Automation and Effective Data Sharing for Metabolite Identification

Richard Lee
Advanced Chemistry Development, Inc.
Toronto, ON, Canada
www.acdlabs.com

Introduction

Fast and accurate identification of metabolites is a major challenge in the study of drug metabolism. Not only do scientists want accurate structure identification, they are typically also burdened with streamlining the process to improve the turnaround time for results. This problem is often exacerbated by the fact that most laboratories have several vendor instruments with their own unique data processing and analysis software resulting in a highly heterogeneous software environment. It is, therefore, problematic for scientists to properly process and analyze data, and makes reporting a monumentally difficult task. In this paper, we present a new approach for manual, semi-automated, or fully automated identification of metabolites.

MetaSense on the ACD/Spectrus Platform is an informatics solution designed with the workflows of scientists responsible for metabolite analysis in mind. It uses the same logical approach in spectral data analysis as an expert would, to associate parent and metabolites to help deduce answers from analytical data.

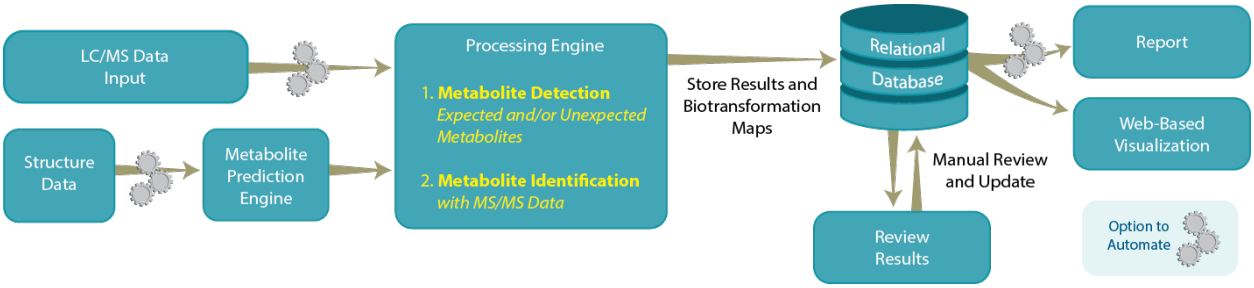

Figure 1: *MetaSense delivers a workflow for the automated identification of metabolites and a solution for sharing knowledge.*

MetaSense provides the following benefits over the current approaches used in metabolite characterization:

- Chemically intelligent data processing and interpretation algorithms help scientists to create and elucidate biotransformation maps of parent structures and their metabolites.
- Information from metabolite prediction is combined with a data-driven approach to provide comprehensive metabolite coverage.
- The burden of a fragmented software environment is reduced with a single software platform designed to support the workflow of metabolite identification.
- All relevant data and results may be captured and shared through a centralized repository with analytical data-centric search capabilities.

Analytical Data Unification in a Homogeneous Software Environment

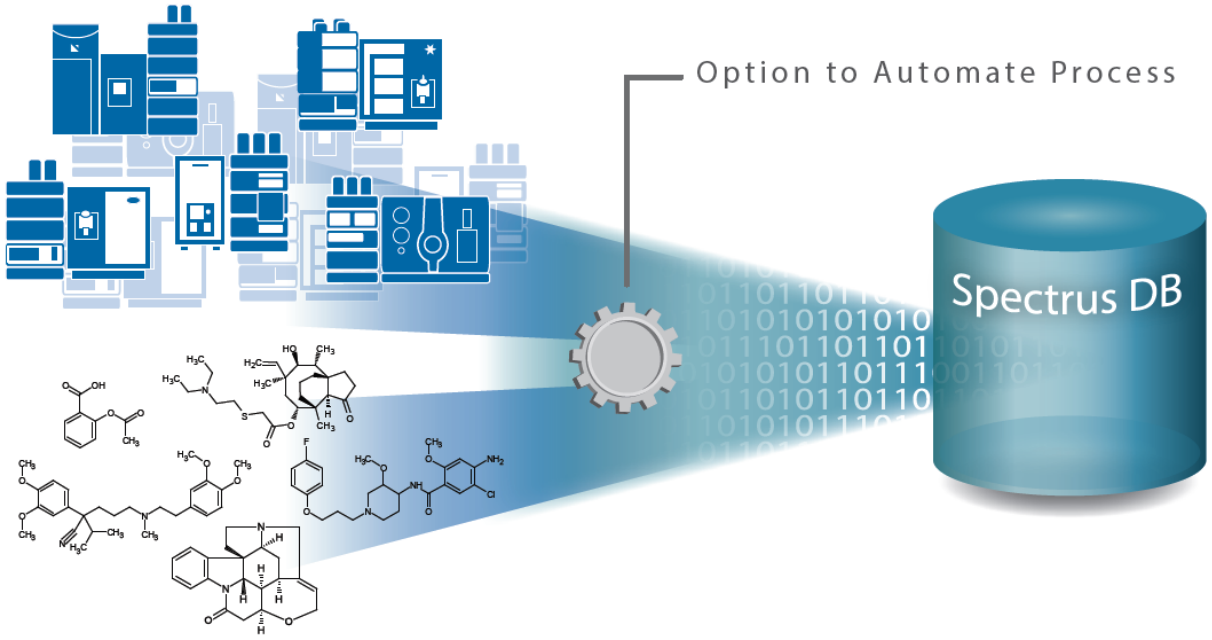

Figure 2: *MetaSense unifies analytical data from various instruments and associated chemical structures into a single environment.*

MetaSense allows data from a variety of instrument vendors to be processed, analyzed, and databased in a uniform software environment. The solution can be configured for full automation, partial supervised analysis, or manual processing with steps including file capture, data processing, and update of results to a centralized repository (Spectrus DB). Regardless of the setup, the solution provides the flexibility for scientists to manually review and revise processing and analysis routines when required.

The initial step imports instrument generated data files into the processing environment (which can be configured for manual or full automation) with corresponding structure files. Once the raw data and structure files are captured, the next step involves metabolite detection and verification consisting of a combination of prediction and data driven analysis. Following data processing, interpreted spectra are uploaded to Spectrus DB and the biotransformation map is automatically created. Scientists can review the full project in one environment with the option to add missed metabolites based on their knowledge.

Metabolite Prediction

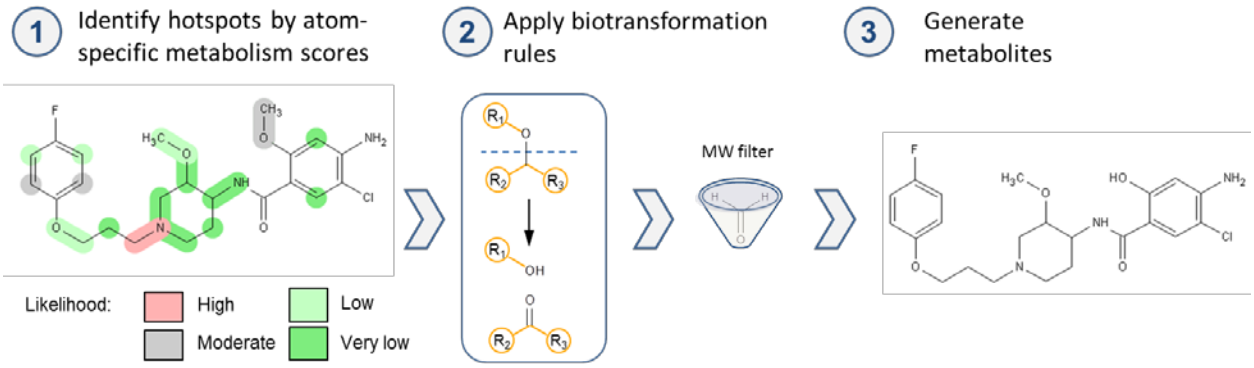

Figure 3: *Phase I and II metabolites may be predicted with MetaSense to help expedite metabolite identification.*

A structure based prediction approach is used to target expected metabolites. Metabolite prediction is performed using an algorithm that consists of several steps. A probabilistic statistical model is used to estimate the likelihood of a metabolic reaction occurring at each potential site of metabolism of the compound of interest. Depending on user-specified settings, the algorithm may run multiple stages of predictions. Metabolites obtained in the first stage may be submitted for further calculations to provide a full metabolic tree that includes Phase I and Phase II metabolites.

Metabolite Detection and Identification

Metabolites are identified via extracted ion chromatograms (XIC). Each XIC may contain several peaks of isobaric metabolites. The site of biotransformation is evaluated using MS/MS spectra from data-dependent acquisition by LC/MS. A combination of common collision-induced dissociation rules and mass shifts of fragment ions are applied to localize the site of biotransformation. Where there is insufficient evidence to support a single site of biotransformation, metabolite structures may be represented using a Markush notation.

Unexpected metabolites are identified using control sample comparison functionality as well as employing fractional mass difference. Unexpected metabolites, characterized by a typical pharmacokinetic profile, are included in the hit list and might be subject of further analyses for structure elucidation.

The automatically generated metabolic scheme, mass spectra, chromatograms, peak areas, and study metadata are stored in a database for future use. While this is an automated process, biotransformation scientists are able to examine the results and make necessary manual changes—such as modifying a structure assignment—based on their experience and understanding of the project.

The overview screen (as shown in Figure 3) provides quick access to all the key information relating to each metabolite. It allows the reviewer to see all relevant and related spectral, chromatographic, and structural data at once so they can check the quality of each automated metabolite assignment. The

metabolism map and base peak chromatogram (BPC) provide the context of the metabolite, while the XIC combined with the MS/MS spectra provide information regarding each assignment. The kinetic profiles for identified metabolites are provided in a table format as well as a plot to monitor the stability of the parent compound or the onset and growth of its metabolites.

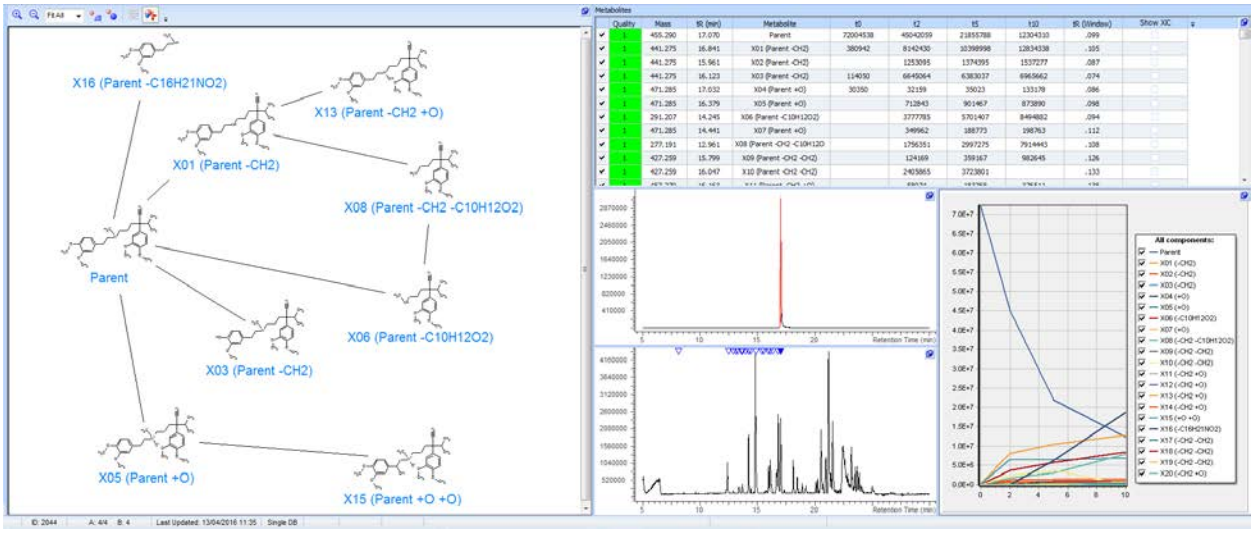

Figure 4: *A database record for the metabolites of Verapamil in MetaSense. The record provides the full biotransformation map, and live spectral and chromatographic data associated with the relevant metabolite. Peak areas are tracked through the entire incubation study and reflected in the stability/kinetic plot.*

Mass shifts of fragment ions and the common collision-induced dissociation rules are applied to localize the site of biotransformation. Possible metabolite structures are analyzed against the experimental MS/MS spectra of the detected metabolite and parent to determine the correct structure. Based on this comparison, score values are calculated for each metabolite for user review (Figure 4).

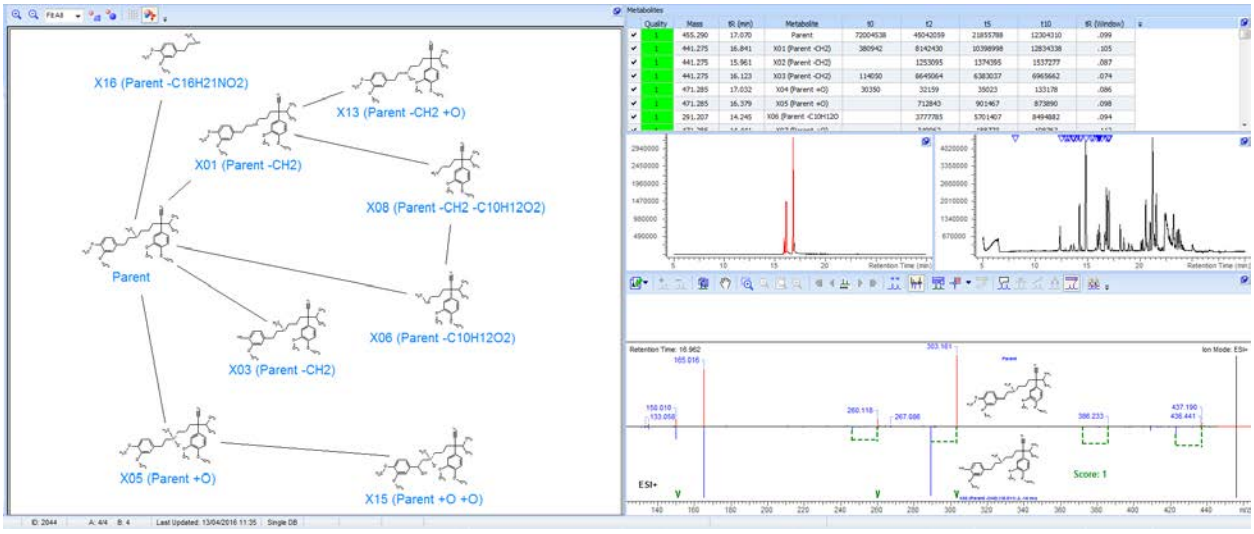

Figure 5: *Structure oriented view allows scientists to review MS/MS spectra from the parent and putative metabolites to assess structural correctness.*

Where the observed data does not conclusively support a single metabolite isomer, Markush notations are used. Chemical intelligence is also employed to exclude chemically unfeasible biotransformation steps from the metabolism map.

Conclusion

MetaSense enables metabolism studies, conducted on data from various instrument vendors, to be accelerated without compromising the quality of interpretation. The configurability of the software lends itself to supporting a number of different workflows including high-throughput where the system can be configured for automated file capture and automated batch processing. Studies that may otherwise have been avoided due to laborious data evaluation—for example, a combination of quantitative studies where larger numbers of samples are measured (e.g., determination of intrinsic clearance), and qualitative studies (e.g., identification of metabolites)—may be undertaken for a more complete biotransformation map and greater confidence in results.