

Small Molecule Real-Time Library Search

William Barshop, Jesse Canterbury, Brandon Bills, Seema Sharma, Thermo Fisher Scientific, 355 River Oaks Parkway, San Jose, CA, USA, 95134

Abstract

Purpose

- To guide instrument acquisition decisions by consideration of the similarity of experimentally acquired spectra to the contents of a spectral library on the Thermo Scientific™ Orbitrap IQ-X™ Tribrid™ mass spectrometers.
- To improve the analytical capability of Orbitrap IQ-X Tribrid mass spectrometers by allowing the instrument to make informed decisions about which analyte ions will yield the most useful information, within the context of the current analyte, matrix, and experimental objectives.

Methods

- Real-Time Search infrastructure within the Tune 3.5 instrument control software has been extended to facilitate spectral library searching for small molecules, called Real-Time Library Search.
- Search scores already available in Thermo Scientific™ Compound Discoverer™ software and Thermo Scientific™ mzCloud database are used for real-time analysis of mass spectra with Thermo Scientific™ mzVault spectral library access.

Results

- Real-Time Library Search generates familiar measures of similarity and identification confidence during data acquisition.
- The calculated scores can be used in various method architectures to guide acquisition behaviors in real-time.

Introduction

Primary Challenge

- To enable comparisons of experimental spectra to library spectra in real-time.
- To leverage score outputs of Real-Time Library Search to guide decisions during acquisition.
- Previously, Real-Time Search has been focused for use within the realm of proteomics, with peptide-centric search tools employing databases of *in silico* mass spectra.

Methods

- The Real-Time Search infrastructure was modified to enable library search functionalities.
- Real-Time Library Search can use both custom mzVault libraries, or local copies of the mzCloud MS² libraries in the mzVault .db format.
- ddMS² scans are sent from the instrument to the search service, and the scored results are returned to the instrument filter

Summary

- Real-Time Library Search is available on the Orbitrap IQ-X Tribrid MS and Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometers with the Instrument Control Software 3.5 release.



mzCloud Spectral Library Support



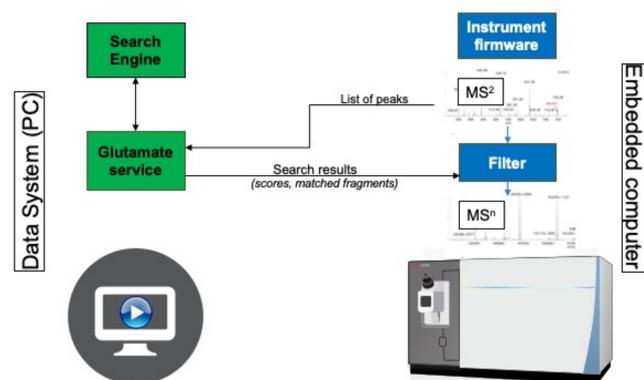
mzVault Search Backend



Cosine and Confidence scores for identifications

Real-Time Library Search Background Services & Communication

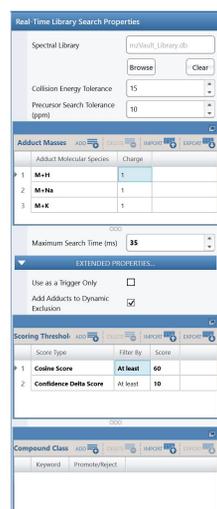
- Real-Time Library Search utilizes a backend service, called Glutamate, to process each MS² scan generated during acquisition. In turn, it produces and returns search results based on user provided filter parameters, and the design of the method.
- Upon receipt of the search results, the instrument will apply the relevant method logic to determine if the subsequent scan action will be executed and which peaks may be selected.
- If the search results satisfy the thresholds within the respective instance of the Real-Time Library Search filter, the MS² scan can trigger the subsequent MS³ scan as defined within the method.



The Real-Time Library Search User Interface

Configurable Parameters

- Custom mzVault spectral libraries and local copies of the mzCloud MS² libraries (.db format) can be utilized by Real-Time Library Search.
- Collision Energy Tolerance and precursor search tolerance (ppm) are used as +/- value thresholds for filtering the spectral library candidates during search.
- Adducts enumerated in the "Adduct Masses" molecular species definition table are considered in the adduct offset query searching. When the filter passes its defined scoring thresholds, the recalculated adducted masses can be added to dynamic exclusion.
- Maximum search duration can be specified for Real-Time Library Search, after which elapsed time the search will abort to limit overly long search times. During search execution, scan acquisition runs in parallel and will continue as normal.
- By default, only peaks from the experimental spectrum which match the top library hit are considered for the next MSⁿ scan. When "Use as a Trigger Only" is enabled, all peaks from the experimental spectrum are considered.

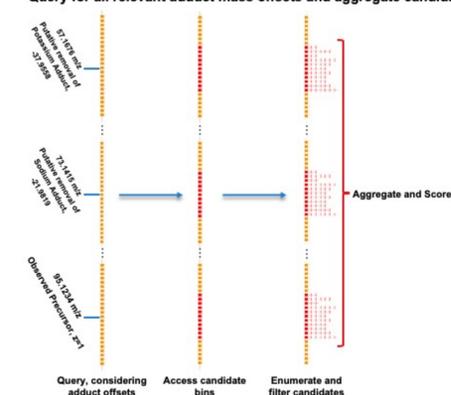


Small Molecule Real-Time Search Execution Logic

Searching and Scoring Scheme:

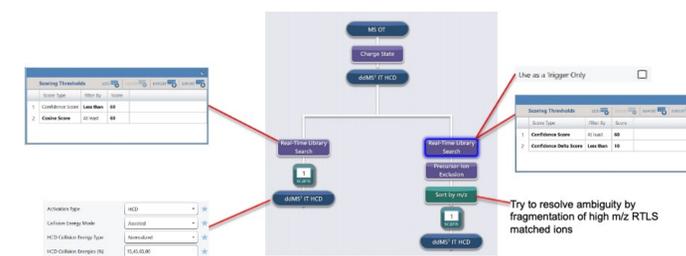
- At the start of method execution, the spectral library is indexed into memory and filtered for relevant polarity, activation mode, collision energy tolerance range, and analyzer type
- For each scan, query the spectral library at each relevant adduct offset(s) to accumulate the complete candidate spectra set
- Candidate-Query Similarity Scoring: Generate cosine similarity scores for each selected library spectrum
- Rank candidates by Cosine Score
- The highest scoring spectrum from each of the top two scoring compounds are used to calculate an identification confidence score
- Calculate Delta scores for the best match to the second-best compound ID for cosine and confidence values

Query for all relevant adduct mass offsets and aggregate candidates:



An Example of a Branched Real-Time Library Search Method

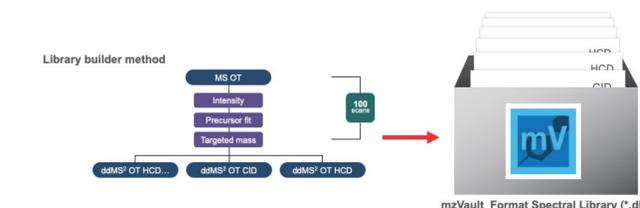
- This example demonstrates the large experimental flexibility of Real-Time Library Search. Pre-built method templates are available.
- Do nothing if search results are unambiguous & confident (high cosine/high confidence/high confidence delta)
- Try to find a better Collision Energy (CE) if match appears under-fragmented via Assisted CE (high cosine/low confidence)
- If high scorers are ambiguous (low confidence delta score), try to elucidate substructures of high m/z fragments



mzVault Library Generation

Users of Real-Time Library Search may create or obtain libraries in three ways:

- Fragmentation spectra can be acquired for compounds of interest via the library builder method template. After acquisition, custom compound libraries are assembled within the mzVault application, directly from the raw data.
- Existing spectral libraries may be converted to the .db mzVault format from the MassBank Record Format.
- Leverage downloadable copies of the mzCloud MS² fragmentation library, which contains over 1.7 million positive mode MS² spectra, and over 760,000 negative mode MS² spectra. They are available on the Thermo Fisher Scientific Flexnet Operations Portal (Flexera, <https://thermo.flexnetoperations.com/>).



Real-Time Library Search Extends Intelligent Mass Spectrometry for Small Molecules

<p>Real-Time Library Search In-Acquisition Decision-making</p> <ul style="list-style-type: none"> Library Directed MSⁿ – Target compound classes or confident/non-confident identifications with flexible threshold logic Control Scan Decisions – Multiple instances of the Real-Time Library Search filter with independent settings control complex acquisition logic and downstream scans 	<p>AcquireX workflows Cross-Acquisition Decision-making</p> <ul style="list-style-type: none"> Dig Deeper – Exclude background ions in complex matrices to enhance sampling of precursors of interest Automatic Rejection – Comprehensive sampling of low abundance components by exclusion of previously fragmented precursors 	<p>Post-Acquisition Analysis Comprehensive LC-MS Interrogation</p> <ul style="list-style-type: none"> Multi-sample Analysis – Compile and compare acquisitions and LC-MS feature intensities Computational Workflows – Precursor reassignment, RT alignment, fragmentation tree-based searching, elemental composition assignment for unknowns, pathway analyses
---	--	---

Conclusions

Real-Time Library Search Extends Intelligent Acquisition on the Orbitrap IQ-X Tribrid MS

- Real-Time Library Search provides spectral similarity measures and identification confidence scores for molecular species upon which method acquisition decision-making can be based during method execution.
- Real-Time Search has been extended to integrate spectral-library search functionality, called Real-Time Library Search technology.
- Pre-built method templates will be available for specific applications, such as Met-IQ data acquisition. Met-IQ may be used to identify unknown primary metabolites or transformation products based on similarity to known small molecule analyte of interest, such as pro-drugs.
- Real-Time Library Search is built upon mzVault's spectral library access and mzCloud's scoring to accelerate single-spectrum queries.
- Create custom libraries using mzVault, or use downloadable copies of the mzCloud MS² spectral library for use with Real-Time Library Search (Flexera, <https://thermo.flexnetoperations.com/>) PO000115-na-en06/21