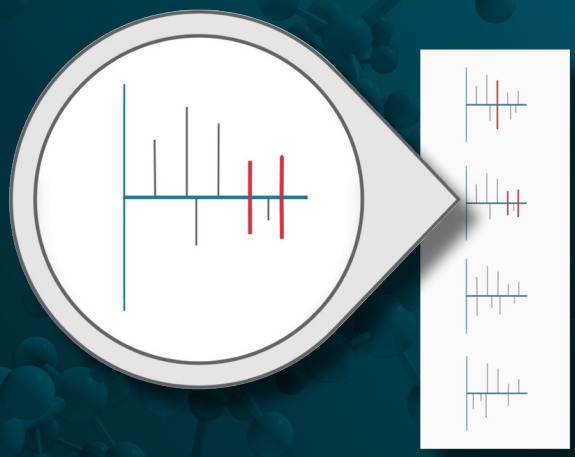
thermo scientific

SMART NOTE

mzLogic Data Analysis Algorithm



Accelerate small-molecule unknown identification

When there is no spectral library match for your small-molecule unknown compound, how can you use your data to confidently assign a structure?

Many small-molecule analyses, from metabolism studies and drug analyte analysis, to food and environmental contaminants, rely on the use of comprehensive spectral libraries to aid in unknown identification. However, subtle compound transformations (e.g., designer drugs or new impurities) can result in no spectral match. Subsequent searches of extensive online structural libraries can result in hundreds of potential structure matches, but without spectral data, confirming potential structural matches remains a challenge.

The mzLogic algorithm allows comparison of experimental data to the extensive, chemically diverse, fully curated MS^n fragment ion information in the mzCloud™ online library to rank putative structures, increasing efficiency and confidence in small-molecule unknown identification (Figure 1).

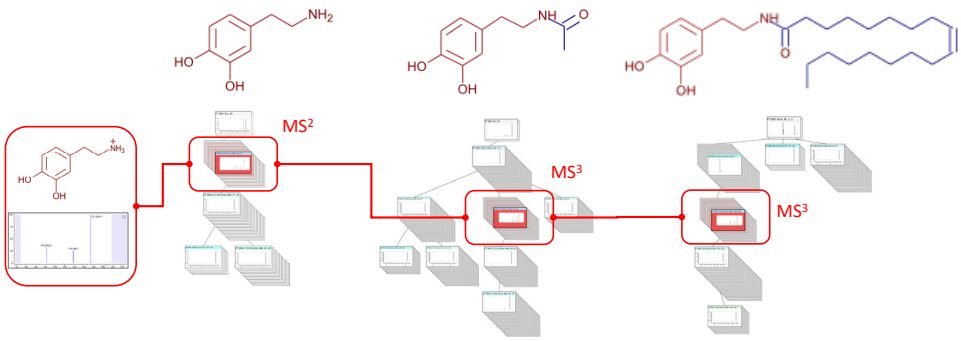


Figure 1. High quality, fully curated and annotated mass-spectral fragmentation libraries such as mzCloud can be interrogated to identify common substructural information to aid in unknown compound identification.

ThermoFisher SCIENTIFIC

What are the limitations of efficient small-molecule unknown identification?

From degradants and impurities, natural products, forensics, pharmaceutical metabolite identification, and environmental to metabolomics applications, one of the most challenging tasks for an analyst is compound identification and turning vast quantities of data into actionable, shareable knowledge.

How do libraries and databases aid in unknown identification?

There are many publicly accessible online databases containing hundreds to millions of chemical structures, which can be used to aid in unknown identification. Despite the large numbers of chemical structures these databases contain, the information associated with the structures can vary significantly. Some databases only contain structures and metadata, whereas others also contain structurally annotated spectral information, including fragmentation spectra.

One important consideration for libraries containing spectral information is the quality of the data. Where spectral data are included, there could be significant mass errors, incorrect substructure assignments, or even no substructure assignments.

mzCloud is the largest (in terms of total spectra and data per compound) publicly available online mass spectral fragmentation library. mzCloud also contains the largest number of high-resolution accurate-mass (HRAM) spectra, including exhaustive high-resolution MS/MS and multi-stage MS^n spectra.
Each entry contains considerable metadata and, most importantly, has been extensively curated (i.e., filtered, recalibrated, averaged, and annotated) to provide absolute confidence in quality of library contents.

How can libraries and databases be used to identify unknowns when the unknowns are not present in either?

Despite the vast amount of known structures available to search against, there will always be cases where there are no plausible matches. Or, when searching against the elemental composition of a precursor, many thousands of potential candidates may be returned.

Analysts looking for impurities and degradants, metabolites, and designer drugs frequently encounter true unknowns in their samples.

The extensive chemical diversity covered by mzCloud, combined with the significant amount of fully annotated MS/MS and MS^n fragmentation data acquired in both positive and negative ionization modes and at various collision energies for multiple adducts, means that there is a wealth of information against which unknown compound fragment spectra can be searched.

What is precursor ion fingerprinting?

Precursor ion fingerprinting (PIF) identifies substructure information by comparing unknown fragmentation spectra to the product ion spectra contained in mzCloud.

The fully characterized fragmentation structures in mzCloud are matched against an unknown's fragmentation spectra to build knowledge using real fragmentation data, and therefore substructure information, to aid in identifying the unknown's precursor structure (Figure 2).

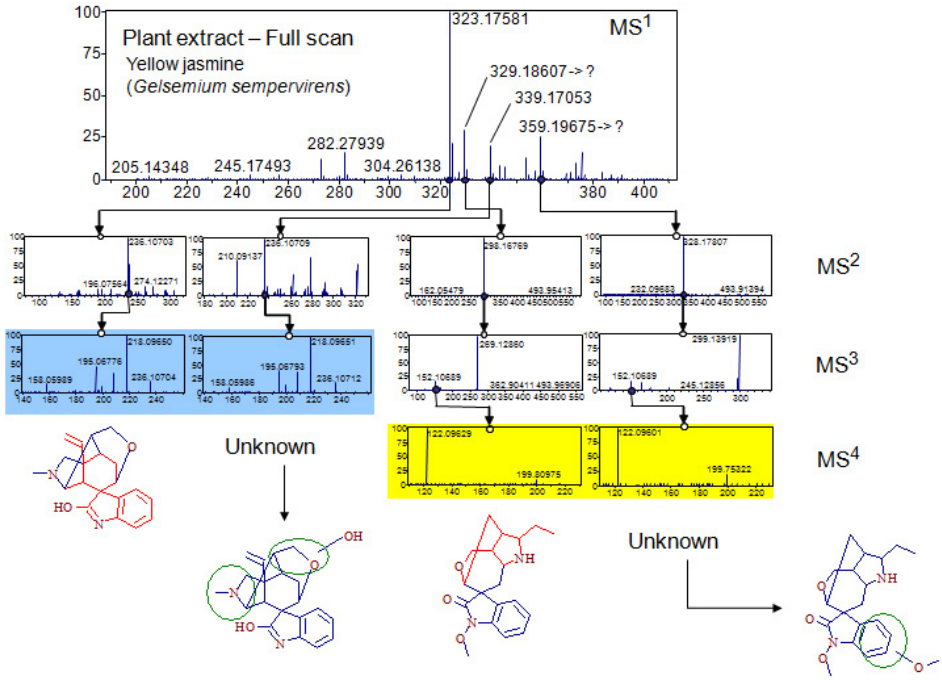


Figure 2. An example of using multi-stage fragmentation information and precursor ion fingerprinting to provide substructure information.

mzLogic compares mzCloud similarity matches against database hits, looking for maximum common substructure. mzLogic ranks database hits based on structure overlap, leveraging real fragmentation information to rank structure candidates for true unknowns.

How does mzLogic help identify unknowns more efficiently?

Using elemental composition and isotopic distribution based upon HRAM precursor *m/z* values to search against numerous online resources (such as ChemSpider) can produce hundreds of potential structures. mzLogic rapidly takes the potential structures from the online search, and combines this with the knowledge gained from precursor ion fingerprinting. The experimental fragment and substructure information are then automatically ranked based upon the maximum amount of common substructure overlap from both forward and reverse similarity searches of the fragment ions.

Using mzLogic, the initial list of potential structures is rapidly reduced to a rank-ordered list of putative candidates (Figure 3).

Compared to other search tools, the power of mzLogic is based upon chemical and substructural diversity, as well as the quality of data in mzCloud. Real-world fragmentation data that have been exhaustively curated and annotated are automatically mapped to the experimental unknown fragmentation data and ranked for each unknown, all in a matter of seconds.

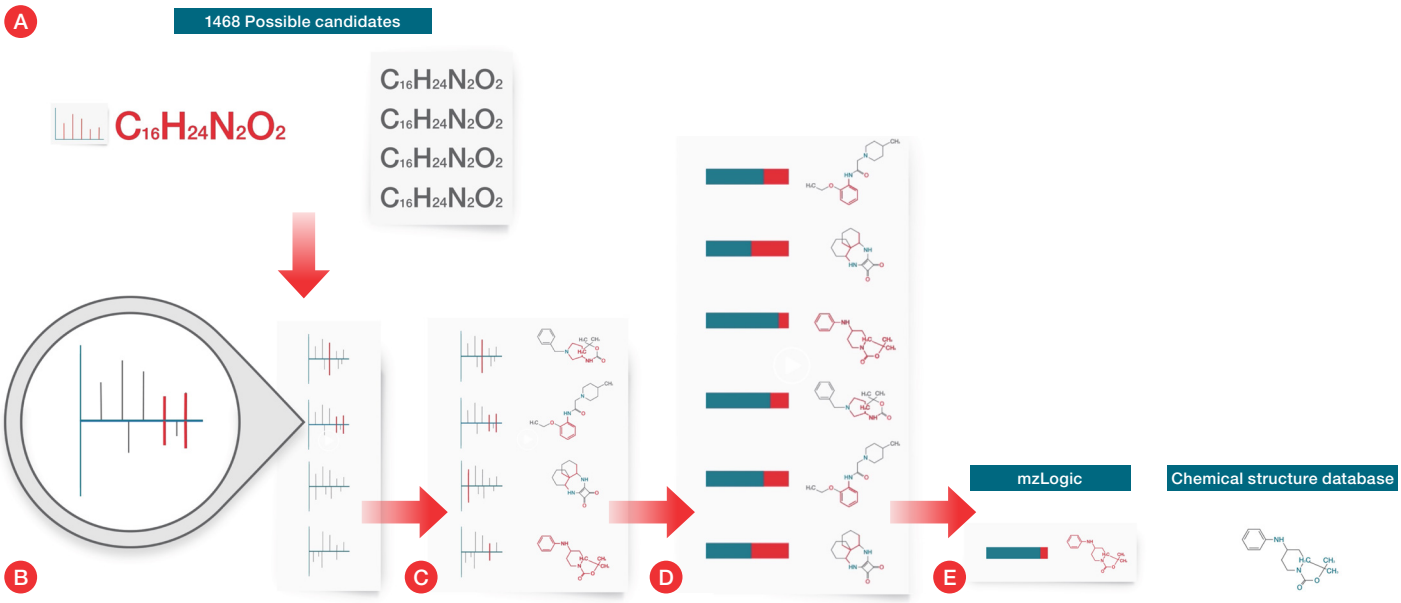


Figure 3. mzLogic reduction of putative candidates. a) Searching online databases with an elemental composition can return many hundreds of possible candidates. b) mzLogic uses the extensive high-quality fragmentation information from mzCloud to identify fragments which are similar. c) Structurally similar fragments are rapidly identified using forward and reverse searches and then d) ranked based upon their measure of similarity. e) The ranking reduces the many hundreds of potential candidates to a significantly smaller list of putative structures. Visualization tools provide explanations for the ranking.

Mass Frontier software provides a rational understanding of fragmentation schemes using the proprietary mass spectral interpretation algorithms of the HighChem Fragmentation Library. Mass Frontier software includes the fragmentation mechanisms published in more than 95% of all peer-reviewed scientific journals.

How are top-ranked putative candidates confirmed?

The use of mzLogic can be fully automated within Thermo Scientific™ Mass Frontier™ software and Thermo Scientific™ Compound Discoverer™ software, allowing analysis of acquired product ion spectra to provide either confirmed matches, or confidently proposed structures. Once candidates have been ranked in order of structural similarity using mzLogic, the fragmentation sub-structures used for ranking can be easily viewed, increasing confidence in putative compound identification.

When absolute confidence for proposed structures is required, Mass Frontier software can be used to enable greater insights through in-depth visualization and understanding of fragmentation pathways using extensive MS^n data. The Fragments and Mechanism tool simulates unimolecular dissociation of the proposed structures, including rearrangements and predicted fragments, and can be used to explain the experimental fragmentation data through fragment annotations to help identify which candidates are indeed plausible.

This approach adds greater certainty to putative structure identification through the deep fragmentation and characterization information provided by the high quality data and extensive spectral trees in mzCloud.

How to access mzLogic data analysis algorithm

- Purchase or upgrade to Compound Discoverer software version 3.0 or greater
- Purchase or upgrade to Mass Frontier software version 8.0 or greater

Find out more at **thermofisher.com/mzLogic**

ThermoFisher SCIENTIFIC