respective class in Azaindole/Indazole/indole model

Name	Class
MDMB-FUB7AICA	Azaindole/Indazole/Indole
MDMB-CHM7AICA	Azaindole/Indazole/Indole
MMB-CHM7AICA	Azaindole/Indazole/Indole
5-Fluoro MMB-P7AICA	Azaindole/Indazole/Indole
5-Fluoro AB-7-PAICA	Azaindole/Indazole/Indole
5-Fluoro-PCN	Azaindole/Indazole/Indole
CUMYL-CBMINACA	Azaindole/Indazole/Indole
ADB-FUBINACA	Azaindole/Indazole/Indole
5-Fluoro-THJ	Azaindole/Indazole/Indole
MN-18	Azaindole/Indazole/Indole
5-Fluoro-AEB	Azaindole/Indazole/Indole
AB-PINACA	Azaindole/Indazole/Indole
SDB-005	Azaindole/Indazole/Indole
AB-CHMINACA	Azaindole/Indazole/Indole
AB-FUBINACA	Azaindole/Indazole/Indole
MA-CHMINACA	Azaindole/Indazole/Indole
5-Fluoro-7-APAICA	Azaindole/Indazole/Indole
MO-CHMINACA	Azaindole/Indazole/Indole
4-Fluoro-AMB	Azaindole/Indazole/Indole
5-Fluoro-ADB	Azaindole/Indazole/Indole
5-Fluoro-ADB-PINACA	Azaindole/Indazole/Indole
5-Fluoro-ADB-PINACA isomer 2	Azaindole/Indazole/Indole
5-Fluoro-MN-18	Azaindole/Indazole/Indole
AB-PINACA N-(3-fluoropentyl) isomer	Azaindole/Indazole/Indole
ADB-CHMINACA	Azaindole/Indazole/Indole
ADB-PINACA	Azaindole/Indazole/Indole

AKB48	Azaindole/Indazole/Indole
AMB	Azaindole/Indazole/Indole
FUB-AMB	Azaindole/Indazole/Indole
MDMB-CHMINACA	Azaindole/Indazole/Indole
MDMB-FUBINACA	Azaindole/Indazole/Indole
PX 2	Azaindole/Indazole/Indole
THJ	Azaindole/Indazole/Indole
THJ-018	Azaindole/Indazole/Indole
THJ2201	Azaindole/Indazole/Indole
ADB-BINACA	Azaindole/Indazole/Indole
CUMYL-THPINACA	Azaindole/Indazole/Indole
Adamantyl-THPINACA	Azaindole/Indazole/Indole
4-Fluoro MDMB-BUTINACA	Azaindole/Indazole/Indole
5-Fluoro MDMB-PICA metabolite 7	Azaindole/Indazole/Indole
5-Fluoro-PB-22	Azaindole/Indazole/Indole
5-Fluoro-SDB-006	Azaindole/Indazole/Indole
5F-NNEI 2'-naphthyl isomer	Azaindole/Indazole/Indole
AB-001	Azaindole/Indazole/Indole
ADBICA	Azaindole/Indazole/Indole
AM1220	Azaindole/Indazole/Indole
AM1241	Azaindole/Indazole/Indole
AM1248	Azaindole/Indazole/Indole
AM2201	Azaindole/Indazole/Indole
AM2201 8-quinolinyl carboxamide	Azaindole/Indazole/Indole
AM2232	Azaindole/Indazole/Indole
AM2233	Azaindole/Indazole/Indole
AM679	Azaindole/Indazole/Indole
AM694	Azaindole/Indazole/Indole
APICA	Azaindole/Indazole/Indole
BB-22	Azaindole/Indazole/Indole
CUMYL-PICA	Azaindole/Indazole/Indole
FDU-PB-22	Azaindole/Indazole/Indole
FUB-144	Azaindole/Indazole/Indole
FUB-PB-22	Azaindole/Indazole/Indole
JWH-007	Azaindole/Indazole/Indole
JWH-011	Azaindole/Indazole/Indole
JWH-015	Azaindole/Indazole/Indole
JWH-016	Azaindole/Indazole/Indole
JWH-018	Azaindole/Indazole/Indole
JWH-019	Azaindole/Indazole/Indole
JWH-020	Azaindole/Indazole/Indole
JWH-072	Azaindole/Indazole/Indole
JWH-073	Azaindole/Indazole/Indole
JWH-073	Azaindole/Indazole/Indole
JWH-073 2-methylnaphthyl analog	Azaindole/Indazole/Indole

JWH-073 4-methylnaphthyl analog	Azaindole/Indazole/Indole
JWH-081	Azaindole/Indazole/Indole
JWH-081-N-(cyclohexylmethyl) analog	Azaindole/Indazole/Indole
JWH-098	Azaindole/Indazole/Indole
JWH-122	Azaindole/Indazole/Indole
JWH-122 N-(4-pentenyl) analog	Azaindole/Indazole/Indole
JWH-167	Azaindole/Indazole/Indole
JWH-175	Azaindole/Indazole/Indole
JWH-180	Azaindole/Indazole/Indole
JWH-182	Azaindole/Indazole/Indole
JWH-200	Azaindole/Indazole/Indole
JWH-203	Azaindole/Indazole/Indole
JWH-210	Azaindole/Indazole/Indole
JWH-249	Azaindole/Indazole/Indole
JWH-251	Azaindole/Indazole/Indole
JWH-398	Azaindole/Indazole/Indole
JWH-424	Azaindole/Indazole/Indole
MAM2201	Azaindole/Indazole/Indole
MDMB-CHMICA	Azaindole/Indazole/Indole
MMB2201	Azaindole/Indazole/Indole
NM2201	Azaindole/Indazole/Indole
NNE1	Azaindole/Indazole/Indole
PB-22	Azaindole/Indazole/Indole
Pravadoline	Azaindole/Indazole/Indole
PX 1	Azaindole/Indazole/Indole
RCS-4	Azaindole/Indazole/Indole
RCS-4-C4 homolog	Azaindole/Indazole/Indole
RCS-8	Azaindole/Indazole/Indole
SDB-006	Azaindole/Indazole/Indole
STS-135	Azaindole/Indazole/Indole
UR-144	Azaindole/Indazole/Indole
UR-144 N-heptyl analog	Azaindole/Indazole/Indole
XLR11	Azaindole/Indazole/Indole
XLR11 N-(4-pentenyl) analog	Azaindole/Indazole/Indole
XLR12	Azaindole/Indazole/Indole
5F-AMBICA	Azaindole/Indazole/Indole
5-Fluoro-ADBICA	Azaindole/Indazole/Indole
5-Fluoro-AB-001	Azaindole/Indazole/Indole
5-Fluoro-CUMYL-PICA	Azaindole/Indazole/Indole
4-Chloro MDMB-BUTICA	Azaindole/Indazole/Indole
MPP-PICA	Azaindole/Indazole/Indole
5-Fluoro EDMB-PICA	Azaindole/Indazole/Indole
5-Fluoro EMB-PICA	Azaindole/Indazole/Indole
CUMYL-CBMICA	Azaindole/Indazole/Indole
5-Chloro MDMB-PICA	Azaindole/Indazole/Indole

Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Azaindole/Indazole/Indole
5-Chloro AMB-PICA	Azaindole/Indazole/Indole
5-Bromo AMB-PICA	Azaindole/Indazole/Indole
JWH-145	Naphthoylpyrrole
JWH-147	Naphthoylpyrrole
JWH-307	Naphthoylpyrrole
JWH-309	Naphthoylpyrrole
JWH-368	Naphthoylpyrrole
JWH-369	Naphthoylpyrrole
JWH-370	Naphthoylpyrrole
AM1235	Naphthoylpyrrole
JWH-022	Naphthoylpyrrole
JWH-030	Naphthoylpyrrole
JWH-031	Naphthoylpyrrole

Table13. List of compounds with respective classes of Naphthyl versus other head groups model

Name	Class
5F-NNEI 2'-naphthyl isomer	Naphthyl
AM1220	Naphthyl
AM1235	Naphthyl
AM2201	Naphthyl
AM2232	Naphthyl
JWH-007	Naphthyl
JWH-011	Naphthyl
JWH-015	Naphthyl
JWH-016	Naphthyl
JWH-018	Naphthyl
JWH-019	Naphthyl
JWH-020	Naphthyl
JWH-022	Naphthyl
JWH-030	Naphthyl
JWH-031	Naphthyl
JWH-072	Naphthyl
JWH-073	Naphthyl
JWH-073	Naphthyl
JWH-073 2-methylnaphthyl analog	Naphthyl
JWH-073 4-methylnaphthyl analog	Naphthyl
JWH-081	Naphthyl
JWH-081-N-(cyclohexylmethyl) analog	Naphthyl
JWH-098	Naphthyl
JWH-122	Naphthyl
JWH-122 N-(4-pentenyl) analog	Naphthyl

JWH-145	Naphthyl
JWH-147	Naphthyl
JWH-175	Naphthyl
JWH-180	Naphthyl
JWH-182	Naphthyl
JWH-200	Naphthyl
JWH-210	Naphthyl
JWH-307	Naphthyl
JWH-309	Naphthyl
JWH-368	Naphthyl
JWH-369	Naphthyl
JWH-370	Naphthyl
JWH-398	Naphthyl
JWH-424	Naphthyl
MAM2201	Naphthyl
NM2201	Naphthyl
MN-18	Naphthyl
THJ-018	Naphthyl
THJ2201	Naphthyl
THJ	Other Head Groups
5-Fluoro MDMB-PICA metabolite 7	Other Head Groups
5-Fluoro-PB-22	Other Head Groups
5-Fluoro-SDB-006	Other Head Groups
AB-001	Other Head Groups
ADBICA	Other Head Groups
AM1241	Other Head Groups
AM1248	Other Head Groups
AM2201 8-quinolinyl carboxamide	Other Head Groups
AM2233	Other Head Groups
AM679	Other Head Groups
AM694	Other Head Groups
APICA	Other Head Groups
BB-22	Other Head Groups
CUMYL-PICA	Other Head Groups
FDU-PB-22	Other Head Groups
FUB-144	Other Head Groups
FUB-PB-22	Other Head Groups
JWH-167	Other Head Groups
JWH-203	Other Head Groups
JWH-249	Other Head Groups
JWH-251	Other Head Groups
MDMB-CHMICA	Other Head Groups
MMB2201	Other Head Groups
NNE1	Other Head Groups
PB-22	Other Head Groups

Pravadoline	Other Head Groups
PX 1	Other Head Groups
RCS-4	Other Head Groups
RCS-4-C4 homolog	Other Head Groups
RCS-8	Other Head Groups
SDB-006	Other Head Groups
STS-135	Other Head Groups
UR-144	Other Head Groups
UR-144 N-heptyl analog	Other Head Groups
XLR11	Other Head Groups
XLR11 N-(4-pentenyl) analog	Other Head Groups
XLR12	Other Head Groups
5F-AMBICA	Other Head Groups
5-Fluoro-ADBICA	Other Head Groups
5-Fluoro-AB-001	Other Head Groups
5-Fluoro-CUMYL-PICA	Other Head Groups
4-Chloro MDMB-BUTICA	Other Head Groups
MDMB-FUB7AICA	Other Head Groups
MDMB-CHM7AICA	Other Head Groups
MPP-PICA	Other Head Groups
MMB-CHM7AICA	Other Head Groups
5-Fluoro MMB-P7AICA	Other Head Groups
5-Fluoro EDMB-PICA	Other Head Groups
5-Fluoro EMB-PICA	Other Head Groups
5-Fluoro AB-7-PAICA	Other Head Groups
CUMYL-CBMICA	Other Head Groups
5-Chloro MDMB-PICA	Other Head Groups
Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Other Head Groups
5-Chloro AMB-PICA	Other Head Groups
5-Bromo AMB-PICA	Other Head Groups
5-Fluoro-PCN	Other Head Groups
CUMYL-CBMINACA	Other Head Groups
ADB-FUBINACA	Other Head Groups
5-Fluoro-THJ	Other Head Groups
5-Fluoro-AEB	Other Head Groups
AB-PINACA	Other Head Groups
SDB-005	Other Head Groups
4-Fluoro MDMB-BUTINACA	Other Head Groups
AB-CHMINACA	Other Head Groups
AB-FUBINACA	Other Head Groups
MA-CHMINACA	Other Head Groups
5-Fluoro-7-APAICA	Other Head Groups
MO-CHMINACA	Other Head Groups
4-Fluoro-AMB	Other Head Groups
5-Fluoro-ADB	Other Head Groups

5-Fluoro-ADB-PINACA	Other Head Groups
5-Fluoro-ADB-PINACA isomer 2	Other Head Groups
5-Fluoro-MN-18	Other Head Groups
AB-PINACA N-(3-fluoropentyl) isomer	Other Head Groups
ADB-CHMINACA	Other Head Groups
ADB-PINACA	Other Head Groups
AKB48	Other Head Groups
AMB	Other Head Groups
FUB-AMB	Other Head Groups
MDMB-CHMINACA	Other Head Groups
MDMB-FUBINACA	Other Head Groups
PX 2	Other Head Groups
CUMYL-PINACA	Other Head Groups
ADB-BINACA	Other Head Groups
CUMYL-THPINACA	Other Head Groups
Adamantyl-THPINACA	Other Head Groups
MDMB-BUTINACA	Other Head Groups
ADB-4en-PINACA	Other Head Groups
ADB-PHETINACA	Other Head Groups
THQ-PINACA	Other Head Groups
iPDMB-FUBINACA	Other Head Groups
4-Cyano ADB-BUTINACA	Other Head Groups
Benzyl-4-cyano BUTINACA	Other Head Groups
4-Fluoro ABUTINACA	Other Head Groups
5F-EDMB-PINACA	Other Head Groups
ADB-BUTINACA	Other Head Groups
MMB-4en-PINACA	Other Head Groups
4-Cyano MDMB-BUTINACA	Other Head Groups
MDMB-3en-BUTINACA	Other Head Groups
4-Fluoro MDMB-BUTINACA	Other Head Groups

Table 14. List of compounds with respective classes of FUB versus other Tail groups model

Name	Class
FDU-PB-22	FUB
FUB-144	FUB
FUB-PB-22	FUB
MDMB-FUB7AICA	FUB
ADB-FUBINACA	FUB
AB-FUBINACA	FUB
FUB-AMB	FUB
MDMB-FUBINACA	FUB
iPDMB-FUBINACA	FUB

5-Fluoro MDMB-PICA metabolite 7	Other Tail Groups
5-Fluoro-PB-22	Other Tail Groups
5-Fluoro-SDB-006	Other Tail Groups
5F-NNEI 2'-naphthyl isomer	Other Tail Groups
AB-001	Other Tail Groups
ADBICA	Other Tail Groups
AM1220	Other Tail Groups
AM1235	Other Tail Groups
AM1241	Other Tail Groups
AM1248	Other Tail Groups
AM2201	Other Tail Groups
AM2201 8-quinolinyl carboxamide	Other Tail Groups
AM2232	Other Tail Groups
AM2233	Other Tail Groups
AM679	Other Tail Groups
AM694	Other Tail Groups
APICA	Other Tail Groups
BB-22	Other Tail Groups
CUMYL-PICA	Other Tail Groups
JWH-007	Other Tail Groups
JWH-011	Other Tail Groups
JWH-015	Other Tail Groups
JWH-016	Other Tail Groups
JWH-018	Other Tail Groups
JWH-019	Other Tail Groups
JWH-020	Other Tail Groups
JWH-022	Other Tail Groups
JWH-030	Other Tail Groups
JWH-031	Other Tail Groups
JWH-072	Other Tail Groups
JWH-073	Other Tail Groups
JWH-073	Other Tail Groups
JWH-073 2-methylnaphthyl analog	Other Tail Groups
JWH-073 4-methylnaphthyl analog	Other Tail Groups
JWH-081	Other Tail Groups
JWH-081-N-(cyclohexylmethyl) analog	Other Tail Groups
JWH-098	Other Tail Groups
JWH-122	Other Tail Groups
JWH-122 N-(4-pentenyl) analog	Other Tail Groups
JWH-145	Other Tail Groups
JWH-147	Other Tail Groups
JWH-167	Other Tail Groups
JWH-175	Other Tail Groups
JWH-180	Other Tail Groups
JWH-182	Other Tail Groups

JWH-200	Other Tail Groups
JWH-203	Other Tail Groups
JWH-210	Other Tail Groups
JWH-249	Other Tail Groups
JWH-251	Other Tail Groups
JWH-307	Other Tail Groups
JWH-309	Other Tail Groups
JWH-368	Other Tail Groups
JWH-369	Other Tail Groups
JWH-370	Other Tail Groups
JWH-398	Other Tail Groups
JWH-424	Other Tail Groups
MAM2201	Other Tail Groups
MDMB-CHMICA	Other Tail Groups
MMB2201	Other Tail Groups
NM2201	Other Tail Groups
NNE1	Other Tail Groups
PB-22	Other Tail Groups
Pravadoline	Other Tail Groups
PX 1	Other Tail Groups
RCS-4	Other Tail Groups
RCS-4-C4 homolog	Other Tail Groups
RCS-8	Other Tail Groups
SDB-006	Other Tail Groups
STS-135	Other Tail Groups
UR-144	Other Tail Groups
UR-144 N-heptyl analog	Other Tail Groups
XLR11	Other Tail Groups
XLR11 N-(4-pentenyl) analog	Other Tail Groups
XLR12	Other Tail Groups
5F-AMBICA	Other Tail Groups
5-Fluoro-ADBICA	Other Tail Groups
5-Fluoro-AB-001	Other Tail Groups
5-Fluoro-CUMYL-PICA	Other Tail Groups
4-Chloro MDMB-BUTICA	Other Tail Groups
MDMB-CHM7AICA	Other Tail Groups
MPP-PICA	Other Tail Groups
MMB-CHM7AICA	Other Tail Groups
5-Fluoro MMB-P7AICA	Other Tail Groups
5-Fluoro EDMB-PICA	Other Tail Groups
5-Fluoro EMB-PICA	Other Tail Groups
5-Fluoro AB-7-PAICA	Other Tail Groups
CUMYL-CBMICA	Other Tail Groups
5-Chloro MDMB-PICA	Other Tail Groups
Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Other Tail Groups

5-Chloro AMB-PICA	Other Tail Groups
Indole	Other Tail Groups
5-Bromo AMB-PICA	Other Tail Groups
5-Fluoro-PCN	Other Tail Groups
CUMYL-CBMINACA	Other Tail Groups
5-Fluoro-THJ	Other Tail Groups
MN-18	Other Tail Groups
5-Fluoro-AEB	Other Tail Groups
AB-PINACA	Other Tail Groups
SDB-005	Other Tail Groups
4-Fluoro MDMB-BUTINACA	Other Tail Groups
AB-CHMINACA	Other Tail Groups
MA-CHMINACA	Other Tail Groups
5-Fluoro-7-APAICA	Other Tail Groups
MO-CHMINACA	Other Tail Groups
4-Fluoro-AMB	Other Tail Groups
5-Fluoro-ADB	Other Tail Groups
5-Fluoro-ADB-PINACA	Other Tail Groups
5-Fluoro-ADB-PINACA isomer 2	Other Tail Groups
5-Fluoro-MN-18	Other Tail Groups
AB-PINACA N-(3-fluoropentyl) isomer	Other Tail Groups
ADB-CHMINACA	Other Tail Groups
ADB-PINACA	Other Tail Groups
AKB48	Other Tail Groups
AMB	Other Tail Groups
MDMB-CHMINACA	Other Tail Groups
PX 2	Other Tail Groups
THJ	Other Tail Groups
THJ-018	Other Tail Groups
THJ2201	Other Tail Groups
CUMYL-PINACA	Other Tail Groups
ADB-BINACA	Other Tail Groups
CUMYL-THPINACA	Other Tail Groups
Adamantyl-THPINACA	Other Tail Groups
MDMB-BUTINACA	Other Tail Groups
ADB-4en-PINACA	Other Tail Groups
ADB-PHETINACA	Other Tail Groups
THQ-PINACA	Other Tail Groups
4-Cyano ADB-BUTINACA	Other Tail Groups
Benzyl-4-cyano BUTINACA	Other Tail Groups
4-Fluoro ABUTINACA	Other Tail Groups
5F-EDMB-PINACA	Other Tail Groups
ADB-BUTINACA	Other Tail Groups
MMB-4en-PINACA	Other Tail Groups
4-Cyano MDMB-BUTINACA	Other Tail Groups

MDMB-3en-BUTINACA	Other Tail Groups
4-Fluoro MDMB-BUTINACA	Other Tail Groups

Table 15. List of compounds with respective classes of BUT versus other Tail groups model

Name	Class
4-Fluoro MDMB-BUTINACA	BUT
4-Fluoro ABUTINACA	BUT
4-Fluoro MDMB-BUTINACA	BUT
MDMB-BUTINACA	BUT
4-Cyano ADB-BUTINACA	BUT
Benzyl-4-cyano BUTINACA	BUT
ADB-BUTINACA	BUT
4-Cyano MDMB-BUTINACA	BUT
MDMB-3en-BUTINACA	BUT
FDU-PB-22	Other Tail Groups
FUB-144	Other Tail Groups
FUB-PB-22	Other Tail Groups
MDMB-FUB7AICA	Other Tail Groups
ADB-FUBINACA	Other Tail Groups
AB-FUBINACA	Other Tail Groups
FUB-AMB	Other Tail Groups
MDMB-FUBINACA	Other Tail Groups
iPDMB-FUBINACA	Other Tail Groups
5-Fluoro MDMB-PICA metabolite 7	Other Tail Groups
5-Fluoro-PB-22	Other Tail Groups
5-Fluoro-SDB-006	Other Tail Groups
5F-NNEI 2'-naphthyl isomer	Other Tail Groups
5F-AMBICA	Other Tail Groups
5-Fluoro-ADBICA	Other Tail Groups
5-Fluoro-AB-001	Other Tail Groups
5-Fluoro-CUMYL-PICA	Other Tail Groups
5-Fluoro MMB-P7AICA	Other Tail Groups
5-Fluoro EDMB-PICA	Other Tail Groups
5-Fluoro EMB-PICA	Other Tail Groups
5-Fluoro AB-7-PAICA	Other Tail Groups
5-Fluoro-PCN	Other Tail Groups
5-Fluoro-THJ	Other Tail Groups
5-Fluoro-AEB	Other Tail Groups
5-Fluoro-7-APAICA	Other Tail Groups
4-Fluoro-AMB	Other Tail Groups
5-Fluoro-ADB	Other Tail Groups
5-Fluoro-ADB-PINACA	Other Tail Groups

5-Fluoro-ADB-PINACA isomer 2	Other Tail Groups
5-Fluoro-MN-18	Other Tail Groups
5F-EDMB-PINACA	Other Tail Groups
AB-001	Other Tail Groups
ADBICA	Other Tail Groups
AM1220	Other Tail Groups
AM1235	Other Tail Groups
AM1241	Other Tail Groups
AM1248	Other Tail Groups
AM2201	Other Tail Groups
AM2201 8-quinolinyl carboxamide	Other Tail Groups
AM2232	Other Tail Groups
AM2233	Other Tail Groups
AM679	Other Tail Groups
AM694	Other Tail Groups
APICA	Other Tail Groups
BB-22	Other Tail Groups
CUMYL-PICA	Other Tail Groups
JWH-007	Other Tail Groups
JWH-011	Other Tail Groups
JWH-015	Other Tail Groups
JWH-016	Other Tail Groups
JWH-018	Other Tail Groups
JWH-019	Other Tail Groups
JWH-020	Other Tail Groups
JWH-022	Other Tail Groups
JWH-030	Other Tail Groups
JWH-031	Other Tail Groups
JWH-072	Other Tail Groups
JWH-073	Other Tail Groups
JWH-073	Other Tail Groups
JWH-073 2-methylnaphthyl analog	Other Tail Groups
JWH-073 4-methylnaphthyl analog	Other Tail Groups
JWH-081	Other Tail Groups
JWH-081-N-(cyclohexylmethyl) analog	Other Tail Groups
JWH-098	Other Tail Groups
JWH-122	Other Tail Groups
JWH-122 N-(4-pentenyl) analog	Other Tail Groups
JWH-145	Other Tail Groups
JWH-147	Other Tail Groups
JWH-167	Other Tail Groups
JWH-175	Other Tail Groups
JWH-180	Other Tail Groups
JWH-182	Other Tail Groups
JWH-200	Other Tail Groups

JWH-203	Other Tail Groups
JWH-210	Other Tail Groups
JWH-249	Other Tail Groups
JWH-251	Other Tail Groups
JWH-307	Other Tail Groups
JWH-309	Other Tail Groups
JWH-368	Other Tail Groups
JWH-369	Other Tail Groups
JWH-370	Other Tail Groups
JWH-398	Other Tail Groups
JWH-424	Other Tail Groups
MAM2201	Other Tail Groups
MDMB-CHMICA	Other Tail Groups
MMB2201	Other Tail Groups
NM2201	Other Tail Groups
NNE1	Other Tail Groups
PB-22	Other Tail Groups
Pravadoline	Other Tail Groups
PX 1	Other Tail Groups
RCS-4	Other Tail Groups
RCS-4-C4 homolog	Other Tail Groups
RCS-8	Other Tail Groups
SDB-006	Other Tail Groups
STS-135	Other Tail Groups
UR-144	Other Tail Groups
UR-144 N-heptyl analog	Other Tail Groups
XLR11	Other Tail Groups
XLR11 N-(4-pentenyl) analog	Other Tail Groups
XLR12	Other Tail Groups
4-Chloro MDMB-BUTICA	Other Tail Groups
MDMB-CHM7AICA	Other Tail Groups
MPP-PICA	Other Tail Groups
MMB-CHM7AICA	Other Tail Groups
CUMYL-CBMICA	Other Tail Groups
5-Chloro MDMB-PICA	Other Tail Groups
Methyl (S)-2-(1H-indole-3-carboxamido)-3,3-dimethylbutanoate	Other Tail Groups
5-Chloro AMB-PICA	Other Tail Groups
Indole	Other Tail Groups
5-Bromo AMB-PICA	Other Tail Groups
CUMYL-CBMINACA	Other Tail Groups
MN-18	Other Tail Groups
AB-PINACA	Other Tail Groups
SDB-005	Other Tail Groups
AB-CHMINACA	Other Tail Groups
MA-CHMINACA	Other Tail Groups

MO-CHMINACA	Other Tail Groups
AB-PINACA N-(3-fluoropentyl) isomer	Other Tail Groups
ADB-CHMINACA	Other Tail Groups
ADB-PINACA	Other Tail Groups
AKB48	Other Tail Groups
AMB	Other Tail Groups
MDMB-CHMINACA	Other Tail Groups
PX 2	Other Tail Groups
THJ	Other Tail Groups
THJ-018	Other Tail Groups
THJ2201	Other Tail Groups
CUMYL-PINACA	Other Tail Groups
ADB-BINACA	Other Tail Groups
CUMYL-THPINACA	Other Tail Groups
Adamantyl-THPINACA	Other Tail Groups
ADB-4en-PINACA	Other Tail Groups
ADB-PHETINACA	Other Tail Groups
THQ-PINACA	Other Tail Groups
MMB-4en-PINACA	Other Tail Groups