Application Note

Metabolomics

Agilent

Trusted Answers

# Comprehensive Accurate Mass Metabolomics Library and Its Evaluation in Targeted and Nontargeted Data Analysis Workflows


**Authors**

Luis Valdiviez,
Shunyang Wang,
Wasim Sandhu, Honglian Ye,
and Oliver Fiehn
West Coast Metabolomics
Center, University of California,
Davis, CA.

Sofia Nieto
Agilent Technologies, Inc.



## Abstract

A confident annotation and identification of metabolites is essential to provide biological context in metabolomics studies. It helps to understand which biochemical pathways are affected by disease states and candidate drugs, and to advance the fundamental biological studies. Comprehensive spectral libraries help to address the critical challenge of identifying metabolites that participate in biological pathways, thus aiding in biological interpretation. This application note describes the new Agilent Fiehn accurate mass metabolomics personal compound database and library (PCDL), as well as its characterization and evaluation in both targeted and nontargeted screening workflows using blood plasma samples. The availability of this accurate mass GC/MS library for metabolomics applications allows investigators to expand the scope, sensitivity, and reliability of the metabolite identifications in complex biological matrices.

## Introduction

GC/MS has long been established as a valuable tool for targeted and untargeted metabolite profiling. It is well known for its ruggedness, sharp peak shapes, and for generating reproducible electron ionization (EI) spectra that can be used for metabolite identification. More recently, with the advent of high-resolution accurate mass time-of-flight (TOF) instrumentation, investigators have even more powerful resources to identify a higher number of biologically relevant molecules with more confidence.

GC/MS mass spectral libraries have been proven to be essential tools that enable quick compound identification with minimal effort.[1-4] Both retention time/index and accurate mass information provide more evidence that helps to confirm or reject compound identity and effectively decrease the rate of false positives.[5,6] Accurate mass GC/MS libraries are particularly useful for applications that involve difficult matrices or where the compounds of interest have a complex elemental formula—metabolomics being one of these applications. To help improve reliability and increase throughput in metabolomics applications, an accurate mass retention time locked (RTL) EI library of compounds that are generally found in biological matrices has been created. The library contains over 900 entries and represents a wide range of compound classes.

The accurate mass metabolomics PCDL has been tested using both target and nontarget screening workflows with plasma samples. Also, it is demonstrated that the accurate mass metabolomics PCDL can be used to identify unique and overlapping metabolites in a variety of tissue samples.

## Experimental

### Sample preparation

Derivatization of metabolites is a common approach to improve the volatility of polar metabolites and make them more amenable to GC/MS. Here, metabolite standards, as well as blood plasma and tissue extracts, were derivatized by methoximation followed by silylation with MSTFA + 1% TMCS as described elsewhere.[7] D27 myristic acid was added to every sample as an internal standard before the derivatization. The metabolites were extracted from plasma and tissues using acetonitrile:isopropanol:water (3:3:2). The extracts were dried, derivatized as described previously, and further used to evaluate accurate mass library screening workflows.

### Data acquisition and data processing

EI spectra for individual standards or mixtures of metabolites, known to separate well chromatographically, were acquired using the accurate mass high-resolution Agilent 7250 GC/Q-TOF system. All the data including standards and extracts were acquired in EI mode at 70 eV. Data acquisition parameters are described in Table 1. The retention indices (RIs) for each compound in the accurate mass metabolomics PCDL were calculated based on fatty acid methyl esters (FAMEs) as well as alkanes. Accurate mass EI fragments were converted to the theoretical *m/z* using Agilent MassHunter Qualitative Analysis software version 10. Then, the spectra were imported into the accurate mass metabolomics PCDL using the Agilent PCDL Manager software version 8.0. The PCDL was further amended with additional metadata and chemical structure information using Agilent ChemVista 1.0, a new library management software application. The data were processed using Agilent MassHunter Quantitative Analysis software (including the Unknowns Analysis software tool) versions 10.2 and 11.1.

**Table 1.** Data acquisition parameters.

| Parameter | Value |
|---|---|
| MS | Agilent 7250 GC/Q-TOF |
| GC | Agilent 7890B GC |
| Column | Agilent J&W DB-5ms Ultra Inert<br>30 m × 0.25 mm × 0.25 µm, DuraGuard, 10 m |
| Inlet | Split/splitless inlet, 4 mm Agilent Ultra Inert inlet liner, single taper |
| Injection Volume | 1 µL |
| Injection Mode | Splitless |
| Inlet Temperature | 280 °C |
| Oven Temperature Program | 50 °C for 0.5 min; 10 °C/min to 325 °C, 10 min hold |
| Carrier Gas | Helium |
| Column Flow | 1 mL/min constant flow |
| Transfer Line Temperature | 280 °C |
| Quadrupole Temperature | 150 °C |
| Source Temperature | 200 °C |
| Electron Energy | 70 eV |
| Emission Current | 5 µA |
| Spectral Acquisition Rate | 5 Hz |
| Mass Range | *m/z* 50 to 1,200 |

## Results and discussion

### Construction of the accurate mass metabolomics PCDL

The compounds for the accurate mass metabolomics PCDL were selected to make sure that most of the GC/MS-amenable metabolites that participate in primary metabolic pathways were included. Additionally, to expand the scope and utility of the library, over 250 secondary metabolites and xenobiotics that are commonly found in biological matrices were added as well.

To convert each fragment ion *m/z* from the acquired spectra into the theoretical *m/z*, an automated annotation of fragment formula was performed in MassHunter Qualitative Analysis software. The automated annotations were reviewed and corrected manually, if necessary, prior to exporting the spectrum into the PCDL (Figure 1). Remaining metadata such as chemical structures, structural identifiers (InChI String, InChIKey, SMILES), and various database identifiers (CAS, PubChem IDs, etc.) were added within the ChemVista Library Manager software. The accurate mass metabolomics PCDL for GC/Q-TOF includes RTs as well as both Kovats and FAMEs-based RIs. The current version of the library contains over 900 accurate mass EI spectra and over 670 unique compounds.

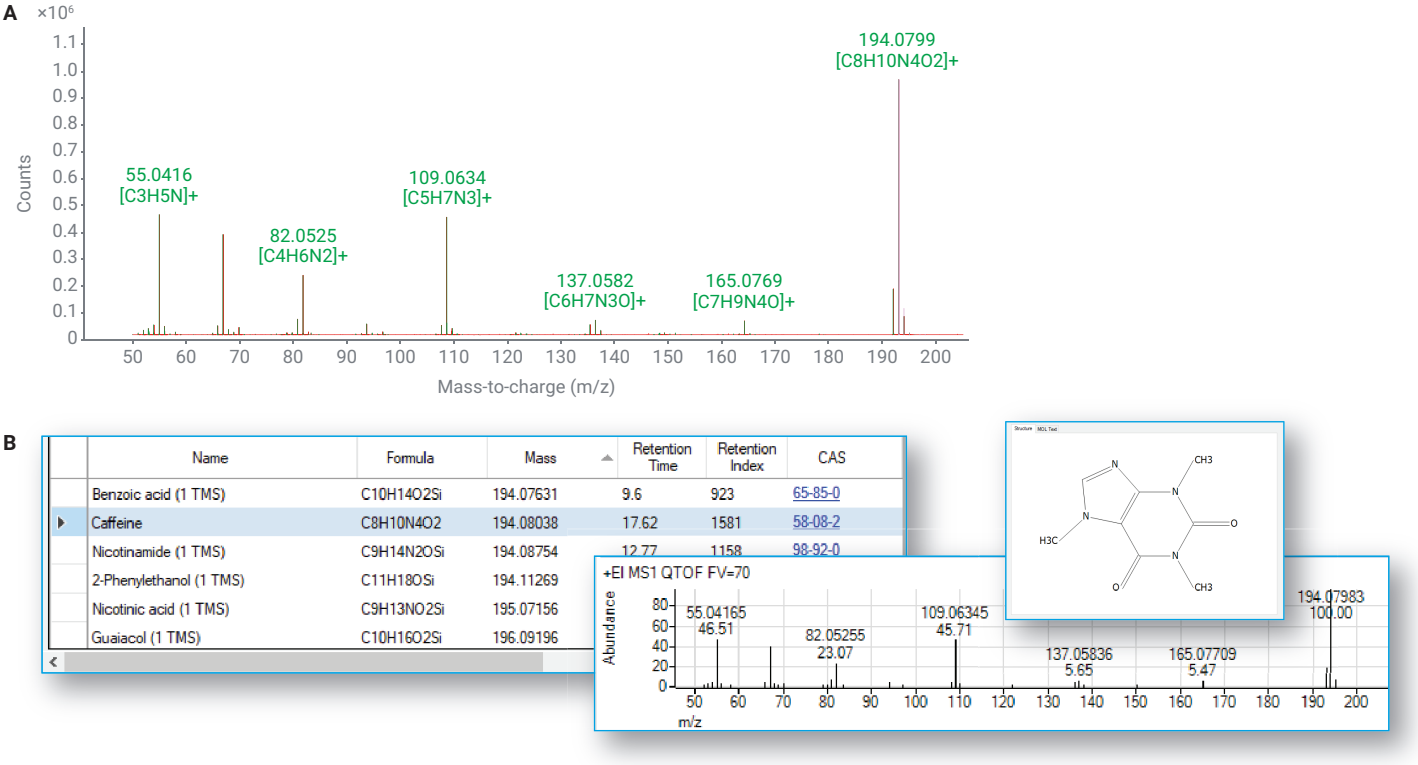


**Figure 1.** Creation of the PCDL. (A) Automated fragment formula annotation for subsequent conversion of the accurate *m/z* of the entire spectrum to the theoretical *m/z*. (B) The Agilent Fiehn accurate mass metabolomics PCDL shown in the Agilent PCDL Manager software, where metadata can be edited.

## Comparison of the accurate mass metabolomics PCDL to a unit mass GC/MS metabolomics library

The overlap between the new accurate mass metabolomics PCDL and the existing unit mass Agilent Fiehn GC/MS metabolomics library (Fiehn.L) is shown in Figure 2. Note that entries for metabolites that can form multiple derivatization states are considered as one compound entity in this comparison. While there are almost 500 entities in common between the two libraries, nearly 200 compounds are unique to the accurate mass PCDL. Therefore, both libraries are complementary regarding the compound coverage. Although the metabolomics PCDL provides the additional benefits of the accurate mass and Kovats RIs, the unit mass metabolomics library, which is based on FAMEs RIs, can also be used to search the GC/Q-TOF data.

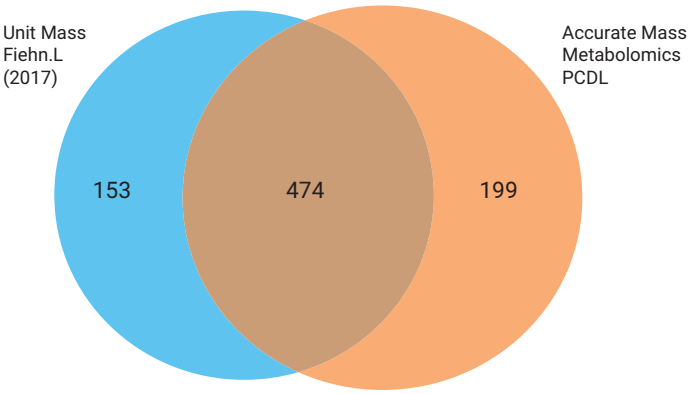


**Figure 2.** Comparison between the number of the unique compound entities in the accurate mass metabolomics PCDL and the unit mass Fiehn GC/MS metabolomics library.

Unlike the unit mass GC/MS metabolomics library, the accurate mass metabolomics PCDL contains the full acquisition spectra that covers ions of higher mass range (up to *m/z* 1,200). Since the unit mass Fiehn GC/MS metabolomics library is based on single-quadrupole data, and the ion transmission efficiency for quadrupole instruments is relatively low, the acquired data had been limited to *m/z* 600. TOF instruments have more efficient ion transmission at higher *m/z* compared to quadrupole systems, which provided an advantage when acquiring the data for the derivatized metabolites, since many of their spectra contain ions above *m/z* 600 and sometimes up to *m/z* 1,000 and higher. By detecting these additional high *m/z* masses, which tend to be unique identifiers, the accurate mass GC/Q-TOF can provide additional data points for spectral matching. Figure 3 shows examples of such compound spectra from both the accurate mass metabolomics PCDL and the unit mass metabolomics library, and highlights the specificity added by the enhanced mass range of the GC/Q-TOF. As seen from the above example, a wide data acquisition mass range for metabolomics applications and libraries increases the confidence in the compound identification.

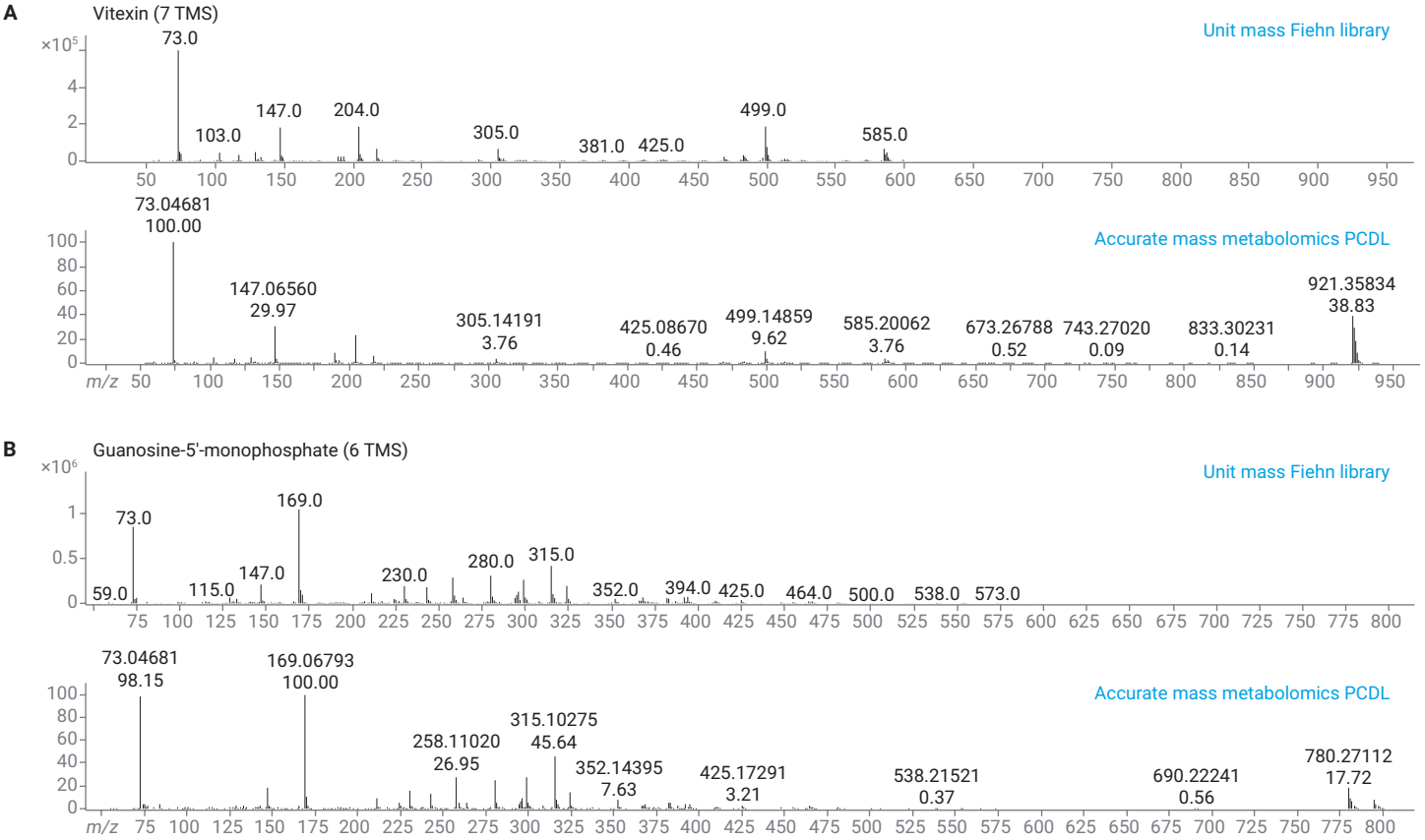


**Figure 3.** Examples of library spectra of the compounds with high molecular weight and high fragment *m/z*.

## Compound class coverage of the accurate mass metabolomics PCDL and its characterization using various tissue extracts

The current version of the accurate mass metabolomics PCDL contains a comprehensive variety of GC-amenable compound classes, the distribution of which is shown in Figure 4. The most prominent compound groups include carboxylic acids, amino acids, and carbohydrates, which are critical to understanding basic biological building blocks and energy metabolism.

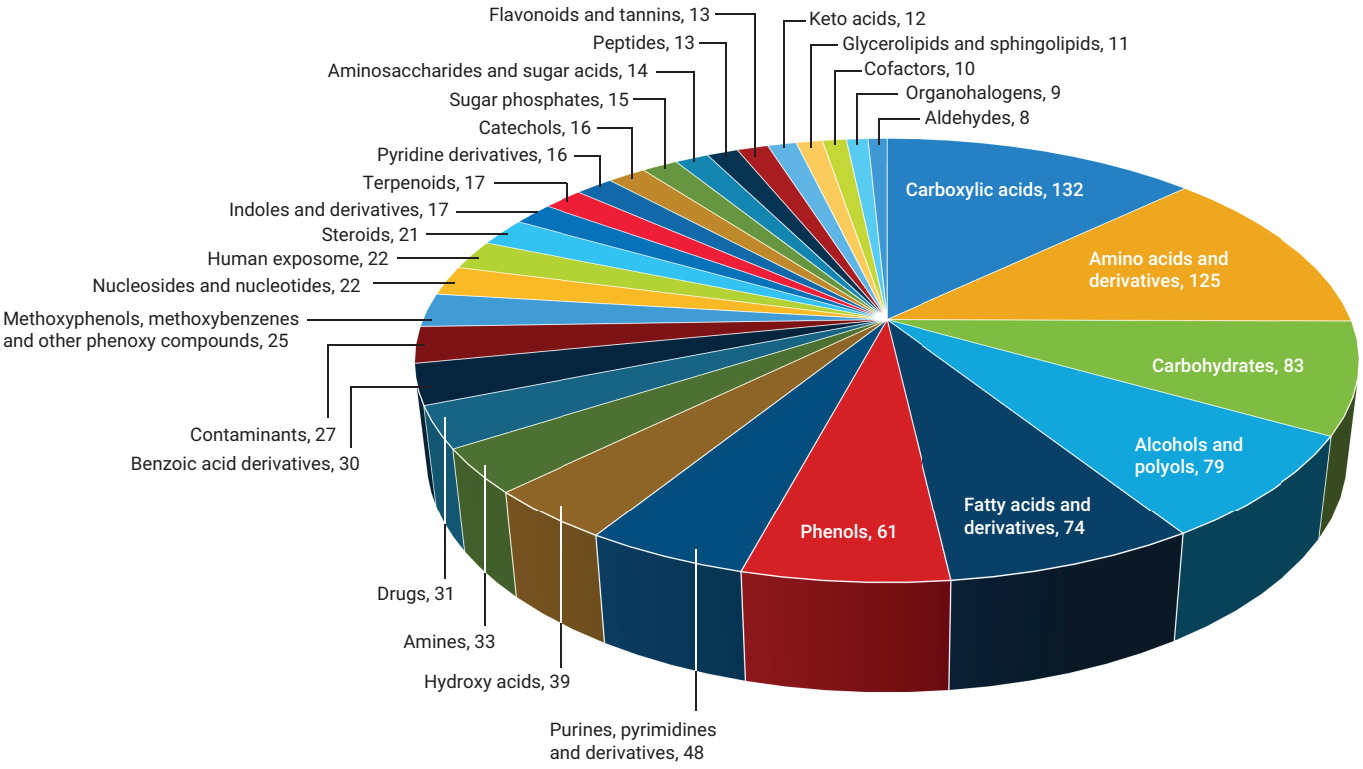


**Figure 4.** Compound classes included in the Agilent Fiehn accurate mass metabolomics PCDL.

Over 200 compounds were identified in five mouse tissues that included brain, liver, kidney, plasma, and serum using the accurate mass metabolomics PCDL and the Unknowns Analysis tool. As shown in the Venn diagram (Figure 5), 86 metabolites were found in common between all the tissues analyzed. However, at least a few compounds have been identified uniquely in each tissue (Table 2). Not surprisingly, the largest pair-wise overlap, comprising mostly organic acids, has been found between plasma and serum (Table 3). Another prominent pair-wise overlap between the different tissues was observed between liver and kidney and dominated by oligosaccharides and phospho-metabolites (Table 3).

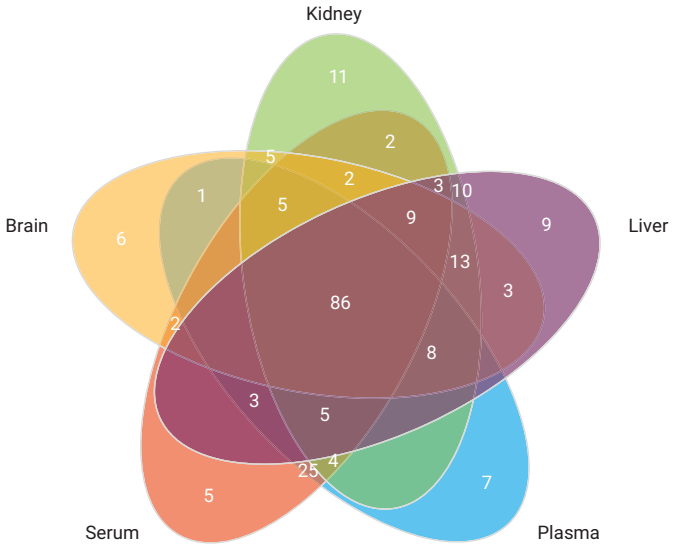


**Figure 5.** Metabolites overlap in various mouse tissues, identified by the Agilent Fiehn accurate mass metabolomics PCDL.

**Table 3.** List of library hits from the Agilent Fiehn accurate mass metabolomics PCDL that were uniquely identified in the following pairs of mouse tissues.

| Plasma-Serum | Kidney-Liver |
|---|---|
| 2-Ketoisovaleric acid | Glyceraldehyde |
| 1,3-Propanediol | Pyrophosphate |
| 2-Furoic acid | Dihydroxyacetone phosphate |
| 2-Hydroxybutyric acid | D-Glucose-6-phosphate |
| 2-Hydroxy-3-methylbutyric acid | Xanthosine |
| 2-Aminobutyric acid | Cellobiose 1 |
| 2-Keto-3-methylvaleric acid 1 | Cellobiose 2 |
| 2-Keto-3-methylvaleric acid 2 | Raffinose |
| 2-Ketoisocaproic acid | Maltotriose 1 |
| Benzoic acid | Maltotriose 2 |
| Nicotinic acid | |
| D-Threitol | |
| 4-Hydroxyphenylacetic acid | |
| Lauric acid | |
| Phthalic acid | |
| D-Sorbitol | |
| 3-Indoleacetic acid | |
| 3-Indolepropionic acid | |
| Mono-2-ethylhexyl phthalate | |
| Stearic acid | |
| Behenic acid | |
| D-Trehalose | |
| 2-Stearoylglycerol | |
| β-Tocopherol | |
| α-Tocopherol | |

**Table 2.** List of library hits from the Agilent Fiehn accurate mass metabolomics PCDL that were uniquely identified in various mouse tissues.

| Brain | Kidney | Liver | Plasma | Serum |
|---|---|---|---|---|
| 2'-Deoxycytidine | N-Acetyl-L-alanine | Cytosine | N-Methyl-L-proline | 2,3-Butanediol |
| 3-Aminopyridin-2(1H)-one | β-Cyano-L-alanine | Homogentisic acid | N-Acetylglycine | Pyruvic acid |
| Thymine | L-Methionine sulfoxide | Heptadecenoic acid | 1,2,4-Benzenetriol | Indole-3-lactate |
| N-Acetyl-L-aspartic acid | D-Galactose | 5-Methyluridine | β-Glycerophosphate | Serotonin |
| L-Ascorbic acid | Galacturonic acid | Sophorose 1 | Sorbose 1 | Dioctyl phthalate |
| 5'-Deoxy-5'-(methylthio)adenosine | Glucuronic acid | Sophorose 2 | Arachidic acid | |
| | H-Pro-Hyp-OH | Melibiose 1 | Bis(2-ethylhexyl) phthalate | |
| | L-Cystine 1 | Melibiose 2 | | |
| | L-Cystine 2 | β-Sitosterol | | |
| | Mannose-6-phosphate | | | |
| | Fructose-6-phosphate | | | |

## Evaluation of the accurate mass metabolomics PCDL in nontarget and target screening workflows using blood plasma samples

In metabolomics, both targeted and nontargeted approaches are frequently used. Therefore, the new accurate mass metabolomics PCDL was evaluated in both workflows using derivatized plasma extracts. The nontargeted workflow was performed in Unknowns Analysis software, where the built-in ExactMass tool (Figure 6) helped to efficiently eliminate false positives. To better understand potential benefits of the accurate mass metabolomics PCDL in a nontargeted approach, a comprehensive unit mass NIST17 library was tested in addition to the accurate mass metabolomics PCDL.

The target screening workflow is incorporated into the MassHunter Quantitative Analysis software and is based exclusively on accurate mass libraries; therefore, the accurate mass metabolomics PCDL was used in this case. The visualization of the target screening results is shown in Figure 7. An advantage of the target screening approach implemented in MassHunter Quantitative Analysis software is that all the method parameters can be set up for each compound individually, allowing for significant flexibility. The Screener summary view allows users to quickly review the results for further method optimization to minimize false positives and false negatives.

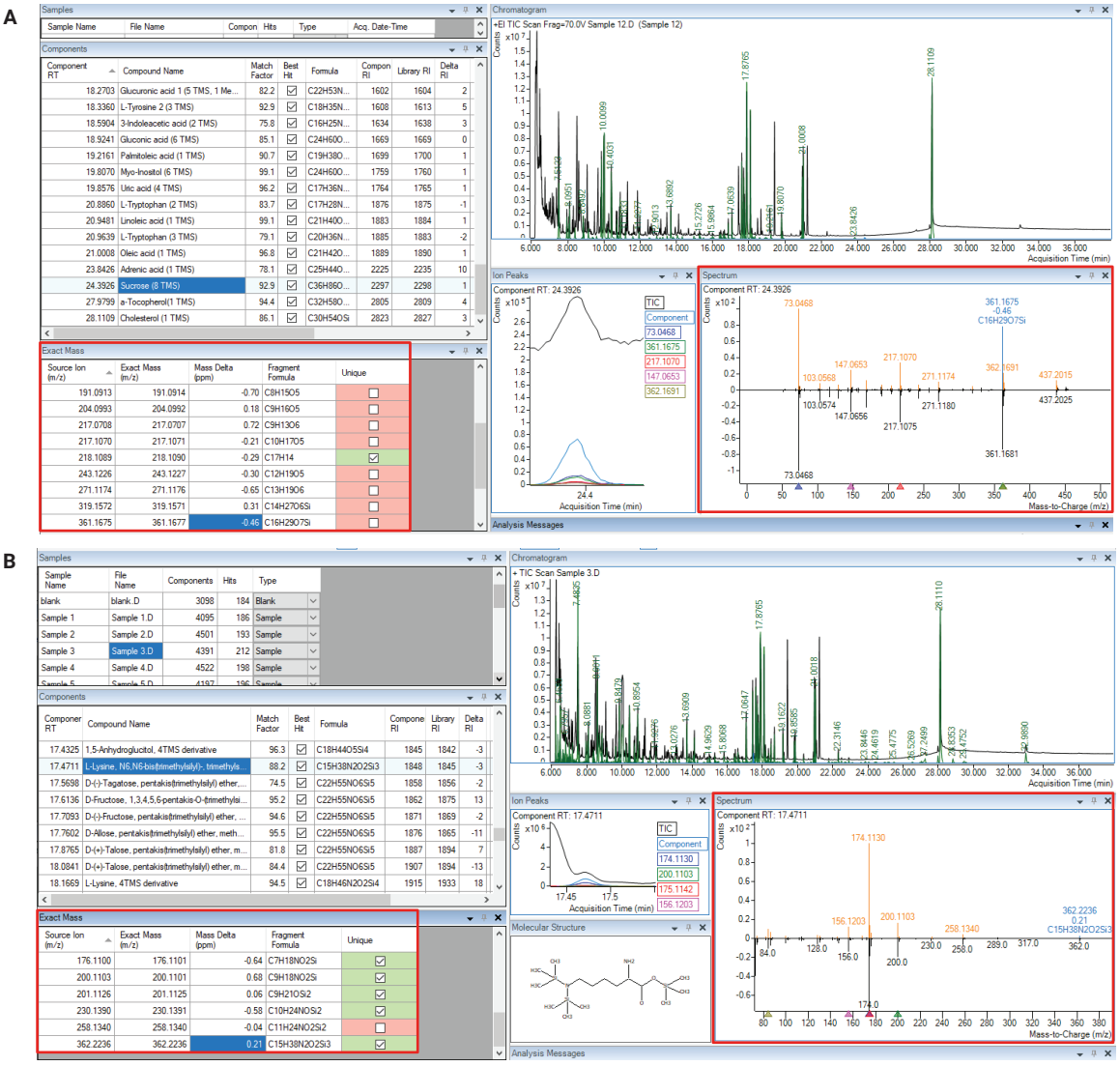


**Figure 6.** Nontargeted analysis performed in the Agilent Unknowns Analysis software using (A) the Agilent Fiehn accurate mass metabolomics PCDL and (B) the unit mass NIST17 library. ExactMass results are displayed in the outlined tables and the mirror plots. Fragment ions of compound spectra consistent with the library hit formula are highlighted in orange and blue.

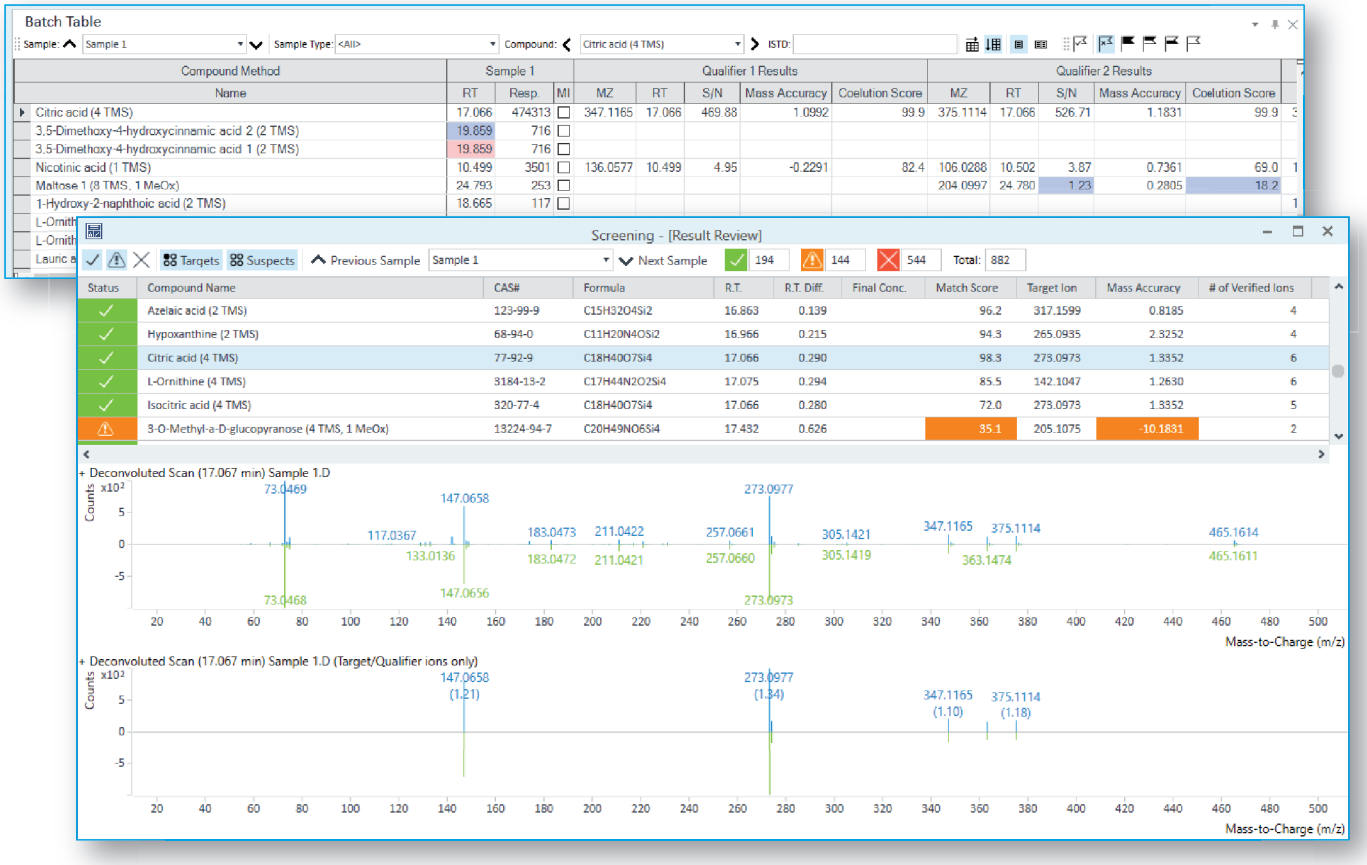


**Figure 7.** Target screening results visualization. The detailed Quantitative Analysis table containing complete information for each compound is shown at the top. The Screener summary window that allows to quickly review the most meaningful results is displayed at the bottom.

The results obtained using the NIST library were filtered out based on the accurate mass information using the ExactMass tool to eliminate false positives. A few examples are illustrated in Figure 8. Around 36 to 48 library hits per sample were found to be false positives based on the accurate mass information. This represented approximately 30% of total hits.

Using the accurate mass metabolomics PCDL (in both target and nontarget approaches) allowed the automatic exclusion of false positives due to the accurate mass discrepancies, since its detection is built into the Screener and the library search tools.

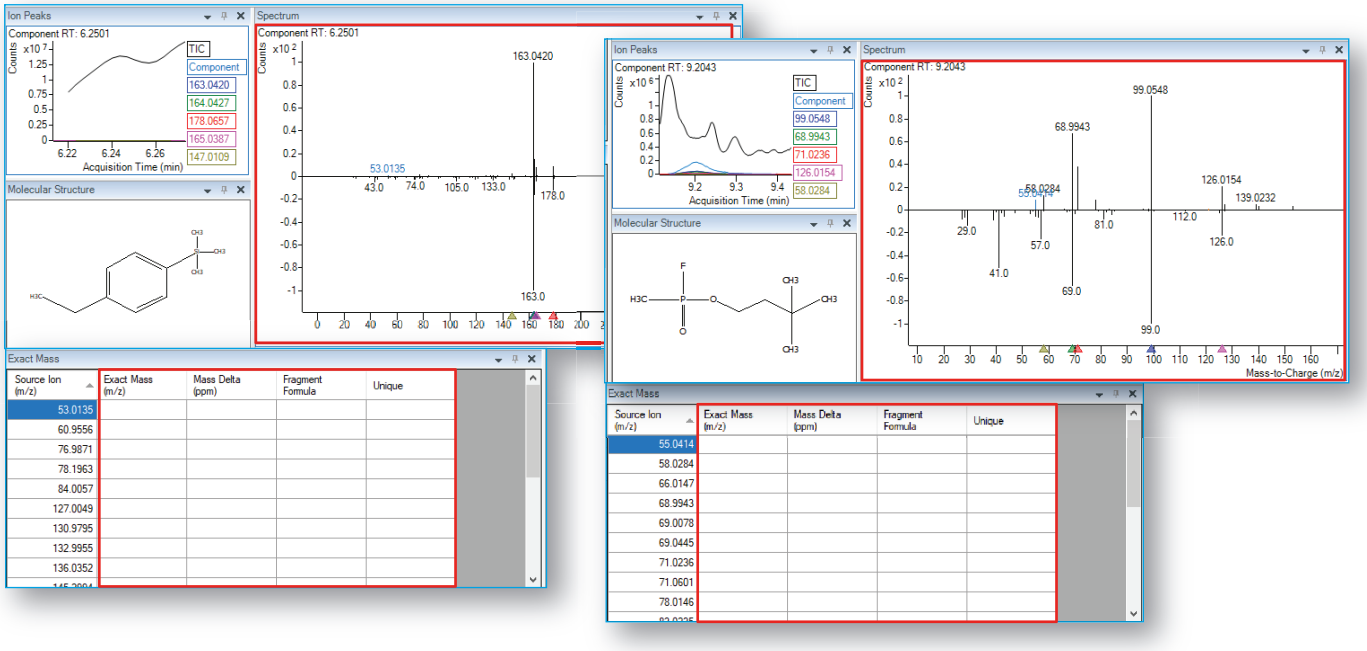


**Figure 8.** Examples of the false positives that can be easily recognized with a help of the ExactMass tool. Notice the empty ExactMass tables and the absence of highlighted ions on the mirror plots. The library search was performed using the NIST17 library.

The number of compounds identified in plasma samples in the target screening approach with the accurate mass metabolomics PCDL and nontarget analysis using both NIST17 and the accurate mass metabolomics PCDL were compared (Table 4). In most cases, a higher number of compounds was identified when using the accurate mass metabolomics PCDL compared to NIST. However, interestingly, target and nontarget approaches were complementary with regard to the specific compounds identified by each method.

A range of polar and nonpolar metabolites involved in primary and lipid metabolism were confidently identified in plasma using the metabolomics PCDL and are listed in Table 5. In the interest of space, only library hits with a high library match score of >80 in both workflows are displayed.

**Table 4.** Number of identified compounds when using different screening approaches with unit and accurate mass libraries (NIST17 and the Agilent Fiehn accurate mass metabolomics PCDL, respectively).

| | Number of Identified Compounds | | |
|---|---|---|---|
| **Sample Name** | **Screener** | **UA PCDL** | **UA NIST** |
| Sample A | 153 | 151 | 140 |
| Sample B | 149 | 146 | 144 |
| Sample C | 151 | 146 | 138 |
| Sample D | 152 | 146 | 143 |
| Sample E | 145 | 138 | 145 |
| Sample F | 141 | 141 | 139 |
| Sample G | 139 | 148 | 146 |
| Sample H | 164 | 144 | 147 |
| Sample I | 146 | 139 | 131 |
| Sample J | 152 | 148 | 136 |
| Sample K | 151 | 146 | 132 |
| Sample M | 168 | 151 | 140 |
| Sample N | 167 | 152 | 157 |

**Table 5.** Compounds identified in plasma using targeted (Screener) and nontargeted (UA) approaches with the Agilent Fiehn accurate mass metabolomics PCDL. The table only shows compounds identified with the library match score of > 80 in both approaches.

| RT | Compound Name | Derivatized Formula | UA Library Match Score | Screener Library Score |
|---|---|---|---|---|
| 6.48 | Boric acid | $C_9H_{27}BO_3Si_3$ | 93.2 | 82.1 |
| 7.33 | L-Lactic acid | $C_9H_{22}O_3Si_2$ | 95.4 | 83.8 |
| 8.39 | 2-Hydroxybutanoic acid | $C_{10}H_{24}O_3Si_2$ | 96.7 | 99.6 |
| 8.78 | p-Cresol | $C_{10}H_{16}OSi$ | 85.3 | 99.6 |
| 8.85 | L-Leucine | $C_9H_{21}NO_2Si$ | 81.3 | 98.5 |
| 8.85 | 3-Hydroxybutyric acid | $C_{10}H_{24}O_3Si_2$ | 95.6 | 81.9 |
| 8.90 | 2-Hydroxy-3-methylbutyric acid | $C_{11}H_{26}O_3Si_2$ | 92.2 | 92.6 |
| 9.58 | 2-Ketoisocaproic acid 2 | $C_{10}H_{21}NO_3Si$ | 94.2 | 97.4 |
| 9.85 | Ethanolamine | $C_{11}H_{31}NOSi_3$ | 97.6 | 98.9 |
| 10.14 | Benzoic acid | $C_{10}H_{14}O_2Si$ | 96.8 | 99.6 |
| 10.21 | L-Serine | $C_9H_{23}NO_3Si_2$ | 98.0 | 98.8 |
| 10.40 | Phosphoric acid | $C_9H_{27}O_4PSi_3$ | 89.3 | 98 |
| 10.78 | L-Proline | $C_{11}H_{25}NO_2Si_2$ | 93.8 | 99.7 |
| 10.89 | Glycine | $C_{11}H_{29}NO_2Si_3$ | 97.8 | 99.6 |
| 11.59 | L-Serine | $C_{12}H_{31}NO_3Si_3$ | 96.0 | 99.2 |
| 11.64 | Nonanoic acid | $C_{12}H_{26}O_2Si$ | 98.2 | 99.5 |
| 12.50 | β-Alanine | $C_{12}H_{31}NO_2Si_3$ | 94.3 | 99.3 |
| 12.90 | Capric acid | $C_{13}H_{28}O_2Si$ | 93.9 | 99.4 |
| 13.55 | O-Acetylsalicylic acid | $C_{13}H_{22}O_3Si_2$ | 87.9 | 99.5 |
| 13.69 | L-Glutamic acid (dehydrated) | $C_{11}H_{23}NO_3Si_2$ | 99.1 | 98.9 |
| 13.81 | L-Glutamate | $C_{11}H_{25}NO_4Si_2$ | 87.3 | 99.1 |
| 14.05 | L-Phenylalanine | $C_{12}H_{19}NO_2Si$ | 95.6 | 98.3 |
| 14.10 | Creatinine | $C_{13}H_{31}N_3OSi_3$ | 86.3 | 99.3 |
| 14.32 | α-Hydroxyglutaric acid | $C_{14}H_{32}O_5Si_3$ | 94.9 | 80 |
| 14.64 | Pyrogallol | $C_{15}H_{30}O_3Si_3$ | 83.3 | 82.7 |
| 14.86 | L-Glutamate | $C_{14}H_{33}NO_4Si_3$ | 90.2 | 99.7 |
| 14.96 | L-Phenylalanine | $C_{15}H_{27}NO_2Si_2$ | 98.1 | 99.1 |
| 14.98 | 4-Hydroxybenzoic acid | $C_{13}H_{22}O_3Si_2$ | 82.9 | 95.2 |
| 15.27 | Lauric acid | $C_{15}H_{32}O_2Si$ | 96.3 | 99.5 |
| 15.49 | Arabinose | $C_{18}H_{45}NO_5Si_4$ | 96.8 | 85 |
| 15.90 | 3-Indoxylsulfate | $C_{14}H_{23}NO_4SSi_2$ | 91.5 | 86 |
| 16.13 | 3-(3-Hydroxyphenyl) propionic acid | $C_{15}H_{26}O_3Si_2$ | 80.9 | 99.6 |
| 16.38 | L-Ornithine | $C_{14}H_{36}N_2O_2Si_3$ | 96.0 | 98.1 |
| 16.48 | Ribonic acid | $C_{20}H_{50}O_6Si_5$ | 93.5 | 93.1 |
| 16.51 | 4-Hydroxy-3-methoxybenzoic acid | $C_{14}H_{24}O_4Si_2$ | 92.7 | 99.8 |
| 16.86 | Azelaic acid | $C_{15}H_{32}O_4Si_2$ | 97.8 | 94.5 |

| RT | Compound Name | Derivatized Formula | UA Library Match Score | Screener Library Score |
|---|---|---|---|---|
| 16.96 | Hypoxanthine | $C_{11}H_{20}N_4OSi_2$ | 91.1 | 93.8 |
| 17.06 | Citric acid | $C_{18}H_{40}O_7Si_4$ | 99.3 | 98.2 |
| 17.07 | L-Ornithine | $C_{17}H_{44}N_2O_2Si_4$ | 95.7 | 87.1 |
| 17.51 | Caffeine | $C_8H_{10}N_4O_2$ | 88.2 | 99.6 |
| 17.57 | D-Tagatose 2 | $C_{22}H_{55}NO_6Si_5$ | 94.4 | 90.4 |
| 17.76 | Mannose 1 | $C_{22}H_{55}NO_6Si_5$ | 98.8 | 99.8 |
| 17.83 | L-Tyrosine 1 | $C_{15}H_{27}NO_3Si_2$ | 84.7 | 84.8 |
| 17.87 | D-Glucose 1 | $C_{22}H_{55}NO_6Si_5$ | 83.6 | 99.8 |
| 18.23 | D-Mannitol | $C_{24}H_{62}O_6Si_6$ | 98.6 | 80.8 |
| 18.26 | Glucuronic acid 1 | $C_{21}H_{50}O_7Si_5$ | 92.7 | 82.1 |
| 18.34 | L-Tyrosine 2 | $C_{18}H_{35}NO_3Si_3$ | 95.0 | 99.5 |
| 18.48 | Sinigrin | $C_{22}H_{44}NO_9S_2Si_4$ | 87.2 | 85.1 |
| 18.48 | Myo-Inositol | $C_{24}H_{60}O_6Si_6$ | 88.0 | 82.5 |
| 18.59 | Indole-3-acetic acid | $C_{16}H_{25}NO_2Si_2$ | 88.5 | 82.9 |
| 18.93 | Gluconic acid | $C_{24}H_{60}O_7Si_6$ | 90.9 | 97.6 |
| 19.21 | Palmitoleic acid | $C_{19}H_{38}O_2Si$ | 96.0 | 99.7 |
| 19.41 | Palmitic acid | $C_{19}H_{40}O_2Si$ | 97.2 | 99.9 |
| 19.41 | Cyclohexanecarboxylic acid | $C_{10}H_{20}O_2Si$ | 98.1 | 85.8 |
| 19.76 | Indole-3-propionic acid | $C_{17}H_{27}NO_2Si_2$ | 83.4 | 84.6 |
| 19.80 | Conduritol-β-expoxide 2 | $C_{18}H_{38}O_5Si_4$ | 93.7 | 86.8 |
| 19.85 | Uric acid | $C_{17}H_{36}N_4O_3Si_4$ | 96.4 | 99.7 |
| 20.58 | Indole-3-lactate | $C_{20}H_{35}NO_3Si_3$ | 89.0 | 93.5 |
| 20.89 | L-Tryptophan | $C_{19}H_{30}N_2O_3Si_2$ | 80.3 | 83.9 |
| 20.96 | N-Acetyl-L-tryptophan | $C_{20}H_{36}N_2O_2Si_3$ | 83.4 | 81.9 |
| 21.00 | Oleic acid | $C_{21}H_{42}O_2Si$ | 97.9 | 99.9 |
| 21.22 | Stearic acid | $C_{21}H_{44}O_2Si$ | 97.2 | 99.2 |
| 22.31 | Adrenic acid | $C_{25}H_{44}O_2Si$ | 85.3 | 92 |
| 24.03 | 1-Monopalmitin | $C_{25}H_{54}O_4Si_2$ | 83.7 | 94.5 |
| 24.39 | Sucrose | $C_{36}H_{86}O_{11}Si_8$ | 94.5 | 96 |
| 24.46 | Behenic acid | $C_{25}H_{52}O_2Si$ | 82.4 | 95.7 |
| 25.15 | Maltose 1 | $C_{37}H_{89}NO_{11}Si_8$ | 91.4 | 98.7 |
| 25.31 | 1-Monoolein | $C_{27}H_{56}O_4Si_2$ | 87.3 | 96 |
| 25.47 | Glyceryl monostearate | $C_{27}H_{58}O_4Si_2$ | 84.9 | 97.4 |
| 25.79 | Squalene | $C_{30}H_{50}$ | 90.1 | 98.9 |
| 27.03 | β-Tocopherol | $C_{31}H_{56}O_2Si$ | 90.6 | 99.7 |
| 27.98 | α-Tocopherol | $C_{32}H_{58}O_2Si$ | 95.2 | 94.7 |
| 28.09 | Cholesterol | $C_{30}H_{54}OSi$ | 88.0 | 99.8 |
| 29.48 | β-Sitosterol | $C_{32}H_{58}OSi$ | 92.2 | 98.6 |

## Conclusion

The creation of the Agilent Fiehn accurate mass metabolomics PCDL for use with GC/Q-TOF enables researchers to make even more confident spectral identification of biologically relevant molecules from metabolomics data. This application note describes the creation and characterization of the comprehensive accurate mass metabolomics PCDL, which covers a range of polar and nonpolar metabolites involved in metabolic pathways, as well as a number of xenobiotics that are typically present in biological matrices. Agilent ChemVista was used during the PCDL creation process, and the PCDL can be readily used with ChemVista for further management and spectral consolidation. The use of the accurate mass metabolomics PCDL was demonstrated using commonly analyzed plasma matrix in both targeted and nontargeted workflows. As compared to a unit mass library, the accurate mass metabolomics PCDL ensured more specific and sensitive detection of metabolites in matrices by incorporating accurate mass fragment spectra and an expanded mass range, as well as enabling the accurate mass-based suspect screening workflow. By incorporating the accurate mass metabolomics PCDL into their workflows, researchers will now be able to take advantage of increased confidence in metabolite identification and further accelerate their ability to make meaningful biological conclusions from their experiments.

## References


1. Schauer, N. *et al.* GC–MS Libraries for the Rapid Identification of Metabolites in Complex Biological Samples. *FEBS Letters*. **2005**, *579(6)*, 1332–1337. DOI: 10.1016/j.febslet.2005.01.029
2. Stein, S. Mass Spectral Reference Libraries: an Ever-Expanding Resource for Chemical Identification. *Anal. Chem.* **2012**, *84(17)*, 7274–7282. DOI: 10.1021/ac301205z
3. Beale, D. J. *et al.* Review of Recent Developments in GC-MS Approaches to Metabolomics-Based Research. *Metabolomics*. **2018**, *14(11)*, 152. DOI: 10.1007/s11306-018-1449-2
4. Grapp, M.; Maurer, H. H.; Desel, H. Systematic Forensic Toxicological Analysis by GC-MS in Serum Using Automated Mass Spectral Deconvolution and Identification System. *Drug Test Anal.* **2016**, *8(8)*, 816–25. DOI: 10.1002/dta.1848
5. Mihaleva, V. V.; *et al.* Automated Procedure for Candidate Compound Selection in GC-MS Metabolomics Based on Prediction of Kovats Retention Index. *Bioinformatics*. **2009**, *25(6)*, 787–794. DOI: 10.1093/bioinformatics/btp056
6. Nieto, S. *et al.* Contaminants Screening Using High-Resolution GC/Q-TOF and an Expanded Accurate Mass Library of Pesticides and Environmental Pollutants. *Agilent Technologies application note*, publication number 5994-1346EN
7. Fiehn, O. Metabolomics by Gas Chromatography-Mass Spectrometry: the Combination of Targeted and Untargeted Profiling. *Curr. Protoc. Mol. Biol.* **2016**, 114: 30.4.1–30.4.32. DOI: 10.1002/0471142727.mb3004s114


**www.agilent.com**




RA44995.6218171296





Printed in the USA, March 29, 2023
5994-5832EN


Agilent
Trusted Answers