Application Note

Clinical Research

Agilent
Trusted Answers

Differences in Metabolic Profiles of Individuals with Heart Failure Using High-Resolution GC/Q-TOF

Authors

Luis Valdiviez and Oliver Fiehn
West Coast Metabolomics Center, University of California, Davis, CA

Sofia Nieto and
Andrew McEachran
Agilent Technologies, Inc.

Abstract

It is critical to study the metabolic foundations of cardiovascular and other diseases to obtain a better understanding of disease biology and the impacts of potential intervention regimens. Key to achieving this is the development of workflows that can provide deep molecular insights. This application note describes a GC/Q-TOF untargeted metabolomics workflow that uses the Agilent accurate mass metabolomics personal compound database and library (PCDL) as well as unit mass libraries for compound identification. It also employs Agilent ChemVista, a new library management software tool, to integrate third-party libraries into the screening workflows. In this study, a metabolic profiling of individuals with heart failure (HF) has been performed to identify underlying mechanisms of this pathology that could be helpful in designing an effective treatment. Taken together, the workflow shown here provides an example to researchers seeking to understand the molecular mechanisms of disease.

Introduction

Heart failure (HF) is a clinical condition that significantly affects quality and duration of life and is a major global public health problem affecting 23 million people worldwide.[1] Approximately half of HF cases present with reduced left ventricular ejection fraction (HFrEF) while another half are characterized by preserved ejection fraction (HFpEF).[2] Both pathologies have similar morbidity and mortality, however, no effective treatment exists for HFpEF.[2] Further understanding of the metabolic profile of HF individuals relative to healthy subjects could shed light on potential therapies. Thus, to facilitate the development of successful therapies for HF in the future, a blood plasma metabolic profiling study of healthy subjects and individuals with both HfrEF and HfpEF was performed.

One of the unique advantages of using GC/MS for metabolic profiling is an opportunity for fast and efficient compound identification due to the availability of GC/MS libraries with a vast number of highly reproducible electron ionization (EI) spectra.[3–6] To help further speed up compound ID and enhance its reliability, the following EI spectral library characteristics are typically desired: a large number of spectra to help identify high numbers of compounds; the presence of retention indices (RIs) to increase versatility; and, in the case of a high-resolution instrument, accurate mass spectra would be desirable to efficiently decrease the false positives rate.

Therefore, a variety of GC/MS EI libraries were used in this study to maximize compound ID potential. This included the accurate mass metabolomics PCDL, NIST23, and spectra from MassBank of North America (MassBank.us)[7] exported into PCDL using the ChemVista software.

The accurate mass metabolomics PCDL used in this application note contains over 900 curated spectra of derivatized compounds including metabolites and xenobiotics commonly found in biological matrices.[8] It also includes RTs for the standard Fiehn metabolomics method and RIs based on both fatty acid methyl esters (FAMEs) and alkanes.[8] Integration of the accurate mass metabolomics PCDL into this workflow provided highly accurate compound annotations through multiple points of confirmation.

In this study, the combination of the accurate and unit mass libraries for metabolomics, MassBank.us, and NIST23 in conjunction with SureMass accurate mass deconvolution proved critical for metabolite identification. A substantial number of metabolites were identified across sample groups, enabling observation of statistical differences in HF and healthy individuals. Together, this workflow and its results can aid understanding of the metabolic signature of HF.

Experimental

Sample preparation

Blood plasma was collected from subjects of HFrEF and HFpEF groups as well as healthy individuals (10 samples per each group). Thirty microliters of each blood plasma sample were extracted using 1 mL of acetonitrile:isopropanol:water (3:3:2), and 450 µL of extract were dried, derivatized by O-methoximation, followed by trimethylsilylation with MSTFA + 1% TMCS as described elsewhere.[9]

Data acquisition and data processing

The data were acquired using the accurate-mass, high-resolution Agilent 7250 GC/Q-TOF system. The GC method was retention-time locked to d_{27} myristic acid as described elsewhere.[10] The data were acquired at 70 eV, and in low energy EI and chemical ionization (CI) modes. MS/MS was further used for structure elucidation of the unknown compounds of interest. Further details of the data acquisition method are shown in Table 1.

Table 1. Data acquisition parameters.

Ionization Mode	EI	CI
MS	Agilent 7250 GC/Q-TOF	
GC	Agilent 8890 GC	
Column	Agilent DB-5ms Ultra Inert, 30 m × 0.25 mm, 0.25 µm with DuraGuard, 10 m	
Inlet	Split/splitless inlet, 4 mm Agilent Ultra Inert inlet liner, single taper	
Injection Volume	0.2 µL	0.5 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	290 °C	
Quadrupole Temperature	150 °C	
Source Temperature	200 °C	280 °C
Electron Energy	70 eV (standard EI), and 15, 12, and 10 eV (low-energy EI)	60 eV
Emission Current	5 µA (standard EI), 0.4 to 0.8 µA (low-energy EI)	30 µA
Spectral Acquisition Rate	5 Hz	
Mass Range	*m/z* 50 to 1,200	

Agilent SureMass deconvolution and library search were performed in Agilent MassHunter Unknowns Analysis 11.1. ChemVista 1.0 was used to export the GC/MS spectra from third-party databases, in this case, MassBank of North America (massbank.us). The accurate mass metabolomics PCDL, unit mass Fiehn metabolomics library, NIST23, as

well as MassBank.us were used to perform initial compound identification. Alkanes and FAMEs retention indexes of all the libraries were used to confirm the compound ID. Statistical analysis was performed in Agilent Mass Profiler Professional (MPP) 15.1. Structural elucidation was performed using the Agilent MassHunter Molecular Structure Correlator (MSC) 8.2 software to map fragments to their proposed substructures.

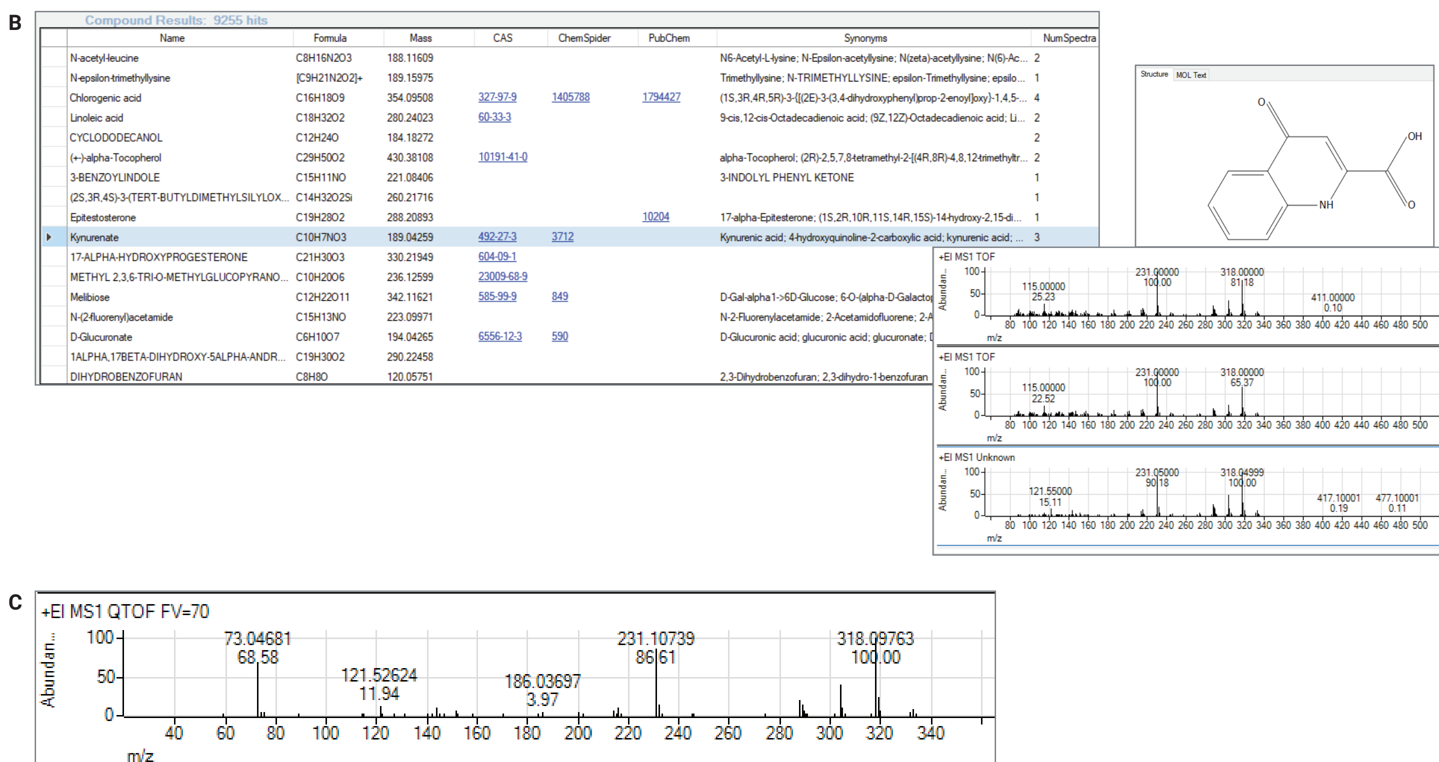
Results and discussion

Including third-party libraries in the metabolomics workflow using ChemVista

ChemVista is a standalone software application that can complement compound identification results produced in metabolomics studies by including third-party libraries in the workflow. ChemVista can manage spectra, RT/RI information, as well as other compound metadata and supports multiple data formats including SDF and PCDL.[11] It can also be used in a process of creating a PCDL to add compound structures and identifiers.

To use third-party library spectra in the current study, all GC/MS spectral data (containing over 18,000 spectra) were downloaded from the MassBank.us webpage in SDF format and imported into ChemVista (Figure 1A). The imported content was then filtered to contain only EI spectra and exported in the PCDL format for further use in the downstream workflow. The newly created MassBank.us PCDL included over 9,000 EI GC/MS compounds with spectra (Figure 1B). Note that one compound entity in the ChemVista output can contain multiple spectra, including those obtained using different GC/MS analyzers, which provides flexibility with respect to a choice of the instrument used to acquire the data. Figure 1C shows a spectrum from the GC/Q-TOF accurate mass metabolomics PCDL for the same compound shown in 1B for comparison.

A



Compound Results: 9255 hits

Name	Formula	Mass	CAS	ChemSpider	PubChem	Synonyms	NumSpectra
N-acetyl-leucine	C8H16N2O3	188.11609				N6-Acetyl-L-lysine; N-Epsilon-acetyllysine; N(zeta)-acetyllysine; N(6)-Ac...	2
N-epsilon-trimethyllysine	[C9H21N2O2]+	189.15975				Trimethyllysine; N-TRIMETHYLLYSINE; epsilon-Trimethyllysine; epsilo...	1
Chlorogenic acid	C16H18O9	354.09508	327-97-9	1405788	1794427	(1S,3R,4R,5R)-3-{[(2E)-3-(3,4-dihydroxyphenyl)prop-2-enoyl]oxy}-1,4,5-...	4
Linoleic acid	C18H32O2	280.24023	60-33-3			9-cis,12-cis-Octadecadienoic acid; (9Z,12Z)-Octadecadienoic acid; Li...	2
CYCLODODECANOL	C12H24O	184.18272					2
(+-)-alpha-Tocopherol	C29H50O2	430.38108	10191-41-0			alpha-Tocopherol; (2R)-2,5,7,8-tetramethyl-2-[(4R,8R)-4,8,12-trimethyltr...	2
3-BENZOYLINDOLE	C15H11NO	221.08406				3-INDOLYL PHENYL KETONE	1
(2S,3R,4S)-3-(TERT-BUTYLDIMETHYLSILYLOX...	C14H32O2Si	260.21716					1
Epitestosterone	C19H28O2	288.20893			10204	17-alpha-Epitesterone; (1S,2R,10R,11S,14R,15S)-14-hydroxy-2,15-di...	1
Kynurenate	C10H7NO3	189.04259	492-27-3	3712		Kynurenic acid; 4-hydroxyquinoline-2-carboxylic acid; kynurenic acid; ...	3
17-ALPHA-HYDROXYPROGESTERONE	C21H30O3	330.21949	604-09-1				
METHYL 2,3,6-TRI-O-METHYLGLUCOPYRANO...	C10H20O6	236.12599	23009-68-9				
Melibiose	C12H22O11	342.11621	585-99-9	849		D-Gal-alpha1->6D-Glucose; 6-O-(alpha-D-Galacto[illegible]	
N-(2-fluorenyl)acetamide	C15H13NO	223.09971				N-2-Fluorenylacetamide; 2-Acetamidofluorene; 2-A[illegible]	
D-Glucuronate	C6H10O7	194.04265	6556-12-3	590		D-Glucuronic acid; glucuronic acid; glucuronate; [illegible]	
1ALPHA,17BETA-DIHYDROXY-5ALPHA-ANDR...	C19H30O2	290.22458					
DIHYDROBENZOFURAN	C8H8O	120.05751				2,3-Dihydrobenzofuran; 2,3-dihydro-1-benzofuran	

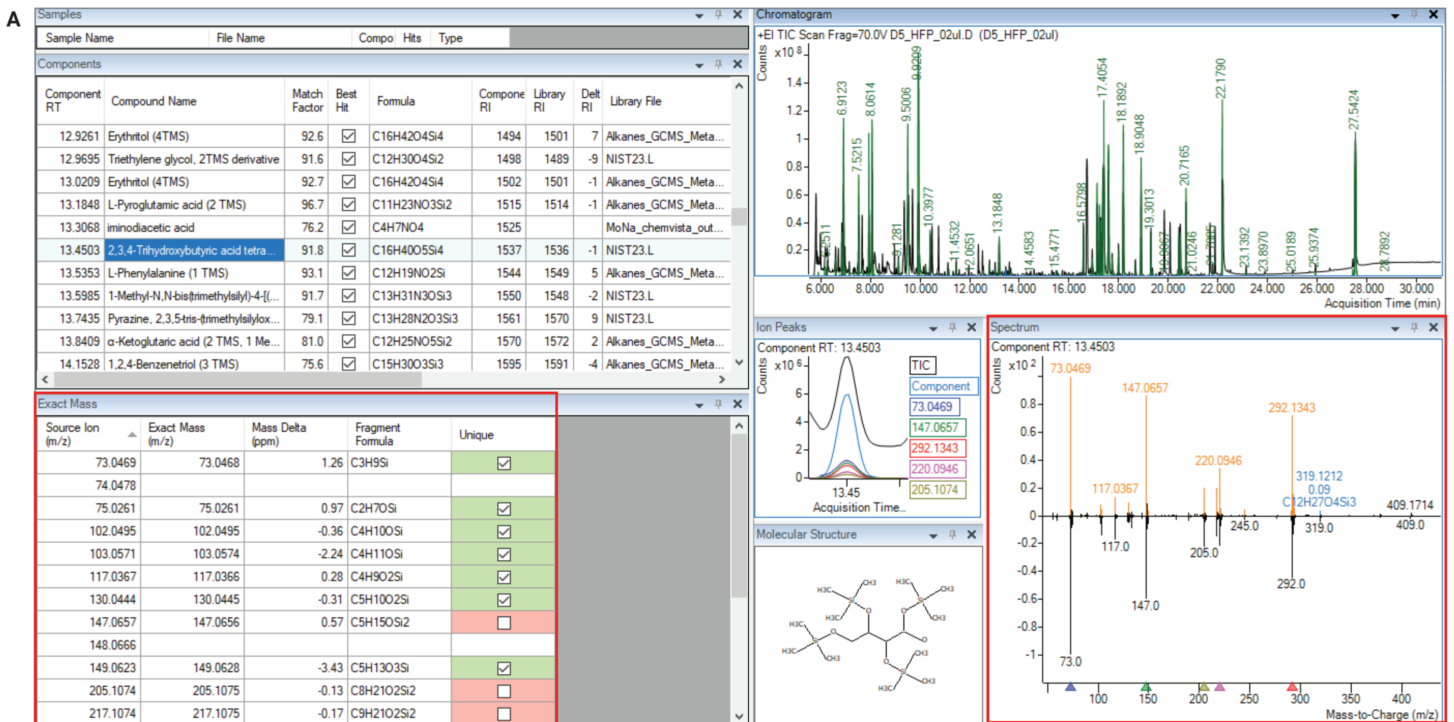
Figure 1. Adding third-party libraries into the metabolomics workflow. (A) Extracting GC/MS EI spectra and metadata from MassBank.us. (B) Agilent ChemVista output for MassBank.us containing EI spectra. (C) GC/Q-TOF spectrum for the same compound (kynurenic acid 2TMS) from the Agilent accurate mass metabolomics PCDL.

Metabolic profiling using a combination of accurate mass and unit mass libraries

To identify metabolites involved in HF pathophysiology, an untargeted metabolic plasma profiling of HF subjects as well as healthy individuals using the 7250 GC/Q-TOF was performed. The chromatographic deconvolution was performed using a SureMass algorithm that is based on profile data and is specifically optimized for complex, high-resolution, accurate-mass EI data to provide high sensitivity and increase dynamic range and mass accuracy of deconvoluted spectra.

All the libraries were used simultaneously in the same Unknowns Analysis method, where the library search parameters can be selected individually for each library to ensure optimal results.

When searching unit mass libraries such as NIST, the ExactMass tool of Unknowns Analysis can help to significantly reduce the number of false positives based on the accurate mass of the acquired data and molecular formula of the library hit. This tool assigns ions of the deconvoluted spectrum with formulas that are the subset of the hit's molecular formula and calculates mass errors for the fragment ions. Two examples are shown in Figure 2, where one of the NIST hits, 2,3,4-trihydroxybutyric acid 4TMS with a library match score of 91.8, exhibited low mass error on all the fragments ions assigned with a subset of the molecular formula (Figure 2A), thus confirming the ID using the accurate-mass information. However, another hit, methacrylamide, with a library match score of 89.1, did not have any ions matching the molecular formula within 10 ppm mass error (Figure 2B), indicating that the library hit was a false positive, thus highlighting the advantage of accurate-mass GC/MS to curate spectral hits. Note that GC/Q-TOF spectra provide excellent spectral library matching with NIST and other libraries that are based on GC single quadrupole spectra, since they do not exhibit the excessive fragmentation characteristics of other accurate mass analyzers.



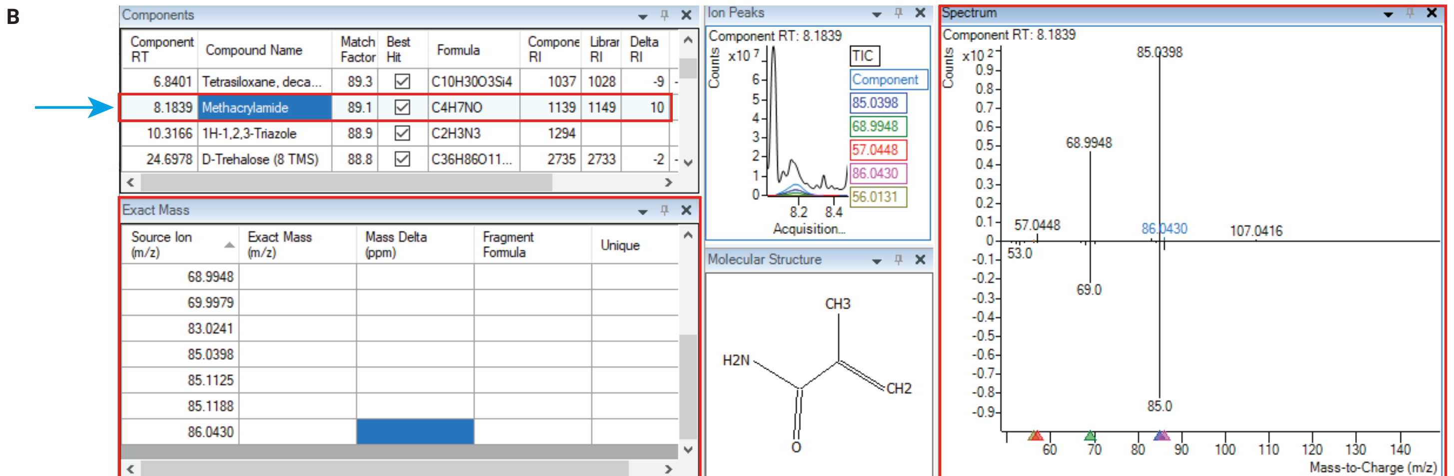


Figure 2. Using the ExactMass feature of Agilent MassHunter Unknowns Analysis for confirmation of the unit mass library hits with accurate mass. (A) Accurate mass spectrum of the deconvoluted compound matches the NIST library hit formula. Mass error for each ion that matches the molecular formula subset within 10 ppm is shown in the ExactMass table, and compound ions in the mirror plot are highlighted. (B) The arrow is pointing to a false positive that can easily be recognized based on the empty ExactMass table.

Both NIST23 and the metabolomics PCDL each produced over 100 hits per sample after blank subtraction and eliminating false positives with either the ExactMass (for NIST) or accurate mass tolerance function that is built into Unknowns Analysis (for the accurate mass PCDL). When searching the data with MassBank.us, the number of hits per sample on average was slightly below 100 (Table 2). The library match score cut-off of 70 was used in the case of both NIST and the metabolomics PCDL, and 75 for MassBank.us. The reason for a higher library match score cut-off in the latter case was that most compounds in MassBank.us did not contain retention indices, and their molecular formulas often do not reflect the derivatization, making it more challenging to filter out false positives.

Table 2. Number of hits detected by each library separately after blank subtraction and elimination of false positives. The last two columns indicate non-overlapping hits identified with either NIST or the metabolomics PCDL.

	Library				
Sample	NIST23	PCDL	MoNA	NIST Unique	PCDL Unique
C1	140	130	95	52	43
C4	124	117	88	46	42
C5	126	115	85	46	37
D5	128	125	74	47	43
D6	138	126	92	55	42
D8	130	118	95	47	35

A substantial fraction of compounds was identified uniquely with either NIST or PCDL (Table 2), strongly suggesting the benefit of using multiple libraries in metabolomics research and highlighting the strengths of software like ChemVista to be able to integrate spectra from multiple sources.

Differential analysis comparing blood plasma metabolites of healthy subjects and individuals with heart failure

Statistical analysis was performed in MPP software, where the differences between the HF subjects and the healthy controls were evaluated. To compare metabolic profiles of HF subjects and healthy individuals, the samples from the subjects of two HF pathologies were grouped together forming a distinct cluster from the controls (healthy individuals) as can be seen on the principal component analysis (PCA) plot (Figure 3).

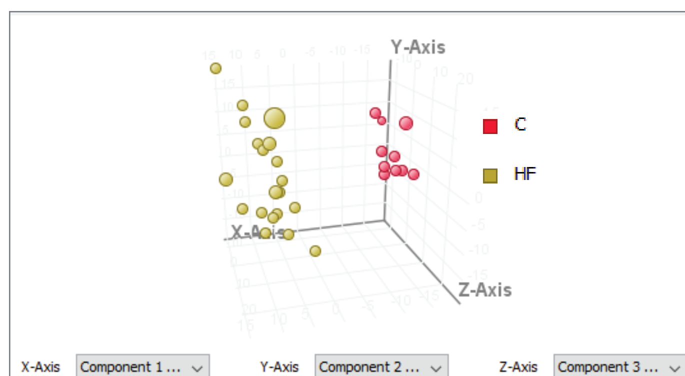


Figure 3. PCA plot showing a clear separation between the clusters of plasma samples from HF (HF) and healthy individuals (C).

Striking differences were observed between metabolic profiles of the HF subjects as compared to the healthy individuals when using the fold change analysis visualized on a volcano plot (Figure 4).

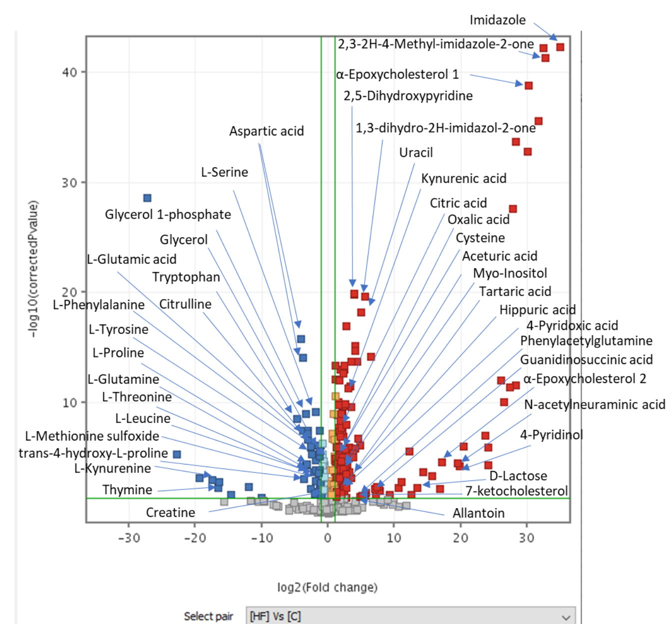


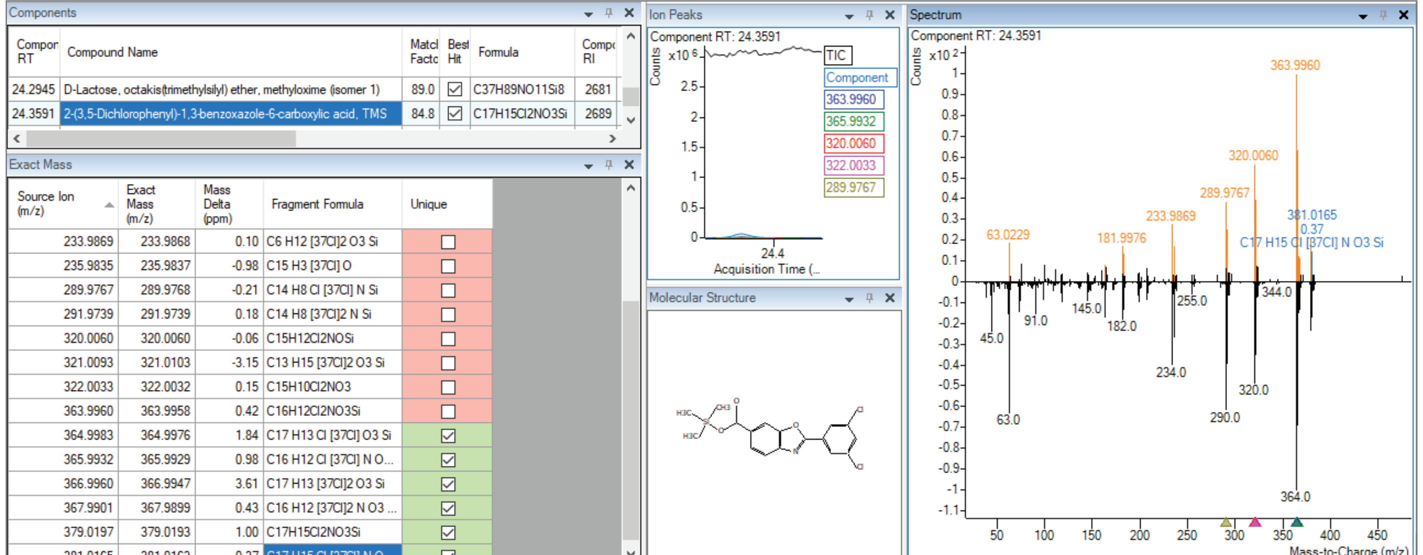
Figure 4. Volcano plot displaying log of fold change (cut-off 2) versus log of p-value (cut-off 0.05) for HF subjects versus healthy individuals.

Notably, the identified metabolites present at higher concentrations in the control samples were predominantly proteinogenic amino acids, while among metabolites identified at higher concentrations in the plasma of HF subjects were organic acids, sterols, and nitrogen-containing compounds, some of which could possibly be metabolites of imidazole-based drugs.

A few additional xenobiotics were identified in the blood plasma of some of the HF subjects, but not the others. Two such examples are shown in Figure 5, where tamafidis, an anti-cardiomyopathy drug, was identified in the plasma of the individual D5 (Figure 5A), and 2-(4-chlorophenoxy)acetic acid (a herbicide) was detected in the plasma of the D6 subject (Figure 5B).

These examples show the benefit of non-targeted analysis using the GC/Q-TOF system in combination with advanced accurate-mass software tools and multiple libraries in providing the capability to expand the scope of the study and enabling the identification of compounds of potential interest with high confidence in a time-efficient manner.

A 2-(3,5-Dichlorophenyl)-1,3-benzoxazole-6-carboxylic acid (tafamidis)



B 2-(4-Chlorophenoxy)acetic acid

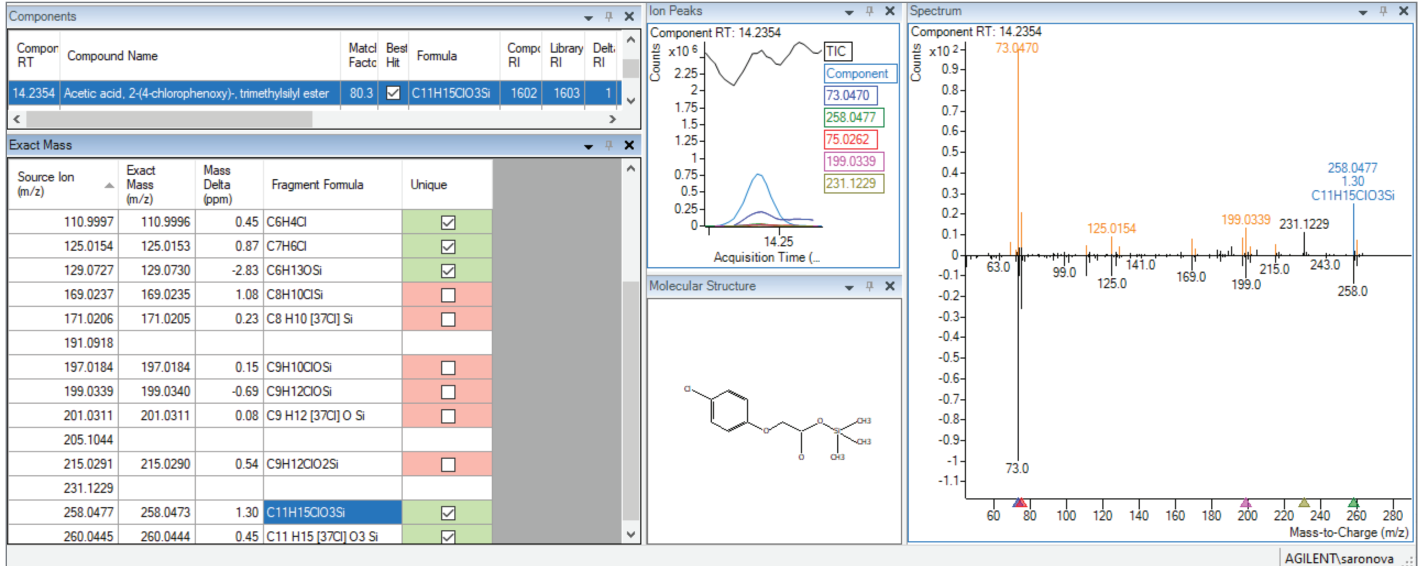


Figure 5. Examples of xenobiotics identified in some individuals with HF. (A) 2-(3,5-Dichlorophenyl)-1,3-benzoxazole-6-carboxylic acid (tafamidis). (B) 2-(4-Chlorophenoxy)acetic acid.

Identification of the unknown metabolites

Despite the use of multiple libraries in the identification workflow, a few compounds present at varied levels between the HF and healthy control groups were unknowns. Therefore, a variety of tools and techniques were used to perform structure elucidation of these compounds. To identify the molecular ion for one of the unknown metabolites present at higher levels in the plasma of HF individuals as compared to the healthy controls as a first step in structure elucidation, low electron energy was used. A 70 eV spectrum of this metabolite, as well as low-energy spectra obtained at 15, 12, and 10 eV for the same compound, are shown in Figure 6. A gradual increase in relative abundance of one of the higher *m/z* ions in the spectrum, 344, was observed, increasing from about 19% to 100% between 70 and 10 eV. This change in relative abundance could possibly be indicative of the ion with *m/z* 344 being a molecular ion.

Figure 6. Identification of the molecular ion using low-electron energy.

Consistent with the low-energy data, positive CI using methane as a reagent gas further confirmed the molecular ion of the unknown based on the presence of $[M+H]^+$ and $[M+C_2H_5]^+$ adducts (Figure 7).

After the confirmation of the molecular ion of the unknown, MSC software was used for structure elucidation. The approach is based on EI MS/MS data, where the candidate molecular ion is used as a precursor, and the accurate-mass MS/MS fragmentation pattern is matched to spectral databases such as ChemSpider, a source of molecular structures (Figure 8). In a complex matrix, such as blood plasma, MS/MS spectra have an advantage for structure elucidation of unknowns over full spectrum acquisition EI data due to significantly lower noise coming from the other spectra that may not be fully resolved chromatographically, thus facilitating the process of unknowns identification.

One of the top-ranked structures, based on the MSC score and number of literature references, was a TMS-derivatized form of pyrimidine-2,4,6-triol.

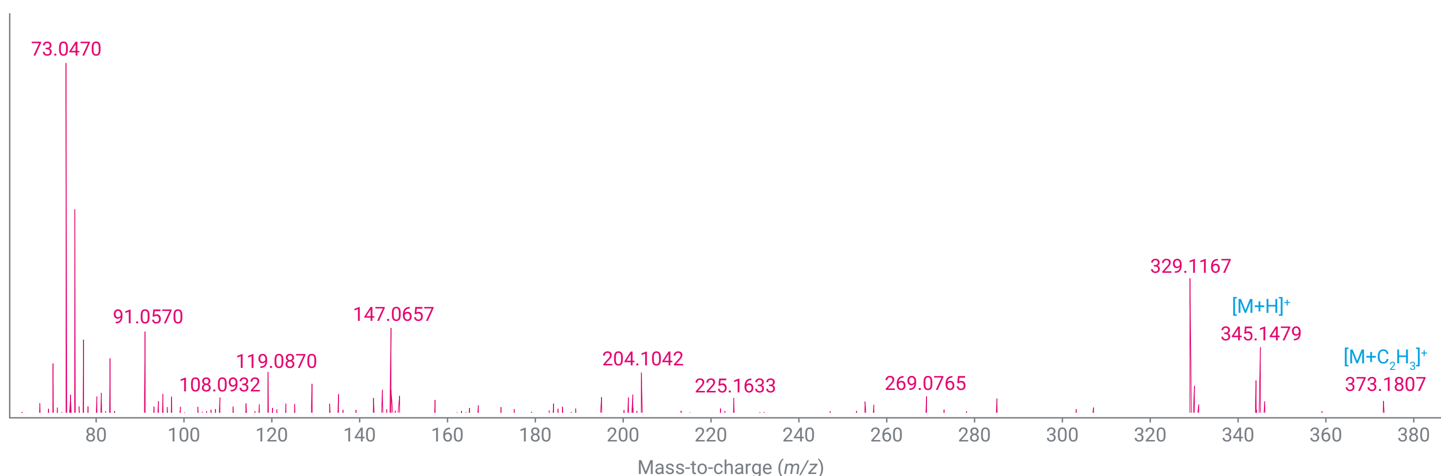
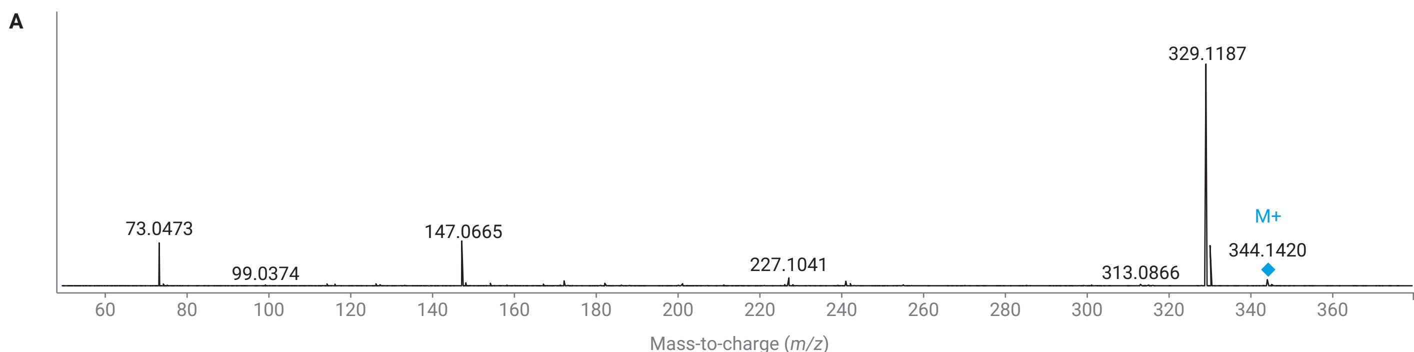


Figure 7. Positive CI spectrum of the unknown metabolite confirmed the molecular ion identified in low-energy experiments.

A



B

48 structures found for rt=14.408:ce20

Sort by # References Sort

Show structures for C13H28N2O3Si3

C13H28N2O3Si3; 461809

Scores MFG=90.6 MSC=

ChemSpider:

More Info...

C13H28N2O3Si3; 481642

Scores MFG=90.6 MSC=

ChemSpider:

More Info...

C13H28N2O3Si3; 481658

Scores MFG=90.6 MSC=

ChemSpider:

More Info...

Structure #1 -- elucidated: 75.0% ions , 98.2% Weight

Display Filt

	Mass	Intensity	Weight(%)	No. of candid.	Best score
1	329.1187	50455.72	78.7	2	90.7
2	147.0665	10204.34	3.2	4	42.2
3	73.0473	9762.28	0.7	1	98.6
4	330.1197	9215.45	14.5	2	55.4
5	227.1041	1790.00	1.3	0	0.0
6	172.0981	1129.09	0.5	0	0.0
7	241.0467	1073.97	0.9	2	95.0
8	148.0672	707.77	0.2	1	66.9

Penalty=1.0 dM=-5.9ppm F.D.S.=91.0 1 Of 9

C12H26N2O3Si3-H Score=90.7

Penalty=9.5 dM=-5.9ppm F.D.S.=91.0 1

C12H30N2O3Si3-5H Score=65.1

Figure 8. Structure elucidation of the unknown in Agilent MassHunter Molecular Structure Correlator software based on EI MS/MS data. (A) EI MS/MS spectrum of the unknown using a tentatively identified molecular ion as a precursor. (B) MSC results using the ChemSpider database, where one of the most probable candidates is shown.

Conclusion

This application note demonstrates the real-world impacts in understanding the metabolic signature of HF. It describes an approach to conduct a metabolomics study using the Agilent 7250 GC/Q-TOF with multiple libraries, including the new Agilent accurate mass metabolomics PCDL as well as third-party libraries managed in Agilent ChemVista software. While this study is focused on the metabolic profiling of heart failure (HF) pathology performed to understand the underlying mechanisms of this condition, it could be easily applied to any number of other health conditions and disease states. Over 40 metabolites, including amino acids, organic acids, and sterols have been identified as potentially responsible for the differences between HF and heathy individuals and thus the researchers in this field will be better armed to understand the metabolic impacts of disease and disease intervention, and potentially be able to rationally design future intervention regimens.

References

1. Funk, M. Epidemiology of Heart Failure. *Crit. Care Nurs. Clin. North Am.* **1993**, *5(4)*, 569–73. PMID: 8297547.
2. Borlaug, B. A.; Paulus, W. J. Heart Failure with Preserved Ejection Fraction: Pathophysiology, Diagnosis, and Treatment. *Eur. Heart J.* **2011**, *32(6)*, 670–679. DOI: 10.1093/eurheartj/ehq426
3. Beale, D. J.; Pinu, F. R.; Kouremenos, K. A.; Poojary, M. M.; Narayana, V. K.; Boughton, B. A.; Kanojia, K.; Dayalan, S.; Jones, O. A. H.; Dias, D. A. 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)*, 816–825. DOI: 10.1002/dta.1848
5. Schauer, N.; Steinhauser, D.; Strelkov, S.; Schomburg, D.; Allison, G.; Moritz, T.; Lundgren, K.; Roessner-Tunali, U.; Forbes, M. G.; Willmitzer, L.; *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
6. Stein, S. Mass Spectral Reference Libraries: an Ever-Expanding Resource for Chemical Identification. *Anal. Chem.* **2012**, *84(17)*, 7274–7282. DOI: 10.1021/ac301205z
7. Wohlgemuth, G.; Mehta, S. S.; Mejia, R. F.; Neumann, S.; Pedrosa, D.; Pluskal, T.; Schymanski, E. L.; Willighagen, E. L.; Wilson, M.; Wishart, D. S.; *et al.* SPLASH, a Hashed Identifier for Mass Spectra. *Nat. Biotechnol.* **2016**, *34*, 1099–1101. https://massbank.us/
8. Valdiviez, L.; Wang, S.; Sandhu, W.; Ye, H.; Fiehn, O.; Nieto, S. Comprehensive Accurate Mass Metabolomics Library and Its Evaluation in Targeted and Nontargeted Data Analysis Workflows. *Agilent Technologies application note*, publication number 5994-5832EN, **2023**.
9. 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
10. Agilent GC/Q-TOF PCDL User Guide. *Agilent Technologies user guide*, publication number D0030675. https://www.agilent.com/cs/library/usermanuals/public/D0030675%20GC-QTOF%20PCDL%20User%20Guide.pdf
11. Agilent ChemVista Library Manager. *Agilent Technologies technical overview*, publication number 5994-5924EN, **2023**.

www.agilent.com

For Research Use Only. Not for use in diagnostic procedures.

RA45230.6108564815

This information is subject to change without notice.

© Agilent Technologies, Inc. 2023

Printed in the USA, November 15, 2023
5994-6858EN

Agilent
Trusted Answers