Talanta 179 (2018) 546–553

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Identification and classification of cathinone unknowns by statistical analysis processing of direct analysis in real time-high resolution mass spectrometry-derived "neutral loss" spectra

Kristen L. Fowble, Jason R.E. Shepard, Rabi A. Musah*

Department of Chemistry, University at Albany-State University of New York, 1400 Washington Ave, Albany, NY 12222, United States

ARTICLE INFO

Keywords:
Direct analysis in real time-high resolution mass spectrometry
Neutral losses
Principle component analysis
Hierarchical clustering analysis
Dendrogram
Ethcathinones, methcathinones, methylenedioxy-containing cathinones, pyrrolidine-containing cathinones, and buphedrones

ABSTRACT

An approach to the rapid determination of the structures of novel synthetic cathinone designer drugs, also known as bath salts, is reported. While cathinones fragment so extensively by electron impact mass spectrometry that their mass spectra often cannot be used to identify the structure, collision-induced dissociation (CID) direct analysis in real time-high resolution mass spectrometry (DART-HRMS) experiments furnished spectra that provided diagnostic fragmentation patterns for the analyzed cathinones. From this data, neutral loss spectra, which reflect the presence of specific chemical moieties, could be acquired. These spectra showed striking similarities between cathinones sharing structural features such as pyrrolidine rings and methylenedioxy moieties. Principle component analysis (PCA) of the neutral loss spectra of nine synthetic cathinones of various types including ethcathinones, those containing a methylenedioxy moiety appended to the benzene ring, and pyrrolidine-containing structures, illustrated that cathinones falling within the same class clustered together and could be distinguished from those of other classes. Furthermore, hierarchical clustering analysis of the neutral loss data of a model set derived from 44 synthetic cathinones, furnished a dendrogram in which structurally similar cathinones clustered together. The ability of this model system to facilitate structure determination was tested using 4-fluoroethcathinone, 3,4-methylenedioxy-α-pyrrolidinohexanophenone (MDPHP), and ethylone, which fall into the ethcathinone, pyrrolidine-containing, and methylenedioxy-containing subclasses respectively. The results showed that their neutral loss spectra correctly fell within the ethcathinone, pyrrolidine-containing and methylenedioxy-containing cathinone clades of the dendrogram, and that the neutral loss information could be used to infer the structures of these compounds. The analysis and data processing steps are rapid and samples can be analyzed in their native form without any sample processing steps. The robustness of the dendrogram dataset can be readily increased by continued addition of newly discovered structures. The approach can be broadly applied to structure determination of unknowns, and would be particularly useful for analyses where sample amounts are limited.

1. Introduction

The United Nations Office on Drugs and Crime (UNODC) has categorized designer hallucinogenic drugs or mixtures that are not currently controlled under previous drug law conventions as "new psychoactive substances" (NPS) [1]. NPS include synthetic cathinones and their structural variants, also known as "bath salts". These designer drugs are readily available for purchase on the internet and are often labelled as "not for human consumption", a designation which not only belies their intended purpose, but shields them from FDA scrutiny. The emergence of synthetic cathinones has imposed unique challenges on crime labs, primarily because law enforcement agencies have been unable to keep abreast of the rapid influx of novel structural variants that appear on the market within days to weeks after earlier generations of new cathinone structures have been scheduled [2–7]. Although the UNODC has the ability to schedule and ban newly discovered psychoactive materials, the process of presenting a compelling case for scheduling can be lengthy, lasting from months to years. A major contributor to this bottleneck is the often necessary precondition that the structure of the unknown NPS be identified. This has proven to be challenging due to the rapid emergence of structural variants and the significant time, equipment and human resources required for full structural characterization [6,7].

Further complicating matters, novel cathinones rarely appear on the

* Corresponding author.
E-mail address: rmusah@albany.edu (R.A. Musah).

https://doi.org/10.1016/j.talanta.2017.11.020
Received 27 July 2017; Received in revised form 13 November 2017; Accepted 15 November 2017
Available online 16 November 2017
0039-9140/

market as purified compounds, but rather as mixtures comprised of cutting agents and diluents, among other components. In order to identify them, the mixture must not only be deconstructed, but the unknown must then be structurally characterized and unequivocally identified. The resources required to routinely embark upon this enterprise for each emergent product are not available to many crime labs. Thus, manufacturers rapidly respond to the scheduling of specific synthetic cathinones by releasing novel structural variants, knowing that there will be few if any repercussions for the marketing of new substances [6,7]. The National Drug Intelligence Center has noted that "no substantial law enforcement or regulatory action has significantly prevented synthetic cathinone products from reaching distributors or consumers" [8]. In the US, the DEA has recently temporarily placed ten synthetic cathinones into Schedule I in 2016 as an extension to their provisional placement in 2014 [9]. However, of the hundreds of novel variants, few have been successfully scheduled by the DEA.

Among the repercussions of the use of these substances are their impacts on human health, as well as the criminal offenses committed by individuals while under their influence. Their devastating physiological effects include hyperthermia, severe psychosis, cardiac and neurological problems and death [2,10–15]. Some of the more recent medical incidents (including fatalities) were attributed to α-pyrrolidinopropiophenone, mephedrone, and 3,4-methylenedioxypyrovalerone, three compounds which possess significant structural differences, while having in common the core cathinone structure [12,14,15]. Although their mechanisms of action remain to be fully elucidated, they are thought to act similarly to amphetamines, which are structurally similar to cathinones, but lack the β-keto carbonyl group. All three aforementioned synthetic cathinones have since been scheduled after lengthy studies by the DEA.

Crime laboratories that do undertake structure elucidation studies of suspected cathinone unknown products rely on a variety of methods. These include colorimetric assays, thin layer chromatography (TLC), liquid chromatography (LC), gas chromatography-mass spectrometry (GC-MS), and Fourier transform infrared (FT-IR) and nuclear magnetic resonance (NMR) spectroscopies, among others [16,17]. Aside from the significant time investment and human resource expertise required, other challenges associated with this effort include the amount of sample needed for analysis, and the inherent attributes of various molecules when subjected to some of these forms of analysis. For example, synthetic cathinones are well-known to fragment so extensively by EI mass spectrometry (the mainstay of most crime labs) that the data generated have limited usefulness [18]. Cathinones typically fragment by α-cleavage resulting in the appearance of prominent peaks representing common iminium and acylium ions, while the molecular ion peak is often absent. The result is that cathinones of very different structure can have remarkably similar but uninformative EI mass spectra [18,19]. Even when MS is coupled with GC, the retention time information acquired is often of limited utility for novel compounds whose retention times have not been established. LC-MS/MS analysis of cathinones is also commonly used to separate and identify unknown cathinones in bath salt mixtures and whole blood samples [20–23]. However, as with GC-MS, new protocols must be developed in response to novel cathinones entering the market.

Spectroscopic techniques, such as IR spectroscopy can alert the analyst to the presence of certain diagnostic structural features that imply the presence of a cathinone, but it does not supply definitive structural information. TLC, particularly if visualization agents are used, can yield information that supports the presence of a cathinone, but it does not provide cathinone identity information for an unknown. NMR, albeit a powerful structure elucidation technique, is uncommon at most crime labs, and requires relatively high concentrations of purified samples that are readily solubilized. This combination of factors remains a formidable barrier to the routine practice by crime labs of structure elucidation of NPS believed to be novel cathinone variants. Given the persistent use of bath salt products, methods need to be developed that are complimentary to the existing techniques already commonly used in crime labs, but which add a new dimension that facilitates rapid elucidation of novel variant structures. Such a method could exploit the advantages offered by GC-MS (e.g. the identification of the components within a mixture that are known compounds), but also complement them by rapidly providing additional more definitive structural information that is normally unavailable when cathinones are subjected to EI-MS.

Direct analysis in real time-high resolution mass spectrometry (DART-HRMS) has been shown previously to provide a rapid means by which to distinguish between different cathinones and cutting agents in mixtures [24,25]. In this work, we illustrate how the unique capabilities of DART-HRMS can be used to elucidate the structures of unknowns by combining: (1) the information content from the neutral loss masses calculated from the difference between protonated cathinone molecules and their fragments; and (2) chemometric processing of the data. This approach would complement the already well-established GC-MS and LC-MS methods used in crime laboratories that provide information on the number of components of a mixture and identify known cutting agents and/or diluents, while the DART-HRMS approach outlined herein would enable more comprehensive identification of the key structural fragments of an unknown, thereby greatly facilitating the process of structure elucidation for cathinone unknowns.

2. Materials and methods

2.1. Chemicals

Forty-seven cathinone standards (listed in Table 1) were purchased from Cayman Chemical Company (Ann Arbor, Michigan, USA).

2.2. DART-HRMS

All mass spectra were acquired using a DART-SVP™ ion source (Ionsense, Saugus, MA, USA) coupled to a JEOL AccuTOF high resolution mass spectrometer (JEOL USA, Inc., Peabody, MA, USA). Cathinone standards were sampled in replicates of ten by dipping the closed end of a melting point capillary (VWR, Radnor, PA, USA) into the solid sample and positioning it for ~5 s in the open air gap between the DART-SVP™ ion source and the mass spectrometer inlet.

2.3. DART-HRMS parameters

The JEOL AccuTOF high resolution mass spectrometer with a resolving power of 6000 (FWHM definition) (JEOL USA, Peabody, MA, USA) was operated in positive ion mode for all experiments. Poly (ethyleneglycol) (PEG; average MW 600) was analyzed with every acquired mass measurement as a standard for accurate mass determinations. Orifice 1 was cycled through 20, 30, 60, and 90 V, and orifice 2 and the ring lens voltages were both held at 5 V. The ion guide voltage was set to 400 V to allow for the detection of ions above m/z 40. The ion source was operated with ultra-high purity helium (Airgas, Albany, NY, USA) at a flow rate of 2 L min^{-1} and gas heater temperature of 350 °C. TSSPro 3 software (Shrader Analytical, Detroit, MI, USA) was used for data processing including averaging, background subtraction, and peak centroiding.

2.4. Neutral mass losses determination

Mass Mountaineer software (Mass-spec-software.com, RBC software, Portsmouth, NH, USA) was used to determine the neutral masses lost during fragmentation caused by raising the orifice 1 voltage to 90 V. Using the Mass Mountaineer software, each of the m/z values of the fragment peaks were subtracted from the peak representing the protonated cathinone. This resulted in an intensity vs m/z value plot featuring the neutral losses, and which constituted the neutral loss or

Table 1
The 47 cathinones analyzed in this study.

Methcathinones	
1	4-Bromomethcathinone
2	3,4-Dimethylmethcathinone
3	2-Fluoromethcathinone
4	4-Fluoromethcathinone
5	Mephedrone
Ethcathinones	
6	2,3-Dimethylethcathinone
7	3,4-Dimethylethcathinone
8	2-Methylethcathinone
9	3-Methylethcathinone
10	2-Ethylethcathinone
11	2-Fluoroethcathinone
12	3-Fluoroethcathinone
13	4-Fluoroethcathinone
Methylenedioxy-containing cathinones	
14	Ethylone
15	Eutylone
16	3,4-Methylenedioxy-*N*-benzylcathinone
17	3,4-Methylenedioxy-5-methylethcathinone
18	3,4-Methylenedioxypyrrolidinobutiophenone (MDPBP)
19	3,4-Methylenedioxypyrrolidinohexanophenone (MDPHP)
20	3,4-Methylenedioxypyrrolidinopropiophenone (MDPPP)
21	3,4-Methylenedioxypyrovalerone
22	3,4-Methylenedioxypyrovalerone metabolite 2
23	Methylone
24	Pentylone
Pyrrolidine ring-containing cathinones	
25	3,4-Dimethoxy-α-pyrrolidinopentiophenone
26	3-Fluoro-α-pyrrolidinopropiophenone
27	4-Fluoro-α-pyrrolidinopropiophenone
28	4-Methoxy-α-pyrrolidinopentiophenone
29	2-Methyl-α-pyrrolidinopropiophenone
30	3-Methyl-α-pyrrolidinopropiophenone
31	Naphyrone
32	Pyrovalerone
33	α-Pyrrolidinoheptiophenone (PV9)
34	α-Pyrrolidinohexiophenone (PV8)
35	α-Pyrrolidinopentiophenone (α-PVP)
36	α-Pyrrolidinopropiophenone (α-PPP)
Buphedrones	
37	*N*-Ethylbuphedrone
38	3-Methylbuphedrone
39	4-Methylbuphedrone
40	4-Methyl-*N*-methylbuphedrone
Unclassified cathinones and amines	
41	Diethylcathinone
42	2,5-Dimethoxy-4-methyl-phenethylamine
43	α-Ethylaminopentiophenone
44	*N*-Ethyl-*N*-methylcathinone
45	Isopentedrone
46	4-Methyl-α-ethylaminobutiophenone
47	Pentedrone

"loss spectrum" for each cathinone.

2.5. Classification and clustering

Mass Mountaineer software was used to perform Principal Component Analysis on generated normalized data. Applied Maths Inc. BioNumerics software version 7.6 (Austin, TX, USA) was used for clustering and classification of the cathinones. Within the BioNumerics software, the similarity coefficient used was the Jaccard index with a minimum height of 20% and a peak matching constant tolerance of 0.5. The clustering analysis method used was the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) with an advanced cluster analysis of Topscore UPGMA.

3. Results

The intent of this study was to develop a method to facilitate determination of the structural identity of novel variants of synthetic cathinones. The structure of the underivatized cathinone skeleton is shown in Fig. 1. Forty-seven cathinones representing five different subclasses (Table 1) were analyzed in this study. They included: (1) methcathinones, having a methyl group attached to nitrogen; (2) ethcathinones, similar to methcathinones, but exhibiting an ethyl group attached to nitrogen; (3) buphedrones, which have an ethyl group at the α-carbon; (4) those cathinones with a methylenedioxy moiety appended to the benzene ring; and (5) pyrrolidine-containing cathinones, in which nitrogen is part of a five membered ring. Other cathinones studied but which did not fall into any of these categories included *N,N*-di-substituted compounds such as diethylcathinone and *N*-ethyl-*N*-methylcathinone, and cathinones with three or more carbons attached to the α-carbon, such as pentedrone.

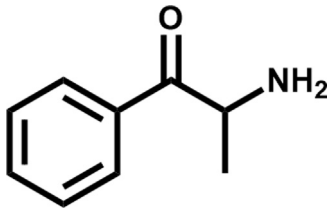

Fig. 1. The structure of cathinone.

In order to confirm that the cathinones were consistently detectable by DART-HRMS, soft ionization spectra acquired at an orifice 1 voltage of 20 V were collected for each of the 47 cathinones. These data, along with the structures of the analyzed compounds, are presented in Supplementary Fig. S1. The spectra of 3,4-methylenedioxy-α-pyrrolidinopropiophenone (MDPPP) and 3,4-methylenedioxy-α-pyrrolidinobutiophenone (MDPBP), shown in Fig. 2 (Panel A) and rendered as a head-to-tail plot, are representative examples. That of MDPPP is dominated by the appearance of the protonated cathinone at *m/z* 248.1274, with no fragmentation. The spectrum of MDPBP shows the expected peak for the protonated molecule at *m/z* 262.1425, with no fragment ions. The similar soft ionization results observed for all 47 cathinones analyzed are shown in Supplementary Fig. S1.

To ascertain whether cathinones of similar structure would exhibit similar mass spectral features upon fragmentation, DART-HRMS spectra were also acquired under in-source collision-induced dissociation (CID) conditions. CID was accomplished by setting the orifice 1 voltage to 90 V. Representative examples of CID spectra of both MDPPP and MDPBP are shown in Fig. 2 (Panel B). The CID fragmentation pattern of MDPPP (top spectrum) showed prominent peaks at *m/z* 98.0953, 147.0430, and 177.0527 along with a peak representing the protonated parent at *m/z* 248.1242. The CID fragmentation pattern of MDPBP (bottom spectrum) featured high intensity peaks at *m/z* 112.1084, 149.0187, 161.0558, and 191.0642, as well as at 262.1389 (representing the protonated parent cathinone). Despite the fact that MDPPP and MDPBP differ only in the length of the alkyl chain, their spectra are visually quite different and they did not share any of the observed prominent peaks. This trend was seen for all other cathinones of similar structure (Supplementary Fig. S2). For example, ethylone and eutylone, two methylenedioxy-containing cathinones differing only in the length of the alkyl chain on the α-carbon, exhibited very different CID spectra, with only one shared peak at nominal *m/z* 65.

Structurally related cathinones would be expected to fragment by a similar mechanism. Accordingly, Mass Mountaineer software was used to generate fragmentation spectra based on neutral loss values. These "virtual" neutral loss fragmentation spectra were created by subtracting the high-resolution (HR) *m/z* values of the fragment ions observed in

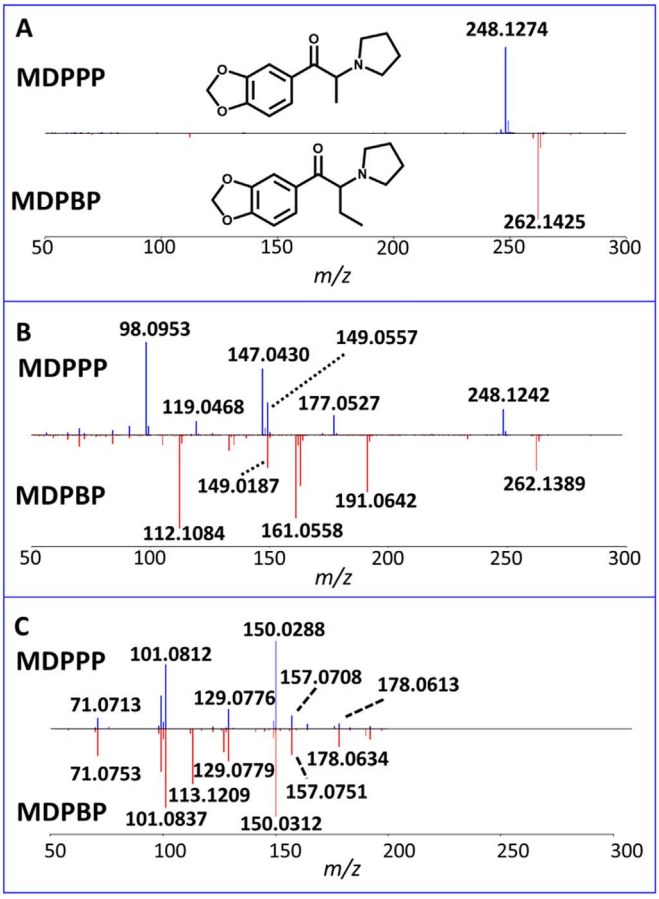

Fig. 2. The soft ionization, CID, and neutral loss spectra of MDPPP and MDPBP. The soft ionization (20 V) spectra of MDPPP (top) and MDPBP (bottom) rendered as a head-to-tail plot (Panel A). The Panel A spectra exhibit the high resolution *m/z* values 248.1274 and 262.1425 representing the protonated cathinones MDPPP and MDPBP, respectively; collision-induced dissociation spectra (Panel B) of MDPPP (top) and MDPBP (bottom) rendered as a head-to-tail plot. Each spectrum in Panel B retains the protonated cathinone peak (nominal *m/z* 248 and 262) while also showing the *m/z* values representing the ions produced during fragmentation; neutral loss spectra (Panel C) of MDPPP (top) and MDPBP (bottom) rendered as a head-to-tail plot. These two structurally similar cathinones share the neutral losses of nominal *m/z* 71, 101, 129, 150, 157, and 178 (Panel C).

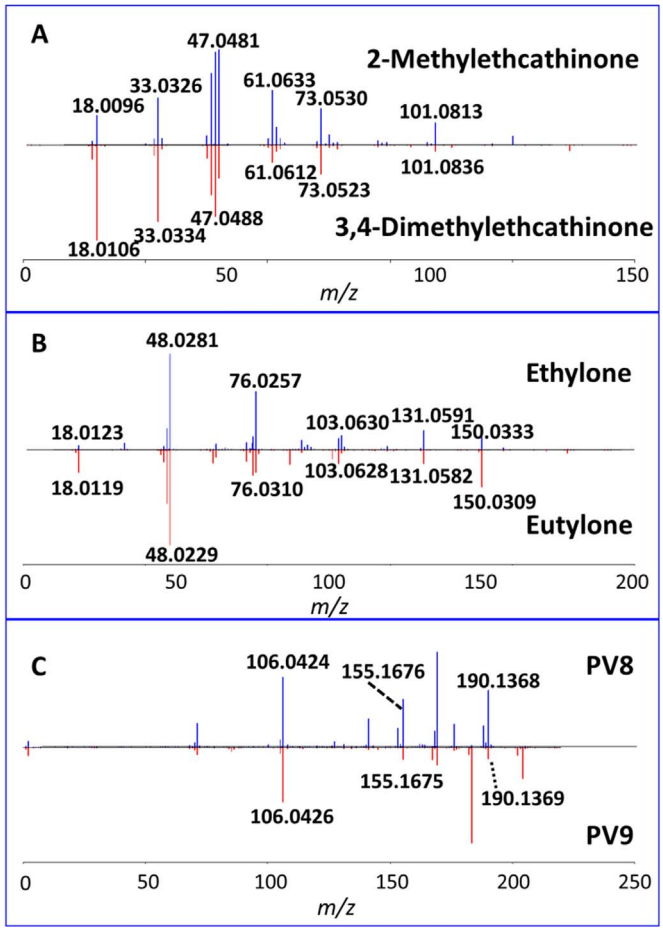

Fig. 3. Neutral loss spectra for cathinone pairs falling within one of the three cathinone subclasses [i.e. ethcathinones (Panel A), methylenedioxy-containing cathinones (Panel B), and pyrrolidine-containing cathinones (Panel C)]. The ethcathinones 2-methylethcathinone and 3,4-dimethylethcathinone share the nominal *m/z* values 18, 33, 47, 61, 73, and 101. Ethylone and eutylone, which share the methylenedioxy moiety, have nominal *m/z* 18, 48, 76, 103, 131, and 150 in common. In the pyrrolidine-containing subclass, PV8 and PV9 have shared neutral losses of nominal *m/z* 106, 155, and 190. The structures of each cathinone are shown in Supplementary Fig. S1.

the CID spectrum of each cathinone from the mass of the protonated cathinone. This resulted in a new spectrum which showed the HR neutral masses of the species lost during fragmentation. The peak intensities shown in the neutral loss spectra are equivalent to the intensities of the fragment peaks in the CID spectra from which they were derived. As examples, the neutral loss spectra of MDPPP and MDPBP are shown in Fig. 2 (Panel C). Importantly, the two cathinones share common HR neutral losses represented by nominal *m/z* 71, 101, 129, and 150. The loss of 71 corresponds to the pyrrolidine ring. The peak at *m/z* 101 represents the combined loss of the pyrrolidine ring and formaldehyde from the methylenedioxy moiety. Nominal *m/z* 150 is the methylenedioxy benzoyl group. All of these structural features are shared by MDPPP and MDPBP, but this information is not readily available from GC-EI-MS analysis of these compounds. Similar results were observed with the other cathinones analyzed in this study that shared similar structural features (Supplementary Fig. S3).

Comparison of the neutral loss spectra of the other cathinones from each of the classes represented in this study was then conducted. The results, rendered as head-to-tail plots are shown in Fig. 3. Panel A shows the neutral loss spectra of two ethcathinones: 2-methylethcathinone and 3,4-dimethylethcathinone; two methylenedioxy cathinones: ethylone and eutylone are shown in Panel B; and two pyrrolidine-containing cathinones: PV8 and PV9 are featured in Panel C. For the ethcathinones (Panel A), common neutral losses included nominal *m/z* 18, 33, 47, 61, 73, and 101. For the methylenedioxy cathinones (Panel B), the shared neutral losses were nominal *m/z* 18, 48, 76, 103, 131, and 150. The pyrrolidine-containing cathinones (Panel C) had nominal *m/z* 106, 155, and 190 in common. Overall, the neutral losses observed for cathinones of the same class were remarkably similar, while those of different classes were not. This was confirmed by multivariate statistical analysis processing of the data.

To determine if the neutral loss data could be used to categorize cathinones according to class, the neutral loss spectra of a small subset of nine cathinones were subjected to principle component analysis (PCA) which was accomplished using Mass Mountaineer software. To generate the plot, eight feature masses derived from the neutral loss spectra of nine cathinones (representing three different classes) were used (Table 2). The training set was comprised of ten normalized spectral replicates of each of the nine cathinones. The resulting PCA plot, illustrated in Fig. 4, showed that the nine cathinones fell into one of three cathinone subclasses: ethcathinones (blue circles), methylenedioxy-containing cathinones (red squares), and pyrrolidine ring-containing cathinones (green triangles). Although each of the cathinones could be distinguished from others within the same class, this plot also illustrated that this small subset of cathinones clustered based on shared structural features. The first three principal components accounted for 65% of the variance. However, the three subclasses were clearly

Table 2
Feature masses and compositions used for the principal component analysis (PCA) plot shown in Fig. 4.

m/z	Loss formula
18.0104	H_2O
33.0310	CH_5O
46.0462	C_2H_6O
48.0577	C_2H_8O
71.0783	C_4H_9N
76.0253	$C_2H_4O_3$
106.0382	C_7H_6O
150.0336	$C_8H_6O_3$

separated from one another. Leave-one-out cross-validation (LOOCV) was 98.9% with only one misclassification: one replicate of 2-methylethcathinone classified as a methylenedioxy-containing cathinone. The data points representing the ethcathinone subclass showed variation along all principle components (PC) 1, 2, and 3. The points corresponding to the methylenedioxy cathinones clustered with small variations along PC2 and larger variations on PC1 and PC3. The green triangles, representing the pyrrolidine-containing subclass, varied along PC1 and slightly on PC2, but not along PC3. The successful separation of the three distinct classes of cathinones showed proof of principle that the neutral loss spectra could possibly be used for classification of structural variants of synthetic cathinones.

To expand on this proof of principle, the entire spectral dataset (44 cathinones) was also processed by an unsupervised multivariate statistical analysis method, hierarchical clustering, using Applied Maths Inc. BioNumerics software version 7.6. This approach did not rely on the prior selection of feature masses, but rather utilized the entire neutral loss spectrum of each cathinone. Ten replicates for each cathinone were used. The similarity coefficient used was the Jaccard index with the hierarchical clustering analysis method being UPGMA with an advanced cluster analysis algorithm of Topscore UPGMA. The resulting linear dendrogram is shown in Fig. 5. Each cathinone class, including ethcathinone, methcathinone, buphedrone, methylenedioxy-containing, and pyrrolidine-containing, was assigned a different color (red, purple, light blue, green, and yellow respectively). Several cathinones clustered together based on their shared structural features. The pyrrolidine-containing cathinones (yellow) dominate most of the left half of the dendrogram. Within the two large clusters of pyrrolidine ring-containing cathinones were two di-substituted nitrogen cathinones that were not classified as being part of a subclass (diethylcathinone and *N*-ethyl-*N*-methylcathinone). The methcathinones, indicated in purple, were split into two separate groups located near the left-hand side and middle right of the dendrogram. The ethcathinones, in red, were localized to the furthermost right end of the dendrogram, while the methylenedioxy-containing cathinones (shown in green) covered a large portion of the right half of the dendrogram. The buphedrones, in light blue, appeared in clusters interspersed throughout the dendrogram.

In order to determine whether the dendrogram could be used to facilitate the determination of the structures of cathinone unknowns, the system was tested by external validation using three "unknowns". Three cathinones that were not members of the model dataset, including 4-fluoroethcathinone, 3,4-methylenedioxy-α-pyrrolidinohexanophenone (MDPHP), and ethylone which fall into the ethcathinone, pyrrolidine-containing, and methylenedioxy-containing subclasses respectively, were processed in a similar fashion to that which was done for the model dataset. The resulting neutral loss spectra were then imported into the model dendrogram using BioNumerics software. In every case, these unknowns fell into the area represented by the subclass to which they belonged. The unknowns are illustrated in the dark blue-shaded areas of the dendrogram in Fig. 6. The 4-fluoroethcathinone replicates fell in the area of ethcathinones, MDPHP replicates fell in the area of the pyrrolidine-containing cathinones, and ethylone replicates fell in the area of the methylenedioxy-containing cathinones.

The areas of the dendrogram into which the unknowns fell, along with their respective soft ionization and CID spectra, were then used to piece together their structures. The approach to structure elucidation is illustrated with the determination of the structure of MDPHP. Its neutral loss spectrum clustered within the pyrrolidine-containing cathinone area of the dendrogram (Fig. 6). Its chemical formula, $[C_{17}H_{23}NO_3 + H]^+$, was determined from the HR protonated mass acquired under soft ionization conditions (*m/z* 290.1768) (Fig. 7A). The CID spectrum (Fig. 7B) was used to determine the neutral loss spectrum (Fig. 7C). Briefly, the fragment ions produced with an orifice 1 voltage of 90 V were subtracted from the protonated cathinone (*m/z* 290.1732). The resulting values were then converted to the neutral loss spectrum shown in Panel C. The loss of *m/z* 71.0763 corresponds to the pyrrolidine ring, the loss of *m/z* 101.0880 corresponds to a combined loss of the pyrrolidine ring and formaldehyde from the methylenedioxy moiety, and the loss of *m/z* 150.0322 corresponds to the loss of the benzene ring (fused with the methylenedioxy moiety) and the carbonyl. Nominal *m/z* 141.1535 could represent loss of the pyrrolidine ring and a four-carbon alkyl chain. Piecing these together, the structure of MDPHP could be tentatively determined, with the alkyl chain possibilities being *n*-butyl, *t*-butyl or *sec*-butyl. Similar steps were taken to determine the structures of the other two unknowns, 4-fluoroethcathinone and ethylone (Supplementary Figs. S4-S5).

4. Discussion

This study was inspired by the need to develop methods that can be used to triage drug samples in the crime lab and quickly determine their identity based on the similarity of their structures to those of previous generations of related compounds. With the influx of novel synthetic cathinones on the drug market, this technique could be used to quickly characterize structural variants of the basic cathinone scaffold, and guide subsequent more definitive structural determinations. Furthermore, this technique could be applied to other drug classes, such as amphetamines, if databases similar to that described here are

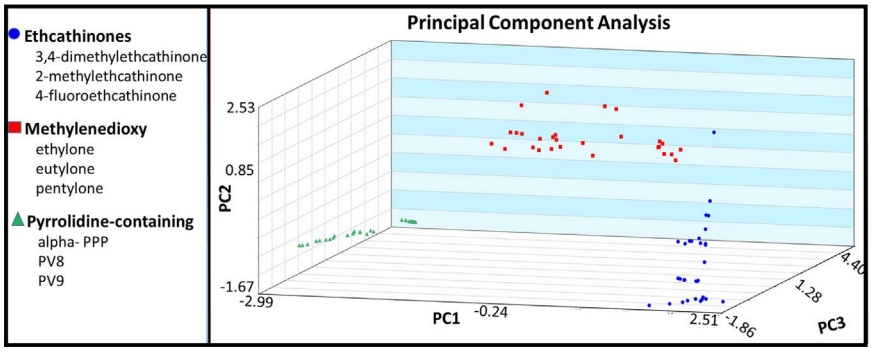

Fig. 4. Principal component analysis plot featuring nine cathinones (3,4-dimethylethcathinone, 2-methylethcathinone, 4-fluoroethcathinone, ethylone, eutylone, pentylone, α-PPP, PV8 and PV9) which fall within the three subclasses ethcathinones, methylenedioxy-containing cathinones, and pyrrolidine-containing cathinones (indicated in blue, red, and green, respectively). Three principal components account for 65% of the variance with a leave-one-out cross-validation of 98.9%. As shown, the ten replicates of each cathinone cluster closely with other cathinones within the same subclass.

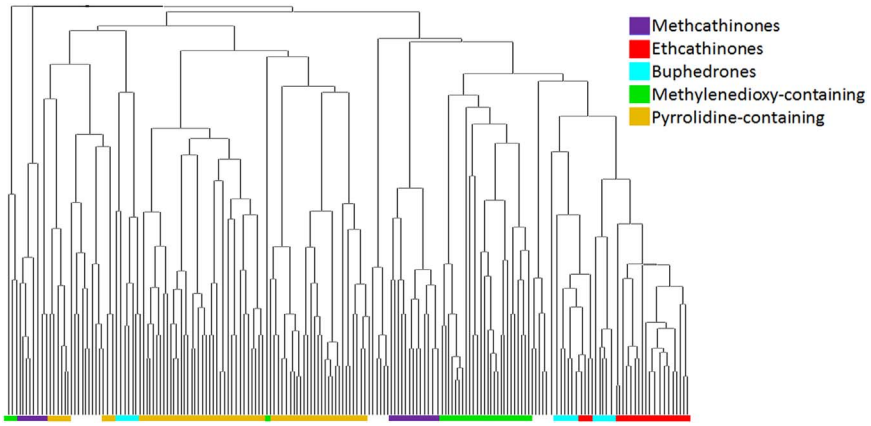

Fig. 5. A hierarchical clustering analysis dendrogram featuring the neutral loss spectra of the 44 cathinones studied. The cathinones were analyzed in replicates of ten. Each cathinone subclass was assigned a color: methcathinones in purple; ethcathinones in red; buphedrones in blue; methylenedioxy-containing cathinones in green; and pyrrolidine-containing cathinones in yellow. Several cathinones of the same subclass clustered together, implying their structural similarity. The areas without color represent cathinones which did not fall into one of the specified subclasses.

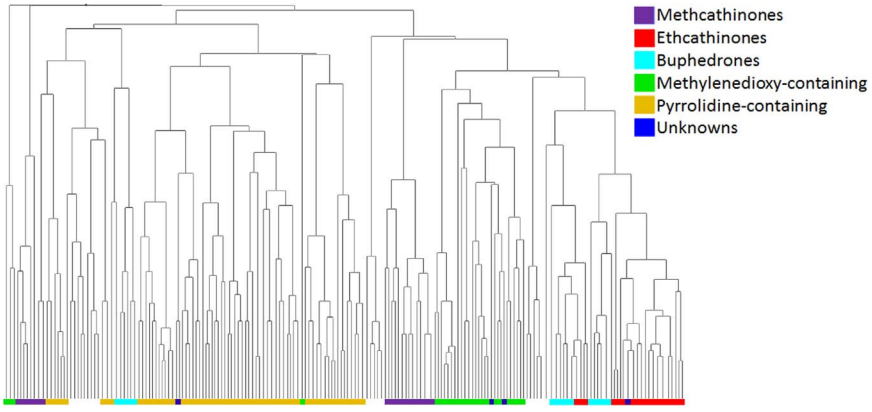

Fig. 6. Hierarchical clustering analysis dendrogram of the neutral loss spectra of the 44 cathinones included in the model dataset, with the addition of three "unknowns". Three replicates of each unknown were entered into the software and their position within the dendrogram is indicated in dark blue. Methcathinones, ethcathinones, buphedrones, methylenedioxy-containing cathinones, and pyrrolidine-containing cathinones are distinguished by the colors purple, red, light blue, green and yellow, respectively. The unknown MDPHP fell within the pyrrolidine-containing cathinone section at the middle left of the dendrogram; ethylone fell within the methylenedioxy-containing cathinone section (highlighted green area of the dendrogram); and 4-fluoroethcathinone fell within the red ethcathinone section in the right-hand area of the dendrogram.

compiled. The successful approach reported here involved chemometric processing of DART-HRMS-derived neutral loss data. Unlike the standard analytical methods often employed in crime labs, DART-HRMS requires very little sample. The method, including sample analysis and data processing, is also very rapid, and the sample can be analyzed in its native form with no sample pretreatment steps.

The approach permits the categorization of an unknown into a subclass, which in turn enables the extraction of structural information that could be used to identify the variations made to the cathinone scaffold (Fig. 1). As anticipated, cathinones within the same family fragmented by similar pathways. Thus, the extracted neutral loss information could be used not only to classify structures based on their shared features, but their unique fragment peaks could also be used to rapidly determine the limited number of viable structures for the unknown. Although the information content inherent in the neutral losses observed in mass spectra has always been used in their interpretation, the approach taken here was to not only generate mass spectra based solely on the neutral losses, but also to subject these spectra to statistical analysis approaches that permitted extraction of structural information.

Comparisons of the CID and neutral loss spectra of cathinones within the same family confirmed that the neutral loss spectra were far more useful than the CID spectra in enabling successful classification of a small set of structurally similar cathinones, as shown in the PCA plot (Fig. 4). For the PCA plots, the feature masses chosen were a combination of those that were common to the cathinones, and others that were unique to each of them. Although there are slight variations in the spectral replicates in the PCA plot, these variations are much smaller within a subclass than those between structurally different subclasses. This is demonstrated by the high leave-one-out cross validation value obtained for the plot. To cover the widest range of cathinone variations, 44 cathinones falling into five different subclasses were used to build the model dataset that was subjected to hierarchical clustering analysis. Some of the cathinones were members of two subclasses. For example, a few of the pyrrolidine-containing cathinones contained a methylenedioxy moiety as well. These included MDPPP and MDPBP. In the hierarchical clustering analysis dendrogram, the pyrrolidine ring-containing cathinones were observed to cluster together, regardless of the presence or absence of a fused methylenedioxy benzene ring. A previous study investigated the fragmentation mechanisms of methylenedioxy-containing cathinones including MDPBP [20]. In that study, the authors determined that due to steric factors, pyrrolidine-containing cathinones do not lose the methylenedioxy group, as do other cathinones of the same class [20]. This could explain why these cathinones, which fall into two subgroups, cluster as pyrrolidine-containing rather than methylenedioxy-containing cathinones. However, under CID conditions for the DART-HRMS, the neutral losses do provide insight into the presence of a methylenedioxy moiety, and allow elucidation of a correct structure for these compounds, as shown in Fig. 7.

As illustrated in Fig. 5, two un-classified cathinones and one buphedrone-classified cathinone appeared within the pyrrolidine ring-containing cathinone area. Although these three compounds, namely diethylcathinone, *N*-ethyl-*N*-methylcathinone, and 4-methyl-*N*-methylbuphedrone do not contain a pyrrolidine ring, they do have a di-substituted nitrogen. Given that a pyrrolidine ring is comprised of a di-substituted nitrogen in which the two substituents are connected (thereby forming a ring), the appearance of the three aforementioned compounds in the pyrrolidine area of the dendrogram is not surprising and may not actually be a "misclassification" per se. It may simply reflect fragmentation of these very similar structures by the same mechanism, leading to similar neutral losses.

At the right-hand side of the dendrogram are two separate groups of buphedrones. They include 3-methylbuphedrone, 4-methylbuphedrone, *N*-ethylbuphedrone, and 4-methyl-α-ethylamino-butiophenone. In the

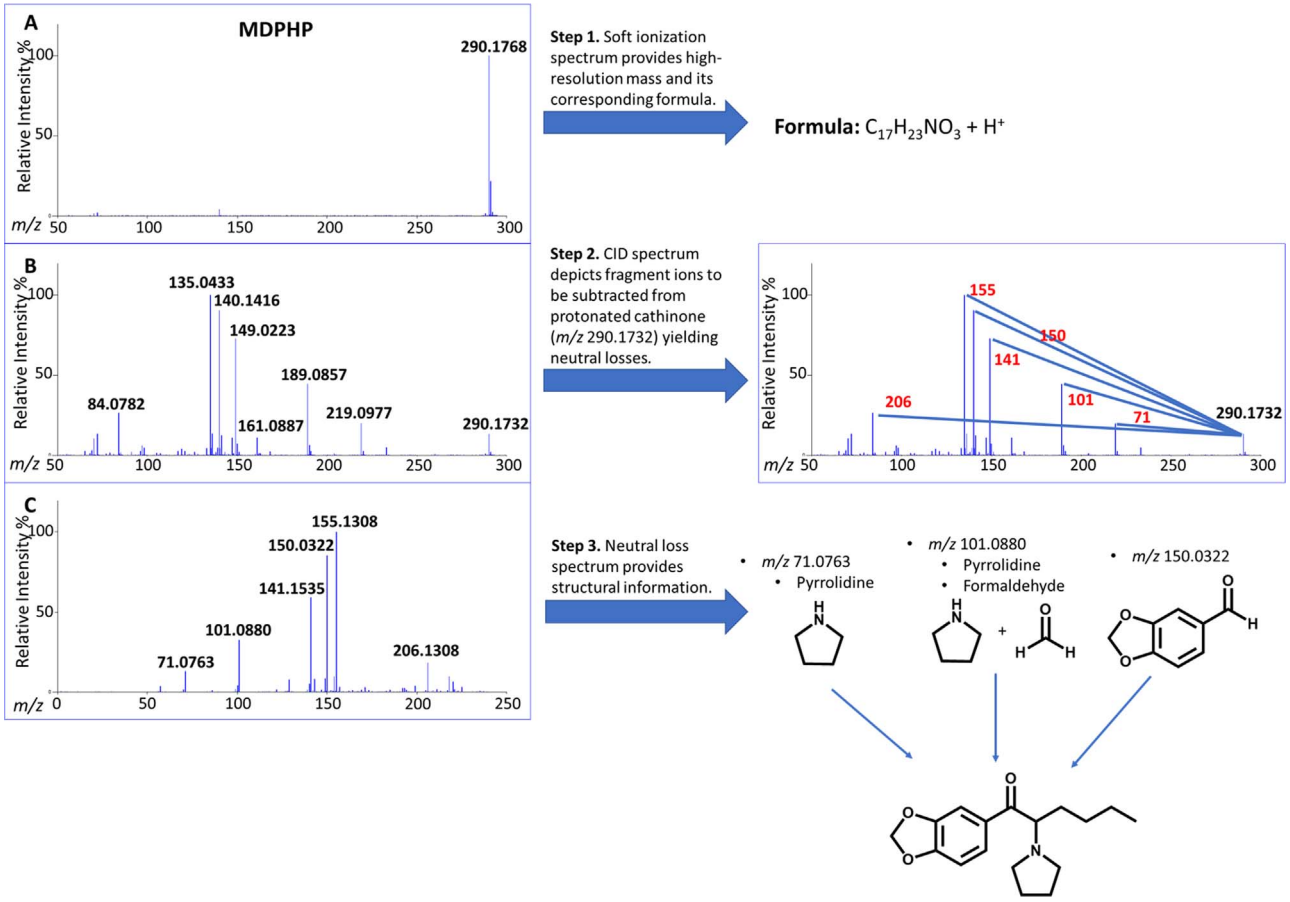

Fig. 7. Steps taken to deduce the structure of the "unknown" MDPHP. The soft ionization spectrum (Panel A) shows the protonated cathinone *m/z* 290.1768 which is consistent with the formula $[C_{17}H_{23}NO_3 + H]^+$. The CID spectrum (Panel B) retains the protonated cathinone but also includes fragment ions. The fragment ions are subtracted from the protonated cathinone mass and used to calculate the neutral masses lost during fragmentation. Panel C, the neutral loss spectrum, allows for the determination of structural characteristics of the unknown.

dendrogram, these cathinones are surrounded by ethcathinones. The only difference between the two subclasses (i.e. buphedrones and ethcathinones) is the location of the ethyl group either on the α-carbon or the nitrogen, respectively. Thus, the similarity between the two could explain why the subclasses clustered together.

The utility of this method for the determination of the possible structures of unknowns was confirmed by using it to investigate the structures of three cathinones that were not included in the model dataset. The determination of the structures required the use of the soft ionization mass spectra, CID mass spectra, and results obtained from the hierarchical clustering analysis dendrogram. For each cathinone unknown, all the required mass spectra were obtained in less than 1 min in a single experiment using negligible amounts of the sample in its native form, which is ideal in a forensic context. The model dataset can be continuously expanded to include the increasing number of characterized synthetic cathinone variants as they become known. The expansion of the cathinone dendrogram database would include an increasing number of subclasses, and make identification of unknowns by this method not only feasible but more rapid.

Synthetic cathinones are often found as mixtures comprised of multiple additional components including stimulants, synthetic cathinones, or cutting agents. GC-MS experiments can be used to determine the number and identity of the components of a mixture, while the technique described here can be complementary and extend the power of the analyst in not only identifying cutting agents and diluents (by GC-MS), but also by providing definitive structural information that can facilitate cathinone unknown identification. GC-MS is currently limited in this regard due to extensive fragmentation that occurs when cathinones are analyzed by this method. DART-HRMS could also provide independent confirmation of the components of mixtures that have been identified in GC-MS experiments.

5. Conclusions

The feasibility of elucidating the core structural features of cathinones of unknown identity by chemometric processing of DART-HRMS-derived neutral loss spectra was demonstrated. The approach relies on the fact that while conventional cathinone analysis methods such as GC-MS provide little structural information because cathinones tend to undergo extensive fragmentation, CID DART-HRMS spectra exhibit features that allow similarities between related cathinone structures to emerge. *The power of our approach is in the extraction of the dark matter of neutral loss information and rendering it as "dark matter spectra," followed by statistical analysis processing of it for the identification of unknowns.* The neutral loss spectra derived from the fragmentation spectra show striking similarities between cathinones sharing structural features such as pyrrolidine rings and methylenedioxy moieties. These neutral losses can be employed to organize cathinones into subclasses based on shared structural features, as illustrated by both PCA and hierarchical clustering analysis. The combination of the soft ionization, in-source CID, and neutral loss spectra can be used to construct the structures of unknown cathinones. The method can be applied to a range of cathinone structures, and the robustness of the dendrogram dataset described here can be readily increased by continued addition of newly discovered structures. The approach provides a means for forensic laboratories to triage in-coming samples and make informed

decisions regarding subsequent more definitive analyses to which samples should be subjected. This method could contribute to the reduction of crime lab backlogs that accrue due to the time, financial and human resource constraints that limit the efficiency of crime labs in analyzing the increasing number of synthetic cathinones and other sample unknowns.

Acknowledgements

The authors thank Dr. Robert Cody (JEOL USA) for helpful discussions. This project was supported by Award Numbers 2013-DN-BX-K041 and 2015-DN-BX-K057 awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. The opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect those of the Department of Justice.

Funding

This work was supported by the US National Science Foundation [grant # 1429329] as well as the United States National Institute of Justice [grant # 2013-DN-BX-K041].

Conflicts of interest

There are no conflicts of interest to declare.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2017.11.020.

References

[1] The Challenge of New Psychoactive Substances, United Nations Office on Drugs and Crime, Vienna, Austria, 2013.
[2] M.E. Nelson, S.M. Bryant, S.E. Aks, Emerging drugs of abuse, Emerg. Med. Clin. North Am. 32 (1) (2014) 1–28.
[3] M.H. Baumann, E. Solis Jr., L.R. Watterson, J.A. Marusich, W.E. Fantegrossi, J.L. Wiley, Baths salts, spice, and related designer drugs: the science behind the headlines, J. Neurosci. 34 (46) (2014) 15150–15158.
[4] L.A. Johnson, R.L. Johnson, R.B. Portier, Current "legal highs", J. Emerg. Med. 44 (6) (2013) 1108–1115.
[5] C.D. Rosenbaum, S.P. Carreiro, K.M. Babu, Here today, gone tomorrow…and back again? A review of herbal marijuana alternatives (K2, Spice), synthetic cathinones (bath salts), kratom, *Salvia divinorum*, methoxetamine, and piperazines, J. Med. Toxicol. 8 (1) (2012) 15–32.
[6] K.A. Seely, A.L. Patton, C.L. Moran, M.L. Womack, P.L. Prather, W.E. Fantegrossi, A. Radominska-Pandya, G.W. Endres, K.B. Channell, N.H. Smith, K.R. McCain, L.P. James, J.H. Moran, Forensic investigation of K2, Spice, and "bath salt" commercial preparations: a three-year study of new designer drug products containing synthetic cannabinoid, stimulant, and hallucinogenic compounds, Forensic Sci. Int. 233 (1–3) (2013) 416–422.
[7] K.G. Shanks, T. Dahn, G. Behonick, A. Terrell, Analysis of first and second generation legal highs for synthetic cannabinoids and synthetic stimulants by ultra-performance liquid chromatography and time of flight mass spectrometry, J. Anal. Toxicol. 36 (6) (2012) 360–371.
[8] Synthetic Cathinones (bath salts): An Emerging DomesticThreat, U.S. Department of Justice, Johnston, PA, 2011.
[9] Schedules of Controlled Substances: Extension of Temporary Placement of 10 Synthetic Cathinones in Schedule I of the Controlled Substances Act, Department of Justice, 2016, pp. 11429–11431.
[10] D. Lee, C.W. Chronister, J. Hoyer, B.A. Goldberger, Ethylone-related deaths: toxicological findings, J. Anal. Toxicol. 39 (7) (2015) 567–571.
[11] J.M. Prosser, L.S. Nelson, The toxicology of bath salts: a review of synthetic cathinones, J. Med. Toxicol. 8 (1) (2012) 33–42.
[12] C.M. White, Mephedrone and 3,4-methylenedioxypyrovalerone (MDPV): synthetic cathinones with serious health implications, J. Clin. Pharmacol. 56 (11) (2016) 1319–1325.
[13] J. Wojcieszak, D. Andrzejczak, A. Woldan-Tambor, J.B. Zawilska, Cytotoxic activity of pyrovalerone derivatives, an emerging group of psychostimulant designer cathinones, Neurotox. Res. 30 (2) (2016) 239–250.
[14] C.J. Woloshchuk, K.H. Nelson, K.C. Rice, A.L. Riley, Effects of 3,4-methylenedioxypyrovalerone (MDPV) pre-exposure on the aversive effects of MDPV, cocaine and lithium chloride: implications for abuse vulnerability, Drug Alcohol Depend. 167 (2016) 121–127.
[15] T.H. Wright, C. Harris, Twenty-one cases involving alpha-pyrrolidinovalerophenone (alpha-PVP), J. Anal. Toxicol. 40 (5) (2016) 396–402.
[16] Recommended Methods for the Identification and Analysis of Synthetic Cathinones in Seized Material, United Nations Office on Drugs and Crime, 2015.
[17] D. Zuba, Identification of cathinones and other active components of 'legal highs' by mass spectrometric methods, Trends Anal. Chem. 32 (2012) 15–30.
[18] S. Kerrigan, M. Savage, C. Cavazos, P. Bella, Thermal degradation of synthetic cathinones: implications for forensic toxicology, J. Anal. Toxicol. 40 (1) (2016) 1–11.
[19] S. Kerrigan, Improved Detection of Synthetic Cathinones in Forensic Toxicology Samples: Thermal Degradation and Analytical Considerations, U.S. Department of Justice, 2015.
[20] E. Fornal, Identification of substituted cathinones: 3,4-methylenedioxy derivatives by high performance liquid chromatography-quadrupole time of flight mass spectrometry, J. Pharm. Biomed. Anal. 81-82 (2013) 13–19.
[21] E. Fornal, Study of collision-induced dissociation of electrospray-generated protonated cathinones, Drug Test. Anal. 6 (7–8) (2014) 705–715.
[22] E. Fornal, A. Stachniuk, A. Wojtyla, LC-Q/TOF mass spectrometry data driven identification and spectroscopic characterisation of a new 3,4-methylenedioxy-N-benzyl cathinone (BMDP), J. Pharm. Biomed. Anal. 72 (2013) 139–144.
[23] L. Mercolini, M. Protti, M.C. Catapano, J. Rudge, A.E. Sberna, LC-MS/MS and volumetric absorptive microsampling for quantitative bioanalysis of cathinone analogues in dried urine, plasma and oral fluid samples, J. Pharm. Biomed. Anal. 123 (2016) 186–194.
[24] R.A. Musah, R.B. Cody, M.A. Domin, A.D. Lesiak, A.J. Dane, J.R.E. Shepard, DART–MS in-source collision induced dissociation and high mass accuracy for new psychoactive substance determinations, Forensic Sci. Int. 244 (2014) 42–49.
[25] J. LaPointe, B. Musselman, T. O'Neill, J.R. Shepard, Detection of "bath salt" synthetic cathinones and metabolites in urine via DART-MS and solid phase microextraction, J. Am. Soc. Mass Spectrom. 26 (1) (2015) 159–165.