

Sum Formula Calculation and Identification of a Bacterial Metabolite with $m/z > 1100$

Wiebke Lohmann, Aiko Barsch, Gabriela Zurek

Bruker Daltonik GmbH, Fahrenheitstraße 4, 28359 Bremen, Germany

Introduction

In the course of identification and structural characterization of compounds, the unambiguous sum formula calculation is an essential step. With increasing molecular mass, the number of possible sum formulae is increasing exponentially, so that already for m/z values around 500 it is not possible anymore to get just one sum formula suggestion from an accurate mass measurement only, even if a mass accuracy of 0.1 ppm is assumed (Fig. 1).

Calculation constraints like ranges for meaningful H/C ratios, number of rings/double bonds and careful selection of elements allowed for sum formula calculation will reduce the number of suggestions significantly and help to exclude chemically unlikely molecules from the result list. However, for m/z 1000 at 0.1 ppm mass accuracy more than 2200 sum formulae would still be calculated (>10.000 for mass accuracy 1 ppm).

On the other hand, the combination of accurate mass information with additional information like isotopic pattern derived both from full scan and MS/MS spectra can significantly extend the m/z range for reliable sum formula suggestion to values higher than 1000.

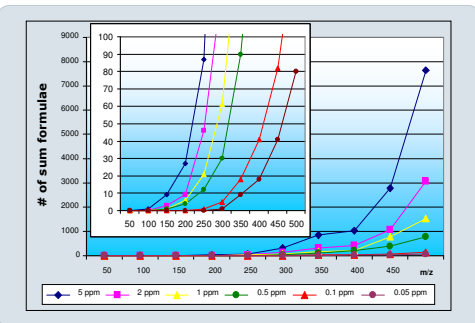


Fig. 1: Theoretical number of calculated sum formulae in dependence on compound mass and mass accuracy (search range: C₂₀₀H₄₀₀N₅₀O₅₀S₅P₅F₅Cl₅Br₅I₅, even electron configuration, no further constraints).

Methods

HPLC: Ultimate 3000 Rapid Separation LC ("RSLC", Dionex).

Column: Zorbax XDB-C8 2.1x50 mm, 1.8 μm (Agilent).

Column temp. 30 °C. **Mobile phase:** A = H₂O, B = MeOH (each 0.1% HCOOH). **Gradient:** linear gradient 85 – 99.5% in 3 min, hold 5 min. **Flow rate:** 0.2 mL/min. **Injection volume:** 1 μL.

Sample: bacterial metabolite, provided by CVUA Stuttgart. **MS: micrOTOF-Q II, maXis** (ESI-Qq-TOF MS, UHR-TOF MS, Bruker Daltonik GmbH).

Ionization: ESI(+). **Scan range:** m/z 50-1300. **Scan mode:** Fullscan, MSMS 1171.

Calibration: external: sodium formate clusters injected at the beginning of each chromatographic run. For maXis: additional internal calibration, m/z 1221.990637 (Hexakis(1H,1H,4H-hexafluorobutyl)phosphazine, C₂₄H₁₉O₆N₃P₃F₃₆).

A solution containing a bacterial metabolite that gave an MS signal at about m/z 1171 was analyzed using a reversed phase gradient separation on an UHPLC system interfaced to an ESI-TOF-MS or an ultrahigh-resolution ESI-TOF-MS system. Full scan data was acquired in ESI positive mode (scan range 50-1300 m/z). An additional run in MS/MS mode was performed, acquiring full scan MS/MS spectra for a precursor mass of m/z 1171. Sodium formate solution was injected as external calibration standard at beginning of each chromatographic run. Sum formula suggestions were calculated using DataAnalysis™ 4.0 SR2 (Bruker Daltonik GmbH) applying different settings.

Results

Full scan mass spectra on the micrOTOF-Q II gave an exact mass value of m/z 1170.7112 for the bacterial metabolite with a mass accuracy of 5 ppm according to instrument specification for external mass calibration. Additional MS signals of lower intensity allowed for the characterization of the observed ion as ammonium adduct (presence of [M+H]⁺, sodium and potassium adducts, Fig. 2a). Allowing only C, H, N, O as elements 30 sum formula suggestions were calculated using the 5 ppm mass accuracy window (Fig. 3a).

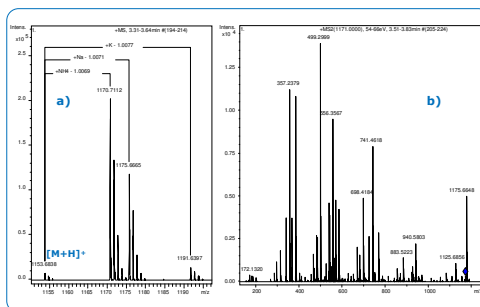


Fig. 2: Fullscan (a) and MS/MS spectrum (b) of the metabolite. The adduct pattern observed in the fullscan spectrum allows for a clear assignment of the [M+H]⁺ signal.

The data evaluation software allowed the rating of results according to the matching of experimental and theoretical isotope patterns. Taking isotope pattern information into account the number of meaningful suggestions was reduced to about 12 (Fig. 3a).

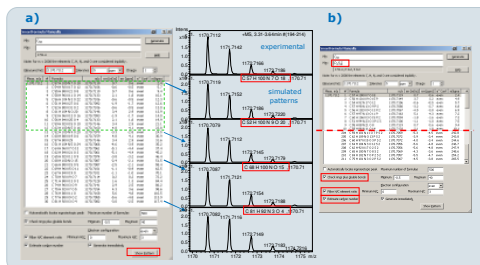


Fig. 3: Sum formula calculations for m/z 1170.7112 using a calculation window of 5 ppm and chemical constraints: The number of sum formula suggestions depends extremely on the predefined element range.

If additional elements were allowed, the number of theoretical suggestions increased rapidly to several hundred, even if rules were applied, which filter the results for example for validity of the nitrogen rule, electron configuration, reasonable C/H ratios or a minimum number of C atoms according to the intensity of the second isotope (Fig. 3b).

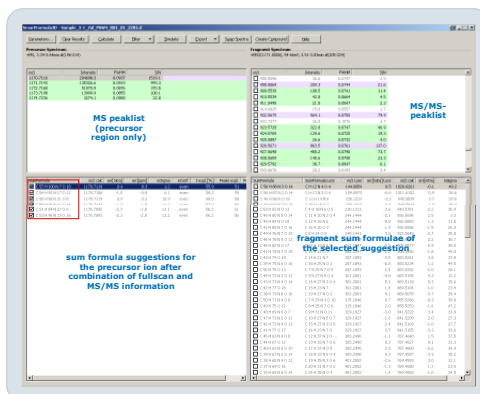


Fig. 4: Principle of SmartFormula3D™ and result for micrOTOF-Q data (5 ppm calculation window): only six sum formula candidates left.

The additional tool Smart Formula 3D™ allowed for the verification of sum formula suggestions by combining full scan MS and MS/MS data, taking also the fragments' accurate mass and isotope pattern information into account. Only sum formulae remain, which can explain a specified minimum number of fragment signals. This reduced the sum formula candidate list to six (Fig. 4).

Only one of these five sum formulae could be found in online available databases (like ChemSpider), using the CompoundCrawler software tool (Fig. 5). Thus the bacterial metabolite could putatively be identified as cereulide.

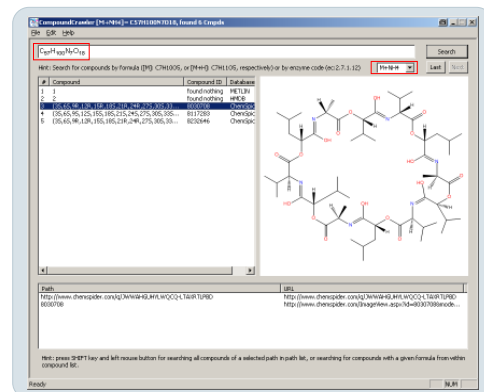


Fig. 5: Online search result for the sum formula suggestions from SF3D: only for one compounds and structures can be assigned.

An additional similar analysis using an UHR-TOF-MS system (maXis) showed the benefit of enhanced mass accuracy (1 ppm for internal mass calibration), since this approach only left one sum formula in addition to the cereulide formula for combined mass, isotope and fragment information (Fig. 6).

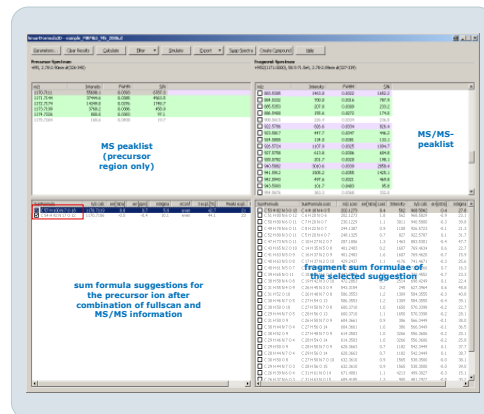


Fig. 6: Principle of SmartFormula3D™ and result for maXis data (1 ppm calculation window): only two sum formula candidates left.

Conclusions

- Combination of information from MS and MS/MS spectra (accurate mass and isotope pattern) together with a query in online available databases allows for sum formula generation and identification of compounds even at high m/z values > 1000.

Acknowledgment

Many thanks to Roland Perz (CVUA Stuttgart, Germany) for providing this sample!