Increased Throughput and Coverage for the Annotation of Saponins Using a Structure-based MS^n Approach on a Tribrid Orbitrap Mass Spectrometer

Reiko Kiyonami, Caroline Ding, Seema Sharma and Andreas Huhmer
Thermo Fisher Scientific, San Jose

ABSTRACT

Purpose: Develop a structure-based MS^n workflow to collect more structurally relevant fragment ion information for the annotation of saponins from traditional Chinese medicines with increased coverage and confidence.

Methods: All the MS and MS^n data were collected with a Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer. We collected MS^2 spectra on precursor ions detected in the survey MS scan within a 1.2 second cycle time. Higher order MS^n $^{(3\text{-}4)}$ spectra were only collected when the instrument detected the sugar neutral loss in MS^2 and/or MS^3 data. The MS^n spectral tree data were processed using Thermo Scientific™ Mass Frontier™ 8.0 and Thermo Scientific™ Compound Discoverer™ 3.1 software.

Results: The structure-based MS^n workflow was able to annotate 127 saponin compounds from a traditional Chinese medicine, Sanqi.

INTRODUCTION

Saponins are major components of Chinese medicines and exert various pharmacological effects such as cardiovascular protective activity and anticancer activity. Plus, they could reduce the side-effects of radiotherapy and chemotherapy. The comprehensive annotation of saponins from various Chinese medicines remains challenging because of the limited availability of authentic standards and the structural diversity of this class of compounds. Taking advantage of high-resolution MS and MS^n capability offered by the Tribrid Orbitrap mass spectrometer, we developed a neutral loss-dependent MS^n data acquisition method in which MS^2 data is constantly collected and further followed by higher order FTMS^n if a sugar neutral loss is detected from the MS^2 data. The collected MS^n tree data were used to annotate the saponin class compounds which contain the basic saponin structures. The ChemSpider™ database and custom databases are further used for final saponin structure annotation.

MATERIALS AND METHODS

Sample Preparation

1 g of Sanqi powder purchased from a store was extracted using 10 ml of 80% methanol. The sample was filtered using a 0.2 µm filter, dried using a Speed Vac, reconstituted into 1 ml of 80% water, and further filtered with a 3K centrifugal filter.

HPLC Conditions

A Thermo Scientific™ Vanquish™ UHPLC system performed separations. Mobile phase A was water with 0.1% formic acid and mobile phase B was methanol with 0.1% formic acid. The column was a Thermo Scientific™ Hypersil Gold™ column (2.1 x 150mm, 1.9µm) operated at 45 ˚C with a flow rate of 200 µL/min. Separation of compounds was carried out with gradient elution profile: 0 min, A:B 99.5:0.5; 1 min, A:B 90:10; 10 min, A:B 70:30, 18 min, A:B 50:50, 22 min, A:B 1:99; total 30 min. The injection volume was 2 µL. The Deep Scan mode of Acqure X intelligent acquisition was used with 5 iterative MS^n injections.

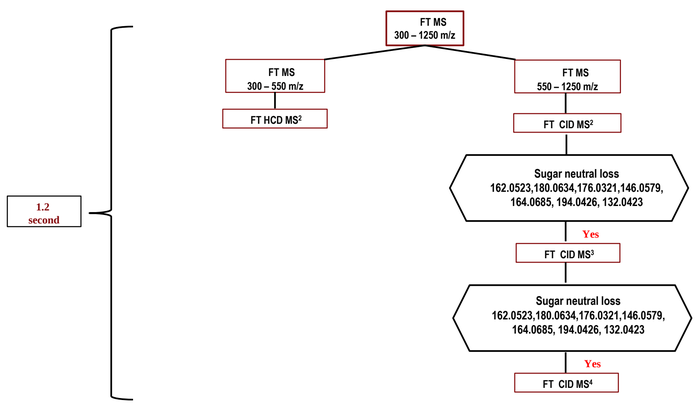
MS Conditions

All the data was collected on anOrbitrap ID-X Tribrid mass spectrometer. The mass spectrometer set up is shown in Table 1. For the precursor ion mass range between 300 – 550 m/z, data dependent HCD MS2 scans were collected. For the precursor ion mass range between 550– 1250 m/z, an intelligent neutral loss molecular-dependent MS^n approach was used, in which an HRAM full MS scan was followed by CID MS2 scans. The product ions generated from each MS2 scan are monitored by instrument and an MS^3 scan is triggered if one or multiple pre-defined neutral sugar molecular were detected from an MS^2 scan. An MS^4 scan is further followed if pre-defined neutral sugar molecular were detected from the MS^3 scan. Figure 1 shows the flowchart of the developed neutral loss molecular-dependent MS^n data acquisition instrument method.

Table 1. Oribitrap ID-X instrument set up.

ESI source	Orbitrap ID-X
Sheath gas 35	Pos ion (300-1250 amu)
Aux gas 5	MS: R=60K (FWHM at m/z 200)
Spray volt. 3.4 kV	MSn: R=15K (FWHM at m/z 200)
RF-Lens 40	Cycle time: 1.2 second
Cap. temp. 300°C	MS^2 Isolation width: 1.6 Da
Heater temp. 300°C	MS^n Isolation width: 1.6 Da (MS2) → 2.0 Da (MSn)

Figure 1. Flow chart of sugar neutral loss triggered high order MS^n data acquisition on the Orbitrap ID-X MS instrument. The method is available as an instrument method template on the Orbitrap ID-X to allow easy set up of the MS^n method.



Data Processing

Mass Frontier 8.0 software and Compound Discoverer 3.1 software were used for data processing. Multiple databases were employed in the processing workflow including mzCloud™ spectral library, ChemSpider database, and custom databases for unknown saponin structure annotation.

RESULTS

Saponin class compound annotation using MS^n spectral tree data with Mass Frontier 8.0 software

The Sanqi sample was analyzed using the neutral loss triggered MS^n workflow. The MS^n $^{(up\ to\ 4)}$ spectral tree data were collected on unknown compounds consisting of sugar modifications generating ions resulting from sugar neutral loss in the MS^2 fragmentation stage. The collected MS^n tree data were searched against the mzCloud MS^n spectral library in batch mode using the subtree search tool of Mass Frontier 8.0 software (Figure 2). The unknown compounds with exact MS^n spectral tree match against the library references were annotated. However, for most unknown compounds, only partial MS^n tree matched with the library references, resulting in only partial saponin basic structure annotation of these unknown compounds. Figure 3 shows one example of partial structure annotation. The MS^n spectral tree of an unknown compound with molecular weight 782.4816 did not have an exact library spectral tree match. But its MS^3 spectrum matched the MS^3 spectrum a saponin library reference. As a result, the partial structure of the unknown compound could be annotated as the precursor structure which generated the MS^3 spectrum of the saponin library reference and the unknown compound was annotated as one of saponin class compounds. Figure 4 shows another example to annotate the unknown compound with molecular weight 962.5450 as a saponin class compound. The annotated saponin class compounds with sub-structure annotation were exported as a mass list with Compound Discoverer software format for pursuing full structure annotation using the software tools provided in the Compound Discoverer 3.1.

Figure 2. MS^n spectral tree data search using Mass Frontier 8.0 software.

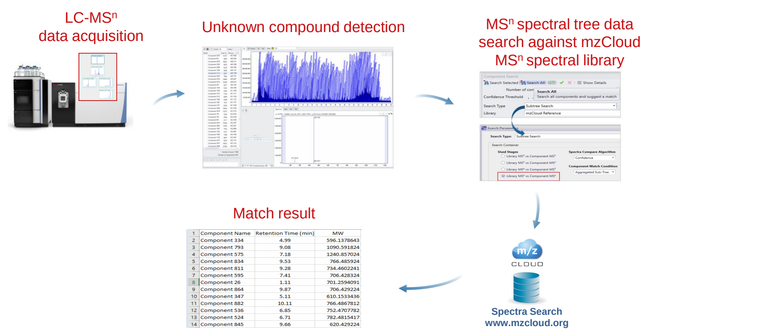


Figure 3. Sub-structure annotation of unknown compound (MW: 782.4816) using MS^n spectral tree data search result.

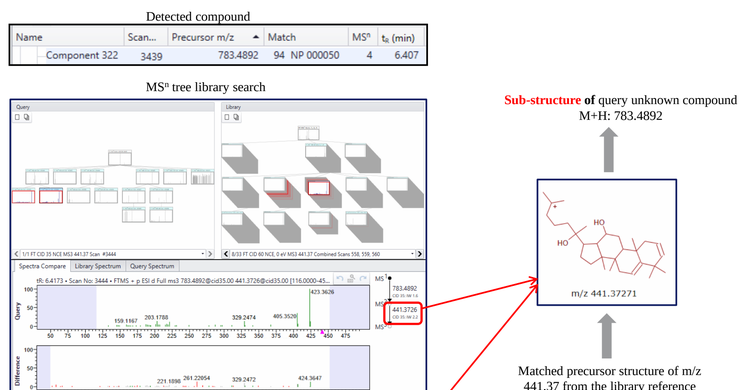
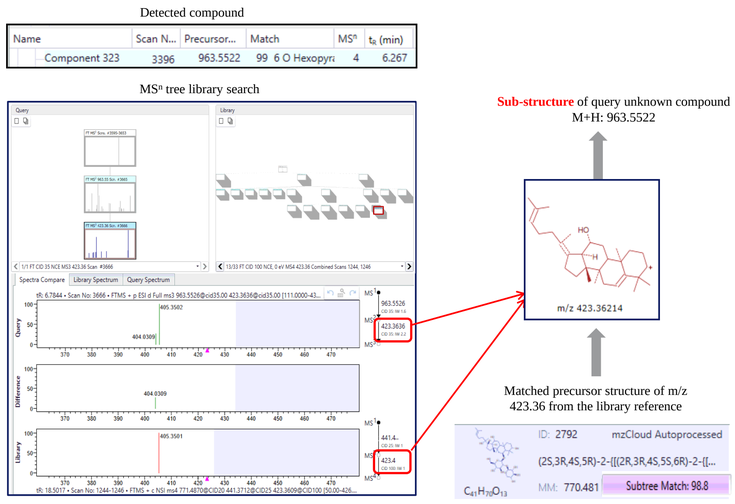


Figure 4. Sub-structure annotation of unknown compound (MW: 962.5450) using MS^n spectral tree data search result.

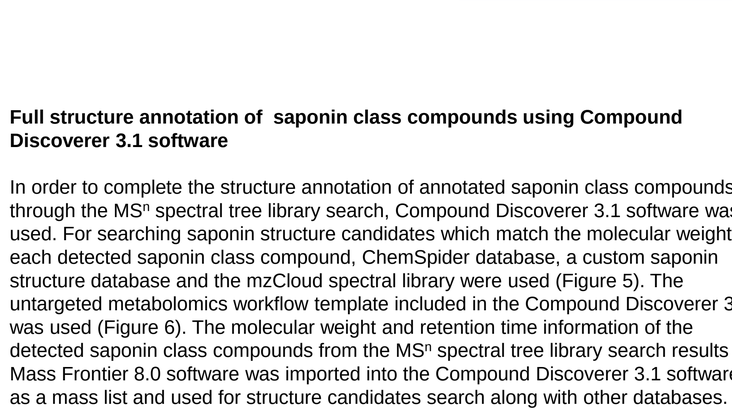


Full structure annotation of saponin class compounds using Compound Discoverer 3.1 software

In order to complete the structure annotation of annotated saponin class compounds through the MS^n spectral tree library search, Compound Discoverer 3.1 software was used. For searching saponin structure candidates which match the molecular weight of each detected saponin class compound, ChemSpider database, a custom saponin structure database and the mzCloud spectral library were used (Figure 5). The untargeted metabolomics workflow template included in the Compound Discoverer 3.1 was used (Figure 6). The molecular weight and retention time information of the detected saponin class compounds from the MS^n spectral tree library search results vs Mass Frontier 8.0 software was imported into the Compound Discoverer 3.1 software as a mass list and used for structure candidates search along with other databases.

Figure 5. Full structure annotation of detected saponin class compounds using Compound Discoverer 3.1 software.

Figure 6. Data processing workflow tree used for structure candidate search.



The Compound Discoverer 3.1 software determines the molecular weight of each detected unknown compound and searches the annotated saponin class compound mass list to see if there is a match with one of the annotated saponin class compounds with both molecular weight and retention time. If there is a match, then the structure search results from ChemSpider database and custom saponin structure database for this saponin class compound are examined for proposing its structure candidates. Sub-structure annotation result is considered in this process.

FISh (Fragment Ion Search) score is calculated per proposed structure candidate and used for ranking which structure is closest to the real structure of the saponin class compound.

With this approach, the detected saponin class compound (MW: 782.4816) was annotated as (3β,9β,12α,16β)-12,16-Dihydroxy-24-oxo-9,19-cyclolanostan-3-yl 2-O-(6-deoxy-α-L-mannopyranosyl)-β-D-glucopyranoside (Figure 7). The detected saponin class compound (MW: 962.5450) was annotated as Mogoroside III-A1 (Figure 8).

In total, 127 saponin class compounds were successfully annotated from the Sanqi sample.

Figure 7. The detected saponin class compound (MW: 782.4816) was annotated as (3β,9β,12α,16β)-12,16-Dihydroxy-24-oxo-9,19-cyclolanostan-3-yl 2-O-(6-deoxy-α-L-mannopyranosyl)-β-D-glucopyranoside using a database search and the FISh ranking tool.

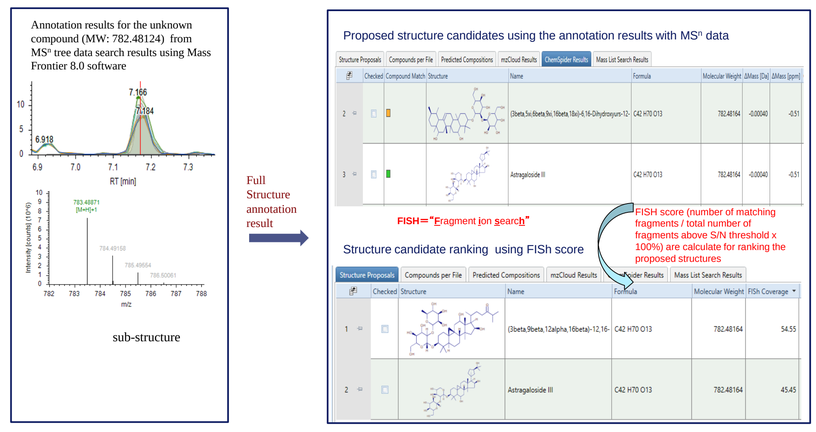
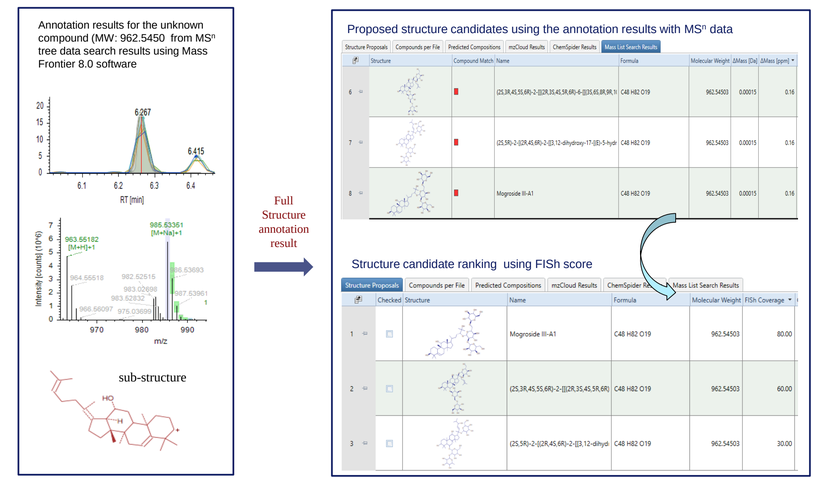


Figure 8. The detected saponin class compound (MW: 962.5450) was annotated as Mogoroside III-A1 using a database search and the FISh ranking tool.



CONCLUSIONS

- A structure-based MS^n workflow on an Orbitrap ID-X mass spectrometer was developed for saponin class compound annotation from a Traditional Chinese Medicine, Sanqi sample.
- The workflow significantly increased the numbers of annotated saponin compounds with limited library references by using MS^n spectral tree data to detect the saponin basic structure of unknown compounds.
- The capability of building a custom structure database allows saponin compound annotation beyond the ChemSpider database.
- In total, 127 saponin compounds were annotated rapidly in the Sanqi sample.

TRADEMARKS/LICENSING