m/z Cloud – A New Type of Mass Spectral Library for the Identification of Unknowns even if the Unknown Compound is not Present in the Library

Robert Mistrik[1]; Juraj Lutisan[1]; Mark Sanders[2]; Yingying Huang[2]; Rose Herbold[2]; Tim Stratton[2]
[1]HighChem, Ltd., Bratislava, SLOVAKIA; [2]Thermo Fisher Scientific, San Jose, CA

Advanced Mass Spectral Database

Annotated Spectral Peaks, Fragment Structures, Resolution and Accuracy per Peak, Spectral Trees, Precursor Ion Fingerprinting, Substructure Identification, HR Search Algorithms, Relational Database

Introduction

Analysis of numerous small molecules and structural assignments of individual metabolic components is a major bottleneck in various areas of metabolomics. In mass spectrometry widely used library search systems are designed to identify compounds represented in the reference library. If the unknown compound is not represented in the library, the compound cannot be identified by this method. Here we present a new type of mass spectral library providing the functionality required for elucidation of unknowns even if compounds are not present in the library.

FIGURE 1. Precursor Ion Fingerprinting

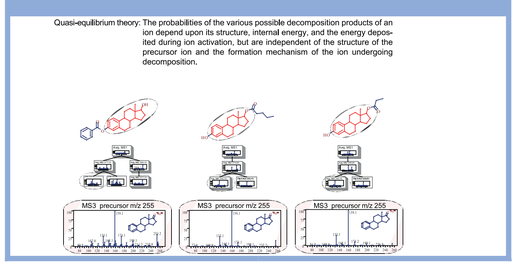


FIGURE 2. Library technology based on spectral ion trees

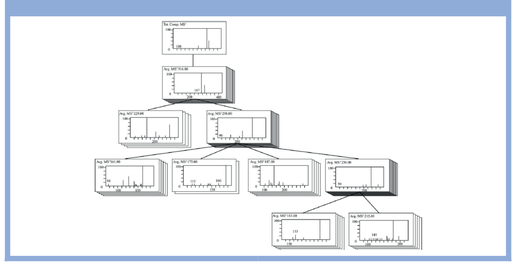


FIGURE 3. Newly developed algorithms allow comparison of high resolution product ion spectra

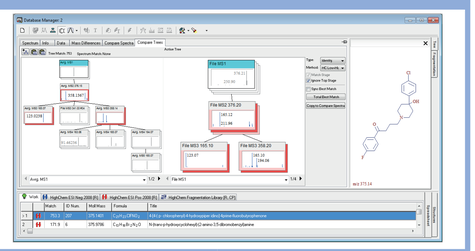
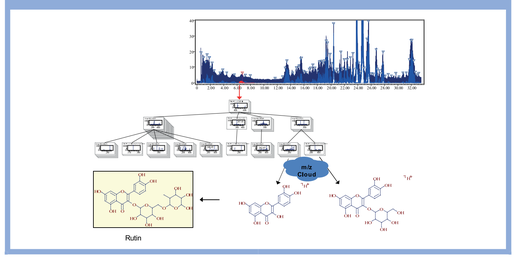


FIGURE 4. Structurally characterized product ion spectra allow identification of substructures in unknowns



Precursor Ion Fingerprinting

A few years ago, Precursor Ion Fingerprinting (PIF) was developed (Figure 1.). This innovative approach identifies substructural information through the comparison of product ion spectra of structurally related compounds. Structural information is derived by utilizing previously characterized ion structures stored in reference libraries of tandem mass spectral data and matching them with unknown product ion spectra. PIF is a very powerful technique that heavily depends upon libraries containing spectra of precursor ions of various chemical classes acquired at various experimental conditions. Here we present a new type of spectral library providing the functionality required for elucidation of unknowns even if compounds are not present in the library.

Novel Spectral Database

Even if the acquisition process follows a standardized experimental protocol, reproducible product ion spectra are difficult to achieve since the probabilities of various possible ion decomposition products depend on its internal energy that can differ for identical ions if derived from different parent compounds. Since the standardized criteria cannot be met completely, several product ion spectra for the same precursor ion should be acquired using various experimental conditions (collision energies, isolation widths, etc.) to compensate for possible differences between reference and investigated spectra. Various mathematical methods have been developed to harness the spectral dissimilarities to allow correct identification of fragment structures. The idea of this project is to create a library of comprehensive spectral ion trees based on structurally characterized product ion spectra to enable the identification of substructures in unknowns (Figure 4.). Each structurally characterized product ion spectrum will contain the precursor ion m/z value, a list of product ion m/z values with mass accuracies, corresponding absolute and relative intensities, ion polarity, charge state, the structure of the precursor ion, and the structure of the parent molecule.

For the assignment of fragment structures to a precursor ion in the process of creating structurally characterized product ion spectra, it is extremely beneficial to have high-resolution spectra since the accurate m/z values of precursor and product ions greatly reduce the number of possible molecular formulas for fragment structures. Also, the determination of the structural arrangement for the elucidated molecule benefits from exact mass measurements by constraining the elemental composition of the elucidated molecule and consistently validating the calculated mass of recognized fragment structures and accurate m/z values of precursor and product ions.

The m/z Cloud public domain database (Figure 5) aims to provide complete library technology based on spectral ion trees to enable elucidation of unknowns using the precursor ion fingerprinting method.

m/z Cloud Consortium

For the purpose of library creation, precursor ions (isolated, activated, and dissociated ions) can be selected using various criteria; however, the more product ion spectra that are generated, the more comprehensive a library can be built. Spectra and associated experimental data are organized into a logical structure called spectral ion trees (Figure 2.), where branches represent precursor ions and nodes the corresponding product ion spectra. The library is implemented in a relational database that will be accessible through a public domain web site of an emerging consortium named "m/z Cloud". Since the consortium is predominantly oriented towards high-resolution, accurate mass spectra, the database design, spectral management, and library search algorithms require a completely new architecture compared to traditional spectral databases (Figure 3).

FIGURE 5. m/z Cloud architecture

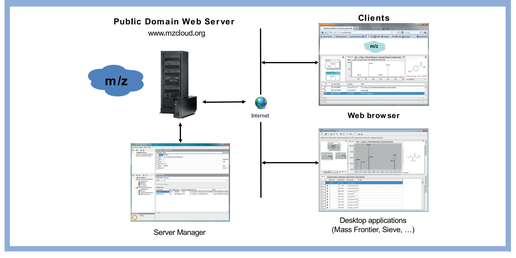
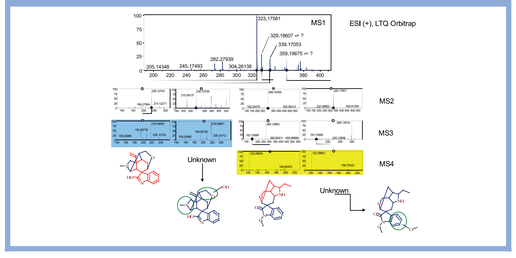


FIGURE 6. Substructure identification using structurally identical fragment



Spectral trees contributors welcomed