In-house built accurate mass spectral library for GC/ToF-based metabolomics analyses

Manuela Cerdán-Calero*, **Juan C. García-Cañaveras and Agustín Lahoz**
Biomarkers and Precision Medicine Research Unit.
Instituto de Investigación Sanitaria - Hospital La Fe, Valencia, Spain
*Corresponding author: mcerdan86@gmail.com (M. Cerdán)

1st GC/Q-TOF Forum
Barcelona. 24 & 25 November, 2016

Agilent Technologies

Introduction

Gas chromatography coupled with mass spectrometry (GC/MS) is one of the most widely used technologies for qualitative and quantitative analysis of small molecules, including untargeted metabolomic approaches. Compared with traditional single quadrupole systems, modern quadrupole time of flight (Q-ToF) systems offer better sensitivity and selectivity and allow for accurate mass determination.

Despite the continuous improvement in GC/MS technology, current commercially available spectral libraries only provide unit mass spectra thus hampering to take the maximum profit of modern mass spectrometers in terms of data processing. To help improve identification of metabolites, an in-house built accurate mass EI spectral library (with extension .cdb) has been created using commercially available standards of metabolites of interest within the biomedicine research field. This library contains characteristic mass spectrum, chromatographic retention time (RT), formula of significant characteristic fragment (or molecular) ions, among other relevant information.

The capabilities of the in-house built accurate mass spectral library were tested and compared to commercially available unit mass spectral libraries (i.e. NIST and Fiehn) by the use of different software and options within Agilent MassHunter tools.

Experimental

Sample Preparation

Analytical standard mixture solutions were prepared to reach a final concentration of 1-10 ppm and vacuum-dried. Metabolite extraction and protein precipitation of pooled serum samples (with or without the addition of a mixture of 70 metabolite standards at a concentration of 1ppm) was performed by the addition of 3 volumes of methanol to 100 µL of sample. After centrifugation the clean supernatant was collected and vacuum-dried. Dried samples were derivatized by methoximation using 20 µL of a MEOX solution (20 mg/mL in pyridine) followed by silytation using 80 µL of a 1% TMCS solution in MSTFA (Figure 1).

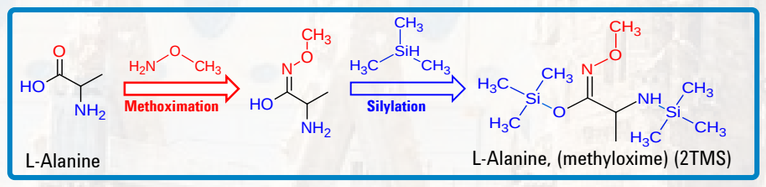

Figure 1. TMS Derivatization procedure.

GC/MS Analytical Conditions

GC-MS analysis was performed using an Agilent 7890A gas chromatograph coupled to an Agilent 7200 accurate mass high resolution GC/Q-ToF. Separation was performed using an Agilent DB-5ms + DG capillary column (30 m x 0.25 mm i.d., 0.25 µm film thickness + 10 m Duraguard) using Helium as carrier gas. Mass analysis was operated on EI conditions, recording data in full-scan mode at 70eV in a mass range of m/z 50 to 600. Ion source, quadrupole and transfer line temperatures were 250 ºC, 150ºC and 290 ºC respectively.

Data Processing Software Tools

In a first step, MassHunter Qualitative Analysis (B.07.00 SP2) software tools were used to automatically convert accurate mass of fragment ions into theoretical masses for all the abundant fragments in the spectrum and to import the spectra into accurate mass metabolomics Personal Compound Database and Library (PCDL). The identity and actual mass of the identified fragments were confirmed by the use of ACDS/MS Fragmenter (v12.01) software. A characteristic molecular or fragment ion was selected for each metabolite/standard.

In last step, the metabolite screening with the accurate mass metabolomics library was performed in two different software: i) MassHunter Qualitative Analysis, using option Find Compounds by Formula (mass of characteristic molecular or fragment ion, RT and fragment spectra confirmation were required); and ii) Unknowns Analysis standalone tool of MassHunter Quantitative Analysis (B.07.01).

Results and Discussion

In-House Built Accurate Mass Metabolomics GC/Q-ToF Library

The metabolites to be included in the in-house built accurate mass spectral library were selected based on a double criteria: i) biological relevance within the biomedicine research field; ii) feasibility for TMS derivatization. Based on that criteria almost 200 commercially available analytical standards covering a wide range of chemical structures and biological functions (Figure 2) were derivatized and analyzed using the GC/Q-ToF analytical platform. For each standard the following information was included in the spectral library: i) retention time; ii) annotated accurate mass EI spectra, for each compound the included spectra contained the actual fragment relative distribution but for each relevant fragment its theoretical mass and formula was calculated and corrected using in silyco fragmentation software ACDS/MS Fragmenter, and iii) characteristic fragment or molecular ion. The selection of the characteristic ion is of utmost relevance as it facilitates the posterior data processing and thus it has among other characteristics it has be account for possible co-eluting or similar (even isobaric) metabolites.

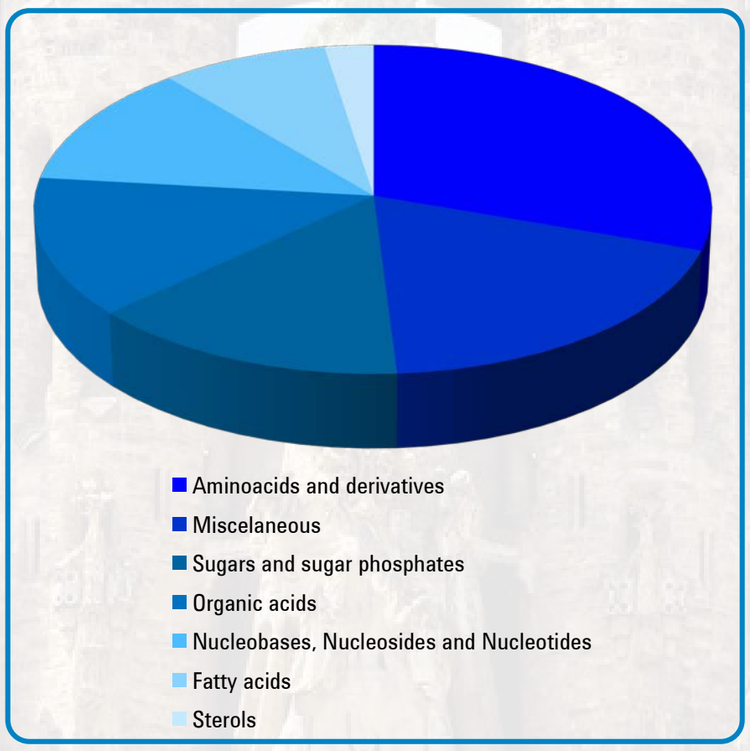

Figure 2. Classification of the compounds added to the Accurate Mass Metabolomics Library.

Accurate Mass Metabolomics Screening Using different MassHunter tools and the Accurate Mass Metabolomics PCD compared to commercially available unit mass spectral libraries

To test the capabilities of the in-house built accurate mass spectral library an analytical standard mixture solution (70 compounds at 1ppm) and pooled human serum samples (with or without the addition of the above-mentioned standard mixture solution) were prepared and analyzed. Moreover two different data analysis strategies were compared. On one hand Unkwown Analysis deconvolution followed by library identification using either commercially available unit mass spectral libraries or the in-house built mass spectral library. And on the other hand, and only using the in-house built library due to format compatibility issues, Qualitative Analysis software (Table 1).

Library	Unknown Analysis				Qualitative Analysis
	NIST	Fiehn		In-house	In-house
	No RT	No RT	RT	RT	RT
Total Hits	3826	239	235	75	136
Standards	43	54	63	33	64
Spiked Serum	61	73	79	29	79
Serum	45	51	60	15	49

Table 1. Performance comparison of the different data analysis strategies using different types of samples.

As expected, and due to size issues, NIST and Fiehn libraries provided a higher number of hits compared to the in-house built library. However, when comparing the performance of the different spectral libraries either with the standards solution or the increase in the number of identified compounds in spiked serum samples compared to non-spiked serum samples, the best results were obtained with Qualitative Analysis software using the in-house built spectral library. The use of the in-house built spectral library not only allows for a more accurate identification of the metabolites thus minimizing false-positive identifications but also allows to use the Find-by-Formula option in Qualitative Analysis software (Figure 3). Such alternative, that can not be performed using actual commercially available unit mass libraries, performed a directed search of the data via the use of the characteristic ions that simplifies data processing and extremely accelerates data processing. However, the number of metabolites included in the in-house built library to expand its identification capabilities.

Figure 3. Accurate mass metabolomics screening in MassHunter Qualitative Analysis using option Find by Formula

Conclusions

- An accurate mass spectral in-house built metabolomics PCDL library containing 200 compound entries has been created using TMS derivatization, GC/Q-ToF analysis and MassHunter software tools.
- The in-house built library provides a more accurate identification of metabolites of interest and accelerates data processing thank to its compatibility with Find-by-Formula function of MassHunter Qualitative Analysis.