Poster Note PN-33

Boosting compound identification confidence

by exploiting all HRAM spectral information integrating accurate mass, true isotopic pattern, in-source fragmentation, MS/MS fragmentation, and retention time

Introduction

Confident compound identification is still a major bottleneck in metabolomics. While there are ongoing efforts to refine the definitions and levels of metabolite identifications that were first proposed by the MSI initiative (1-4), it is clear that higher levels of confidence can be reached by joining accurate measurement technology, orthogonal molecular features, and sophisticated software tools. Here we present a single integrated software solution for enhancing the confidence in identifications at different levels: molecular formula, compound class, structure, or verified targeted identification (**see Fig. 1**). This highly integrated functionality is implemented in the new MetaboScape® 2.0 software.

Authors

Frederik Walter, Marcus Persicke, Jörn Kalinowski
Bielefeld University, Bielefeld, Germany

Nikolas Kessler, Stephan Mävers, Matthias Szesny, Aiko Barsch, Heiko Neuweger
Bruker Daltonik GmbH, 28359 Bremen, Germany

Keywords	Instrumentation and Software
Metabolomics	MetaboScape 2.0
Compound identification	SmartFormula 3D
In-silico fragmentation	MetFrag
Spectral Library	Bruker HMDB Metabolite Library

Molecular Formulas

Classically, molecular formula generation is based on a single precursor mass, using the indicated mass tolerance, and a limited set of elements (typically CHNOPS). SmartFormula™ allows to additionally consider further properties and prior knowledge:

- True Isotopic Pattern™ (TIP)
- "Golden Rules" (5)
  - Upper formula
  - Element ratios
  - Senior and Lewis rule
  - Element count probability
- In-source fragmentation

SmartFormula 3D provides an intelligent interface to allow users to annotate all monoisotopic peaks of an MS/MS spectrum with fragment formulas (**see Fig. 2**).

Structural Candidates Search

To obtain candidate structures for previously generated molecular formulas, the CompoundCrawler™ tool is available within MetaboScape. This tool queries a customizable set of private or public structural databases, including KEGG, PubChem, ChemSpider, and ChEBI. Matching structures can then be assessed using *in-silico* fragmentation.

MetFrag *In-Silico* Fragmentation

The MetFrag (6,7) *in-silico* MS/MS fragmentation algorithm is integrated in MetaboScape 2.0 and can be applied to assess and score structural candidates from CompoundCrawler search results to acquired MS/MS spectral information. Explained fragments are highlighted in the original compound structure (**see Fig. 3**).

MS/MS Spectral Libraries

Users can create or import MS/MS spectral libraries. These can then be used to annotate compounds, if the spectral similarity meets the user-defined thresholds. An interactive spectral library editor is embedded into the novel MetaboScape 2.0 software.

Targeted Analyte Lists

Compounds with known retention times are easily annotated using so-called Analyte Lists. Optionally, compounds can be linked to entries of spectral libraries, in order to include MS/MS spectral matching. Analyte Lists thus allow to integrating MS, MS/MS, and RT information for confident dereplication.

Conclusions

- Comprehensive set of tools for compound annotation at all levels of specificity
- Integrating HRAM, true isotopic pattern, MS/MS fragmentation, and retention time to boost confidence in compound ID
- Integrating all tools to create application-specific workflows:
  - SmartFormula 3D embedded
  - CompoundCrawler embedded
  - MetFrag embedded
  - MS/MS spectral libraries matching and management
  - AnalyteLists for targeted compound identification

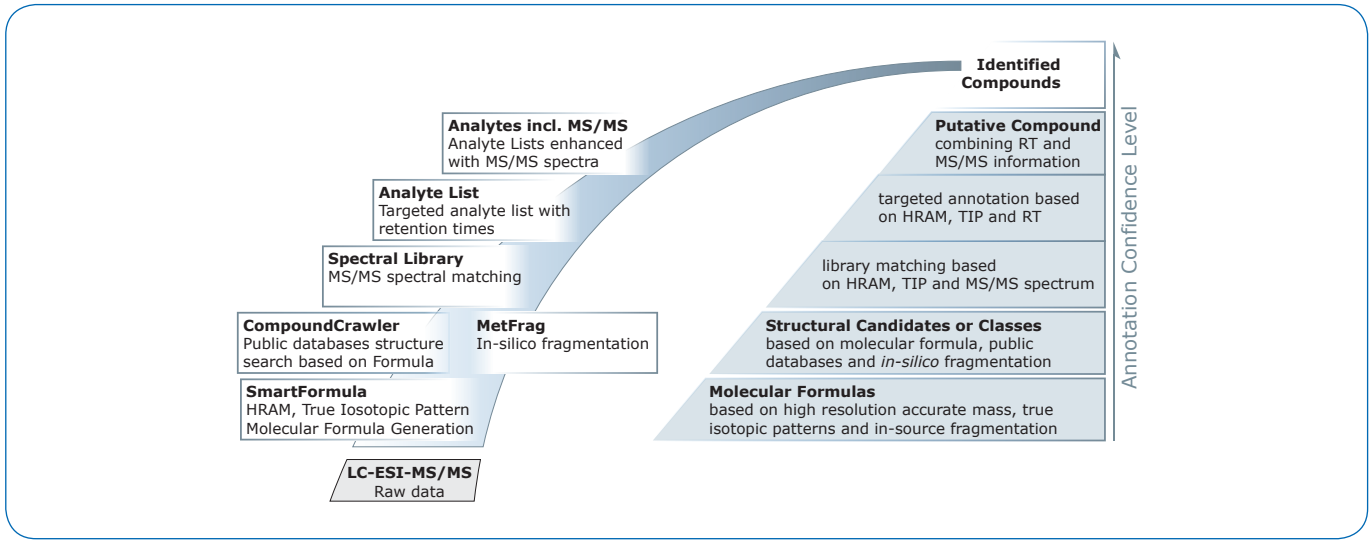

Fig. 1 In MetaboScape 2.0, compound identification is supported by a variety of (semi-)automated tools, which are highly integrated and enable the creation of annotations at increasing levels of specificity and confidence.

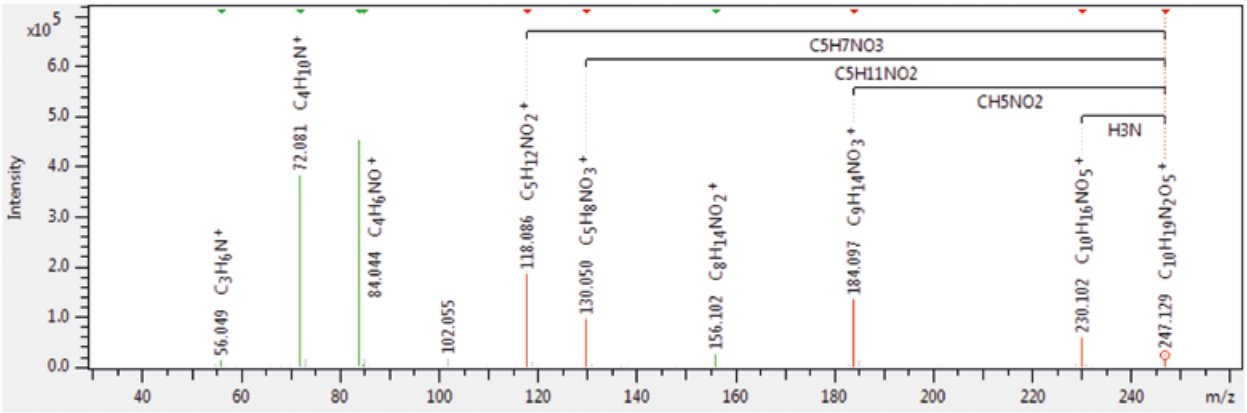

Fig. 2 An example MS/MS spectrum of gamma-glutamyl valine in Smart-Formula 3D, highlighting characteristic fragments corresponding to the losses of NH_3 and CH_5NO_2. These are indicators for gamma-glutamyl dipeptites,rather than alpha-glutamyl dipeptides (8).

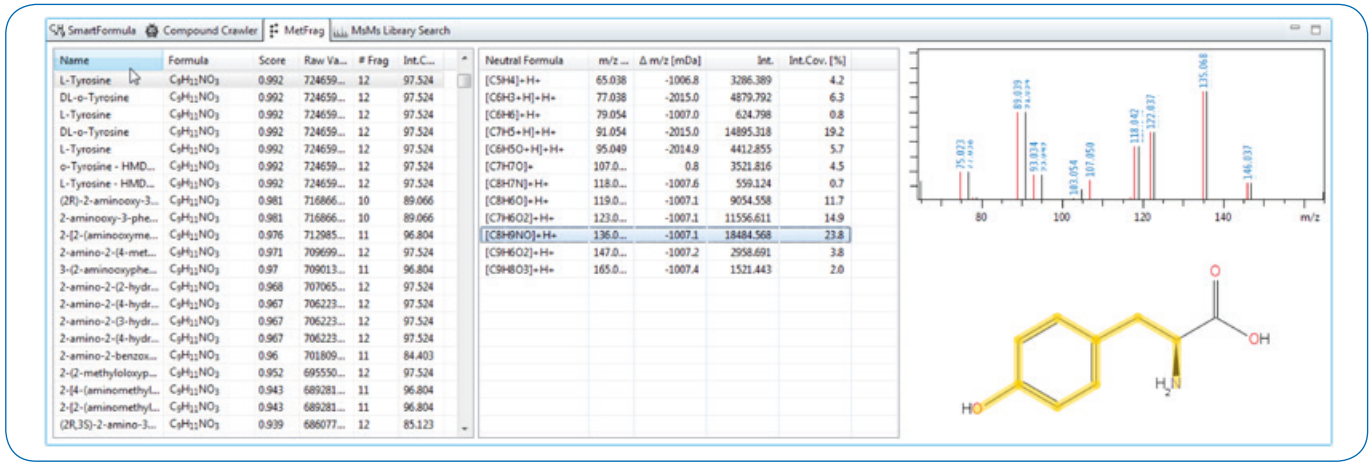

Fig. 3 Screenshot of a MetFrag result in MetaboScape. For the precursor 182.0811 m/z SmartFormula generated three possible molecular formulas. Structural candidates for these were determined using CompoundCrawler (PubChem and custom MetaboScape AnalyteDB). For Tyrosine 97.52 % of the MS/MS intensity could be covered with explained fragments. The annotation was confirmed by an excellent Spectral Library (Bruker HMDB Metabolite Library) match and could be explained in SmartFormula 3D.

References

[1] Fiehn et al.; Metabolomics 2007, 3:175-178
[2] Sumner et al.; Metabolomics 2007, 3:211-221
[3] Sumner et al.; Metabolomics 2014, 6:1047-1049
[4] Schymanski et al.; Environ. Sci. Technol. 2014, 48:2097–2098
[5] Kind and Fiehn; BMC Bioinformatics 2007, 8:105
[6] Wolf et al.; BMC Bioinformatics 2010, 11:148
[7] Ruttkies et al.; J. Cheminf. 2016, 8:3
[8] Harrison; J Mass Spectrom. 2013, 38:174-187

For research use only. Not for use in diagnostic procedures.

Bruker Daltonik GmbH

Bremen · Germany
Phone +49 (0)421-2205-0

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

ms.sales.bdal@bruker.com - www.bruker.com