LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-PROC-774099

Method Development for the Identification of Trichothecenes: Mass Spectral Library Matching and Determination of Unknown Mycotoxins

M. C. Prieto Conaway, M. Dreyer, T. H. Corzett, B. P. Mayer, A. M. Williams

May 13, 2019

67th Conference on Mass Spectrometry and Allied Topics, ASMS
Atlanta, GA, United States
June 2, 2019 through June 6, 2019

Disclaimer

Method Development for the Identification of Trichothecenes: Mass Spectral Library Matching and Determination of Unknown Mycotoxins

Maria C. Prieto Conaway[1], Mark Dreyer[1], Todd H. Corzett[1], Brian P. Mayer[1] and Audrey M. Williams[1]
[1]Forensic Science Center, Lawrence Livermore National Laboratory, Livermore, CA, USA, 94551

ABSTRACT

Purpose: To devise a workflow for mycotoxin structure elucidation without the use of reference standards using a commercial data analysis software and open source mass spectral library.

Methods: High resolution accurate mass (HRAM) infusion data were acquired and curated for adding to mass spectral libraries. A software package for mass spectral and chromatographic data evaluation with similarity and substructure search algorithms was tested.

Results: Two trichothecene compound classes were studied using HRAM LC-MS-MS^2, *in-silico* fragmentation and precursor ion fingerprinting. Substructures present in an unknown compound were matched against MS-MSn spectral library and databases. The suggested structure offers a path forward for unknown compound identification[1].

INTRODUCTION

Trichothecenes are mycotoxins that share a common tricyclic 12,13-epoxytrichothec-9-ene (EPT) moiety. Over 200 trichothecenes exist and are classified into 4 groups (types A-D) depending on the substitution pattern of EPT and structural similarities. As types A and B are commonly found in crops, screening and quantitative mass spectrometric (MS) methods are abundant in the literature; however, these methods almost exclusively rely on reference standards for confirmation.

For some exposure scenarios, whether natural or nefarious, rapid and confident identification of the exact trichothecene is required. Reference standards, however, may not be available for all trichothecenes on acceptable timescales or cost. Our goal is to develop liquid chromatography-mass spectrometry (LC-MS) methods for mycotoxin identification that do not require reference standards for confirmation.

MATERIALS AND METHODS

Sample Preparation

Representative type A and B trichothecene standards (Table 1) were purchased from Cayman Chemical (Ann Arbor, MI).

Solid samples were dissolved to 10 µg/mL with acetonitrile and working solutions of 500 ng/mL were prepared in 2:1:1 H_2O:acetonitrile:methanol for infusion experiments.

A mix containing 7 type A trichothecenes (Table 1) was prepared in 50/50 MeOH:H2O at 100 ng/mL for all except HT-2 which was at 300 ng/mL. 10uL were injected into the LC-MS.

Experimental Methods

Flow infusion data were collected as per Thermo Scientific mzCloud Standard Operating Procedure (https://www.mzcloud.org/). A Thermo Scientific Q Exactive HF-X platforms was used. A 3 min. MS experiment is composed of a full MS spectrum at a resolving power setting of 240,000 (FWHM at m/z 200) followed by MS/MS with higher-energy collisional dissociation (HCD) energies from 10 to 100% in triplicates (a total of 30 MS^2 spectra). Two ionization methods were used: heated electrospray ionization (HESI) and atmospheric pressure chemical ionization (APCI) in positive ion mode. Mobile phases were A: H_2O/10mM Ammonium Formate/10% Methanol, B: Methanol/10% H_2O. Twenty µL were injected using an autosampler (Thermo Scientific Vanquish), through a frit into the MS.

Liquid chromatography mass spectrometry (LC-MS): Thermo Scientific Vanquish Horizon LC system with a Hypersil Gold aQ column (100 X 2.1 mm, 1.9µm). Mobile phases A: H_2O/10mM Ammonium Formate/0.1% Formic acid and B: Methanol/10mM Ammonium Formate/0.1% Formic acid. **Gradient**: a 12 min. gradient, 0-100% B from 1 to 8 min, 2 min at 100% B and equilibration for 2 min at 0%B. A data dependent MS^2 (ddMS^2) experiment was conducted on a sample containing 7 type A trichothecenes.

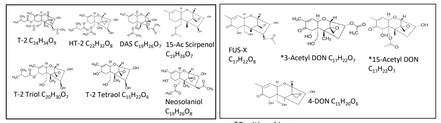
Data Analysis

In silico fragmentation tools, substructure similarity search algorithms (Thermo Scientific Mass Frontier 8.0) and mass spectral library searches (HighChem mzCloud) were tested with HRAM mass spectrometry data collected on a Thermo Scientific Q Exactive HF-X mass spectrometer.

The workflow for creating a trichothecene mass spectral library with data analysis in Mass Frontier was: 1) collect infusion mass spectra of individual compounds, 2) create *.mol structures in the Structure Editor module, 3) predict MS fragments in the Fragments and Mechanisms module, 4) use the Curator module to filter and recalibrate the MS spectra by removing noise, extraneous MS and other artifacts, 5) automatically annotate experimental spectra with predicted fragments, 6) save curated compound to local user library.

RESULTS

Table 1. Type A (left panel) and type B (right panel) trichothecenes.



*Positional isomers

Building Mass Spectral Libraries - Infusion Data Acquisition

Curation improves the quality of data prior to adding to a reference mass spectral library[2]. The goal was to build local libraries and potentially adding new compounds to mzCloud.

Table 2 shows the complementarity of HESI and APCI ionization techniques. The most intense precursor by HESI was the [M+Na]$^+$ adduct, followed by [M+NH4]$^+$ (Table 2, a)). For APCI, the dominant precursors were the [M+NH4]$^+$ and [M+H]$^+$ adducts. In APCI, ionization occurs through ion/molecule reactions at atmospheric pressure making it less susceptible to matrix interferences from salts. In heated electrospray nebulized droplets are desolvated by heat and it is non uncommon to have preformed ions from solution. T-2 tetraol did not ionize by APCI, potentially due to being more thermally labile.

Table 2 a) and b). Infusion data showing areas under the curve (AUC) for all trichothecene standards analyzed by either a) HESI or b) APCI.

a)

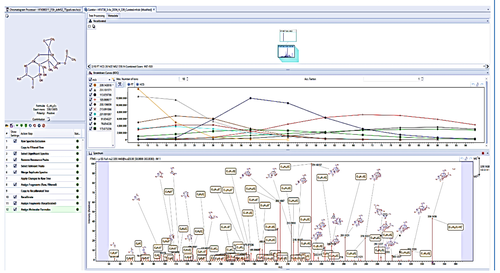
HESI Type A	[M+Na]+	Area EIC* [M+Na]+	[M+NH4]+	Area EIC* [M+NH4]+
T-2	489.2095	5.73E+09	484.2541	1.49E+09
HT-2	447.1989	1.88E+09	442.2435	5.51E+08
DAS	389.1571	7.80E+09	384.2017	2.58E+09
T-2 triol	405.1884	1.55E+09	400.2330	2.94E+08
T-2 Tetraol	321.1309	2.56E+08	316.1755	4.23E+07
Neosolaniol	405.1520	2.24E+09	400.1966	5.33E+08
15-Ac Scirpenol	347.1465	1.57E+09	342.1911	3.62E+08
HESI Type B				
FUS-X	377.1207	1.92E+09	372.1653	1.21E+08
3-Ac DON	361.1258	3.33E+08	356.1704	1.70E+07
15-Ac DON	361.1258	2.20E+09	356.1704	1.66E+08
4 DON	319.1152	8.22E+08	314.1598	1.92E+07

b)

APCI Type A	[M+NH4]+	Area EIC [M+NH4]+	[M+H]+	Area EIC [M+H]+
T-2	484.2541	3.29E+08	467.2280	1.33E+07
HT-2	442.2435	1.09E+07	425.2170	2.65E+07
DAS	384.2017	2.28E+08	367.1750	1.16E+08
T-2 triol	400.2330	2.03E+08	383.2060	6.21E+06
T-2 Tetraol	316.1755	no signal	299.1490	no signal
Neosolaniol	400.1966	3.07E+07	383.1700	5.95E+05
15-Ac Scirpenol	342.1911	7.99E+06	325.1650	2.26E+06
APCI Type B				
FUS-X	372.1653	1.32E+08	355.1390	1.56E+07
3-Ac DON	356.1704	8.89E+05	339.1440	1.12E+08
15-Ac DON	356.1704	2.57E+07	339.1440	1.22E+08
4 DON	314.1598	5.40E+05	297.1330	4.87E+07

*EIC= extracted ion chromatogram within ±10 ppm of precursor m/z.

Figure 1. Curated infusion data for 3-Ac DON (type B trichothecene). Shown below: curation steps and precursor structure (left panel), curated spectral tree (top panel), precursor breakdown curves (middle) and auto annotated MS^2 spectrum from *in silico* predictions (bottom).



DATA ANALYSIS

In Silico and Predictive Fragmentation

Mass Frontier integrates the HighChem Fragmentation Library™, the mzLogic™ structure ranking tool and the Fragment Ion Search (FISh) component extraction tool with the mzCloud mass spectral library enabling fragment prediction capabilities for HRAM and nominal mass data.

The HighChem fragmentation library contains about 216,000 individual chemical reactions and decoded mechanisms gathered from data published in major journals during the past decades[3]. *In silico* fragment prediction in Mass Frontier is based on use of the HighChem fragmentation library and general fragmentation rules.

Components Detection prior to Substructure Searches

Component detection is the automated detection of chromatographic features and extraction of mass spectral signals with deconvolution from closely coeluting components[1,2]. Use the **Joint Components Detection** (JCD) algorithm for LC-MS data, based on the statistical analysis of ion profile maxima. Use the **Direct Infusion Components Detection** (DICD) for infusion data. DICD, a simpler algorithm, is a spectral tree construction utility. Use the **Fragment Ion Search** (FISh) detection with LC-MS data when looking for parent drug modifications, metabolites and impurities in pharmaceutical drug research or for unknown compound identification, as shown in this work.

Similarity and Substructure Searches

Similarity and substructure searches use mzCloud and/or compound databases and relates MS fragmentation data from an unknown compound to known structures[4]. mzCloud is an extensively curated mass spectral library of high-resolution tandem spectra arranged by means of spectral trees (MS, MS/MS and MSn). The MS^2 nodes contain spectra at different collision energies, from various precursor m/z values, at different precursor isolation widths and using different activation schemes (*e.g.*, collision-induced dissociation (CID) or HCD). In contrast, compound databases are much larger repositories of chemical structures with information that includes compound naming, elemental composition, molecular mass and cross references among databases, but no or little MS data.

mzLogic, is a structure ranking tool that combines results from reference spectral library searching (in similarity mode) and compound database searching. The chemical database search taps into MolGate™, an online collection of 40 million chemical structures from public sources.

Figure 2. mzCloud results for HT-2 trichothecene, which does not exist in mzCloud. Top hits are T-2 and T-2 triol trichothecenes (in red square). Difference spectrum between HT-2 (experimental) and T-2 (library) show shared MS/MS fragments.

Figure 3 a), b). Mass Frontier similarity/substructure search window (a) and results (b) for HT-2 MS^2 spectrum against my curated library containing 7 type A and 4 type B trichothecenes. Top hit is HT-2, as expected.

a)

b)

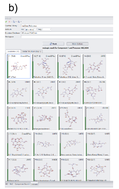
DATA ANALYSIS

Similarity and Substructure Searches (continued)

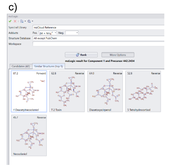
Figure 4 a) b) and c). Similarity and structure ranking using the mzLogic algorithm. a) Conglomerate of compound databases used, b) ranked mzLogic results: top hit is HT-2, c) other similar type A trichothecenes structures suggested.

a)

b)



c)



Data Dependent MS^2 LC-MS Experiment – Type A trichothecene Mix

A data dependent MS^2 experiment takes a full MS every few seconds and selects the most intense peaks for MS^2, dynamically excluding the top MS peaks for a determined time period, so fragmentation can be obtained on progressively lower abundant peaks. Five of the seven type A trichothecenes present in the sample were identified when analyzed against mzCloud. The two misses were T-2 tetraol and HT-2. Although T-2 tetraol is in mzCloud, it eluted very early and was missing from the chromatography in this particular experiment. HT-2 is not among the type A trichothecenes in mzCloud. A similarity search using mzLogic was then used for the unidentified component at m/z 442.2427 (HT-2, [M+NH4]$^+$). Fig. 5 shows the result correctly matching the an unidentified chromatogram component to HT-2.

Figure 5. ddMS2 data and mzLogic results for sample containing 7 type A trichothecenes. mzLogic similarity search ranks component at m/z 442.2427 as HT-2 after mzCloud search missed it.

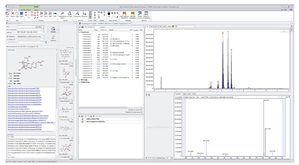
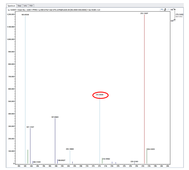
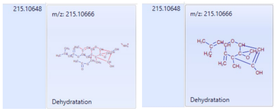


Figure 6 a), b), c). ddMS^2 experiment and Fragment Ion Search component detection. A basic structure (T-2 toxin) and potential modifications (*e.g.*, acetylation, hydroxylation) are defined a priori. (a) MS/MS for an unidentified component after FISh analysis, b) when m/z 215.1065 peak is selected, the sub structure identified by FISh is shown.

a)



b)



c)

Fragment structures can be shown: b) as they fit in the parent structure (red delineation) or c) by themselves.

CONCLUSIONS

- High resolution (resolving power settings of 240k for MS and 30-45k for MS^2) and accurate mass (≤2 ppm) mass spectra were acquired for the creation of user libraries. Two type A and 4 type B trichothecenes, not currently available in the mzCloud mass spectral library, will be added under the 'Natural Toxins' category.
- *In silico* fragmentation with automated annotation is a very useful tool to quickly label most fragments in complex MS spectra. A chemical formula is provided if a peak is not annotated.
- The components detection algorithms in Mass Frontier were successfully used to tackle and simplify mass spectra of increasing complexity. Subsequently, components search algorithms were used for similarity and substructure searches.
- Infusion MS and MS^2 data of individual compounds was the simplest type of data tested with the various algorithms and databases. Untargeted data dependent MS^2 acquisition was the most complex MS data tested in this work.
- Next we aim to analyze a sample containing 14 trichothecenes and compare both untargeted ddMS^2 and data independent acquisition (DIA) experiments. DIA queries the whole m/z range in windows of ~30 amu, thus doing MS^2 over the whole mass range of interest. It has been successfully used for compiling list of shared fragments within different classes of mycotoxins[5].
- A Precursor Ion Scan mode experiment in a triple quadrupole MS is also of interest for follow-up experiments. In this case, known shared fragments of trichothecenes can be used to detect previously unknown modified forms from a fungus culture[6].
- Future work will challenge this workflow for the proper identification of trichothecenes with an unknown sample in matrix.

REFERENCES

1. Peake, D (2018). High-resolution compound identification in metabolomics: a review of current practices. Thermo Scientific Review Paper 65356.

2. Ding, C et al (2018). A full-featured data curation solution for comprehensive and high quality HRAM MS/MS and MS^n library building. 66th ASMS conference, San Diego, CA.

3. Thermo Mass Frontier User Guide, Software version 8.0, Thermo Fisher Scientific, 2018.

4. Stratton, T et al (2018). Improved ranking of putative candidates through a hybrid *In Silico* / real fragmentation technique. 66th ASMS conference, San Diego, CA.

5. Jia, W et al (2019). Multiplexing data independent untargeted workflows for mycotoxins screening on a quadrupole-Orbitrap high resolution mass spectrometry platform. *Food Chemistry 278*: 67-76.

6. Gonzalez-Jartin, JM et al (2018). Detection of new emerging type-A trichothecenes by untargeted mass spectrometry. *Talanta 178: 37-42.*

ACKNOWLEDGEMENTS

M Prieto Conaway would like to thank Caroline Ding and Tim Stratton from Thermo Fisher Scientific for useful discussions in the analysis of data by Mass Frontier and the acquisition of data for library building, respectively. This work was funded by the US Department of Homeland Security.

TRADEMARKS/LICENSING/DISCLAIMER

LLNL-POST-966576 This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.