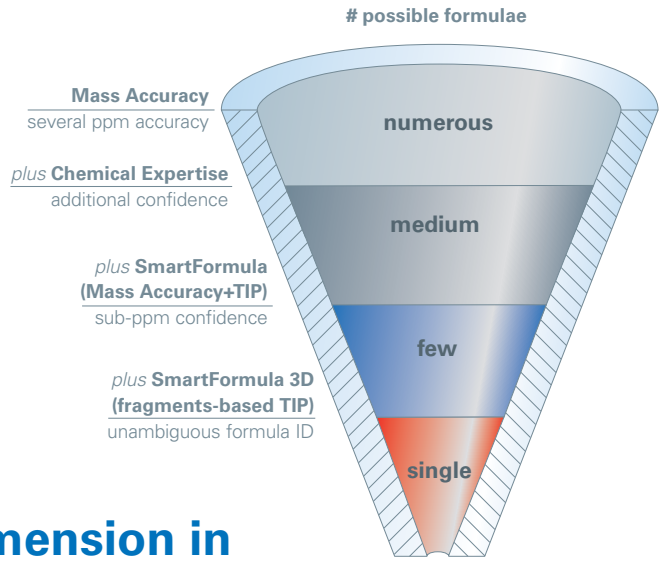


Technical Note TN-26

SmartFormula 3D – the new Dimension in Substance Identification – From Mass Spectrum to Chemical Formula

The identification of an unknown chemical substance by mass spectrometry alone is still a challenge even for experienced scientists. A mass accuracy of 2 ppm, which is typical for the latest benchtop time-of-flight mass spectrometers, is insufficient for unambiguous identification of larger molecules. It has been shown recently that even a mass accuracy of 0.1ppm (which no mass spectrometer can achieve) cannot guarantee a clear assignment of the molecular formula. For identification of metabolites it has been calculated that mass spectrometers with 3ppm mass accuracy and 2% error for isotopic abundance pattern outperform mass spectrometers with less than 1ppm mass accuracy or even 0.1ppm mass accuracy (calculated from simulations) which do not include the isotopic pattern in the calculation of the molecular formulae [1]. Quite often other information is needed for a distinct identification such as NMR data or elemental analysis if the isotopic pattern is not considered in the calculation. Collecting this information (even if possible) requires time and money. It would be much easier if the necessary information for compound identification could be extracted from just mass spectra.

Authors

Dr. Verena Tellstroem, Dr. Ralf Dunsbach
Bruker Daltonics GmbH, Bremen, Germany

Keywords	Instrumentation and Software
Formula Generation	Bruker Q-TOF
Chemical ID	

Precision in formula generation: True isotopic pattern of fragments

Common structure elucidation algorithms just use the exact mass of a compound. The SmartFormula approach we have shown before also considers the isotopic pattern distribution for MS spectra. In SmartFormula 3D, exact mass and isotopic patterns of the fragment ions are used. Using this algorithm, the number of possible hits is in most cases reduced to 1. The sophisticated algorithm combines all information of MS and MS/MS data, thus dramatically reducing the number of possible formulae. In the first step, the program suggests a list of formulae, based on the MS spectrum. In the next step SmartFormula 3D automatically calculates the sum formulae of the fragments by using exact mass and true isotopic pattern. As the fragment-ions must be a subset of the precursor-ion, usually just a few formulae will remain. All this calculation is done without using reference data or databases therefore it works without restriction also on real unknown compounds. SmartFormula 3D is an outstanding tool which delivers well-defined formulae up to 1 kDa. Even bigger molecules can be identified with unchallenged reliability.

Example: ID of Minocycline

Minocycline is a synthetic drug and belongs to the tetracycline family. It is widely used as bacteriostatic drug against gram positive and gram negative bacteria. Minocycline has a nominal molecular mass of 457 Da.

The MS spectrum shows the protonated species at 458.19 Da. The mass accuracy is 1.2ppm with external calibration.

From the MS spectrum, a list of 15 possible compounds was generated in the first step. The number of phosphorous and sulphur atoms in the molecule was restricted to 2 atoms per element. The tolerance was set to 5ppm. (fig. 2)

By using the MS/MS information (fig. 3), the list of possible formulae can be reduced dramatically. In the present example the 3 most intense peaks in the MS/MS spectrum were used to reduce the list further. The selection of peaks which must be explained can be done manually or automatically via the intensity of the fragment peaks.

The information of the MS/MS spectrum in combination with the true isotopic pattern algorithm on the fragment ions reduces the list of possible formulae for the precursor ion to just 1 (fig. 4).

MS spectrum of minocycline shows precursor ion at m/z 458.19

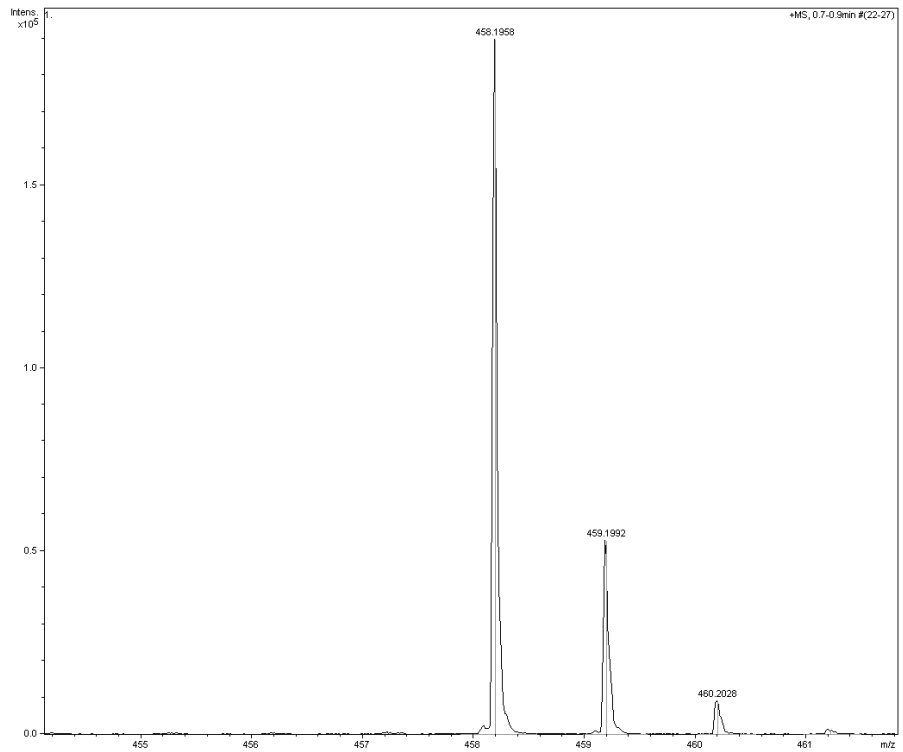


Fig. 1: Mass spectrum of minocycline, external calibration. A micrOTOF-Q II was used delivering 1.2 ppm mass accuracy.

The MS/MS spectrum supplies the user with some basic structural information as well. In the present example the neutral loss of 45 Da shows the elimination of formamide which indicates the presence of an amido group in the molecule. (fig. 5)

The first step: 15 possible sum formulae were identified from the MS spectrum

SmartFormula Manually

Min C_{18} — Generate

Max S_2P_2

C 18-n, P 0-2, S 0-2 — Help

Note: for m < 2000 the elements C, H, N, and O are considered implicitly.

Measured m/z 458.1916 Tolerance 5 ppm Charge 1

#	Mol. Formula	m/z	|err| [ppm]	Sigma Rank
1	C 23 H 28 N 3 O 7	458.1922	1.2	1
2	C 20 H 20 N 13 O	458.1908	1.8	2
3	C 24 H 24 N 7 O 3	458.1935	4.1	3
4	C 21 H 33 N O 8 P	458.1938	4.8	4
5	C 19 H 24 N 9 O 5	458.1895	4.7	5
6	C 19 H 34 N 5 O 2 P 2 S	458.1903	2.9	6
7	C 27 H 30 N 3 P 2	458.1909	1.5	7
8	C 18 H 25 N 11 O 2 P	458.1925	1.8	8
9	C 25 H 33 N O 3 P S	458.1913	0.7	9
10	C 27 H 28 N 3 O 2 S	458.1897	4.3	10
11	C 18 H 33 N 7 O P S 2	458.1920	0.8	11
12	C 20 H 28 N 9 S 2	458.1904	2.8	12
13	C 24 H 32 N 3 O 2 S 2	458.1930	3.1	13
14	C 32 H 28 N S	458.1937	4.5	14
15	C 35 H 24 N	458.1903	2.9	15

Fig. 2: List of compounds, number of C atoms automatically estimated by software

Reducing the formula candidates list by fragments analysis

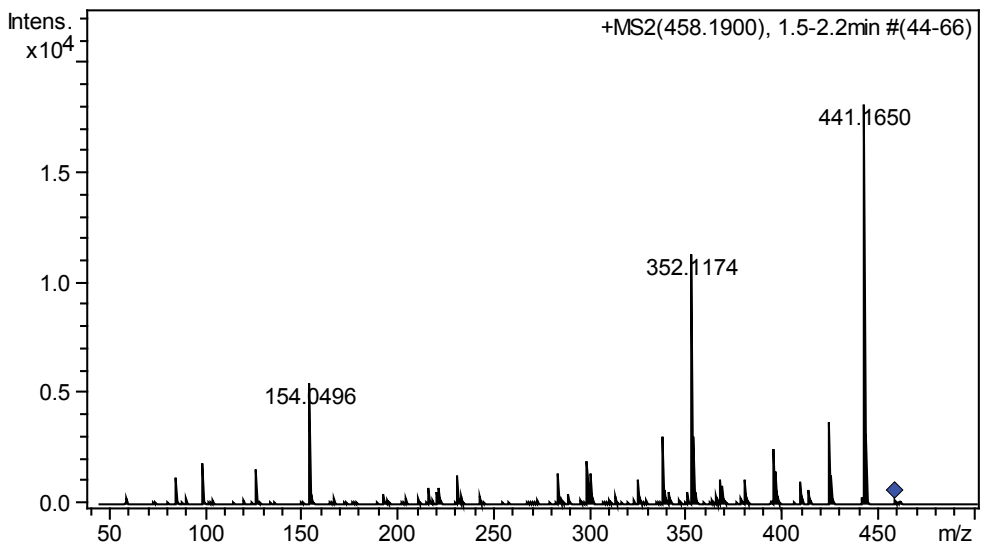
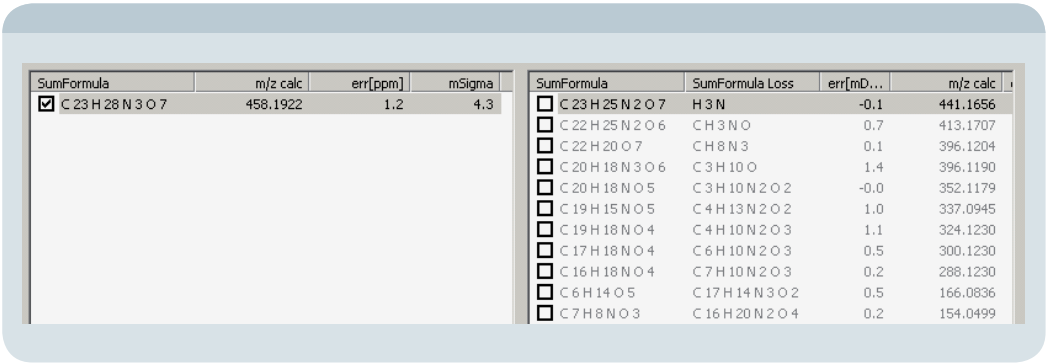


Fig. 3: MS/MS spectrum of monocycline.

Formula ID without Doubt

SmartFormula 3D provides the user with an easy-to-use, powerful tool for the identification of complex molecules. The combination of exact mass, true isotopic pattern, reliable MS/MS spectra and a powerful algorithm, delivers a useful tool for structural identification by mass spectrometry.

A lot of elemental compositions with a molecular weight of 457 Da are possible. Nonetheless the correct sum formula was identified clearly just from the mass spectrum. Also for bigger molecules the software is a helpful tool for unambigous formula generation.



SumFormula	m/z calc	err[ppm]	mSigma
C 23 H 28 N 3 O 7	458.1922	1.2	4.3

SumFormula	SumFormula Loss	err[mD...	m/z calc
C 23 H 25 N 2 O 7	H 3 N	-0.1	441.1656
C 22 H 25 N 2 O 6	C H 3 N O	0.7	413.1707
C 22 H 20 O 7	C H 8 N 3	0.1	396.1204
C 20 H 18 N 3 O 6	C 3 H 10 O	1.4	396.1190
C 20 H 18 N O 5	C 3 H 10 N 2 O 2	-0.0	352.1179
C 19 H 15 N O 5	C 4 H 13 N 2 O 2	1.0	337.0945
C 19 H 18 N O 4	C 4 H 10 N 2 O 3	1.1	324.1230
C 17 H 18 N O 4	C 6 H 10 N 2 O 3	0.5	300.1230
C 16 H 18 N O 4	C 7 H 10 N 2 O 3	0.2	288.1230
C 6 H 14 O 5	C 17 H 14 N 3 O 2	0.5	166.0836
C 7 H 8 N O 3	C 16 H 20 N 2 O 4	0.2	154.0499

Fig. 4: List of sum formulae after reduction by accurate mass and True Isotopic Pattern.

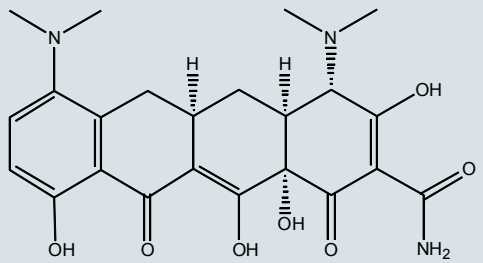


Fig. 5: Chemical structure of minocycline

References

[1] Kind, Fiehn: Metabolomic database annotations via query of elemental composition: Mass accuracy is insufficient even at less than 1 ppm, BMC Bioinformatics 2006, 7:234

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

Bruker Daltonik GmbH

Bremen · Germany
Phone +49 (0)421-2205-0
Fax +49 (0)421-2205-103
ms.sales.bdal@bruker.com

www.bruker.com

Bruker Daltonics Inc.

Billerica, MA · USA
Phone +1 (978) 663-3660
Fax +1 (978) 667-5993
ms.sales.bdal@bruker.com