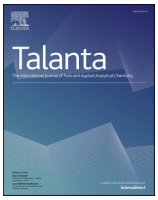
Talanta 246 (2022) 123417

Contents lists available at ScienceDirect

Talanta

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



Revealing the presence of tryptamine new psychoactive substances using fused "neutral loss" spectra derived from DART high-resolution mass spectra

Mónica I. Ventura, Samira Beyramysoltan, Rabi A. Musah *

Department of Chemistry, State University of New York at Albany, 1400 Washington Ave, Albany, NY, 12222, USA

ARTICLE INFO

Keywords:
Tryptamines
DART-HRMS
Neutral loss
Data fusion
Chemometrics
New psychoactive substances

ABSTRACT

A data fusion approach for the rapid extraction of core scaffold information that can be used to facilitate structure determination for new psychoactive substance (NPS) tryptamines is described. The method involves the screening of DART-HRMS data of new tryptamines against a partial least squares-discriminant analysis (PLS-DA) model that predicts the core tryptamine structure class into which the compound can be grouped. The PLS-DA prediction model was created and trained using neutral loss spectra derived from collision-induced dissociation (CID) DART mass spectral analysis of 50 tryptamine structures acquired at 60 V and 90 V, in which the sample groups were revealed by hierarchical clustering analysis (HCA). HCA of the fused neutral loss data clustered the 50 tryptamines into 10 groups based on the identities of the neutral fragments lost during fragmentation. "Leave-one-structure-out" validation of the PLS-DA model gave 100% accuracy, precision, sensitivity, and specificity. For external validation, the ability of the model to classify four compounds that were unfamiliar to it was tested, and the model was found to correctly predict the skeletal framework in each case. The results show proof of concept for how this approach can aide in the identification of new emerging psychoactive compounds.

1. Introduction

Forensic laboratories across the United States continue to grapple with the challenges imposed by the unrelenting rise in the circulation of new psychoactive substances (NPSs), a term that refers to recreationally used, unscheduled products that exhibit psychoactivity [1,2]. Their unregulated status and ready accessibility make them prime targets for misuse [3,4]. Opioids in particular, have rightfully received significant attention, in part because of the high toxicity of some of the compounds that fall under this category, such as fentanyl and its derivatives. However, the focus on opioids has masked the significant problems associated with the emergence of other subsets of NPSs. Tryptamines serve as a case in point. Many are structural derivatives of serotonin or *N, N*-dimethyltryptamine (DMT), and are subject to misuse because of their mind-altering effects [5–7].

While several are scheduled compounds, there are numerous sites within their scaffolds to which structural modifications can be introduced, resulting in the continued emergence of novel variants that retain their psychoactivity [8–10]. There are several bottlenecks to the scheduling of these substances including: (1) the difficulty of rapidly detecting and structurally characterizing emerging compounds; and (2) development of protocols for their routine detection and identification. One obstacle encountered by crime labs that mainly use electron ionization (EI) mass spectral techniques is that closely related structural variants can often exhibit nearly identical EI mass spectral fragmentation patterns, making it challenging to utilize this conventional approach for their definitive identification. Another is that some tryptamines fragment so extensively that their EI fragmentation patterns are rendered minimally informative for identification purposes [3,9,11,12]. A number of studies that have sought to address these issues have been reported. Piorunska-Sedlak and Stypulkowska [13] showed how attenuated total reflectance Fourier transform infrared spectroscopy (ATR-IR) can serve as a good screening method because it is rapid, there is no need for sample pre-treatment, and only a small amount of sample is needed. Nevertheless, it was observed that depending on the solvent, some samples exhibited polymorphic crystalline forms that differed from those of the reference sample, making comparisons challenging. Jones et al. [14] reported the creation of an IR and Raman spectroscopy

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

https://doi.org/10.1016/j.talanta.2022.123417
Received 25 December 2021; Received in revised form 24 March 2022; Accepted 27 March 2022
Available online 1 April 2022

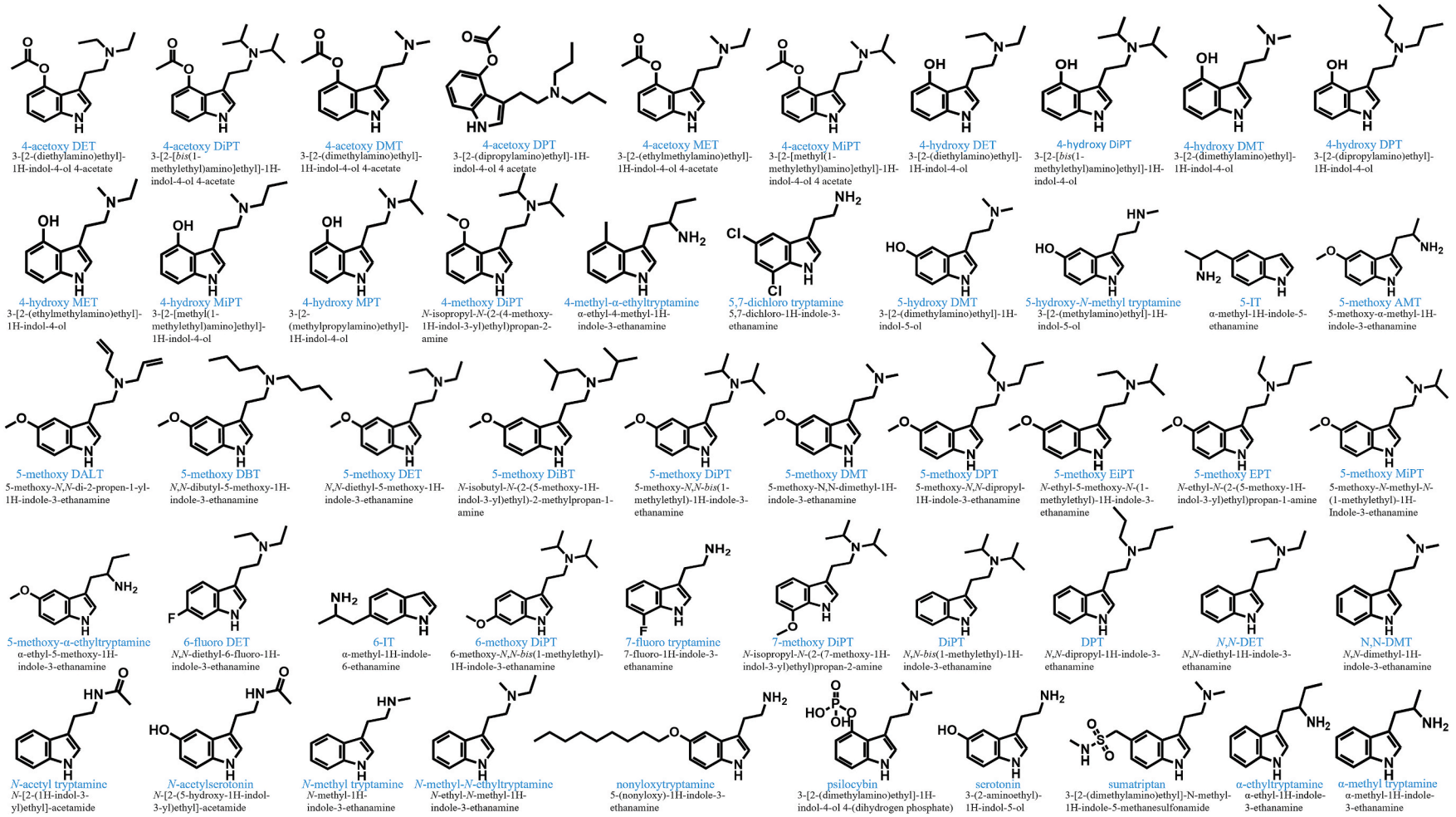


Fig. 1. The structures, common names, and formal names of the tryptamines analyzed in this study.

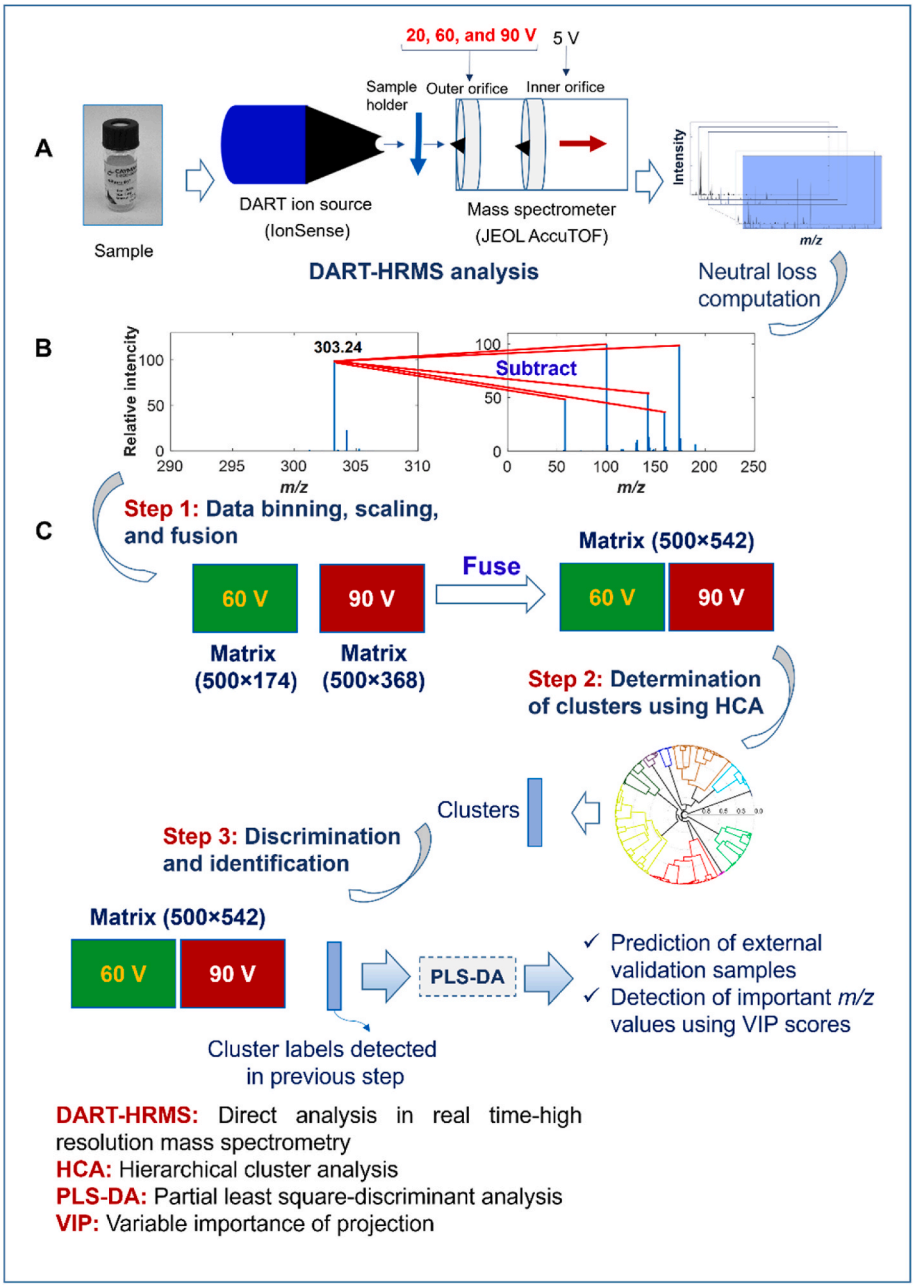


Fig. 2. Data collection and statistical analysis workflow approach for the development of a model to enable prediction of new tryptamine variants. (A): DART-HRMS data collected at 20 V, 60 V and 90 V; (B) Generation of neutral loss spectra; and (C) Multivariate data analysis workflow: Step 1—Conversion of spectra to matrices following binning and normalization; Step 2—HCA analysis of data to define clusters; and Step 3—Creation of PLS-DA model used to discriminate between tryptamine classes.

database of NPSs, and Moorthy and Sisco [15] have created a library search algorithm for NPS identification using DART mass spectra. In these two cases, identification is based on matching of the spectrum of the unknown to representative spectra that are already present in the database [14,15]. Therefore, emerging compounds that are new to the database cannot be identified using these methods. In addition, while a combination of nuclear magnetic resonance (NMR) and DART-MS can be used for compound identification, as was conducted by Marino et al. [16] for the structure determination of synthetic cannabinoids in herbal incense, fairly large amounts of sample that are soluble in a suitable solvent must be available.

In principle, DART-HRMS neutral loss data can be used to circumvent the aforementioned challenges to the identification of NPSs. In previous work, DART-HRMS was used to generate highly informative mass spectra from the analysis of cathinones [2]. The observed peaks enabled the neutral loss or "dark matter" information that is essential to interpreting the spectrum, particularly in terms of elucidating the structures of new compounds, to be extracted. Here, this approach was extended to enable elucidation of new tryptamine structures. However, an added dimension was that neutral loss data derived from collision-induced dissociation (CID) spectra generated at 60 V and 90 V, were fused to generate new virtual spectra which were then subjected to advanced statistical analysis processing.

One approach that has been found to have utility in increasing the prediction accuracy in terms of drawing inferences from chemical data is "data fusion", which improves results by combining the outputs of multiple analytical techniques [17–20]. This technique was exploited in this study through the fusion of DART-HRMS 60 V and 90 V neutral loss data, which served to increase the number of mass spectral peaks that could be used for structure determination. Accordingly, using statistical analysis techniques described herein, tryptamines were clustered into groups according to the structural information embedded in their fused neutral loss data. This allowed for the skeletal framework of the structures in each group to be revealed. The observed clusters were then used

as the input for the creation of a supervised classification model that could identify unknown tryptamine structures. Thus, the developed method can be used to screen emerging tryptamines against the clusters in order to identify their skeletal frameworks and structures. This novel approach represents a significant advancement over current methodologies for the identification of novel structural variants of drugs through: (1) generation of alternative fragmentation patterns that enable retention of the protonated precursor typically not seen in their EI mass spectra; (2) utilization of data fusion which serves to provide greater amounts of interpretable information and increase the accuracy of the results; and (3) application of machine learning to fused neutral loss spectra to create a prediction model to aid in the structure elucidation of unknown NPSs with a reporting of the statistical level of certainty.

2. Materials and methods

2.1. Materials

Fifty tryptamine standards were purchased from Cayman Chemical Company (Ann Arbor, Michigan, USA) for creating the training set for the statistical model: α-ethyl-4-methyl-1H-indole-3-ethanamine; α-ethyl-5-methoxy-1H-indole-3-ethanamine; *N*-ethyl-*N*-methyl-1H-indole-3-ethanamine; α-ethyl-1H-indole-3-ethanamine; 5,7-dichloro-1H-indole-3-ethanamine; α-methyl-1H-indole-5-ethanamine; 5-methoxy-α-methyl-1H-indole-3-ethanamine; α-methyl-1H-indole-6-ethanamine; 7-fluoro-1H-indole-3-ethanamine; 5-(nonyloxy)-1H-indole-3-ethanamine; 3-(2-aminoethyl)-1H-indol-5-ol; α-methyl-1H-indole-3-ethanamine; 3-[2-(methylamino)ethyl]-1H-indol-5-ol; *N*-methyl-1H-indole-3-ethanamine; *N*,*N*-dipropyl-1H-indole-3-ethanamine; *N*,*N*-*bis*(1-methylethyl)-1H-indole-3-ethanamine; *N*,*N*-diethyl-1H-indole-3-ethanamine; *N*,*N*-dimethyl-1H-indole-3-ethanamine; *N*-isopropyl-*N*-(2-(4-methoxy-1H-indol-3-yl)ethyl)propan-2-amine; 5-methoxy-*N*,*N*-di-2-propen-1-yl-1H-indole-3-ethanamine; *N*,*N*-dibutyl-5-methoxy-1H-indole-3-ethanamine; *N*,*N*-diethyl-5-methoxy-1H-indole-3-ethanamine; 5-methoxy-*N*,*N*-dimethyl-1H-indole-3-ethanamine; 5-methoxy-*N*,*N*-dipropyl-1H-indole-3-ethanamine; *N*-isobutyl-*N*-(2-(5-methoxy-1H-indol-3-yl)ethyl)-2-methylpropan-1-amine; 5-methoxy-*N*,*N*-*bis*(1-methylethyl)-1H-indole-3-ethanamine; *N*-ethyl-*N*-(2-(5-methoxy-1H-indol-3-yl)ethyl)propan-1-amine; *N*-ethyl-5-methoxy-*N*-(1-methylethyl)-1H-indole-3-ethanamine; 5-methoxy-*N*-methyl-*N*-(1-methylethyl)-1H-indole-3-ethanamine; 6-methoxy-*N*,*N*-*bis*(1-methylethyl)-1H-indole-3-ethanamine; *N*-isopropyl-*N*-(2-(7-methoxy-1H-indol-3-yl)ethyl)propan-2-amine; *N*-[2-(5-hydroxy-1H-indol-3-yl)ethyl]-acetamide; *N*-[2-(1H-indol-3-yl)ethyl]-acetamide; 3-[2-(diethylamino)ethyl]-1H-indol-4-ol; 3-[2-(dimethylamino)ethyl]-1H-indol-4-ol; 3-[2-(dipropylamino)ethyl]-1H-indol-4-ol; 3-[2-[*bis*(1-methylethyl)amino]ethyl]-1H-indol-4-ol; 3-[2-(ethylmethylamino)ethyl]-1H-indol-4-ol; 3-[2-(methylpropylamino)ethyl]-1H-indol-4-ol; 3-[2-[methyl(1-methylethyl)amino]ethyl]-1H-indol-4-ol; 3-[2-(dimethylamino)ethyl]-1H-indol-4-ol 4-(dihydrogen phosphate); *N*,*N*-diethyl-6-fluoro-1H-indole-3-ethanamine; 3-[2-(dimethylamino)ethyl]-*N*-methyl-1H-indole-5-methanesulfonamide; 3-[2-(diethylamino)ethyl]-1H-indol-4-ol 4-acetate; 3-[2-(dimethylamino)ethyl]-1H-indol-4-ol 4-acetate; 3-[2-(dipropylamino)ethyl]-1H-indol-4-ol 4-acetate; 3-[2-[*bis*(1-methylethyl)amino]ethyl]-1H-indol-4-ol 4-acetate; 3-[2-(ethylmethylamino)ethyl]-1H-indol-4-ol 4-acetate; 3-[2-[methyl(1-methylethyl)amino]ethyl]-1H-indol-4-ol 4-acetate; and 3-[2-(dimethylamino)ethyl]-1H-indol-5-ol. The structures of these compounds, their assigned abbreviations and formal names are shown in Fig. 1. To assess the prediction model's ability to identify tryptamine classes, the following four additional tryptamines were purchased from Cayman Chemical Company (Ann Arbor, Michigan, USA) for external validation: 3-(2-(allyl(methyl)amino)ethyl)-1H-indol-4-yl acetate; 3-[2-(methylpropylamino)ethyl]-1H-indol-4-ol 4-acetate; 3-(2-(allyl(methyl)amino)ethyl)-1H-indol-4-ol; and 3-[2-(dimethylamino)ethyl]-1H-indol-4-ol 4-propanoate.

2.2. Instrumentation

A DART-SVP ion source (IonSense, Saugus, MA, USA) interfaced with a JEOL AccuTOF mass spectrometer (JEOL USA, Peabody, MA, USA) was used to collect mass spectral data in positive-ion mode. The optimized instrument parameter settings were as follows: helium gas flow rate, 2.0 L/min; gas temperature, 350 °C; DART ion source grid voltage, 50 V; ring lens voltage, 5 V; orifice 1 voltage, 20 V, 60 V, and 90 V; orifice 2 voltage, 5 V; and peak voltage, 400 V (to detect m/z values $\geq$ 40). The samples were analyzed by dipping the closed end of a melting point capillary tube into the sample and presenting the coated surface to the open-air space between the mass spectrometer inlet and ion source. Spectra were collected at a rate of one spectrum per second over the mass range m/z 40–1000. Each sample was analyzed by DART-HRMS three times, and the three spectra were averaged to generate a single representative spectrum. This was repeated ten times in order to produce ten replicates of each sample. PEG 600 (Sigma-Aldrich, St. Louis, MO, USA) was used as a mass calibrant. TSSPro 3 software (Shrader Analytical, Detroit, MI, USA) was used for data processing including averaging, background subtraction, and peak centroiding. The workflow for data collection by DART-HRMS is shown in Fig. 2A. The soft ionization spectra (collected at 20 V) for the 50 tryptamines, along with their structures are presented in Figure S1. To assess the reproducibility of the results, DART-HRMS analyses of 4- acetoxy MPT were performed by three different individuals in one day and by one individual on two different days, one year apart.

2.3. Neutral loss spectra generation

Mass Mountaineer software (RBC Software, Portsmouth, NH, USA) was used to determine the masses lost during the fragmentation that was induced when the orifice 1 voltage was raised from 20 V to 60 V, and then to 90 V. For the spectra generated at each voltage, the m/z values above a 0.5% relative peak intensity were subtracted from the peak representing the protonated tryptamine precursor (as illustrated in Fig. 2B). This furnished "neutral loss" spectra in which the peaks observed represented the high-resolution masses lost during fragmentation, and whose relative intensities were equivalent to those of the peaks in the 60 V or 90 V spectra from which they were derived. Figures S2 and S3 display the neutral loss spectra representative of the 50 tryptamines at 60 V and 90 V, respectively.

2.4. Multivariate statistical analysis

The neutral loss spectra in the form of text files were imported into MATLAB 9.3. 0, R2019a software (The MathWorks, Inc., Natick, MA, USA). Each text file was comprised of two columns, with the first containing the m/z values and the second containing the corresponding peak intensities. In order to reveal patterns that could serve as the basis for classifying tryptamine structures, two procedures (shown in Fig. 2C) were introduced. First, an unsupervised exploratory method was applied to assess whether patterns might be present, and second, a supervised method was applied to create a model for discrimination.

Following DART-HRMS analysis and generation of the corresponding neutral loss spectra, the neutral loss data were aligned in a matrix according to common m/z values with a bin width of 20 mmu and a relative intensity cut-off value of 0.5% (Fig. 2C, Step 1). The optimal bin width and relative intensity cutoff were determined by iterative application of agglomerative hierarchical clustering analysis (HCA) using different bin widths and relative intensity cutoffs. This treatment resulted in two matrices with dimensions of 500 × 174 and 500 × 368 (i.e., number of samples × number of m/z values) for the neutral loss spectra collected at orifice 1 voltages of 60 V and 90 V, respectively. The 60 V and 90 V spectrum of each sample was normalized by the maximum intensity and converted to percentages. Since the neutral loss 60 V and 90 V matrices had the same number of samples, they could be fused row

wise, meaning that the two matrices were fused together to make one matrix with the 60 V data on the left and the 90 V data on the right, in order to create a matrix with a dimension of 500 × 542, which could be subjected to further multivariate statistical analysis, including HCA and creation of a classification model. As shown in Fig. 2C—Step 2, HCA, an unsupervised method, was applied to the resulting data matrix in order to reveal the presence of common patterns that might be used as a basis for classifying the data and predicting tryptamine unknowns. In this process, a correlation measure was used for assessing the proximity between objects in the dataset. Then, the objects were grouped into agglomerative hierarchal clusters using an unweighted pair group method with arithmetic mean (UPGMA). Clusters were defined by cutting branches off the dendrogram. A distance threshold of 70% of the maximum linkage was found to be optimal for the height cut-off value. If the height of a node was less than the cutoff value, all leaves at or below the node were grouped into a cluster. The optimum distance threshold and distance metrics were revealed by changing the threshold and metrics, and evaluating the quality of the resulted clusters using a supervised model, which was trained by the fused matrix and the identified cluster labels. The supervised method examines how well the clusters were separated, and therefore, good clustering should result in good classification performance. Probabilistic partial least square-discriminant analysis (PLS-DA) [21] with 11 latent variables was used to create a supervised model for discrimination of tryptamines in accordance with the detected cluster labels. This is illustrated in Fig. 2C—Step 3. The optimal latent variables for the PLS-DA model were computed using iterative application of different numbers of latent variables to see which led to higher accuracy during validation. Validation of the PLS-DA model was performed using the "leave-one-structure-out" strategy (i.e., in each step of the validation, the replicates of one tryptamine structure were left out of the training process and then subsequently predicted using the trained model). To provide a measure of reliability of classification for each prediction, probability density function and Bayesian decision theory were applied to calculate a vector of the posterior probabilities as output, in which each coordinate corresponded to one class. For assignment of samples to each class, a probability threshold was computed for each class using the prediction results of training samples and the *Bayesian* discrimination threshold [21,22]. The sample was assigned to a class if the PLS-DA prediction output demonstrated a probability higher than that of the class threshold. The optimized discrimination model was then used for prediction of external validation samples.

A specific advantage of PLS-DA is its ability to reveal the *m/z* values that are most significant for enabling discrimination between classes through the determination of variable importance in projection (VIP) scores. Ten distinct PLS-DA discrimination models were explored. Each was used to train a one-vs-all binary model (i.e., a particular class of tryptamines was assigned a label of 1 and all the other classes were assigned a label of 0 in the training process). As such, each PLS-DA model revealed the variables that were most important for discrimination of one class from the other classes. The final VIP scores ended up being the average of the VIP scores from the ten PLS-DA models. Masses with VIP scores higher than 1 were deemed important for enabling discrimination between classes [23].

3. Results and discussion

3.1. Generation and processing of DART mass spectra

This study was conducted to develop a simple and rapid method to enable detection and identification of emerging tryptamine structures. Previous work has demonstrated the utility of extracting neutral loss information from DART mass spectra of synthetic cathinone unknowns in stitching together plausible candidate structures [2]. Here, this principle is extended to enable elucidation of new tryptamine structures. However, the approach differed from that in the cathinone work in that

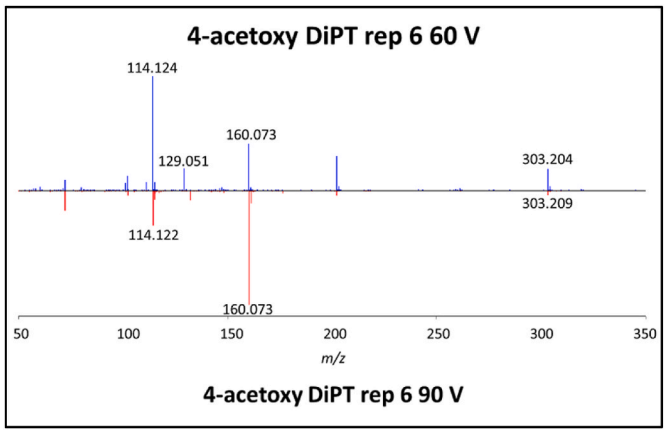


Fig. 3. DART-HRMS-derived neutral loss spectra of 4-acetoxy DiPT analyzed at 60 V (top) and 90 V (bottom).

for cathinones, neutral loss spectra collected under CID conditions at 90 V were utilized, while in this work, CID spectra generated at 60 V *and* 90 V were fused to generate new virtual spectra which were then subjected to advanced statistical analysis processing. The steps taken to achieve this aim are presented in Fig. 2. For each of the tryptamines, spectra in replicates of 10 were collected at 20 V, 60 V and 90 V (Fig. 2A). Analyzing the tryptamines at 20 V resulted in soft ionization spectra, meaning that by and large, only the protonated precursor peaks were observed and there was no fragmentation. 4-Acetoxy DiPT serves as a representative example. With the formula $C_{18}H_{26}N_2O_2$, its calculated monoisotopic mass is 302.199. Its DART mass spectrum acquired at 20 V in positive ion mode (see Figure S1) shows the unfragmented protonated precursor as a single peak at *m/z* 303.207 (within 5 mmu of the calculated mass). When the voltage was raised, fragments that appeared at the expense of the protonated precursor peak were detected (Fig. 3). The higher the voltage, the greater the fragmentation and the lower the intensity of the protonated precursor. Therefore, the 90 V spectrum shows the greatest fragmentation, with a protonated precursor peak of lower intensity when compared to that in the 60 V spectrum. However, it is important to note that even at 90 V, the protonated precursor is still detected, which is generally not the case for EI mass spectra (i.e. the molecular ion in these spectra is often missing). Thus, the head to tail plot of 4-acetoxy DiPT at 60 V(top) and 90 V (bottom) (Fig. 3) shows that these spectra exhibit fragments that appear at the expense of the protonated precursor. Examples include the fragments at nominal *m/z* 114 and 160. While the 60 V and 90 V spectra share several *m/z* values, there were other masses that were unique and appeared as a function of the voltage that was applied. For example, the head to tail plot shows that there is a peak at nominal *m/z* 129 in the 60 V (top) spectrum that is not present in the 90 V (bottom) spectrum. After data collection, the spectra were then processed by subtracting the *m/z* values of each of the product ions from the protonated precursor to yield new neutral loss spectra. This step is summarized in Fig. 2B, and the spectra that were generated are presented in Figures S2 and S3 for the 60 V and 90 V spectra, respectively. In these spectra, the peaks represent high resolution masses of lost structural elements where the relative abundances are based on the fragment peaks from which they were derived.

3.2. Determination of the presence of clustering of tryptamine spectra, indicative of common structural features

To summarize the data and assess whether the fused spectral data might exhibit patterns that could serve as the basis for being able to classify "like" structures, potential similarities between the neutral loss spectral data were determined using the "correlation" metric. This resulted in the correlation matrix shown in Fig. 4, which illustrates the

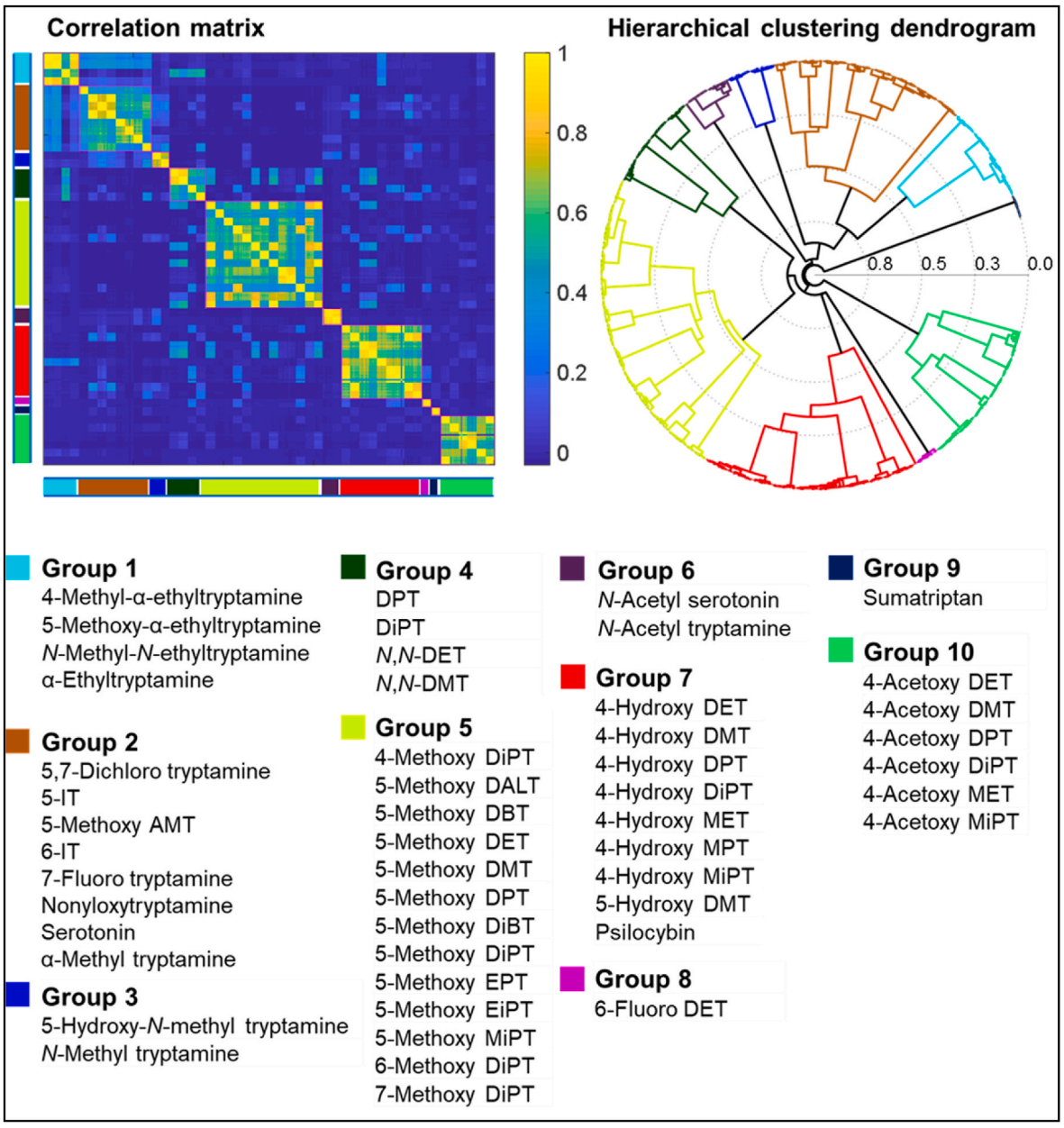


Fig. 4. Correlation matrix showing the computed correlations between molecules, plotted along both the x- and y-axes and arranged by similarity (where yellow corresponds to the highest similarity and blue corresponds to the lowest) that were subjected to hierarchical clustering analysis (HCA). The dendrogram resulting from HCA and the ten groups that emerged are also shown (see full list of compound structures, names and their corresponding abbreviations in Fig. 1).

computed correlations plotted along both the x and y axes, with samples arranged by similarity. It shows the correlations (with values spanning the range 0–1) between the two neutral loss spectra using a color gradient. The color gradient extends from dark blue to yellow, where the darkest shade of blue corresponds to the least similarity (i.e., a value of zero) and the brightest shade of yellow corresponds to the highest similarity (i.e., a value of 1). From the matrix, 10 groups emerged whose neutral loss spectra were similar enough that they were grouped together and distinguished from one another. These are labeled groups 1–10 (Fig. 4) and color-coded as follows: aqua (group 1); brown (group 2); blue (group 3); dark green (group 4); chartreuse (group 5); purple (group 6); red (group 7); magenta (group 8); navy (group 9) and green (group 10). The matrix revealed that the neutral loss data of like tryptamines were highly correlated, based on the bright yellow colors appearing along the diagonal which showed that replicates of each sample always perfectly correlated with itself. Moderate correlations between spectra are also revealed. For example, reading the plot from left to right, the spectral data represented by the aqua color, are shown to be moderately correlated to the spectra represented by the brown color, based on the light blue shading in the off-diagonal area where the two intersect. Furthermore, some of the spectra represented in the aqua area are also correlated to data represented in red on the x-axis.

Next, the data in the correlation matrix served as the input for hierarchical clustering analysis to reveal clustering that might be used as a basis for classifying the data and predicting the structures of tryptamine unknowns. The resulting dendrogram is also presented in Fig. 4. The axial axis in the dendrogram represents the amount of dissimilarity between clusters, and each leaf node displays a replicate of a tryptamine compound. The membership of the ten distinct groups (see full list of compound structures, their names and their corresponding abbreviations in Fig. 1) and their assigned numbers and colors are similar to those in the correlation matrix and are as follows: Group 1 (aqua) includes 4-

methyl-α-ethyl tryptamine, 5-methoxy-α-ethyltryptamine, *N*-methyl-*N*-ethyl tryptamine, and α-ethyltryptamine; Group 2 (brown) includes 5,7-dichloro tryptamine, 5-IT, 5-methoxy AMT, 6-IT, 7-fluoro tryptamine, nonyloxytryptamine, serotonin, and α-methyl tryptamine; Group 3 (blue) includes 5-hydroxy-*N*-methyl tryptamine and *N*-methyl tryptamine; Group 4 (dark green) includes DPT, DiPT, *N*,*N*-DET, and *N*,*N*-DMT; Group 5 (chartreuse) includes 4-methoxy DiPT, 5-methoxy DALT, 5-methoxy DBT, 5-methoxy DET, 5-methoxy DMT, 5-methoxy DPT, 5-methoxy DiBT, 5-methoxy DiPT, 5-methoxy EPT, 5-methoxy EiPT, 5-methoxy MiPT, 6-methoxy DiPT, and 7-methoxy DiPT; Group 6 (purple) includes *N*-acetyl serotonin and *N*-acetyl tryptamine; Group 7 (red) includes 4-hydroxy DET, 4-hydroxy DMT, 4-hydroxy DPT, 4-hydroxy DiPT, 4-hydroxy MET, 4-hydroxy MPT, 4-hydroxy MiPT, 5-hydroxy DMT, and psilocybin; Group 8 (magenta) includes 6-fluoro DET; Group 9 (navy) includes sumatriptan; and Group 10 (green) includes 4-acetoxy DET, 4-acetoxy DMT, 4-acetoxy DPT, 4-acetoxy DiPT, 4-acetoxy MET, and 4-acetoxy MiPT. The dendrogram shows that five groups (1, 2, 4, 5, and 10) are at approximately the same distance on the axial axis, indicating the similarity between them.

3.3. Assessment of the structural basis of the hierarchical clustering analysis-revealed groupings

The groupings that emerged from the clustering results revealed that there are common structural features and fragmentation pathways that are shared by the members of each group which yield similar high-resolution *m/z* values in their neutral loss spectra, and which are thereby highly correlated. The results also made apparent that there are several highly correlated structures whose similarity to one another was not intuitively apparent. For example, the three compounds *N*-methyl-*N*-ethyltryptamine, *N*-methyl tryptamine and *N*,*N*-DET are all mono-substituted indoles in which the substituent at position 3 is comprised of a 2-carbon chain which has at its terminus a nitrogen atom with one or two alkyl groups (see Fig. 1). Based on their highly similar structures, it might be surmised that they would have been grouped together. Yet, the results show that these three compounds were clustered in Groups 1, 3, and 4, respectively. Each of these compounds appears in its respective group based on the similarity of the *m/z* values in its neutral loss spectrum to the other members of the group. For example, Group 1 members all possess a nitrogen-containing three carbon unit (with *m/z* 59.07) that is lost as a neutral fragment from the indole-containing segment of the molecule. The members of Group 5, which contains the greatest diversity of structures, all possess two substituents: a methoxy moiety on the benzene ring (whose position varies), and an ethylamine substituent at position 3 of the indole ring with a range of *N*-alkyl attachments. While the structures are diverse, common fragmentation pathways result in neutral loss spectra that have *m/z* values at nominal 101, 116, 161, and 203, which were important in classifying group 5 and distinguishing its members from those of other groups.

3.4. Creating a structure classification model for tryptamine class prediction

Overall, the correlation matrix and clustering results showed that each of the groups exhibited common features which in principle could serve as a basis for structure classification. To accomplish this, the groupings resulting from HCA were used as the input for PLS-DA. The "leave-one-structure-out" validation strategy was performed to evaluate the PLS-DA model using varying numbers of latent variables. This treatment revealed that eleven latent variables resulted in the lowest error in prediction. The resulting 3D PLS-DA score plot for latent variables 1, 2, and 3 is presented in Figure S4. The score plot shows that each group was discriminated from others. Table S1 shows the confusion matrix that describes the performance of the PLS-DA discrimination model. Each column of the matrix represents the samples in the predicted class, while each row presents the samples in the actual class. The values along the diagonal illustrate the number of samples for which the predicted label matched the true label, while the non-zero values on the off-diagonal represent those that were incorrectly classified or unclassified by the discrimination model. Using class assignment thresholds that were defined by the *Bayesian* discrimination threshold, the model resulted in accuracy, error, multi-label assignment, and not-assigned rates of 1.00, 0.00, 0.05 and 0.04, respectively, in the leave-one-structure-out validation. It should be noted that multi-label and not-assigned predictions were not considered in the accuracy and error calculations.

Table S2 shows the prediction sensitivity, specificity and precision of the discrimination model. Sensitivity and specificity are measures of the 1-false negative rate and 1-false positive rate, respectively. As shown in the confusion matrix (Table S1), the off-diagonal entries for predictions in G2-G6 and G10 are zero, which reflects higher correct prediction rates, and all three figures of merit (i.e. sensitivity, specificity, and precision) are 1.00 for those classes (Table S2). This indicates the model to be good for predicting the classes of emerging structures. Table S3 shows the resulting prediction probabilities of the assignments of tryptamine structures to the classes. Of the 50 tryptamines, all except five (i.e., 5-methoxy-α-ethyltryptamine, *N*-methyl-*N*-ethyltryptamine, 6-fluoro DET, sumatriptan, and 4-hydroxy MET with their probabilities shown in red in Table S3) were correctly classified (i.e. probability of one). Using class assignment thresholds, 5-methoxy-α-ethyltryptamine was assigned to both Group 1 and Group 2 with probabilities of 1.00 and 0.52, respectively. *N*-Methyl-*N*-ethyltryptamine was classified in Group 1 and Group 4 with an equal probability of 1.00 and 1.00, respectively. This is not surprising since all of the compounds in Group 4 and *N*-methyl-*N*-ethyltryptamine have a fragment at nominal *m/z* 131 (Table S4-S5). 4-Hydroxy MET was placed into Group 1 and Group 7 with probabilities of 1.00 and 1.00, respectively. Given the close structural similarity between *N*-methyl-*N*-ethyltryptamine and 4-hydroxy MET, this dual classification was also not surprising, since they both have the *N*-methyl-*N*-ethylamine moiety in their structures, resulting in a fragment at *m/z* 59. 6-Fluoro DET and sumatriptan had a probability of 0.00 for assignment to each cluster since these two compounds belonged to one-member groups (i.e. Groups 8 and 9, respectively). This result was also unsurprising, since the structural characteristics of these compounds deviate quite significantly from the remaining compounds in the training set. Overall, the results confirm that the ability of the model to facilitate accurate prediction of tryptamine classes is dependent upon the similarities of the neutral losses observed from fragmentation of the molecule.

3.5. Assessment of markers important for discrimination between groups of structures

A specific advantage of PLS-DA is that the *m/z* values that are most important for class discrimination can be determined through the evaluation of variable importance in projection (VIP) scores. The VIP score calculation results for the PLS-DA model are presented in Fig. S5. The solid blue points in Fig. S5-Panel A display the VIP scores (x-axis) for *m/z* values (y-axis) that have a score >1. The higher the VIP score, the greater the impact the corresponding *m/z* value has in enabling discrimination. For example, for both the 60 V and 90 V spectra, *m/z* 161.08 has the highest VIP score, indicating that it has the highest impact on discrimination. The heatmaps shown in Fig. S5-Panel B represent the averaged relative abundances of the most impactful neutral loss *m/z* values in each group. The color gradient of the heatmap extends from dark blue to brick red, with the darkest shade of blue indicative of a relative abundance of 0 and the darkest shade of brick red representative of a relative abundance of 100 (on an arbitrary scale). The heatmaps facilitate visualization of how the indicated *m/z* values are associated in different groups, which can be helpful in detecting unique markers. For example, the heatmap results show that the *m/z* value with the highest VIP score for the 60 V data (i.e. *m/z* 161.08) has a

Table 1
The neutral losses observed at 60 and 90 V for the compounds classified into Groups 1 through 10.

Clusters	Neutral loss in 60 V	Neutral loss in 90 V
Group 1	59.07	59.07
Group 2	16.02 and 17.03	17.03 and 31.05
Group 3	30.04, 31.04, 43.04 and 45.06	31.05, 32.05, 43.05, 44.05, 45.06 and 58.07
Group 4	131.07	131.08
Group 5	161.08	161.08
Group 6	17.03, 43.04, 45.06, 58.03, 59.04 and 60.04	17.03, 43.05, 58.03, 59.04, 60.05, 61.06, 86.06 and 88.07
Group 7	147.07	147.07
Group 8	73.09, 149.07 and 161.08	87.11, 101.12, 120.11, 149.07, 161.08 and 177.10
Group 9	45.06, 95.00, 149.07 and 238.08	31.05, 59.07, 138.05, 139.06 and 140.07
Group 10	189.08	189.08

high relative abundance for Group 5 compounds and low relative abundances (around zero) for other groups. Therefore, *m/z* 161.08 is a unique marker that enables the model to distinguish Group 5 molecules from those in the other groups. A similar observation was made for the 90 V data, where *m/z* 161.08 remained a strong marker for Group 5. The relative intensities of the *m/z* values in the 60 V and 90 V neutral loss spectra that were found to be most highly ranked in enabling discrimination between classes are presented in Table S4 and Table S5, respectively. From the neutral loss data presented in the heatmap in Fig. S5 and and the corresponing data shown in Table S4-S5, a number of trends became apparent. First, there are neutral loss values that appear within a given group that are observed for most but not all of the compounds that appear in that group. For example, for the 90V data (Table S5), the neutral loss of *m/z* 90.08 appears for 6 of the 9 compounds in Group 7 (i.e. 4-hydroxy DET; 4-hydroxy DMT; 4-hydroxy MPT; 4-hydroxy MiPT; 5-hydroxy DMT and psilocybin), but not for the other three (4-hydroxy DPT; 4-hydroxy DiPT and 4-hydroxy MET).

Second, there are *m/z* values that are specific to several groups. For example, for the 60 V neutral loss data (Table S4), *m/z* 17.03 is found in groups 1, 2, and 6. Similarly, for the 90 V data (Table S5), *m/z* 59.07 appears in every group except Groups 6 and 9.

Third, there was a subset of neutral losses that were observed for each compound in a given group. These are listed in Table 1 for the 60 V and 90 V data. For example, *m/z* 59.07 is found in all of the compounds in Group 1 in the 60 V and 90 V data. Such masses were found to be important in enabling distinctions to be made between the groups.

These results show that neutral loss spectra derived from DART-HRMS can enable facile classification of tryptamine structures. They also reveal the identities of the skeletal frameworks that were the basis of the ability of the model to distinguish between classes. These are presented in Fig. 5 for Groups 1 through 10.

3.6. Evaluation of the PLS-DA model by external validation using novel compounds

The usefulness of approaches such as those described here lies in part in their potential to correctly classify compounds that are new to the prediction model. Thus, to assess the model's prediction ability, an external validation was performed to determine if tryptamines that were not used in the creation of the model could be correctly classified. Four compounds were tested: 3-(2-(allyl(methyl)amino)ethyl)-1H-indol-4-yl acetate (4-acetoxy MALT); 3-[2-(methylpropylamino)ethyl]-1H-indol-4-ol 4-acetate (4-acetoxy MPT); 3-(2-(allyl(methyl)amino)ethyl)-1H-indol-4-ol (4-hydroxy MALT); and 3-[2-(dimethylamino)ethyl]-1H-indol-4-ol 4-propanoate (4-propanoyloxy DMT). These compounds all contain an ethylamine substituent at the 3-position of the indole scaffold, and a substituent at position 4. They differ in terms of the substituents attached to the exocyclic ring nitrogen and the benzene ring. Each external validation tryptamine was analyzed by DART-HRMS in replicates of 10. Representative neutral loss spectra acquired at 60 V and 90 V for the tryptamines used for external validation are displayed in

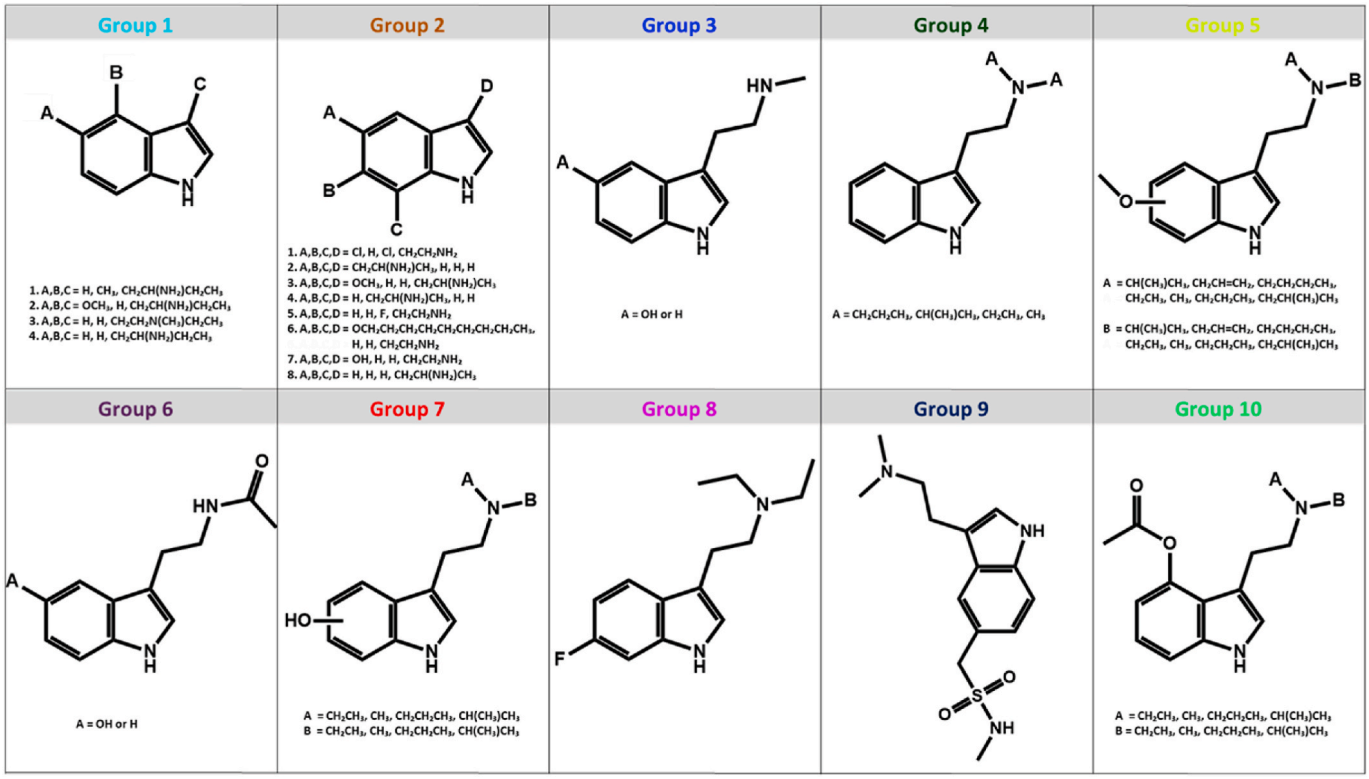


Fig. 5. Skeletal frameworks for each of the 10 classes that emerged from PLS-DA.

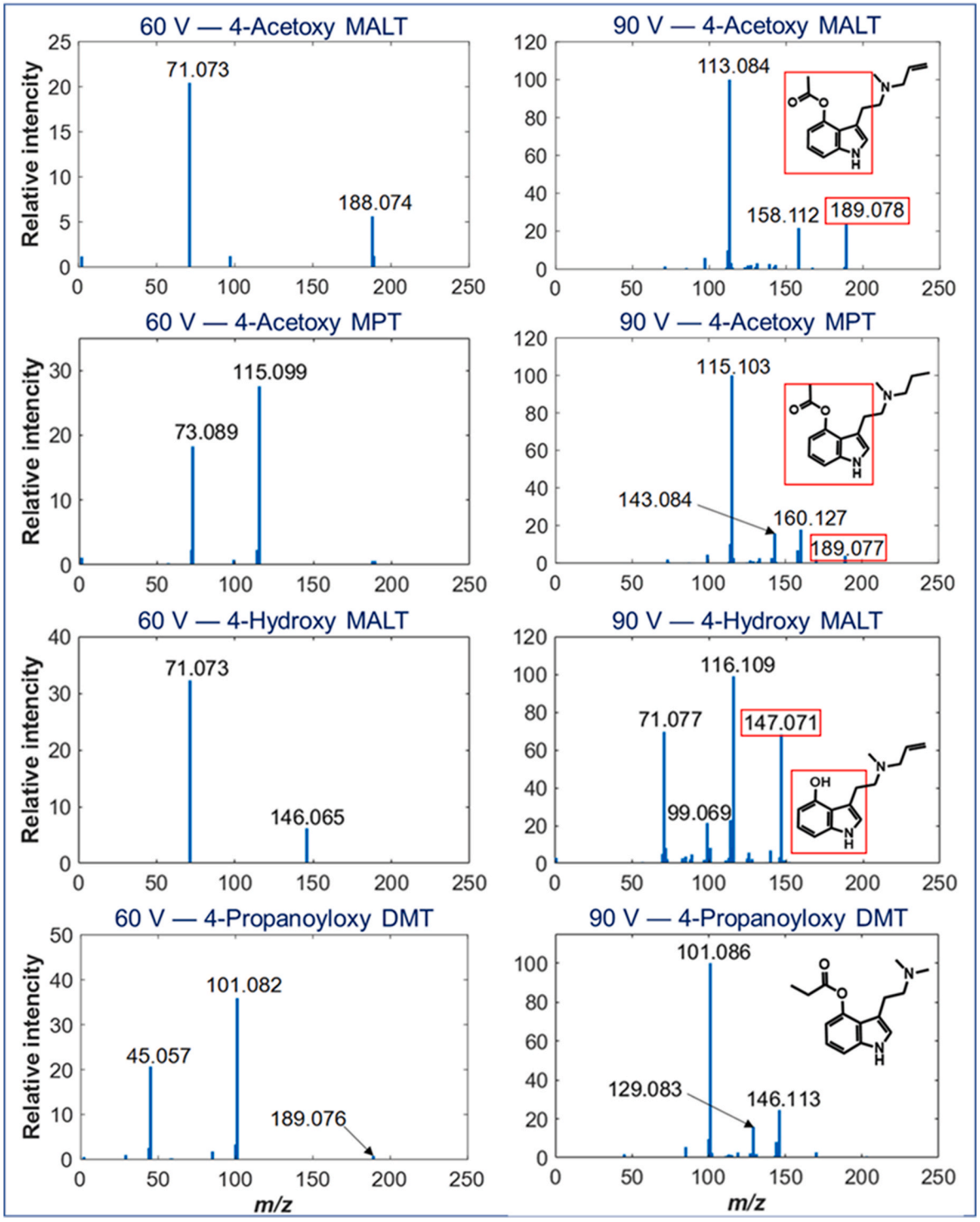


Fig. 6. Neutral loss spectra of the four tryptamines: 4-acetoxy MALT, 4-acetoxy MPT, 4-hydroxy MALT, and 4-propanoyloxy DMT, that were used for the external validation at 60 V and 90 V. The red boxes show the *m/z* values in the neutral loss spectra that are markers for specific groups according to the PLS-DA results. Masses *m/z* 189.08 and 147.07 shown in the red boxes are markers of groups 10 and 7, respectively. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

Fig. 6. From the neutral loss spectra presented and Table 1, it can be seen that three of the tryptamine spectra contain *m/z* values in the neutral loss 90 V spectra that the PLS-DA model revealed are common in the spectra of compounds in specific groups. 4-Acetoxy MALT and 4-acetoxy MPT exhibited a peak at *m/z* 189.08, which is a common *m/z* value for class 10 members. According to Table S5, the *m/z* value 189.08 also appears in the 90 V spectra of some compounds in groups 5 and 7. The 4-hydroxy MALT neutral loss spectrum exhibited a peak at *m/z* 147.07, a class 7 marker. The results for the screening of the four tryptamines against the previously developed hierarchical clustering dendrogram and correlation matrix, along with their corresponding groups are presented in Figure S6. The probabilities of a compound being assigned to a given class are shown in Table S6. The positioning of the unknowns within the correlation matrix is shown in Figure S6 for the compounds

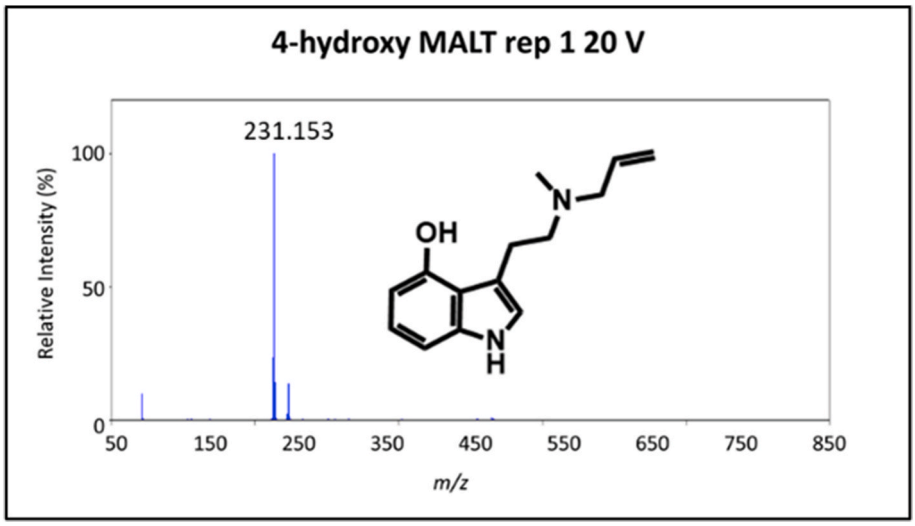


Fig. 7. The 20 V soft ionization spectrum of 4-hydroxy MALT with the protonated precursor labeled at *m/z* 231.153.

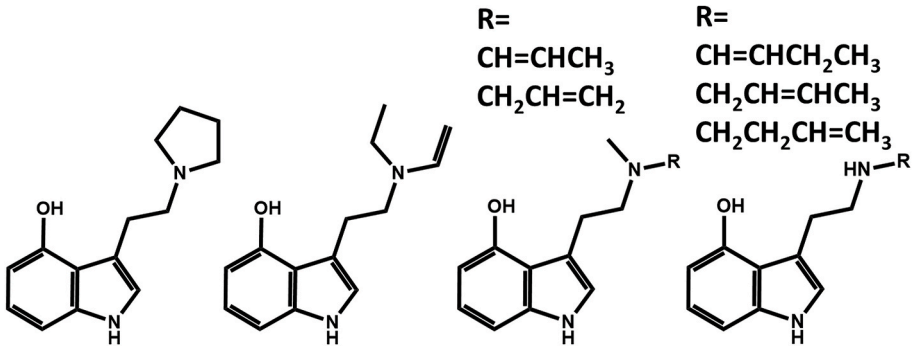


Fig. 8. The possible tryptamine structures of the "unknown". The third structure with the R group –$CH_2CH{=}CH_2$ is the correct structure of 4-hydroxy MALT.

highlighted in the yellow box. Also shown in Figure S6 is where these unknowns fell within the dendrogram, which is indicated with blue shading. 4-Hydroxy MALT fell within the red region of the dendrogram, which represents Group 7, while 4-acetoxy MALT, 4-acetoxy MPT, and 4-propanoyloxy DMT fell within the green region, which represents Group 10. The probabilites presented in Table S6 show that the model was able to correctly classify 35 of the 40 external validation sample replicates. The exceptions were two replicates of 4-acetoxy MALT and three replicates of 4-hydroxy MALT, which the model did not assign to any specific group. However, for all classified replicates, the assignments made were correct.

The results show that the model has the ability to correctly predict detection of new tryptamine structures and reveal indications of their core scaffolds using DART-HRMS data. This enables extraction of their features so that their shared characteristics can be detected (as shown in Table 1). The fusion of the 60 V and 90 V data helped to broaden the range of the fragments that could be interpreted when compared to those detected at either voltage. This effectively enhanced the prediction capacity of the model for detection of new tryptamines.

3.7. Structure elucidation of an unknown Tryptamine

To illustrate how this approach can be utilized to extract information about the core tryptamine scaffold of an unknown, 4-hydroxy MALT will be used as a case in point. As indicated earlier, this molecule was not used to build the PLS-DA prediction model and therefore it can be considered to be an unknown. From the spectrum generated from analysis of the compound at 20 V (Fig. 7), its high-resolution monoisotopic mass (231.153) was found to correspond to the formula $C_{14}H_{19}N_2O$ for the protonated molecule (and thus a formula of $C_{14}H_{18}N_2O$ for the neutral compound). Collection and subsequent fusion of the 60 V and 90 V neutral loss data of this tryptamine and screening the resulting spectrum against the statistical model revealed it to fall into group 7. This is mostly due to the presence of the marker *m/z* 147.071, with the core scaffold revealed to be a hydroxylated indole ring with an *N*-substituted ethylamine appendage (see Fig. 5). The neutral loss fragment that represents this class 7 marker is illustrated in the red box in Fig. 6 and accounts for nine carbons, nine hydrogens, one nitrogen, and one oxygen. When this is subtracted from the molecular formula of $C_{14}H_{18}N_2O$, a balance of five carbons, nine hydrogens, and one nitrogen remain (i.e. C_5H_9N) and this combination of atoms contains one double bond equivalent. Fig. 8 illustrates what the corresponding possibilities are for the substituents on the exocyclic nitrogen, which include the nitrogen as part of a five membered ring; an ethyl and an ethylene substituent; a methyl substituent with various alkenyl substituents; and a hydrogen and a four-carbon alkenyl substituent. In the absence of the predictive model, an analyst conducting a SciFinder search of structures with the molecular formula $C_{14}H_{18}N_2O$ is confronted with a list of approximately 35,000 structures. The approach presented here decreases this number of possibilities by a thousand-fold and provides 7 plausible structures, one of which is correct.

3.8. Data reproducibility

In order to assess the impact that analysis by different individuals and analysis on different days have on intra-laboratory reproducibility

of the prediction results, the DART-HRMS analysis of 4-acetoxy MPT was performed by three different analysts in one day and by a fourth analyst on two different days, but one year apart. The samples were analyzed by DART-HRMS in replicates of 10. The spectra were then subjected to principal component analysis (PCA), the results of which were then examined by Q residuals and Hotelling's T^2 statistic to detect the outliers. Figure S7-A shows a representative PCA scores plot of auto-scaled data (after removing outliers) collected by different analysts on the same day and one analyst on different days. The results show that, except for several replicates acquired by analyst 2, the differences in between multi-day versus multi-person analyses are small. Two approaches were used to quantify the variations: (1) the median of the relative standard deviation (RSD) which, according to Parsons et al. [24], serves as a single summary statistic to reveal reproducibility in metabolomics data; and (2) the Pearson's r coefficient, which is a metric for evaluation of the reproducibility and stability of ambient ionization mass spectral data according to Zhvansky et al. [25].

Overall, the RSD values show greater reproducibility for peaks with higher intensities. The median RSD values for spectra collected for each individual analysis have variations in peak intensities between 10% and 28%, with a value of 24% for analyst-to-analyst variation and 13% for day-to-day variation. The correlation matrix (Figure S7-B) shows the Pearson's r coefficient measure between DART-HRMS-derived neutral loss spectral replicates, revealing the variation between the spectra analyzed by analyst 2. No considerable variation between the chemical profiles that were analyzed by the four analysts on three different days was observed. The image also illustrates the consistency of the results of the analyses, as there is high similarity between the data collected on two consecutive days, and that acquired one year before.

4. Conclusion

The prediction of the skeletal frameworks of NPS tryptamine structures can be accomplished by screening their neutral loss spectra acquired under CID conditions at 60 V and 90 V by DART-HRMS against a PLS-DA prediction model built using the neutral loss mass spectra of 50 tryptamines representing a range of structures. The model revealed 10 groups of tryptamines that were classified based on the similarities of the neutral losses observed when the molecules were fragmented under CID conditions during DART-HRMS analysis. The data generated at 60 V and 90 V were fused in order to expand the range of masses from which structural information could be extracted. "Leave-one-structure-out" validation and the screening of external validation samples (using four tryptamines that were new to the model) were used to assess the prediction capacity of the model. The results showed 100% accuracy, precision, sensitivity, and specificity for the prediction model, as well as 0% error, 5% multi-label assignment, and 4% not-assigned rates. Data from multiple analysts shows that this technique is reproducible. This method provides a rubric for how to facilitate more rapid assessment of the identity of NPS tryptamines when encountering unknowns for which little information on structural identity is initially available.

Credit author statement

MIV performed DART-HRMS experiments and data processing. SB performed statistical analysis. RAM conceived of the project, acquired funding support, supervised its implementation, and interpreted the data. RAM wrote the manuscript. MIV and SB contributed to writing the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The funding support of the U.S. National Science Foundation [Award Number 1429329] and the U.S. National Institute of Justice [Office of Justice Programs, U.S. Department of Justice (DOJ), Award Numbers: 2017-R2-CX-0020 and 2019-DU-BX-0026], as well as the Ford Foundation Initiatives for Women in Science Fellowship is gratefully acknowledged. The opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect those of the DOJ. The contributions of Ms. Allix Coon, Ms. Benedetta Garosi and Ms. Amy Osborne to the DART-MS analysis of 4-acetoxy MPT for determination of data reproducibility are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2022.123417.

References

[1] D. Fabregat-Safont, M. Barneo-Muñoz, F. Martínez-García, J. Sancho, F. Hernández, M. Ibáñez, Proposal of 5-methoxy-*N*-methyl-*N*-isopropyltryptamine consumption biomarkers through identification of in vivo metabolites from mice, J. Chromatogr., A 1508 (2017) 95–105.

[2] K.L. Fowble, J.R. Shepard, R.A. Musah, Identification and classification of cathinone unknowns by statistical analysis processing of direct analysis in real time-high resolution mass spectrometry-derived "neutral loss" spectra, Talanta 178 (2018) 546–553.

[3] A. Wohlfarth, W. Weinmann, S. Dresen, LC-MS/MS screening method for designer amphetamines, tryptamines, and piperazines in serum, Anal. Bioanal. Chem. 396 (7) (2010) 2403–2414.

[4] M.R. Meyer, A. Caspar, S.D. Brandt, H.H. Maurer, A qualitative/quantitative approach for the detection of 37 tryptamine-derived designer drugs, 5 β-carbolines, ibogaine, and yohimbine in human urine and plasma using standard urine screening and multi-analyte approaches, Anal. Bioanal. Chem. 406 (1) (2014) 225–237.

[5] M. Katagi, T. Kamata, K. Zaitsu, N. Shima, H. Kamata, K. Nakanishi, H. Nishioka, A. Miki, H. Tsuchihashi, Metabolism and toxicologic analysis of tryptamine-derived drugs of abuse, Ther. Drug Monit. 32 (3) (2010) 328–331.

[6] T. Kamata, M. Katagi, H. Tsuchihashi, Metabolism and toxicological analyses of hallucinogenic tryptamine analogues being abused in Japan, Forensic Toxicol. 28 (1) (2010) 1–8.

[7] Y. Nakazono, K. Tsujikawa, K. Kuwayama, T. Kanamori, Y.T. Iwata, K. Miyamoto, F. Kasuya, H. Inoue, Simultaneous determination of tryptamine analogues in designer drugs using gas chromatography–mass spectrometry and liquid chromatography–tandem mass spectrometry, Forensic Toxicol. 32 (1) (2014) 154–161.

[8] A. Jovel, A. Felthous, A. Bhattacharyya, Delirium due to intoxication from the novel synthetic tryptamine 5-MeO-DALT, J. Forensic Sci. 59 (3) (2014) 844–846.

[9] S.E. Rodriguez-Cruz, Analysis and characterization of designer tryptamines using electrospray ionization mass spectrometry (ESI-MS), Microgram. J. (2005) 107–112.

[10] R. Kikura-Hanajiri, M. Hayashi, K. Saisho, Y. Goda, Simultaneous determination of nineteen hallucinogenic tryptamines/β-calbolines and phenethylamines using gas chromatography–mass spectrometry and liquid chromatography–electrospray ionisation-mass spectrometry, J. Chromatogr. B 825 (1) (2005) 29–37.

[11] C. Huhn, M. Pütz, N. Martin, R. Dahlenburg, U. Pyell, Determination of tryptamine derivatives in illicit synthetic drugs by capillary electrophoresis and ultraviolet laser-induced fluorescence detection, Electrophoresis 26 (12) (2005) 2391–2401.

[12] S.D. Brandt, R. Tearavarich, N. Dempster, N.V. Cozzi, P.F. Daley, Synthesis and characterization of 5-methoxy-2-methyl-*N, N*-dialkylated tryptamines, Drug Test. Anal. 4 (1) (2012) 24–32.

[13] K. Piorunska-Sedlak, K. Stypulkowska, Strategy for identification of new psychoactive substances in illicit samples using attenuated total reflectance infrared spectroscopy, Forensic Sci. Int. 312 (2020) 110262. https://www.sciencedirect.com/science/article/pii/S0379073820301249?via%3Dihub.

[14] L.E. Jones, A. Stewart, K.L. Peters, M. McNaul, S.J. Speers, N.C. Fletcher, S.E. Bell, Infrared and Raman screening of seized novel psychoactive substances: a large scale study of >200 samples, Analyst 141 (3) (2016) 902–909.

[15] A.S. Moorthy, E. Sisco, A new library-search algorithm for mixture analysis using DART-MS, J. Am. Soc. Mass Spectrom. 32 (7) (2021) 1725–1734.

[16] M.A. Marino, B. Voyer, R.B. Cody, A.J. Dane, M. Veltri, L. Huang, Rapid identification of synthetic cannabinoids in herbal incenses with DART-MS and NMR, J. Forensic Sci. 61 (1) (2016) S82–S91, https://doi.org/10.1111/1556-4029.12932. PMID: 26331594.

[17] M.C. Santos, F. Pereira, Direct analysis of human hair before and after cosmetic modification using a recent data fusion method, J. Braz. Chem. Soc. 31 (1) (2020) 33–39.

[18] K.M. Nunes, M.V.O. Andrade, A.M. Santos Filho, M.C. Lasmar, M.M. Sena, Detection and characterisation of frauds in bovine meat in natura by non-meat

ingredient additions using data fusion of chemical parameters and ATR-FTIR spectroscopy, Food Chem. 205 (2016) 14–22.
[19] T. Trejos, P. Torrione, R. Corzo, A. Raeva, K. Subedi, R. Williamson, J. Yoo, J. Almirall, A novel forensic tool for the characterization and comparison of printing ink evidence: development and evaluation of a searchable database using data fusion of spectrochemical methods, J. Forensic Sci. 61 (3) (2016) 715–724.
[20] C.R. Carneiro, C.S. Silva, M.A. de Carvalho, M.F. Pimentel, M. Talhavini, I. T. Weber, Identification of luminescent markers for gunshot residues: fluorescence, Raman spectroscopy, and chemometrics, Anal. Chem. 91 (19) (2019) 12444–12452.
[21] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models, PLS-DA, Anal. Methods-UK 5 (16) (2013) 3790–3798.
[22] N.F. Pérez, J. Ferré, R. Boqué, Calculation of the reliability of classification in discriminant partial least-squares binary classification, Chemometr. Intell. Lab. Syst. 95 (2) (2009) 122–128.
[23] M. Cocchi, A. Biancolillo, F. Marini, Chapter ten - chemometric methods for classification and feature selection, Compr. Anal. Chem. 82 (2018) 265–299.
[24] H.M. Parsons, D.R. Ekman, T.W. Collette, M.R. Viant, Spectral relative standard deviation: a practical benchmark in metabolomics, Analyst 134 (2009) 478–485.
[25] E.S. Zhvansky, S.I. Pekov, A.A. Sorokin, V.A. Shurkhay, V.A. Eliferov, A. A. Potapov, E.N. Nikolaev, I.A. Popov, Metrics for evaluating the stability and reproducibility of mass spectra, Sci. Rep. 9 (1) (2019) 914.