Novel Real-Time Library Search Driven Data Acquisition Strategy for Identification and Characterization of Metabolites

Brandon Bills[1], Seema Sharma[1], William Barshop[1], Jesse Canterbury[1], Aaron Robitaille[1], Vlad Zabrouskov[1],
[1]Thermo Fisher Scientific, 355 River Oaks Pkwy, San Jose, CA, U.S.A. , 95134

ABSTRACT

Purpose: Identification and structural characterization of novel metabolites in drug discovery or metabolomics experiments is challenging. Multi-level fragmentation (MS^n) based approaches are frequently utilized for facilitating structure assignment of unknown compounds. As each of the MS precursors undergoes MS^n, the instrument cycle time can limit the total number of precursors analyzed in a single run. Here we present a new LC/MS data acquisition strategy, termed Met-IQ, where the decision to perform an MS^n acquisition is automatically made in real-time based on the similarity between an acquired experimental MS^2 spectrum and a spectrum in a reference spectral library. This strategy reduces time wasted collecting MS^n data on irrelevant or well-known compounds and increases the likelihood of obtaining useful structural information on true unknowns.

Methods: A sample of Amprenavir was incubated with human liver microsomes then analyzed on the Thermo Scientific™ Orbitrap IQ-X™ Tribrid™ mass spectrometer (MS). A MS^3 data dependent analysis was carried out with and without Met-IQ to determine if the depth of sampling or number of metabolites with MS^3 spectra improved significantly during a single analytical run.

Results: The number of MS^2 increased nearly 3.5-fold compared to the standard data dependent experiment where MS^3 was triggered for each precursor ion, resulting in identification and structural characterization of 55% more unique metabolites. Furthermore, the MS^3 precursor fragments were selected intelligently, focusing on higher mass fragments of sufficient intensity to maximize acquisition of MS^3 data relevant for structure assignment. This work is for research purposes only.

INTRODUCTION

In applications such as drug development or environmental safety, biotransformations need to be fully characterized to investigate any toxic effects. Targeted analysis can miss unknown metabolites and untargeted analysis can lead to cumbersome analysis of large data sets with low concentration metabolites missed when higher concentration background compounds take up all the cycle time.

Modern mass spectrometers have built-in data acquisition strategies to intelligently guide the instrument. Filters direct the instrument to only consider compounds of sufficient intensity or to target compounds exhibiting a relevant neutral loss. Workflows such as the Thermo Scientific™ AcquireX™ data acquisition workflow direct the instrument to analyze a matrix blank and add those compounds to an exclusion list, so the instrument focuses on likely compounds of interest. These techniques improve the chances of detecting low level metabolites but requires multiple runs which takes additional time for analysis and processing. In addition, a static instrument method may not capture enough information on every compound for full structural characterization.

Real-time decision making allows the instrument to change the scan behavior based on data acquired during the analytical run. The instrument control software (ICSW) takes acquired data, compares it against a spectral library and uses that information to choose subsequent scan behavior during that same run. Here we introduce a small-molecule focused version of real-time searching on the Orbitrap IQ-X Tribrid MS and Thermo Scientific™ Orbitrap Eclipse™ Tribrid MS called Real-Time Library Search (RTLS). The focus is on directing the instrument to acquire additional data on poorly identified compounds. Experimental fragmentation spectra are compared to a spectral library and based on the result, the instrument triggers additional scan behavior, such as MS^n or alternate fragmentation techniques, to get additional structural information about the compound. The new Met-IQ data acquisition strategy uses RTLS to compare all MS^2 spectra against the spectra of the unmetabolized compound for spectral similarities and perform MS^3 scans on likely metabolites.
A sample of the drug Amprenavir metabolized by human liver microsome was analyzed with and without the Met-IQ data acquisition strategy to determine whether the RTLS-guided method improved the detection of metabolites in a single run over an unguided approach.

MATERIALS AND METHODS

Sample Preparation

[1]A sample of 5 µM amprenavir was digested by incubation with human liver microsome with NADPH and GSH.

Test Method

UHPLC-MS analytical runs were 18 minutes long and carried out on a Thermo Scientific™ Vanquish™ Horizon UHPLC system coupled to a Orbitrap IQ-X Tribrid MS. Mass spectra were collected in positive ion mode using data dependent MS/MS ($ddMS^2$). In addition, $ddMS^3$ was collected with and without Real Time Library Search (RTLS).

Data Analysis

Data was processed using Thermo Scientific™ FreeStyle™ software. Known metabolites of Amprenavir were manually checked for within the list of MS^2 and MS^3 spectra. In addition, the total number of MS^2 and MS^3 spectra was tabulated. Compounds were further investigated using Thermo Scientific™ Compound Discoverer™ software.

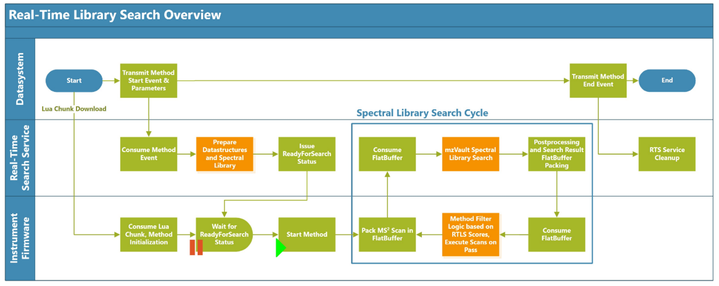
RESULTS

Real-Time Library Search Architecture

At the start of a method that uses RTLS, as shown in figure 1, the user specified RTLS filter parameters from the method are transferred to the Real-Time Search service on the Data System PC. First, the provided Thermo Scientific™ mzVault™ (.db) library is loaded into an in-memory m/z indexed structure, while filtering the library contents based on the method. Library spectra which do not match the RTLS method MS^2 acquisition settings for polarity, activation mode, and analyzer type are omitted. Next, the Real-Time Search service signals its ready status to the instrument firmware and awaits scans from method acquisition. MS^2 spectra generated on the instrument are sent from the instrument firmware to the Real-Time Search service.

Upon receipt, the Real-Time Search service will query the in-memory index for spectral candidates within the defined m/z tolerance, and the collision energy tolerance provided within the RTLS filter parameters. Each of the spectral library entries under consideration are compared with the experimental spectrum by calculation of the cosine similarity and ranked accordingly. Each score threshold can require that values to pass be either greater than or equal to the stated value ("At Least") or below ("Less than") the given value. While the search is running, the instrument will continue to operate normally to collect additional MS^2 scans in accordance with the method design.

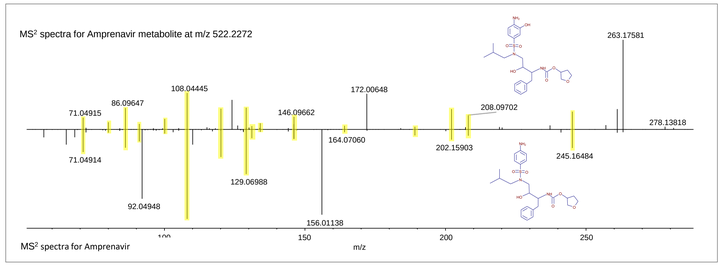
Figure 1. Real-Time Spectral Library Search Overview.



Impact of RTLS on Data

The benefits of Real Time Library Search can be broken down into two main parts. The first benefit is that the instrument precursor-selection logic is more optimized. In traditional data dependent MS^n acquisitions, the instrument conducts MS^n scans on every MS^2 fragment that meets the triggering threshold. This leads to excess time being spent acquiring data on peaks regardless of whether they are likely compounds of interest. The Met-IQ data acquisition strategy makes use of Real-Time Library Search to restrict MS^3 data acquisition to likely targets. For the digested amprenavir sample, the instrument compared each MS^2 spectra to the reference library containing unmetabolized amprenavir mass spectra. When sufficient matched peaks were found, as shown in the example in figure 2, to yield a cosine score of 15 or higher, $ddMS^3$ was performed.

Figure 2. Mirror plot of MS^2 spectra for amprenavir metabolite and Amprenavir. Matching fragments are highlighted in yellow.



Limiting data acquisition to likely compounds of interest reduced the number of triggers for $ddMS^3$ spectra which in turn enabled the instrument to spend more time sampling compounds with dd MS^2. As shown in figure 3, the number of MS^2 collected in a single run increased by more than 3.4 with the use of Met-IQ relative to the DDA control, leading to an increased depth of sampling in a single run.

To evaluate the increased depth of sampling, the m/z values that triggered MS^3 were compared to a list of known amprenavir transformation products. A total of seventeen unique metabolites were found in a single run using the Met-IQ data acquisition strategy compared to only eleven using the traditional data dependent acquisition, an increase of roughly 55%, as shown in table 1.

Figure 3: Number of MS^2 and MS^3 spectra collected with and without Met-IQ.

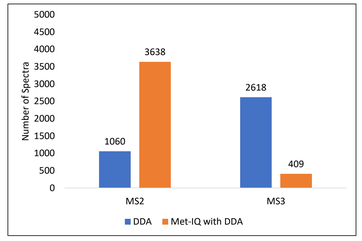


Table 1. Transformation products of Amprenavir with retention time (RT), cosine score from RTLS, and whether a MS^3 scan was triggered with or without Met-IQ.

Transformation	RT [min]	Cosine Score	MS^3 Experiment With Met-IQ	MS^3 Experiment Without Met-IQ
Amide hydrolysis + oxidation	5.3	49	X	X
Sulfonamide hydrolysis	5.4	25	X	X
Amide hydrolysis + oxidation	5.7	48	X	X
Amide hydrolysis + oxidation	6.7	40	X	X
Amide hydrolysis	7	84	X	X
Amide hydrolysis + oxidation	7.2	29	X	-
Oxidation	7.4	53	X	X
Oxidation (+O-2H)	8.1	20	X	X
Dehydration	8.2	53	X	X
Oxidation	8.1	39	X	X
Di-oxidation	8.3	49	X	-
Dehydration	8.4	68	X	-
Dehydration	8.6	56	X	X
Amide Hydrolysis	8.7	45	X	-
Dehydration	8.8	78	X	-
Oxidation	9.2	41	X	X
Dehydration	9.3	42	X	-

The second benefit of the Met-IQ data acquisition strategy is the simplification of post-acquisition data analysis. In addition to increased depth of sampling there is a reduction in the raw data volume and complexity. There is over a six-fold decrease in the number of MS^3 spectra, representing a significant decrease in sampling of unlikely targets (Fig. 3). In addition, each compound that triggered MS^3 had at least one fragment in common with amprenavir, flagging it as a potential metabolite. The real-time library search feature is also capable of flagging MS^2 fragments that aren't shared with the library compound. These fragments can then be selected for MS^3. By focusing on unmatched fragments, this acquisition strategy increases the probability of collecting structural information about unknown portions of the structure. In figure 2, the fragment at m/z 263.1758 for the amprenavir metabolite is at a significantly higher intensity relative to the amprenavir standard spectrum. Looking at the spectrum from Compound Discoverer software in figure 4a, there is no predicted structure for this fragment. Looking at the MS^3 fragmentation spectra of this fragment in figure 4b, multiple sub-structures could be annotated by Compound Discoverer software providing valuable structural information about the metabolite.

Figure 4a. MS^2 spectra of amprenavir metabolite with annotations from Compound Discoverer software

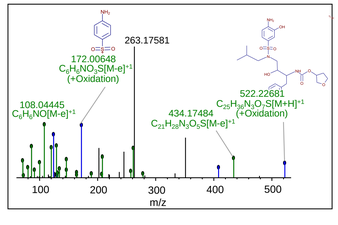
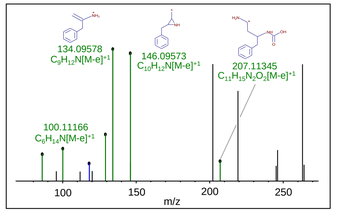


Figure 4b. MS^3 spectra of 263.17581 fragment from Amprenavir metabolite with annotations from Compound Discoverer software



CONCLUSIONS

We have analyzed a sample of metabolized amprenavir using our new Met-IQ data acquisition strategy which uses Real-Time Library Search on Orbitrap IQ-X Tribrid MS (Figure 5) to target compounds for $ddMS^3$ when there are MS^2 spectral similarities to the parent drug. This guided analysis has multiple benefits for the resulting data including:

- Instrument time is spent more optimally by limiting MS^3 analysis to only likely compounds of interest. This led to in increased depth of sampling with a 3.4 factor increase in the total number of MS^2 collected.
- By intelligently focusing instrument time on likely compounds of interest there is an increased probability of finding more potential metabolites in a single run. The number of potential metabolites found in a single run increased by 55% using Met-IQ.
- Using built in features to guide analysis towards higher mass fragments not found in the library, useful structural information about the unknown metabolites was generated in a single run

Figure 5. Orbitrap IQ-X Tribrid MS built specifically for small-molecule unknown analysis

ACKNOWLEDGEMENTS

[1] *Sample of metabolized amprenavir provided by Dr. Shuguang Ma, Genentech*

TRADEMARKS/LICENSING

PO66088 EN0921S