POSTER NOTE

NO. 64895

Creating and utilizing HRAM MS/MS and MS^n libraries for unknown analysis

Marta Kozak,[1]Tim Stratton,[1] Robert Mistrik[2]
[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]HighChem, Bratislava, Slovakia

ABSTRACT

Purpose: To develop a reference library of known chemical entities and utilize it for complex biological sample analysis to reduce complexity for unknown / drug detection.

Methods: A reference library was created (with continuous new compound addition) which utilized both high energy collisional dissociation (HCD) and trap collisional dissociation (CID) to generate both MS/MS and MS^n fragments from several thousand known chemical standards. Known "clean" and "suspect" samples of biological matrices (human urine and plasma) were analyzed by HRAM LC-MS with the subsequent data processed to detect all peaks. Peaks were assigned identification based on fragmentation spectral matches to the library to identify components.

Results: Utilizing this approach, it is possible to reduce the complexity of new component detection (drugs, adulterants, doping agents, etc) by identifying the "components of no interest", namely those of endogenous nature or coming from known environmental contaminants/exposure, personal care products, and foods.

INTRODUCTION

The detection and identification of emerging designer drugs, food or herbal adulterants, or new doping agents is a complex challenge in forensic analysis. The complexity of the biological matrices studied, which can contain many thousands of endogenous components, is increased by the possible presence of many other variable compounds which come from environmental exposure, the use of personal care products, food, packaging, and food ingredients, as well as other sources. The resulting matrix may contain as many as 3,000 to 4,000 unique chemical entities or more in which potential analytes of interest may be hidden. In this work we explore the application of a high resolution, accurate mass (HRAM) fragmentation library of several thousand known compounds as a means to reduce sample complexity by identifying as many components as possible when combined with other potential identification techniques.

MATERIALS AND METHODS

Sample Preparation

Several samples of human plasma and urine containing no illicit drug or adulterant were prepared for analysis. Urine (1.4 mL total volume) was prepared by solid phase extraction (C18) to remove salts and concentrate analytes. The urine was eluted with one volume each of acetonitrile and methanol, evaporated to dryness under N_2, and reconstituted in 280uL of 90:10 water:acetonitrile. Plasma was prepared by precipitation of proteins with 2 volumes of cold acetonitrile followed by centrifugation (10 min, 13,000 RPM). The supernatant was directly injected on column.

Mass Spectrometer Acquisition Conditions

Mass spectrometer: Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap™ mass spectrometer

LC System: Thermo Scientific™ Vanquish™ HPLC system

Samples were injected (5uL) onto the LC system with separation achieved on a 50X2.1, 2.6u column (Accucore™ C18) using a matrix specific gradient of water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). Matrix specific gradients are shown in **Table 1**.

Table 1. LC Gradients

Urine			Plasma		
Time	%A	%B	Time	%A	%B
0	98	2	0	98	2
1	98	2	1	98	2
6	70	30	8	5	95
8	5	95	8.5	5	95
8.5	5	95	9	98	2
9	98	2	10	98	2
10	98	2			

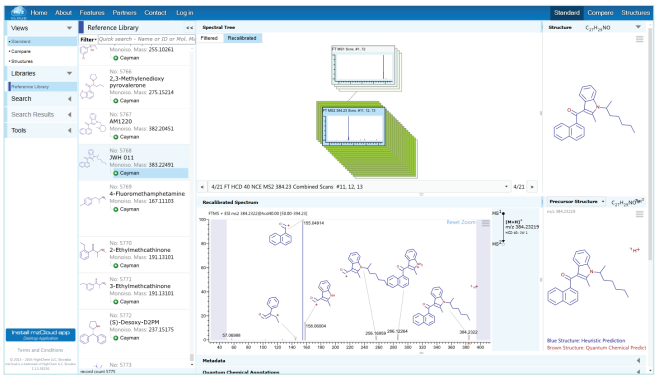
HRAM LC-MS Acquisition – Library Creation

Library spectra were acquired in accordance with the standard operating procedure for mzCloud™ which dictates that fragmentation data be acquired on pure standards through infusion or flow injection with acquisition of MS^2 data for energies between 10% to 100% normalized collision energy (NCE). In addition, many compounds were taken beyond the conditions required in the SOP by going up to 200% NCE. Data was acquired in both positive and negative modes using both electrospray ionization (ESI) with atmospheric pressure chemical ionization (APCI) used for compounds that did not ionize by ESI. In addition, replicate fragmentation spectra (n=3) were acquired at each energy level. In order for acquisition to be triggered, signal stability was confirmed and isolation purity of greater than 90% was required. The isolation purity was rechecked during the acquisition every fifteen scans to confirm the purity of the acquired spectra.

Fragmentation Spectral Library Creation

After acquisition, all fragmentation data was submitted for full curation prior to uploading to the fragmentation library. Briefly, the curation process involves review by a trained chemist/mass spectroscopist with determination of spectral noise using replicate scans to remove electronic and chemical noise from the resulting average spectra for each energy level. The next step involved the recalibration of each spectra using a combination of predicted fragmentation and elemental composition to generate the recalibrated data. Finally, the recalibrated data was annotated for the structure and elemental composition of fragment ions. The data was then uploaded for public use on the mzCloud fragmentation spectral library (mzCloud.org, **Figure 1**) as well as direct searching from the data processing software used for this work. Acquisition of data was performed on Orbitrap™ based mass spectromters including Thermo Scientific Q Exactives and Orbitrap hybrids.

Figure 1. mzCloud.org Library Web Interface



HRAM LC-MS Acquisition – Sample Analysis

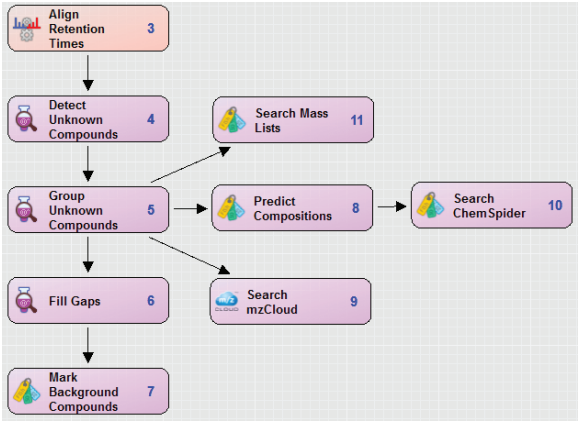
Acquisition of test samples was performed on a Thermo Scientific Q Exactive mass spectrometer (**Figure 2**) with full MS scans at 60,000 resolution (FWHM @ 200 m/z) with data dependent MS^2 triggered from detected components. The MS^2 fragmentation data was acquired at 15,000 resolution using HCD fragmentation (50% normalized collision energy). Data dependent MS^2 provides unambiguous, precursor ion selected MS^2 data which is superior for spectral library searching compared to various independent acquisition approaches. In order to assure a more complete coverage for analytes, the untargeted peak detection data for the suspect and clean samples was compared with any trace level peaks in suspect samples not having matching MS^2 fragmentation acquired in a subsequent inclusion triggered analysis.

Figure 2. Thermo Scientific Q Exactive mass spectrometer and Vanquish UHPLC system.

Data Processing

The acquired data was processed using Thermo Scientific™ Compound Discoverer™ 2.0 software with a workflow built for unknown peak detection, comparison between samples, and identification. Data processing in Compound Discoverer software utilizes customizable workflows which assemble multiple nodes to create an application specific approach. In this way it is clearer how the data flows through the processing. The workflow used for this work is shown in **Figure 3** and includes the retention alignment of data from multiple injections followed by unknown component detection. In this step, peaks are detect and isotopes are grouped to create features. These features are grouped for the different ionization adducts based on retention and mass differences to create components from which an individual molecular weight for each chemical can be determined. This reduces the complexity of the mass spectral data which can have many tens of thousands of "chromatographic features", *m/z* values which have a chromatographic peaks shape, into a list of the relevant compounds.

Figure 3. Processing Workflow in Compound Discoverer 2.0



Detected peaks were compared across the different samples to properly group together the same component detected in multiple samples. Fragmentation spectra from all components was batch searched against the mzCloud fragmentation library online with matches returned for visualization within Compound Discoverer. Fragmentation spectral library matching is not the only tool for determining possible compound identification, however. From the list of relevant compounds, elemental compositions could be predicted. Elemental composition was performed using a suitable open set of elements in use and ranges appropriate for a true unknown detection (C90 H190 Br3 Cl4 N10 O15 P2 S5 F6). The elemental composition algorithm makes use of the high resolution fine isotopic information available in the full scan data to improve the determination of potential elemental compositions. The compositions could be searched against multiple data sources for putative possible components as an additional approach to fragmentation library searching.

RESULTS

Simplifying the Sample Matrix

Often, most of the unknown components in a sample are not of interest in the context of the analysis, such as the detection of new drugs, adulterants, or analogues. To reduce the complexity we can begin binning compounds into broad general categories.

Unknown of no interest – Component frequently observed in the matrix being studied, potentially derived from environmental exposure, personal care products, food, or an unknown endogenous compound. Comparison of suspect samples against historical "clean" samples can help separate these.

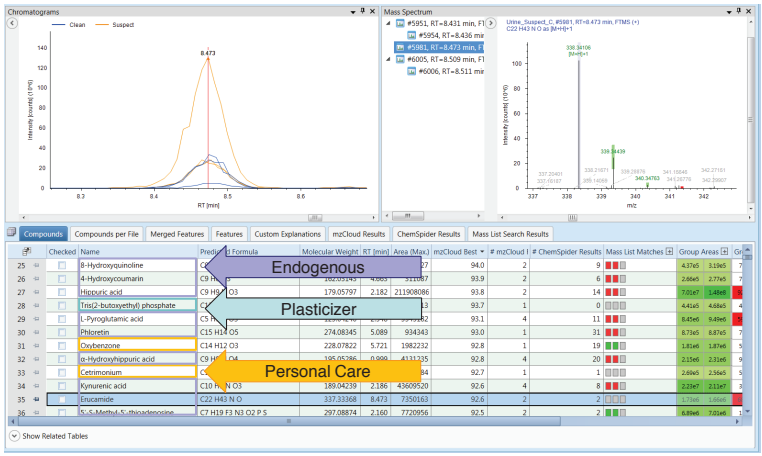
Knowns of no interest – Compounds assigned identifications based on fragmentation library search (with or without retention time) and known to be components which are not interesting in the context of our analysis. Uncovering the compounds in this category is the focus of this work.

Knowns of interest – Putative identifications or similarity results from fragmentation library search against previously known compounds of interest or potentially detected metabolites of known compounds.

Unknowns of interest – Compounds found in suspect samples that have not previously been observed only observed in previous suspect samples.

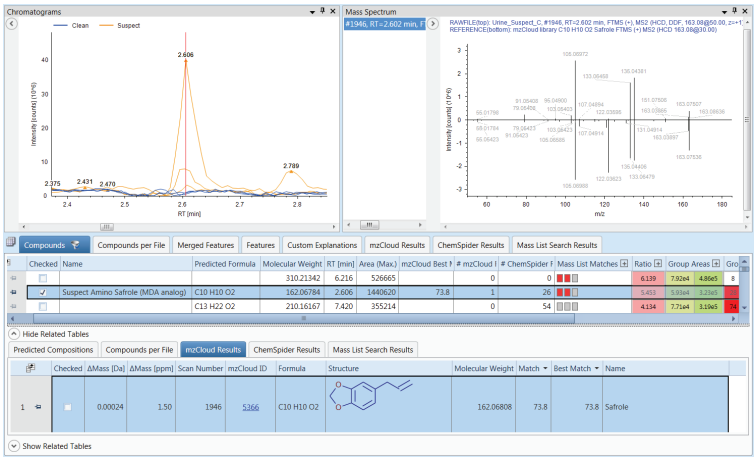
The unknown compounds in all samples of urine and plasma were searched against mzCloud and the identified compounds were reviewed. This allowed the determination of "knowns of no interest" which included endogenous compounds, environmental / food packaging exposure, personal care products, etc. (**Figure 4**).

Figure 4. Identified Components by Fragmentation – Knowns of No Interest



In general, plasma contained several hundred compounds (200-325) while urine contained between 2000-3000 compounds. This is perhaps not surprising given the concentrating nature of the kidney/urine. Fragmentation library matches and putative identifications could be proposed for an average of 25-35% of components in all samples with 12-18% of plasma components being given identification from fragmentation library matches and 8-12% of urine components matching fragmentation library spectra. This resulted in a significant reduction of overall sample complexity. In addition, this approach also helped to identify a putative drug of abuse in a suspect sample through fragmentation similar to safrole with a simultaneous observation of source fragmentation of an amine loss which indicated an amino safrole, potentially an analog of MDA. (**Figure 5**)

Figure 5. Suspect Detected Compound – Amino Safrole / MDA Analog



Determining Potential Components of Interest

An additional approach to determining potential unknowns of interest was the comparison of peak detection from suspect and known 'clean' samples, looking for components with significantly higher detection in the suspect urines and plasmas. This approach detected several components of interest, which may arise from several reasons including:

- Highly variable endogenous components
- Different personal care products used by individuals
- Variations in diet not reflected by the "clean" samples.
- Drugs, adulterants, doping agents.

By scrutinizing the significantly different components detected and performing similarity fragmentation library searches, potential components of interest can be determined. A similarity fragmentation search does not require that the precursor ion mass matches (different, but similar precursor structures will provide similar fragmentation) or will accept lower fragmentation match scores (which may still indicate a degree of parent structure similarity, such as from a positional isomer or analog).

One such component identified in urine was a compound which only showed up in 2 suspect samples (**Figure 6**). The fragmentation was manually submitted to mzCloud for a similarity search (**Figure 7**). The search returned a hit for sufentanil indicating the peak may indeed be low level sufentanil or a sufentanil analog.

Figure 6. Potential Compound of Interest – Significant Change in Suspect vs "Clean"

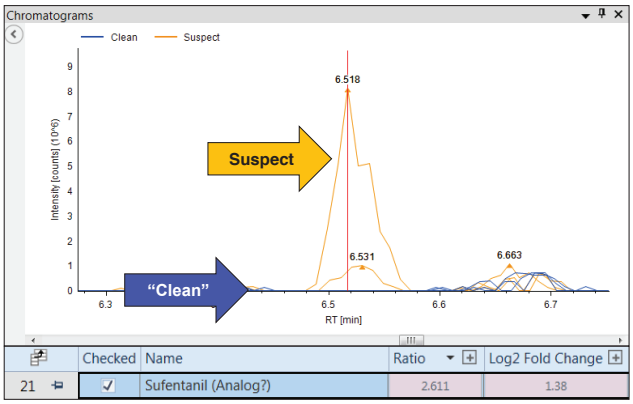
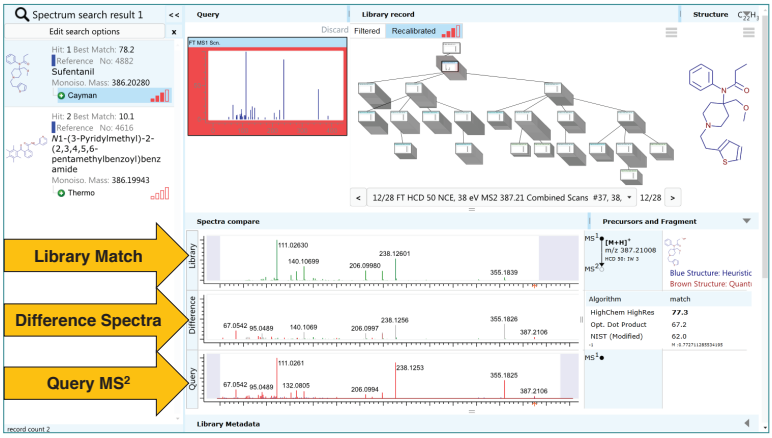


Figure 7. Similarity Fragmentation Search for Unknown of Interest.



CONCLUSIONS

- The utilization of a chemically diverse library can help to reduce biological matrix complexity by providing identification of "Knowns of no interest" in the context of attempts to detect drugs, adulterants, or doping agents.
- The reduction in complexity can be combined with approaches to determine components of interest such as fragmentation similarity or comparison of suspect samples to historical controls to highlight components of interest.

Forensic use only

Find out more at **thermofisher.com**

ThermoFisher SCIENTIFIC

PN64895-EN 11/16S