SMART NOTE

Smart Notes

A

How can the Orbitrap Exploris GC 240 mass spectrometer and Compound Discoverer software lead discovery in metabolomics studies?

For laboratories whose primary focus is metabolomic studies, the capability to comprehensively and accurately characterize and quantify the metabolome is critical. Gas chromatography high-resolution accurate mass (HRAM) mass spectrometry provides supportive information on volatile, semi-volatile, and derivatized polar metabolites across a range of biological, environmental, and clinical sample matrices. This capability in conjunction with liquid chromatography-mass spectrometry provides the ultimate coverage into the metabolome. A complete process to generate meaningful conclusions also includes careful experimental design and conduct prior to sample preparation. The Thermo Scientific™ Orbitrap Exploris™ GC 240 mass spectrometer delivers the very highest data quality for metabolomics to enable researchers to gain unprecedented depth of analysis and to have the flexibility to adapt to evolving research demands, such as biomarker discovery or metabolic pathway identification. With this analytical versatility, the system can provide the deepest insight into metabolomic studies.





High quality raw data is important, but equally so, are informatics solutions to take that data and make clear discoveries between sample groups and to follow up with confident compound identifications resulting in actionable outcomes in an easy-to-use platform. The unrivalled combination of high-sensitivity full-scan HRAM data and intelligent identification using Thermo Scientific[™] Compound Discoverer[™] software addresses these challenges to advance scientific understanding. In this Smart Note, we take a closer look at metabolomics with the Orbitrap Exploris GC 240 MS¹ and Compound Discoverer software.

Metabolomics aims to characterize and quantify the complete small molecule complement, or metabolome, of a biological system. The metabolome consists of a diverse mixture of small molecules, including amino acids, sugars and phosphosugars, and biogenic amines and lipids. Untargeted metabolomics is exceptionally challenging due to the requirement to both identify and quantify hundreds of different compounds with limited a priori knowledge of the metabolites. Therefore, it is advantageous to use a detection system that is not only capable of sensitive detection of specific molecules in an untargeted way but can also provide accurate mass information and spectral information for confident confirmation and structural elucidation of unknowns. Gas chromatography-mass spectrometry (GC-MS) is routinely used for metabolomics applications due to its inherent advantages, especially its easy method development, chromatographic resolution, reproducibility, peak capacity, and extensive commercially available spectral libraries. GC provides excellent chromatographic separation capability for biomarker discovery using untargeted metabolomics. However, it has previously been hampered by the lack of high-end mass spectrometry support providing the dynamic range, accurate mass, and scan rate sufficient to analyze complex biological samples. The polar nature of most central metabolites means that derivatization must be performed to allow effective volatilization and ensure good chromatography. High sample throughput and advanced automation is required for metabolomic analysis, especially for clinical metabolomics. GC-MS using an Orbitrap-based mass spectrometry detector enables ultra-high mass resolution, sub-ppm mass accuracy, a large dynamic range, and a scan rate commensurate with the efficient quantitative analysis of highly complex metabolomic samples. The high resolution, mass accuracy, and scan speed is critical for consistent data deconvolution to permit the detection of species from overlapping TIC peaks. Accurate mass electron ionization (EI) fragment patterns are also suitable

for matching against the widely available NIST and Wiley libraries for tentative compound identification, while providing accurate mass for more in-depth characterization.

Metabolomics with Thermo Scientific Orbitrap Exploris GC 240 MS and Compound Discoverer software

A well-planned untargeted metabolomics study begins with careful study design aimed at eliminating systematic and experimental bias. Second to this comes sample handling, metabolite extraction, and preparation prior to instrumental analysis. As described previously, a key aspect to a successful study is high quality accurate mass data that is processed using a powerful and fast informatics platform. Compound Discoverer software provides this capability to transform high quality data into meaningful results with ease. Starting with feature extraction, statistical analysis, compound identification, and biological pathway analysis, this combination allows researchers to obtain an unprecedented depth of analysis, providing a route to robust scientific discoveries.

Figure 1 details the instrumental and data processing workflow using the Orbitrap Exploris GC 240 MS. Data is first collected using EI in full scan at mass resolving power of 60,000 (FWHM m/z 200) and processed in Compound Discoverer software using the El workflow node. This node performs feature extraction by deconvoluting the data to generate a list of features across all samples. Statistical analysis can also be carried out to identify differences between sample groups or individual samples. For example, these could be differences of study variables such as treatment type, age, origin, etc. These features, once filtered as being statistically significant (a potential biomarker), can then be matched against a nominal mass spectral library such as NIST or Wiley to make a putative compound identification. Further analysis could also be performed using a high-resolution accurate mass library for example Thermo Scientific[™] Orbitrap[™] GC-MS HRAM Metabolomics Library. Further confirmation of the library hit can be obtained by analysis using positive chemical ionization (PCI) on the Orbitrap Exploris GC 240 MS. This additional confirmation in the annotation of putative compounds and metabolites can be achieved by assessing the PCI spectra to predict the elemental composition of the molecular ion by looking at common adducts associated with a reagent gas. In PCI experiments using methane as the reagent gas, three adducts are typically observed: $[M+H]^+$, $[M+C_2H_5]^+$, $[M+C_2H_5]^+$. Alternative PCI reagent gases can be used including isobutane and ammonia.

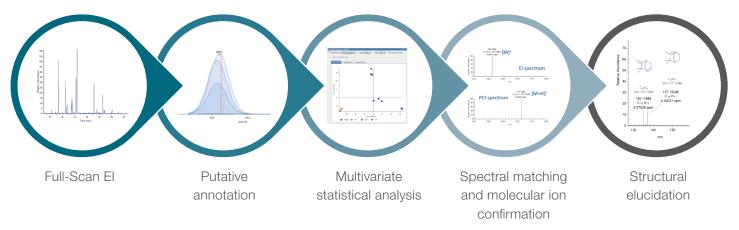
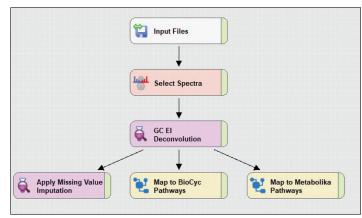


Figure 1. Metabolomics workflow using the Orbitrap Exploris GC 240 MS and Compound Discoverer software, taking high quality data through to confident metabolite identification and confirmation

Figure 2A shows the details of the two nodes available within the Compound Discoverer software specific for processing Orbitrap Exploris GC El and PCI data. The El node includes an initial peak deconvolution to extract the peak features within a data set. These peak features can then be statistically compared to other sample groups prior to compound annotation. The first step in compound identification is to match the clean deconvoluted spectra to existing spectral libraries (Figure 2B). Within the node these can be specified as nominal mass libraries such as NIST or high-resolution accurate mass libraries. The PCI node also included peak feature extraction, but instead of matching to spectral libraries, the molecular ion is identified through mass adducts, and elemental compositions are proposed for the molecular ion using a pre-defined set of elements. With excellent sub-1 ppm mass accuracy expected from Orbitrap MS, the list of possible elemental composition is small, enabling the correct assignment to be made and within a short time frame. Once an elemental composition is proposed, this can then be automatically searched against ChemSpider online database of compounds to make a putative annotation. Absolute confirmation of compound identify will require comparison of retention time and spectra against a certified standard.





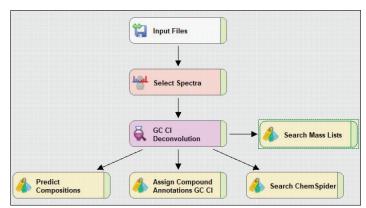
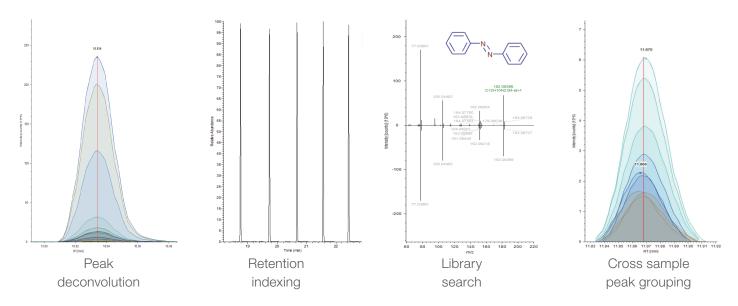


Figure 2A. GC workflows in Compound Discoverer software, covering both EI and PCI HRAM data

GC El workflow





In the results browser of Compound Discoverer software, all the peak detection, annotation, and statistical tools are easily available and clearly displayed to the analyst for review. Figure 3 shows the results browser with El spectral match results to a library entry. Library search index score, ppm mass accuracy of molecular ion, retention index, and high-resolution filtering score are displayed alongside the mirror plot of the library top hit. The extracted ion chromatogram of the compound is also displayed to enable the user to quickly see how this peak intensity varies across all samples within the analytical batch. The sensitivity and selectivity of the Orbitrap Exploris GC 240 MS enables the software to work efficiently, and high quality results with good confidence are displayed to the analyst in this report. For example, in this case the compound is putatively annotated as camphene with 0.7 ppm mass accuracy, a search index of 837 and HRF score 95.3, and retention index delta of 1.

Chromatogra	ams						• 4	× Mass S	pectrum											
Filter By	regano (2/) e (6/6) r: Dregan	120		9.18 9.20 RT[min]	9220	924	926		F7, RT→9220,	Deconvolved Sp	ectrum, FTMS (+)	RAME REFEI 200 (9,0) U	67.0000	20 Deconvolve Camphene, C1 70 0667 70 0667 70 0667 71 00000 77 00000	4 91 05432 94 07343 96 08957 92 00000 91 00000 93 0	854, 07002 92.05213 10 94.07780	5 00994 187 0856 1 197 00000	121 10133 122 10488 122 00000 121 00000 Library	19412466 19512466 19512666 19500000 19500000 V spectrum 140	
	ompound Checked		First Le	evel			Reference m/z				ST Theo. Mol. Mass		Mol. M Total				and the second second second		Ŧ	
67 🖙		Propanal ethyl isopentyl acetal 129.09152 9.205				9.205	71.0492	_	694 C10 H22		174.16143				7.1 680	952			74e6 6.08e6 3.76e6	
68 ቱ		COCCC Camphene			136.12511	9.220	93.0700				136.12465			94.9 95.3 837		953		1.57e7 1.64e7 2.0		
69 🖨		0000 2-Furancarboxylic ad			113.16324	9.348	113.0961						92.9 96.8 707 960						73e6 9.54e6 7.60e6	
70 😐		OOOO Hexane, 1-(hexyloxy			98.10953	9.452	57.0700	-				92.6 94.3 744 966						99e6 2.00e6 2.09e6		
71 🖶		0000 1-Propyl-1-cycloper	itanol		109.10189	9.479	57.0336				128.11957	92.9 95.3 738 967						58e6 9.61e6 9.46e6		
72 🗢		00000 Peak@9.481			84.09397	9.481	86.0681					967						1.48e5 1.42e5		
73 🖶		00000 2-Butene, 1-methox			86.07314	9.485	71.0491					86.	86.07259 89.7 87.2 738			968			87e5 1.02e6 9.73e5	
74 ↔ A Hide Rel	ated Table	Bis(2-methylallyl) ca	rbonate		111.08110	9.493	111.0805	5 2325	437 C9 H14	03	170.09375			86.7 8	5.6 624	968	8.45e5 8	8.50e5 8.62e5 5.8	34e5 6.42e5 5.54e5	
GC El Comp			Second	d Level																
(P	Checked	Structure	Name	CAS Num	Formula	Total Score •	HRF Score	SI Moles	cular Weight	Theo. Mol. Mass	Observed Mol. Mass	ΔMass [ppm]	Library RI	RI Delta	RI Diff[9	6]	M+ In Lib	M+ found	Selected	Elements Found[%]
1 👳		14	Camphene	79-92-5	C10 H16	94.9	95.3	837	136.12520	136.12465	136.12456	-0.69	952		1	0.1	Yes	Yes	True	100.0
2 🖙		A	(-)-beta-Pinene	18172-67-3	C10 H16	94.7	95.3	827	136.12520	136.12465	136.12456	-0.69	986		33	3.5	Yes	Yes	False	100.0

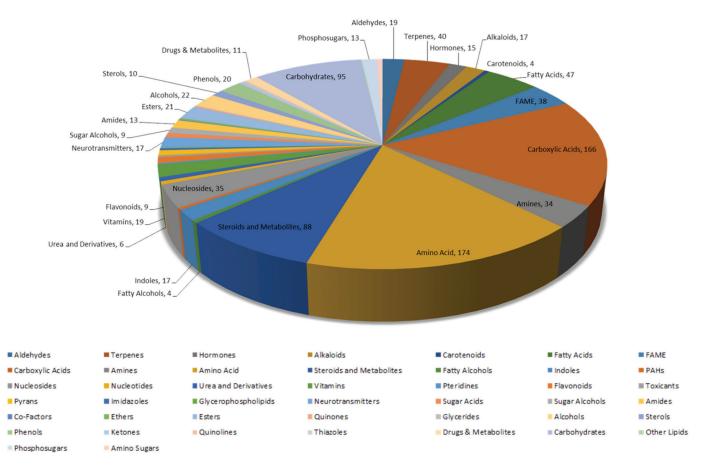
Figure 3. Compound Discoverer library search result browser view. Display shows extracted ion chromatogram across all samples, library match mirror view, list of peaks, and identification results.

Orbitrap GC-MS HRAM Metabolomics Library

With the need to decipher biological complexities, metabolomics analysis demands sophisticated analytical technologies and advanced software solutions. The Orbitrap GC-MS HRAM Metabolomics Library is the first commercially available high-resolution accurate mass metabolomics library for electron ionization GC-MS. It contains more than 900 retention indexed unique entries from more than 800 metabolites, providing a broad coverage of primary and secondary metabolites (including volatiles) in plants, animals, and microbes (Figure 4). When used in combination with powerful Orbitrap GC-MS technology and unique Thermo Scientific software data processing tools, the difficult challenge of metabolite identification in untargeted experiments is easier than ever before. Library entries, many with methoxyamine/MSTFA derivatization were acquired at 60,000 mass resolving power and include Kovats retention index information for additional confirmation.

Positive chemical ionization to annotate the molecular ion

A complimentary approach is to identify unknown compounds or confirm spectral match results using PCI spectral data within Compound Discoverer software. Using the PCI node, the data is deconvoluted to determine peak features across all samples, and these can also be compared to each sample group as shown in Figure 5. The focus in PCI data is to identify molecular ions using expected mass adducts to the molecular ion.



Classification of GC-Orbitrap Metabolomics Library in 44 chemical groups

Figure 4. Orbitrap GC-MS HRAM metabolomics library compound contents details

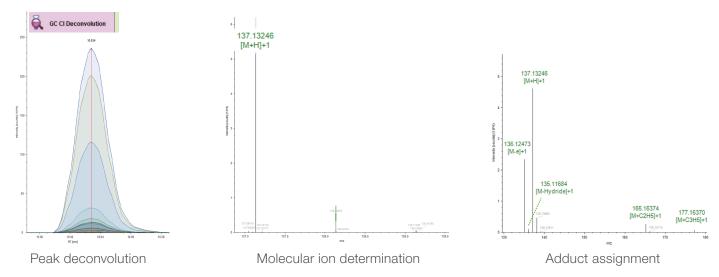


Figure 5. Chemical ionization node functionality with the aim of identifying molecular ion adducts to support unknown identification or compound confirmation from spectral match result. Peak deconvolution is followed by automatic assignment of molecular ion adducts (center and right).

Statistical comparison information

Multivariate statistical analysis is easily accessible within Compound Discoverer software alongside interactive annotation and filtering tools. Figure 6 provides examples of some classical approaches to detect differences between sample groups including PCA, hierarchical cluster, box plots, and volcano plots. PCA is a well-known statistical approach that highlights variation between sample groups and allows visualization of strong patterns in complex datasets. The first step will inform if there are any differences between the sample groups. Ideally, a pooled sample will have been included in the data acquisition, and this should fit around the center point of the PCA to validate the experiment. With confirmation of significant statistical differences, the next action is to determine which peak or peaks contribute to that difference, identifying the markers unique to a sample group. The volcano plot is a powerful tool to identify the features that are unique to a group, and from here the compounds can be identified using the approaches described above.

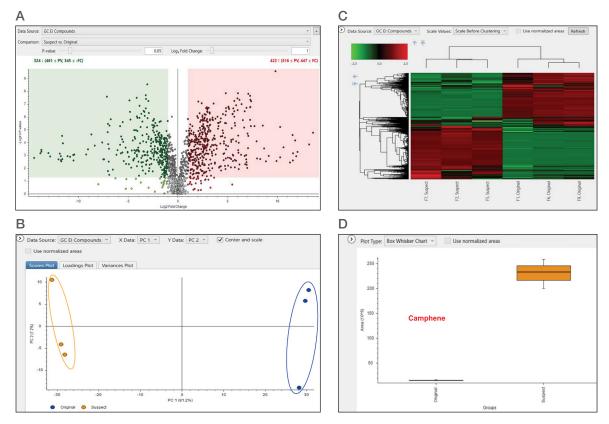


Figure 6. Statistical analysis tools in Compound Discoverer software with El data. (A) Volcano plot, (B) PCA, (C) Hierarchical cluster analysis, (D) Boxplot

Compound Discoverer software will automatically assign $[M+H]^+$, $[M-H]^+$, $[M+C_2H_5]^+$, $[M+C_3H_5]^+$ ions in the spectrum during the deconvolution. If instructed in the workflow, the software will predict the elemental composition of the molecular ion based on these adducts and the elements included within the search parameters. All the results and matched criteria are displayed within the results browser (Figure 7). With sub-ppm mass accuracy, the number of suggested elemental compositions is low and the confidence in assignment high. The end result is that researchers gain unprecedented depth of analysis and speed to result.

For true unknowns with PCI MS² scan, the elemental composition will be screened against the ChemSpider database. Compound Discoverer software calculates a FISh score to generate in silica MS² fragmentation spectra for comparison between the acquired ions and the theoretical fragmentation (Figure 8). The fragments are matched with theoretical pattern and annotated in green. The higher the FISh scores, the higher the confidence in the assignment.

Chromatograms	↓ ↓ × Mass Spe	ectrum					– 4
MZ: 305.10800 File: Bul208_0817 raw (F12) FTMS (+) MS1 0 12.410 12.410 12.410	l I III F1 III F1	12, RT=12.410, Deconvolved 12 #3807, RT=12.410 min, M 12 #3803, RT=12.402 min, M 12 #3808, RT=12.402 min, M <u>*</u> F12 #3808, RT=12.413 min	/S1, FTMS (+)	A, 305.1080@15, -	CI (F13) RF=12.410, Deconvolved O3 P S as [M+6]+1 305 (pieco [M+6]+1 4.10077 4.10077 503 00268 [M-6]+1 303 00268	333.13960 [M+C2H5]+1	346.13963
E 0 1 12.37 12.38 12.39 12.40 12.41 12.42 12.43 12.44 RT [min]	12.45) () () () () () () () () () (307.10376 317.1	335.1 <u>3513</u> 320 330 m/z	34 [M+C3H6]+1 340 35
😤 Checked Tags 🖈 Name RT [min] + Formula	# Adducts MS2	MS Depth Annot. Source +	# ChemSpider Results Annot. ΔI	Mass [ppm] Calc. MW Reference r	m/z Avg TIC Mass List Match	nes 🕂 Area 🕂	
49 🖶 🔲 🔍 💭 Diazinone 12.501 C12 H21 N2 O3 P S	5	2	2	0.64 304.10124 305.108	352 1311013	2.5968	
50 垣 🔲 00000 Omethoate 12.564 C3 H9 O3 P S	2	1	2	0.54 156.00109 185.039	67 19389	3.50e4	
51 ቱ 🔲 00000 2-Adamantan-1-yl-4-methyl-phenol 12.568 C17 H22 O	2	1	10	0.71 242.16724 207.149	25 14154	7.45e3	
52 🗢 🔲 00000 Anthracene 12.624 C14 H10	4	1	10	0.68 178.07837 179.085	65 4002	6.77e3	
53 🗢 🔲 00000 8145294 12.875 C9 H17 CI N3 O3 P S	4	2	1	1.45 313.04213 314.049	941 1146538	1.99e8	
54 😓 🔲 00000 2-amino-4.6-di-tert-butylphenol 12.952 C14 H23 N 0	2	1	28	0.35 221.17804 197.110	075 3813	1.70e3	
55 🗢 🔲 00000 Tetraspiro[cyclopentane-1,3'-pentac 13.101 C28 H40	1	1	6	-0.80 376.31270 215.961	49 1747	1.65e3	
56 H 00000 Chloropyriphos-methyl 14.062 C7 H7 Cl3 N O3 P S	5	2	1	0.07 320.89501 323.899	26 1362083	1.22e8	
Hide Related Tables					-		
	pider Results Mass List Se	arch Results Secon	d Level				
Tags 🛨 Checked Compound Match Formula	Molecular Weight AMass [Da] ∆Mass [ppm] RDBE H	I/C Rank 🔺 # Matched Iso.	# Missed Iso. # Matched Frag.	SFit [%] Pattern Cov. [%]	MS Cov. [%] MSMS Cov. [%]	
1 🖘 00000 🔳 🔳 C12 H21 N2 O3 P S	304.10105 -0.000	032 -1.07 4.0 1	1.8 1 5	0 9	75 99.91	99.91 81.81	
2 ቱ 🔿 🖓 💭 🔲 📕 C11 H17 F N4 O3 S	304.10054 0.000	0.61 5.0 1	1.5 2 5	0 8	79 99.91	99.91 76.72	
3 🗢 😳 😳 🔲 📕 C9 H22 F N2 O4 P S	304.10219 -0.001	47 -4.83 0.0 2	2.4 3 5	0 8	65 100.00	99.91 80.06	
4 🖶 00000 🔲 📕 C14 H16 N4 02 S	304.09940 0.001	4.37 9.0 1	1.1 4 6		57 100.00	99.98 59.38	

Figure 7. GC CI results. Predicted elemental compositions based on accurate mass and isotope pattern. The molecular ion is identified through adduct assignment and elemental compositions proposed and supported with theoretical isotope pattern matching.

Mass Spe	ectru	um																	•
🔺 🗽 F1	13 #	3803, RT	=12.40	2 min, M	S1, FTMS (+)	\mathbf{i}	Diazinon C12 H21 N2 O3 P	s	, FTMS (+), (HCD, DDA, 305.10	180@15, +1)									
), (HCD, DDA, 305	FISh Coverage: 7 Matched,	Unmatched, 13 Ski	oped									```	١
					FTMS (+)		6												
[!*, F13 #3807, RT=12.410 min, MS1, FTMS (+)							6.5		N	2							N-11 5-0-/		
							10		№ • 169.07925						Υ ^ν ν ² σ ^ρ όн,				
						(9,01) (sunns) (10,0		135.09 C8 H11 N2	.09155 N2 [M-e]+1 153.10217					Y	N spon, N o ^P oн	277.07678 C10 H18 N2 O3 P S [M-	e]+1		
							2 3 -	S+OH	114.96120	153,10217		NUMBER				249.04549			
				Then site	⁵ рон он 96.9507	S#0-	P DH				YN OP		N2 O3 P S [M-e]+1	1					
						1 9	H2 O2 P S (1-e]+1 124.98 C2 H6 O2 P	J190		171.02458 C8		231.03416 C8 H12 N2 O2 P S [M-e]+1						
						0 83.11395	82.12959. 1	9.00539	,L	/	/181.13295	2	15.0574	•		¥		\vdash	
							50	10	0	150			200 m/z			250	31	00	/
<				_		>							1102						
💷 GC	CI C	ompoun	ds 🏆	Che	mSpider Result	s Input Files Stu	dy Information												
P		Checked	Tags	+	Name		Formula	Annot. Source	e 🛨 Annot. ∆Mass [ppm] Calc. MW	RT [min] 🔺	# ChemSpi	ider Results	Reference m/z	Avg TIC	Mass List Matches 🛨	# Adducts MS2	MS Depth A	Area 🛨
22 +=	•		00		Diazinone		C12 H21 N2 O3 P S		-1.07	7 304.10073	12.410		2	305.10800	9018762		5	2	9.35e7
23 🖙	•		00	000	m-Xylylenedia	mine	C8 H12 N2		-0.52	2 136.09998	12.419		61	137.10725	656590		3 📕	1	1.35e6
24 🖙	-		00	000	QS6350000		C9 H12 N2		0.35	5 148.10010	12.760		36	149.10738	12538		2	1	7.54e3
25 👳	-		00	000			C6 H9 CI F N5 O3		0.80	253.03800	12.793		1	282.07693	239951		2	1	1.90e5
26 🖙	2		00	000	1-Propyl-4-([4	-(4-propylcyclohexyl)p	C26 H32		0.2	344.25047	12.810		5	185.03961	28098		1 📕	1	5.33e4
27 👳	-		00	000	Amantadine		C10 H17 N		0.78	3 151.13622	12.973		19	152.14349	5473		1 📕	1	2.20e3
• Hide	Rel	ated Tab	es																
Structur	re P	roposals	GC	CI Compo	ounds per File	Predicted Compositi	ions ChemSpider Resul	s Mass List Se	arch Results										
P	(Checked	I Tags	+	Structure		Name	Fo	rmula	Molecular Wei	ight FISh Co	overage	Comments						
					N														
1 +=			Diazinon C		2 H21 N2 O3 P S	304.101	105	63.64											
					Γ Γ Ν	0. 0.													

Figure 8. GC CI results. FISh Scoring for MS² scan. Identification through online database matching is performed and fragments assigned and annotated to mass spectrum.

thermo scientific

Summary

The data quality from the Orbitrap Exploris GC 240 MS combined with the functionality of Compound Discoverer software demonstrate a powerful tool for metabolomic studies. The solution provides an integrated metabolomics approach for targeted analysis, profiling complex samples, and identifying unknown compounds.

- The streamlined GC-EI and GC-CI data processing workflow nodes in Compound Discoverer software allow for extraction, deconvolution, identification of unknown compounds, and multivariate statistical and pathway analysis.
- Principal components and differential analysis tools allow for the detection of differences in the composition of sample groups. After peaks are isolated, they can be quickly screened against existing nominal mass or highresolution accurate mass spectral libraries.

- The high resolving power, consistent sub-1 ppm mass accuracy, and wide dynamic range allow for fast and confident characterization of compounds, regardless of their concentration or matrix complexity.
- Rapid change-over from El (for spectral library search) to softer ionization such as PCI (for molecular ion confirmation using adduct information) and the ability to perform accurate mass MS/MS experiments (for structural elucidation), enable confident compound identification with unprecedented ease.

Reference

 Thermo Fisher Scientific Application Note 10724: Breakthrough performance of the Orbitrap Exploris GC for analytical testing and scientific research applications, 2020. https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/an-10724-gc-msbreakthrough-performance-orbitrap-exploris-an10724-en.pdf

Find out more at thermofisher.com/OrbitrapExplorisGC240

© 2021 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is presented as an example of the capabilities of Thermo Fisher Scientific products. It is not intended to encourage use of these products in any manners that might infringe the intellectual property rights of others. Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representatives for details. **SN000070-EN 0721C**

