**Environmental**

# Comprehensive non-targeted workflow for confident identification of perfluoroalkyl substances (PFAS)

Richard Cochran<sup>1</sup>, Sarah Choyke<sup>2</sup>,  
Collin Meyers<sup>2</sup>, Ralf Tautenhahn<sup>3</sup>

<sup>1</sup>Thermo Fisher Scientific, Bannockburn, IL

<sup>2</sup>Eurofins Environment Testing, Arvada, CO

<sup>3</sup>Thermo Fisher Scientific, San Jose, CA

## Keywords

Perfluoroalkyl substances (PFAS), non-targeted analysis, Orbitrap Exploris 240 MS, HRAM, mass spectrometry, LC-MS, Compound Discoverer software

## Application benefits

- High-resolution accurate-mass (HRAM) data acquired using the Thermo Scientific™ Orbitrap Exploris™ mass spectrometer platform enables confident identification of unknown perfluoroalkyl substances (PFAS).
- Thermo Scientific™ Compound Discoverer™ software (version 3.4) offers the most comprehensive commercially available workflow for PFAS analysis, covering over 40,000 PFAS.
- Compound Discover software enables complete and consistent application of the Schymanski<sup>1</sup> annotation confidence scale, including customizable prioritization of multiple spectral libraries and databases.
- Definition of annotation confidence scales using Compound Discoverer software is useful for other non-targeted applications in addition to PFAS analysis, including extractables and leachables, impurities analysis, food safety, toxicology, and others.

## Goal

To demonstrate a comprehensive workflow for non-targeted PFAS analysis that combines HRAM sample analysis with the Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer and powerful automated data analysis and visualization using Compound Discoverer 3.4 software.

## Introduction

PFAS comprise a class of chemicals that, by definition, are fluorinated substances containing at least one fully fluorinated methyl or methylene carbon atom (without any H/Cl/Br/I atom attached to it), that is, with a few noted exceptions, any chemical with at least a perfluorinated methyl group ( $-\text{CF}_3$ ) or a perfluorinated methylene group ( $-\text{CF}_2-$ ).<sup>2</sup> Many PFAS are persistent, bioaccumulative, and toxic, making them contaminants of concern that are monitored by various regulatory agencies. However, regulatory methods typically use targeted quantitative analysis based on reference standards that do not address the thousands of novel PFAS and their transformation products for which there are no commercially available standards. Non-targeted methods that detect and identify compounds in complex mixtures with little to no prior knowledge about the PFAS present are therefore needed to understand the true extent of contamination and to facilitate source fingerprinting for remediation.

There are multiple fundamental challenges that limit the capability of non-targeted PFAS analysis. This has led to multiple working groups across the scientific community, such as the Best Practices for Non-Targeted Analysis (BP4NTA) group, that focus on non-targeted approaches.<sup>3</sup> First and foremost, the general consensus is that identifications of unknown PFAS need to be made with as much confidence as possible. This requires a high level of selectivity, which is obtained through high-accuracy and high-resolution mass measurements of the monoisotopic masses of the intact molecule and the fragments generated using high energy collision-induced dissociation (HCD) or collision-induced dissociation (CID) or other fragmentation techniques. Second, with the limited availability of authentic standards to generate reference spectral libraries, scientists must rely on other tools and resources such as *in silico* generated spectral libraries, fragment databases, and exhaustive chemical databases.

One of the most important steps in a non-targeted PFAS workflow is to assign a level of confidence in the final PFAS annotation. Schymanski, et al.<sup>1</sup> and Charbonnet, et al. 2022<sup>4</sup> provide a framework for coupling high-resolution accurate-mass analysis with a comprehensive set of libraries, databases, and other PFAS-specific tools to make such assignments. Because of differences in mass analyzer technologies that result in varied mass resolution and mass accuracy, and the assorted use of libraries, databases, and tools, there can be widespread inconsistencies in the assignment of PFAS annotation confidence levels. To evaluate these inconsistencies across

testing laboratories, NIST carried out the Per- and Polyfluoroalkyl Substances–Non-Targeted Analysis Interlaboratory Study (PFAS NTAILS).<sup>5</sup> Thermo Scientific™ Orbitrap™ mass analyzer technology combined with Compound Discoverer<sup>3,4</sup> software addresses annotation disparities by providing sub-1 ppm mass error and ultra-high mass resolution along with the most comprehensive collection of PFAS-specific libraries, databases, and tools and a simplified, automated, and accurate approach to consistently assigning annotation confidence levels.

This application note demonstrates a comprehensive non-targeted PFAS annotation workflow that combines Orbitrap Exploris 240 mass spectrometer sample analysis with powerful automated data analysis and visualization using Compound Discoverer 3.4 software. An analysis of “unknown” samples from the PFAS NTAILS<sup>5</sup> is used to explain how the software leverages PFAS-specific databases, reference spectral libraries, and *in silico* generated spectral libraries covering more than 40,000 PFAS in a unified non-targeted data processing workflow. Software features that enable complete and consistent implementation of the Schymanski annotation confidence scale, including definition of the priority of spectral libraries and databases used for PFAS annotation, are presented. Data comparison and visualization tools that facilitate determination of PFAS sample composition are also covered, including principal component analysis (PCA) and differential analysis plots, mass defect plots for identification of homologous series of PFAS, and molecular networks to view structurally related PFAS.

## Experimental Samples

Three samples from the NIST PFAS NTAILS<sup>5</sup> representing unknowns were analyzed to demonstrate workflow performance. All samples were analyzed as received from NIST without further dilution. Sample A was a mixture of spiked PFAS reference standards in methanol. Each individual PFAS concentration was 0.1  $\mu\text{g/g}$  of solution. The spiked PFAS in the sample are provided in reference 5. Sample B was a mixture of two aqueous film-forming foam (AFFF) formulations—an electrochemical fluorination product and a fluorotelomerization product—diluted in methanol to a total mass concentration of 1,000  $\mu\text{g}$  of AFFF product per 1 g of solution. Sample C was a methanolic extract of AFFF-impacted soil, prepared by extracting approximately 400 g of soil via ultrasonication in 600 mL methanol. The final extract was concentrated 20-fold. The final extract was spiked with an analytical standard of perfluorohexane sulfonamido amine (N-AP-FHxSA) to a nominal concentration of 0.1  $\mu\text{g/g}$ .

## Instrument method

Ultra-high-performance liquid chromatography (UHPLC) was performed using a Thermo Scientific™ Vanquish™ Flex UHPLC system fitted with the PFAS Installation Kit that replaces all wetted materials containing fluoropolymers with PEEK, with the exception of the solvent degasser membrane. Furthermore, a delay column was installed between the solvent pump and the analytical column to shift any background contamination from the solvents to a later retention time in the chromatogram. All chromatographic conditions are listed in Table 1.

**Table 1. Chromatographic conditions**

Parameter	Value																										
Mobile phase A	Water + 20 mM ammonium acetate																										
Mobile phase B	Methanol + 20 mM ammonium acetate																										
Columns	Delay: 30 × 3 mm C18 column Analytical: 50 × 3 mm C18 column																										
Flow rate	0.400 mL/min																										
Column temperature	40 °C																										
Autosampler temperature	10 °C																										
Gradient	<table border="1"><thead><tr><th>Time (min)</th><th>% B</th></tr></thead><tbody><tr><td>0</td><td>10</td></tr><tr><td>0.3</td><td>55</td></tr><tr><td>4</td><td>65</td></tr><tr><td>7.4</td><td>75</td></tr><tr><td>9</td><td>80</td></tr><tr><td>9.5</td><td>85</td></tr><tr><td>11</td><td>90</td></tr><tr><td>12</td><td>95</td></tr><tr><td>13</td><td>99</td></tr><tr><td>13.5</td><td>99</td></tr><tr><td>15</td><td>10</td></tr><tr><td>17</td><td>10</td></tr></tbody></table>	Time (min)	% B	0	10	0.3	55	4	65	7.4	75	9	80	9.5	85	11	90	12	95	13	99	13.5	99	15	10	17	10
Time (min)	% B																										
0	10																										
0.3	55																										
4	65																										
7.4	75																										
9	80																										
9.5	85																										
11	90																										
12	95																										
13	99																										
13.5	99																										
15	10																										
17	10																										

Mass spectrometry analysis was performed using an Orbitrap Exploris 240 mass spectrometer equipped with the Thermo Scientific™ EASY-IC™ (internal calibration) source. The MS ion source parameters are listed in Table 2. The mass spectrometer was operated in the Full Scan (MS<sup>1</sup>) plus data-dependent MS<sup>2</sup> (ddMS<sup>2</sup>) mode using the parameters shown in Table 3. The ddMS<sup>2</sup> mode acquires a full scan (survey scan) followed by a set of ddMS<sup>2</sup> scans during which HCD fragmentation is applied. Stepped normalized collision energies were used to ensure (1) consistent fragmentation of all precursors across the full scan mass range and (2) the collection of rich fragmentation patterns for all MS<sup>2</sup> spectra collected.

**Table 2. MS ion source parameters**

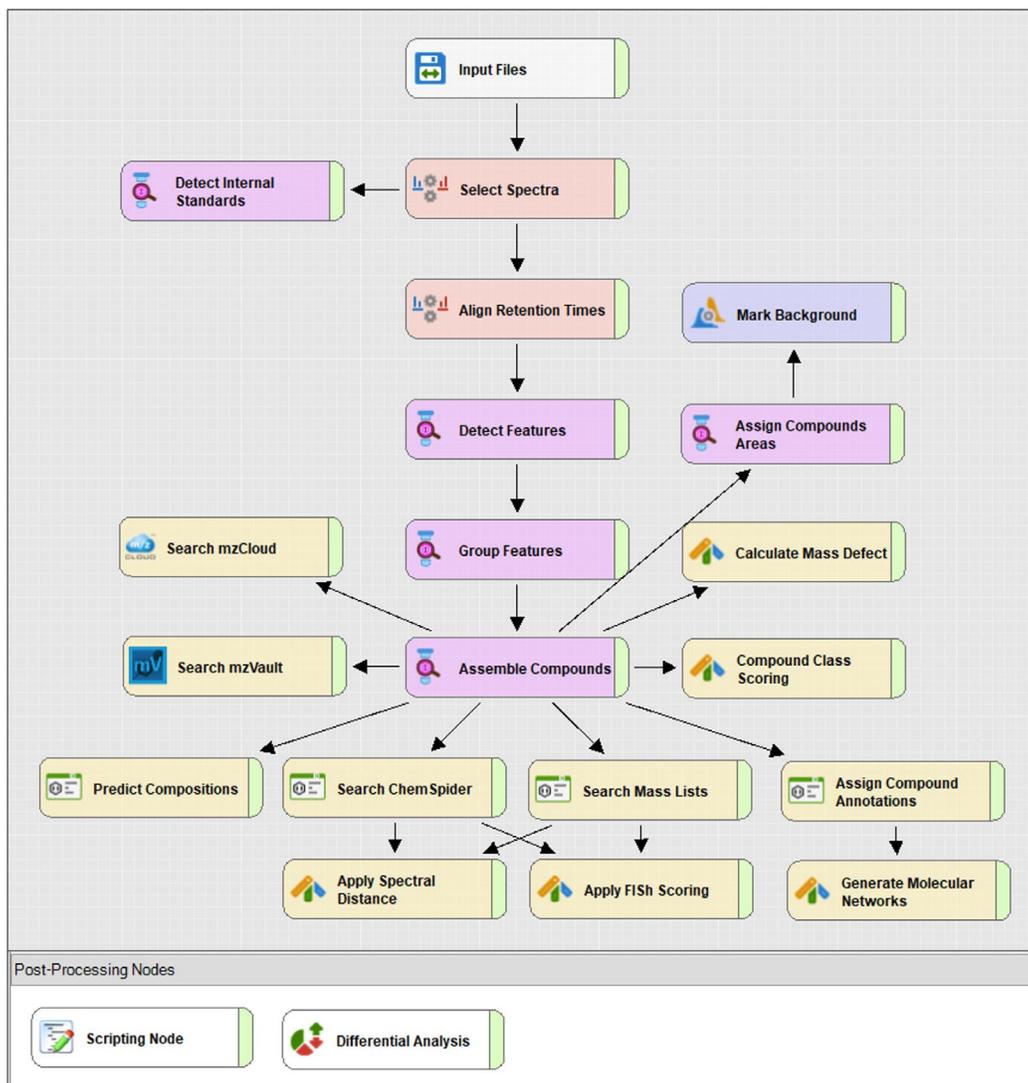
Parameter	Value
Spray voltage	-1,000 V
Vaporizer temperature	300 °C
Ion transfer tube temperature	300 °C
Sheath gas	56 a.u.
Aux gas	10 a.u.
Sweep gas	1 a.u.

**Table 3. MS parameters for full MS-ddMS<sup>2</sup> experiments**

MS parameter	Value
<b>Full Scan</b>	
AGC target	Standard (1e6)
Full Scan (MS <sup>1</sup> ) resolution	240,000 (FWHM at <i>m/z</i> 200)
Full Scan (MS <sup>1</sup> ) mass range	<i>m/z</i> 110–1,200
Lock mass correction	EASY-IC source, scan-to-scan, full scan only
Mild trapping mode	On
RF level (%)	55
<b>ddMS<sup>2</sup> Scan</b>	
Isolation window ( <i>m/z</i> )	1.0
HCD collision energies (normalized, %)	5, 30, 60
MS <sup>2</sup> resolution	30,000 (FWHM at <i>m/z</i> 200)
Maximum injection time	Auto
Intensity filter threshold	50,000 cps
Dynamic exclusion filter	n=3; 6 s exclusion; exclude isotopes

## Data processing and analysis

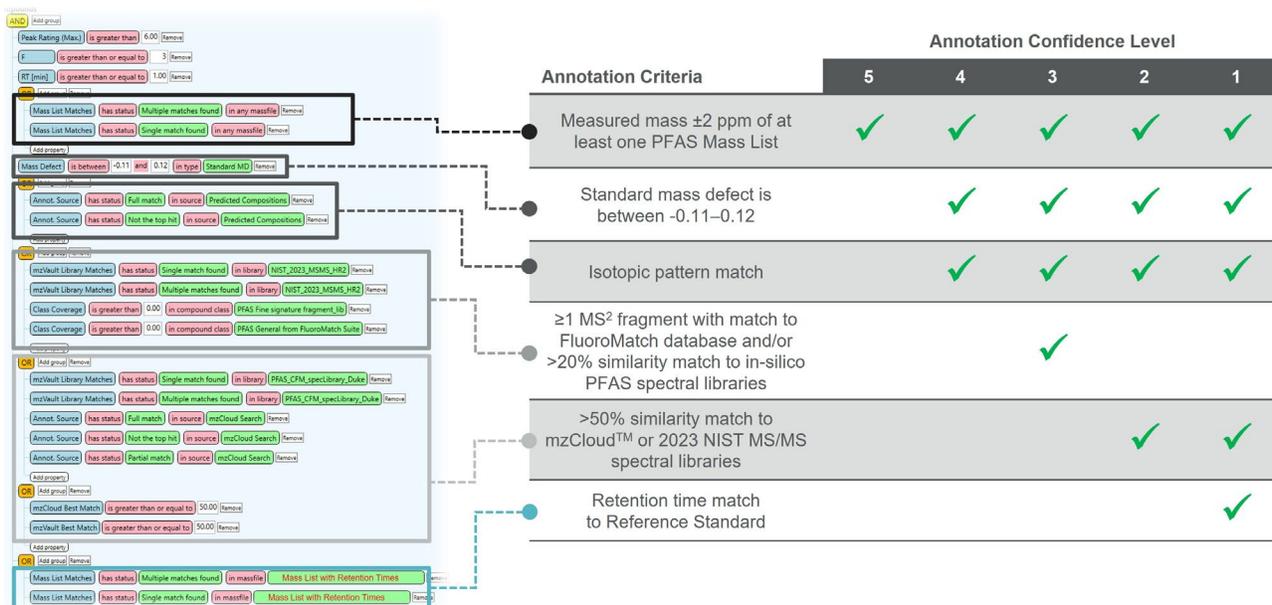
LC-MS data were automatically processed in Compound Discoverer 3.4 software using the workflow shown in Figure 1. The workflow was built in the workflow editor starting with the “PFAS Unknown ID” pre-built template that is supplied with the software. The workflow uses multiple spectral libraries and other resources to automatically process data (Figure 1). A Scripting Node with the “PFAS.R” script is included to calculate values pertaining to the Orthogonal MS<sup>1</sup> data filtering approach developed by Kaufmann, et al.<sup>6</sup> Additional information, including the PFAS.R script, can be downloaded from [mycompounddiscoverer.com](http://mycompounddiscoverer.com). Automatic data processing produces a results file that contains a list of compounds that have been detected through the various workflow nodes shown in Figure 1.



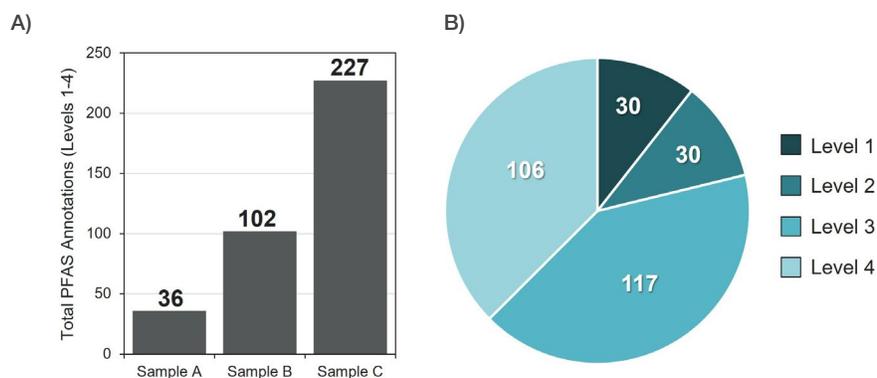
**Figure 1. Compound Discoverer software Workflow Editor showing the steps used to automatically process data files acquired from LC-MS analysis of the study samples.** The workflow was built from the template “PFAS Unknown ID”, which is supplied with the software. The workflow uses multiple spectral fragment libraries, compound databases, PFAS fragmental libraries, elemental formula determinations, and more to annotate compounds. The workflow also includes the post-processing Scripting Node with the “PFAS.R” script to calculate values for the Orthogonal MS<sup>1</sup> data filtering approach developed by Kaufmann, et al.<sup>6</sup> This approach can aid in the identification of certain PFAS. Version 3.4 of Compound Discoverer software also adds a node for detecting isotopically labeled standards, which is useful when analyzing samples that have already been measured using targeted quantification methods which rely on isotopically labeled internal standards (i.e., EPA and ASTM methods).

The initial compounds list generated by Compound Discoverer software includes all compounds it detected, including non-PFAS compounds. This list is filtered so only PFAS are included. Next, the list of PFAS must be evaluated to either approve the annotations automatically assigned or, in some cases, to assign alternative annotations from the resources used. At the same time, annotation confidence levels are assigned. All data reduction and organization steps are performed using the Data Filtering and Tagging features in Compound Discoverer software. As outlined in Figure 2, an annotation confidence level from one to four was manually assigned to each detected compound based on the resources (i.e., library, database, etc.) used to identify it per the approach described by Schymanski, et al.<sup>1</sup> and Charbonnet, et al.<sup>4</sup> Level 1 annotations required all annotation criteria be met (except a match to the Duke University *in silico*-generated PFAS spectral library or the FluoroMatch™ PFAS fragment database), including a retention time match to a reference standard analyzed on the same instrumentation. The

reference standard retention times were in a custom Mass List built and stored in Compound Discoverer software. Level 2 annotations required a match to a reference spectral library, either the Thermo Scientific™ mzCloud™ advanced mass spectral database or 2023 NIST Tandem Mass Spectral Library. Level 4 was a suspect hit without a retention time or spectral library match but with a full match to a PFAS database (either the NIST Suspect List, EPA PFAS Structure List, or the Duke University *in silico* PFAS database) with a measured mass and isotope pattern match for the molecular formula of the suspect PFAS. If the Level 4 PFAS also contained MS<sup>2</sup> fragment information, then a match of at least one measured fragment to the Duke University *in silico* PFAS spectral library<sup>7</sup> or the FluoroMatch PFAS fragment database<sup>8</sup> moved its annotation confidence to Level 3. For complex samples like Samples B and C, applying additional resources such as the Duke University *in silico* PFAS spectral library and FluoroMatch PFAS fragment database greatly expands the scope of PFAS annotation.



**Figure 2.** Detected compounds in the samples were filtered using the Data Filter tool in Compound Discoverer software (revision 3.4). Columns in the Compounds Table (Table 4 provides an example) can be used as variables by which the results can be filtered. Left: filters used to assign Level 1 annotation confidence. A filter set designed for Level 2 annotations would be the same except that the last two filter properties (highlighted in teal) would be removed. Right: relationship between the annotation criteria and confidence levels. The filter sets are available at [mycompounddiscoverer.com](http://mycompounddiscoverer.com).



**Figure 3.** (A) Total number of PFAS detected and annotated at confidence Levels 1–4, per sample; (B) Distribution of all PFAS compounds, across all samples, per annotation confidence Level 1-4

## Results and discussion

### Automated PFAS annotation and assignment of confidence levels

The workflow described here combines the sub-1 ppm mass accuracy and ultra-high mass resolution of the Orbitrap Exploris 240 mass spectrometer with automated data processing and a comprehensive set of PFAS-specific databases, spectral libraries, and *in silico* fragmentation tools in Compound Discoverer 3.4 software to provide confident PFAS annotation. Compound Discoverer software processes the data files first to detect and group mass spectral features, and then assembles these features into individual compounds by combining the features that are adducts and isotopes of the same compound. In addition, Compound Discoverer 3.4 software leverages the mzCloud advanced mass spectral database to identify in-source

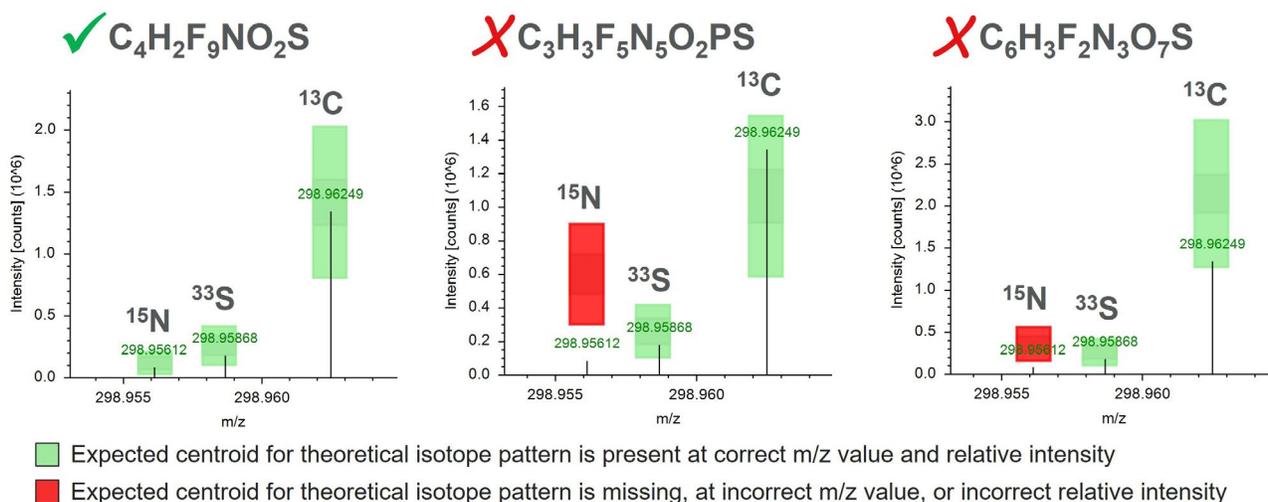
fragments that are then also grouped with the adducts and isotopes of the assembled compound. MS<sup>1</sup> and MS<sup>2</sup> spectra that were acquired closest to assigned retention times of the compounds are also linked to the assembled compound. Following detection and assembly, the compounds are annotated using libraries and databases the user specified in the workflow. This comprehensive, automated workflow enabled annotation of more than 250 PFAS across the three samples analyzed in this study. Figure 3 summarizes the total number of PFAS detected in each sample and the distribution of assigned annotation confidence levels. Due to the ultra-high mass accuracy and resolution of the Orbitrap Exploris 240 mass spectrometer, any detected feature with a monoisotopic mass not matching any of the three Mass Lists had an extremely low probability of being a PFAS. Therefore, only Levels 1–4 are shown.

As shown in Figure 2, Levels 1-4 annotations require confident assignment of elemental formula through both high mass accuracy and full match between the measured and theoretical isotope pattern. For PFAS with heteroatoms that contain multiple naturally occurring isotopes (e.g., N, S, O, etc.), the ability to fully resolve all isotopes is critical in confidently assigning the correct empirical formula. In this work we used the ultra-high mass resolution of the Orbitrap Exploris 240 mass spectrometer, collecting all MS<sup>1</sup> spectra at 240,000 resolution (FWHM at *m/z* 200), enabling fully resolved measurements of the <sup>15</sup>N, <sup>33</sup>S, and <sup>13</sup>C isotopes. With the Predict Composition node, all empirical formulas within 2 ppm mass error were determined. For each formula, the measured and theoretical isotope patterns are compared, providing pattern coverage scores and also annotating the MS<sup>1</sup> spectrum with the position of expected isotopes. Figure 4 shows an example of comparing the measured A+1 isotope pattern against the top three formulas based on Pattern Coverage scores. While all three formulas contain

nitrogen and sulfur, the measured relative intensities of the <sup>15</sup>N and <sup>33</sup>S match to only one of the formulas. This fine isotope structure analysis of <sup>15</sup>N and <sup>33</sup>S isotopes is not possible at lower mass resolutions.

The top 40 PFAS detected in each sample are shown in Tables 4, 5, and 6 with annotation Levels 1-4, with the exception of Sample A, which only contained 36 PFAS. These tables are screenshots of the Compounds table in Compound Discoverer 3.4 software, which offers a simple, easy-to-navigate interface that combines all the information from each resource for every detected feature and includes all the compound annotations that were automatically assigned. The confidence levels listed in the “Tags” column were determined using the criteria shown in Figure 2 and the “Comments” column shows the NIST Suspect ID for each PFAS, if available. Annotations with Level 4 confidence used the following mass list priority: NIST Suspect List, EPA PFAS Structure List, and Duke University *in silico* PFAS spectral library.

Tags	Comments	Name	Formula	RT [min]	Annot. ΔMass [ppm]	Calc. MW	<i>m/z</i>
L4* L3 L2 L1							
○ ○ ○ ○	9	Perfluorobutylsulfonamide	C4 H2 F9 N O2 S	3.259	0.30	298.96634	297.95907



Formula	# Matched Iso.	# Missed Iso.	Pattern Cov. [%]	MSMS Cov. [%]	Molecular Weight	ΔMass [Da]	ΔMass [ppm]	In ChemSpider
C4 H2 F9 N O2 S	7	0	100.00	100.00	298.96625	0.00009	0.30	X
C3 H3 F5 N5 O2 P S	6	1	98.34	100.00	298.96652	-0.00018	-0.60	
C6 H3 F2 N3 O7 S	6	2	97.78	100.00	298.96598	0.00037	1.23	

**Figure 4. Example of using the measured A+1 fine isotope pattern to accurately predict the elemental formula for perfluorobutylsulfonamide.** As the user selects a possible formula within the Predicted Compositions table, Compound Discoverer software annotates the correct position of the expected isotope pattern, providing a simple visualization of how well the measured and theoretical isotope patterns match. A minimum mass resolution of 240,000 (FWHM at *m/z* 200) is required to resolve each of the three <sup>12</sup>C, <sup>15</sup>N, and <sup>33</sup>S isotopes.





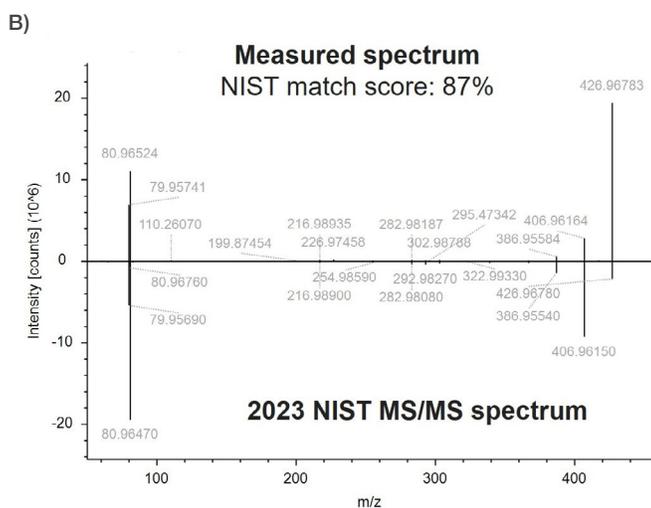
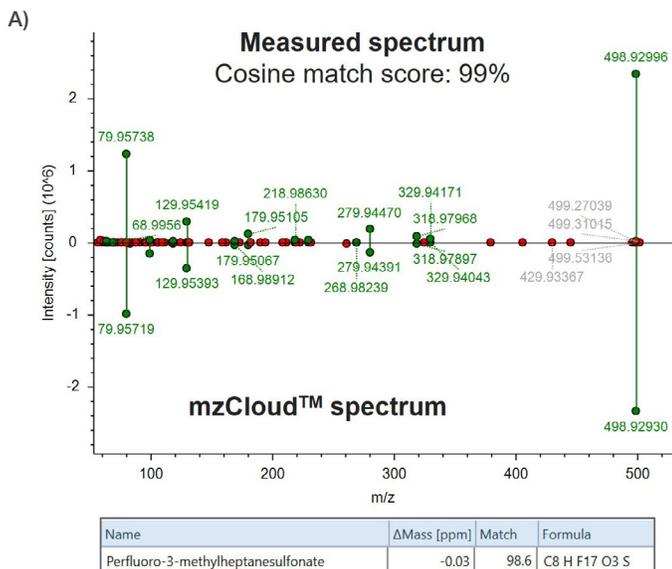
**Table 6. List of top 40 PFAS detected in Sample C with their annotation confidence level displayed in the Compounds table of the Compound Discoverer software interface.** The confidence levels listed in the “Tags” column were determined using the criteria in Figure 2. The list is sorted based on the integrated peak areas of the sample. The “Comments” column shows the NIST Suspect ID for each PFAS, if available.

Tags	Comments	Name	Formula	RT [min]	Calc. MW	m/z	Annot. ΔMass [ppm]	MS2	mzCloud Best Match	mzVault Library Matches NIST 2023 NIST MSMS HRMS NIST CFM Spectral Library Date	Class Coverage PFAS General from FluoroMatch Suite*	Mass List Matches PFAS Fine Signature Fragment Lib PFAS General from FluoroMatch Suite PFAS Branched and Linear Mass List v4 Chemical List PFASSTRACK-2022-04-20 PFAS SuspectDB Date	Mass Defect Standard MDs Relative MD	Annot. Source Predicted Compositions* mzCloud Search ChemSpider Search MassBank Search	Group Areas			
															a b c			
○○○●		3041 Perfluorooctanesulfonic acid	C8 H F17 O3 S	7.046	499.93700	498.92972	-1.00	■	99.8	■	0.87	37.50	-0.063	-126	-0.031	2.70e9	3.52e9	1.73e9
○○○●		2560 Perfluorohexanesulfonamide	C6 H2 F13 N O2 S	5.784	398.95937	397.95210	-1.23	■	94.6	■	0.75	31.25	-0.041	-102	-0.015	4.59e6	4.49e6	
○○○●		3041 Perfluorooctanesulfonic acid	C8 H F17 O3 S	6.624	499.93710	498.92982	-0.79	■	99.8	■	1.49	37.50	-0.063	-126	-0.031	1.41e9	3.62e9	
○○○●		3043 Perfluorohexanesulfonic acid	C6 H F13 O3 S	4.581	399.94352	398.93624	-0.91	■	99.9	■	0.87	37.50	-0.056	-141	-0.031	2.85e9	5.98e9	3.19e9
○○○●		878 Perfluorohexane sulfonamido amine	C11 H13 F13 N2 O2 S	5.718	484.04869	483.04142	-0.66	■	98.2	■	1.24	25.00	0.049	101	0.008	1.22e9	8.34e8	2.79e8
○○○●		756 N-Ethylperfluorooctanesulfonamide	C10 H6 F17 N O2 S	9.065	526.98445	525.97717	-0.62	■	59.8	■	1.37	25.00	-0.016	-30	0.018			2.29e9
○○○●		2559 Perfluorooctanesulfonamide	C8 H2 F17 N O2 S	8.412	498.95315	497.94587	-0.66	■	99.8	■	0.50	25.00	-0.047	-94	-0.015			1.27e9
○○○●		Perfluoro-N-(hexanesulfonyl)hexane-1-sulfonamide	C12 H F26 N O4 S2	10.160	780.89302	779.88574	-0.21	■		■	0.12	6.25	-0.107	-137	-0.031			2.75e6
○○○●		3039 Perfluorodecanesulfonic acid	C10 H F21 O3 S	9.285	599.93086	598.92359	-0.41	■	99.4	■	0.99	25.00	-0.069	-111	-0.031			4.37e6
○○○●		2923 3:2 Fluorotelomer alcohol	C5 H5 F7 O	2.190	214.02266	213.01539	-0.93	■		■	0.50	6.25	0.023	106	0.036			6.27e6
○○○●		2560 Perfluorohexanesulfonamide	C6 H2 F13 N O2 S	5.475	398.95950	397.95222	-0.93	■	94.7	■	0.99	37.50	-0.041	-102	-0.015			5.16e6
○○○●		708 N-Ethyl-N-((tridecafluorohexyl)sulfonyl)glycine	C10 H8 F13 N O4 S	8.007	484.99639	483.98912	-0.52	■		■	1.24	37.50	-0.004	-7	0.027			4.70e6
○○○●		3043 Perfluorohexanesulfonic acid	C6 H F13 O3 S	4.330	399.94364	398.93637	-0.59	■	99.5	■	1.49	50.00	-0.056	-149	-0.031			1.45e6
○○○●		3040 Perfluorononanesulfonic acid	C9 H F19 O3 S	8.225	549.93410	548.92682	-0.37	■	98.9	■	1.12	37.50	-0.066	-120	-0.031			3.79e6
○○○●		2632 Perfluorodecanoic acid	C10 H F19 O2	8.231	513.96692	512.95964	-0.78	■	64.3	■	0.87	25.00	-0.033	-64	0			1.79e5
○○○●		3039 Perfluorodecanesulfonic acid	C10 H F21 O3 S	8.829	599.93106	598.92379	-0.07	■	98.2	■	1.37	31.25	-0.069	-115	-0.031			4.59e5
○○○●		Perfluoro-N-(hexanesulfonyl)hexane-1-sulfonamide	C12 H F26 N O4 S2	9.988	780.89320	779.88593	0.03	■		■	0.25	6.25	-0.107	-137	-0.051			3.47e6
○○○●		3487 Unsaturated perfluorododecane sulfonate	C12 H F23 O3 S	9.600	661.92795	660.92067	0.06	■		■	2.48	19.00	-0.072	-109	-0.031			3.37e6
○○○●		3494 perfluorotridecane sulfonate:perfluorotridecane sulfonate	C13 H F27 O3 S	10.992	749.92164	748.91437	0.16	■		■	2.36	18.75	-0.078	-104	-0.031			3.26e6
○○○●		3325 perfluoropentane sulfonamide	C5 H2 F11 N O2 S	4.454	348.96284	347.95557	-0.62	■		■	0.62	25.00	-0.037	-106	-0.015			3.07e6
○○○●		2559 Perfluorooctanesulfonamide	C8 H2 F17 N O2 S	7.913	498.95332	497.94604	-0.31	■	99.1	■	1.24	56.25	-0.047	-94	-0.015			2.91e6
○○○●		2637 Perfluorooctanoic acid	C8 H F15 O2	5.753	413.97332	412.96604	-0.92	■	66.1	■	0.75	18.75	-0.027	-64	0			1.80e9
○○○●		955 1-Hexanesulfonamide, 1,1,2,2,3,3,4,4,5,5,6,6,6-tridecafluoro-N-	C7 H4 F13 N O2 S	7.837	412.97523	411.96796	-0.68	■		■	1.12	31.25	-0.025	-60	0.002			2.69e6
○○○●		3042 Perfluoroheptanesulfonic acid	C7 H F15 O3 S	5.801	449.94046	448.93318	-0.51	■	97.5	■	0.75	31.25	-0.06	-132	-0.031			7.34e7
○○○●		3044 Perfluoropentanesulfonic acid	C5 H F11 O3 S	3.560	349.94685	348.93958	-0.64	■	96.0	■	0.75	25.00	-0.033	-152	-0.031			6.68e5
○○○●		9 Perfluorobutylsulfonamide	C4 H2 F9 N O2 S	3.259	298.96634	297.95907	0.30	■	98.1	■	0.75	25.00	-0.034	-113	-0.015			1.43e6
○○○●		2643 Perfluorohexanoic acid	C6 H F11 O2	3.488	313.98014	312.97286	0.17	■	79.7	■	0.62	12.50	-0.02	-63	0			1.58e9
○○○●		1054 2-(Heptafluoropropyl)-1,3,5-trimethoxybenzene	C12 H11 F7 O3	3.320	336.05950	335.05223	-0.41	■	81.4	■	0.25	0.00	0.06	177	0.081			2.38e6
○○○●		3493 perfluoroundecane sulfonate	C11 H F23 O3 S	9.744	649.92791	648.92063	0.00	■	99.2	■	1.86	37.50	-0.072	-111	-0.031			2.31e6
○○○●		3495 perfluorotetradecane sulfonate	C14 H F29 O3 S	11.299	799.91871	798.91144	0.48	■		■	2.73	12.50	-0.081	-102	-0.031			2.04e6
○○○●		3493 perfluoroundecane sulfonate	C11 H F23 O3 S	9.520	649.92797	648.92069	0.09	■	98.7	■	2.36	25.00	-0.072	-111	-0.031			1.99e6
○○○●		2630 Perfluoroundecanoic acid	C11 H F21 O2	9.317	563.96398	562.95671	-0.25	■	55.2	■	0.99	25.00	-0.036	-64	0			1.20e6
○○○●		3038 Perfluorododecanesulfonic acid	C12 H F25 O3 S	10.963	699.92478	698.91750	0.08	■	98.3	■	1.99	43.75	-0.075	-107	-0.031			1.91e6
○○○●		3940 perfluoroheptane sulfonamide	C7 H2 F15 N O2 S	7.141	448.95643	447.94916	-0.53	■	99.9	■	0.75	31.25	-0.044	-97	-0.015			1.91e6
○○○●		3940 Perfluorononanesulfonic acid	C9 H F19 O3 S	7.732	549.93416	548.92688	-0.26	■	99.3	■	1.74	50.00	-0.066	-120	-0.031			4.46e6
○○○●		Bis(perfluoropentanesulfonyl)amine	C10 H F22 N O4 S2	8.712	680.89957	679.89229	0.00	■		■	0.25	12.50	-0.1	-147	-0.051			1.85e6
○○○●		3040 Perfluorononanesulfonic acid	C9 H F19 O3 S	7.873	549.93416	548.92688	-0.26	■	99.4	■	1.49	37.50	-0.066	-120	-0.031			1.29e6
○○○●		4484 perfluorotetradecane ether sulfonate	C14 H F29 O4 S	11.463	815.91363	814.90636	0.47	■		■	1.61	18.75	-0.086	-106	-0.034			1.42e6
○○○●		876 Perfluorooctane sulfonamido amine	C13 H13 F17 N2 O2 S	8.411	584.04258	583.03531	-0.07	■		■	1.74	43.75	0.043	73	0.008			1.22e7
○○○●		4369 Unsaturated perfluoro ether heptadecane sulfonic acid	C16 H F31 O4 S	11.822	877.91056	876.90328	0.57	■		■	3.73	31.25	-0.089	-102	-0.033			1.40e6

## Spectral databases enhance annotation confidence

The Compound Discoverer 3.4 software workflow described here leverages the upgraded mzCloud advanced mass spectral database that currently contains MS<sup>n</sup> spectra for more than 120 PFAS and the 2023 NIST Tandem Mass Spectral Library, enabling annotation of hundreds of PFAS compounds with Level 2 confidence or better. Figure 5 shows an example of using these libraries to identify a branched isomer of perfluorooctane sulfonate (with a 99% Cosine match score to a spectrum in the mzCloud advanced mass spectral database) and 6:2 fluorotelomer sulfonic acid (with an 87% NIST match

score to a spectrum in the 2023 NIST Tandem Mass Spectral Library). As previously mentioned, the Duke University *in silico* PFAS spectral library is also valuable for identifying less common PFAS that are not in other spectral libraries. In this case, more than fifteen PFAS were annotated with Level 3 confidence, including the 10:2 fluorotelomer sulfonate shown in Figure 6, using the Duke University library. Match scores for comparisons of measured MS<sup>2</sup> spectra to the Duke University *in silico* PFAS spectral library are also calculated using the NIST search algorithm.



Name	$\Delta$ Mass [ppm]	Match	mzVault Library
3,3,4,4,5,5,6,6,7,7,8,8,8-Tridecafluoro-1-octanesulfonic acid	-0.97	87.3	NIST_2023_MSMS_HR2

Figure 5. Example of using reference spectral libraries to achieve Level 2 annotation confidence. (A) Identification of perfluoro-3-methylheptanesulfonate (branched isomer of perfluorooctane sulfonate) using the mzCloud advanced mass spectral database, and (B) identification of 6:2 fluorotelomer sulfonic acid using the 2023 NIST Tandem Mass Spectral Library.

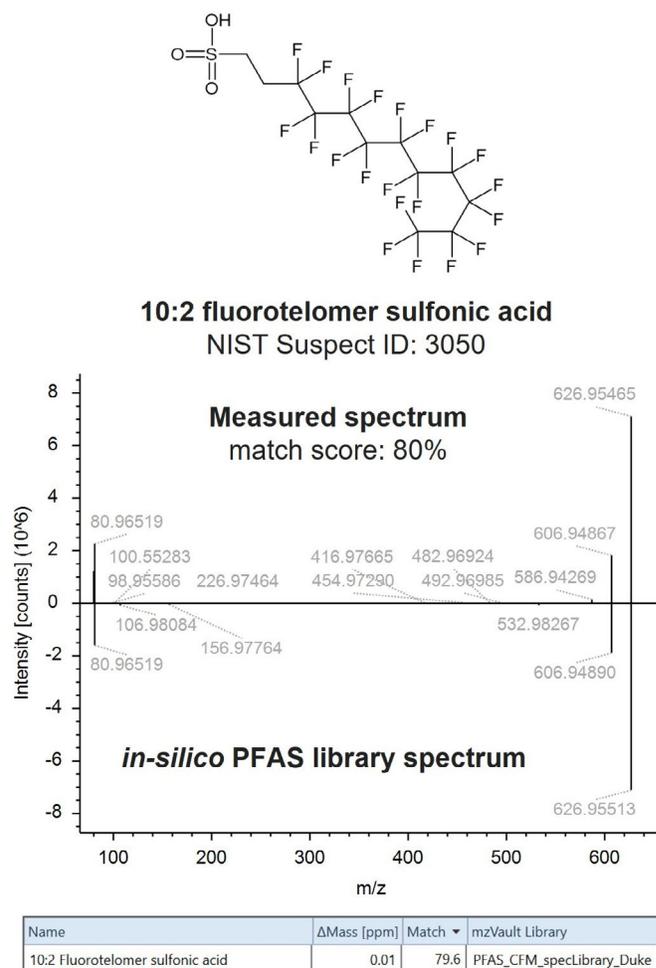


Figure 6. Example that applies the Duke University *in silico* PFAS spectral library to achieve Level 3 annotation confidence for 10:2 fluorotelomer sulfonic acid

Additional Level 3 annotations of less common PFAS not included in reference spectral libraries were possible using the FluoroMatch PFAS fragment database. The FluoroMatch PFAS fragment database is applied in the Compound Class Scoring processing node (Figure 1), which calculates the percentage of fragments in the measured MS<sup>2</sup> spectrum that match the fragments in the FluoroMatch database (within ±5 ppm). Figure 7 shows an example of how the FluoroMatch PFAS fragment database was applied to annotate the PFAS in a homologous series with Level 3 confidence. The series in Figure 7 shows three PFAS, including two isomers of N-[3-(dimethylamino)propyl]-N-[(tridecafluorohexyl)sulfonyl]-beta-alanine.

### Visualization of PFAS results adds insight

In addition to PFAS detection and annotation, Compound Discoverer 3.4 software includes multiple data visualization and statistical comparison options that provide insights about the PFAS composition of samples. For example, the mass defect plotting tool can facilitate the identification of homologous series of PFAS. Figure 8 shows the Kendrick mass defect plot for all PFAS detected and annotated in this study, with the Kendrick formula set to “CF2.” PFAS that are in the same homologous series exhibit the same Kendrick mass defect. An example of a homologous series of bis-perfluoro-N-alkyl sulfonamides

(bis-FASI) that is detected and identified in the AFFF-impacted soil is highlighted in red in Figure 8.

The Orthogonal MS<sup>1</sup> plot was developed by Kaufmann, et al<sup>6</sup> as an alternative for identifying PFAS that contain heteroatoms with negative mass defects (i.e., S, Cl, and Br). In comparison to the standard mass defect plot, the Orthogonal MS<sup>1</sup> plot better distinguishes perfluorinated compounds from non-PFAS organic compounds. The PFAS NTA workflow in Compound Discoverer 3.4 software uses the Scripting mode (Figure 1) with the “PFAS.R” script to automatically calculate the values needed for the plot:

- Estimated number of carbons (eC; based on the observed A1/A0 ratio in the apex MS<sup>1</sup> spectrum of the feature)
- Molecular mass divided by eC (m/C)
- Mass defect divided by eC (md/C)

These values, and the number of fluorine atoms in the assigned formula, are recorded in the Compounds table associated with the results file. Using the Compound Discoverer software Result Chart tool, an Orthogonal MS<sup>1</sup> plot was generated from the calculated md/C and m/C values (Figure 9). Because the m/C and md/C values are calculated based on experimental data rather than the predicted formula, the Kaufmann Plot here slightly differs from those observed in previous studies.<sup>9</sup>

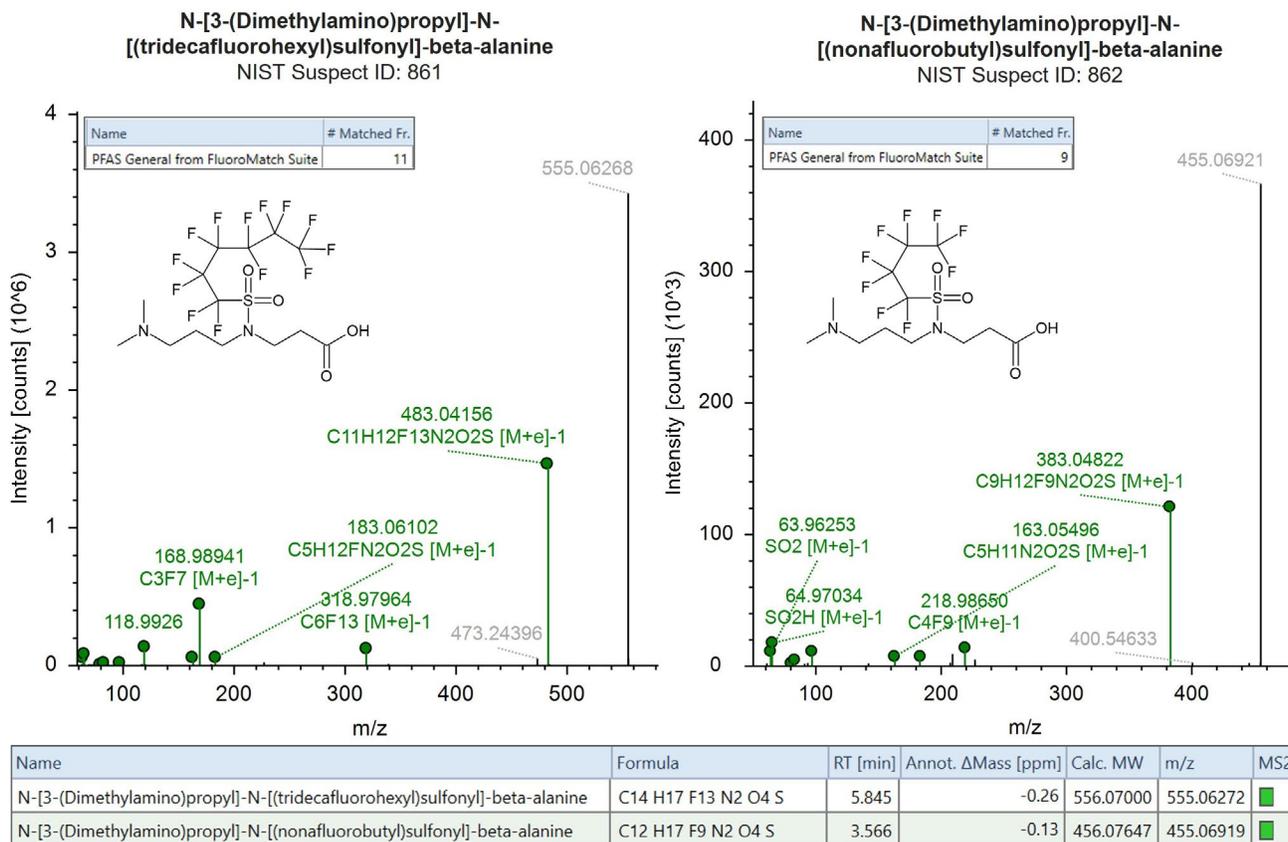
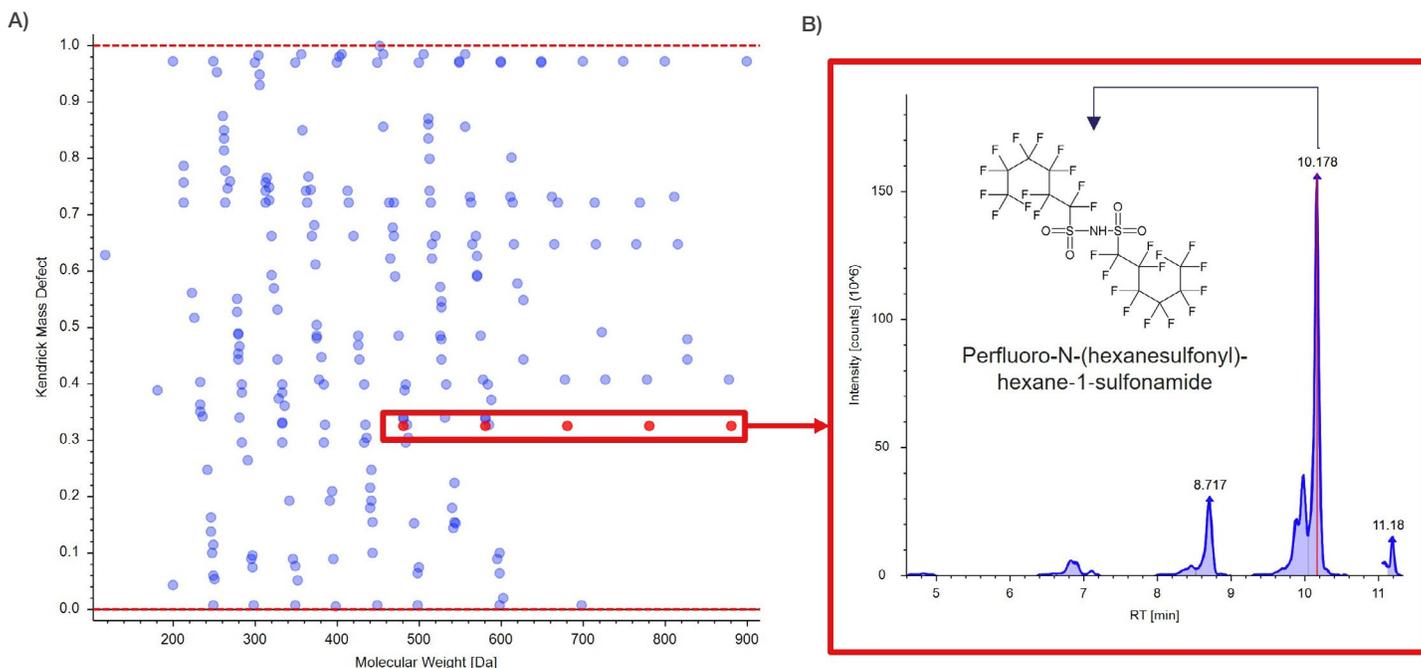
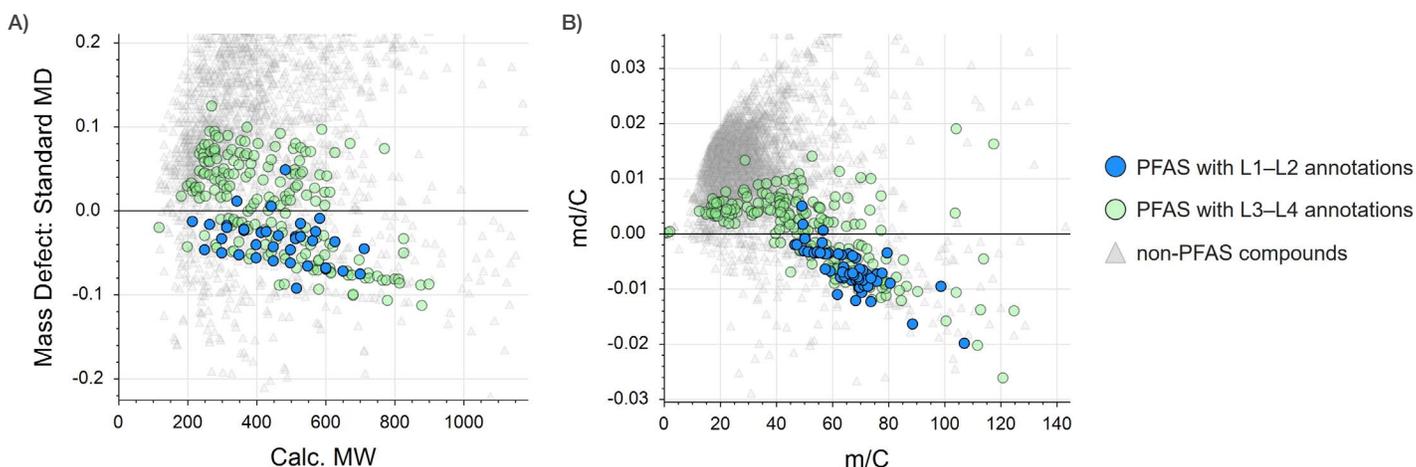


Figure 7. Example application of the FluoroMatch PFAS fragment database to identify compounds in a homologous series with Level 3 annotation confidence. Compound Discoverer software automatically highlights all fragment matches with those in the FluoroMatch PFAS fragment database (in green) and annotates the fragment's formula.



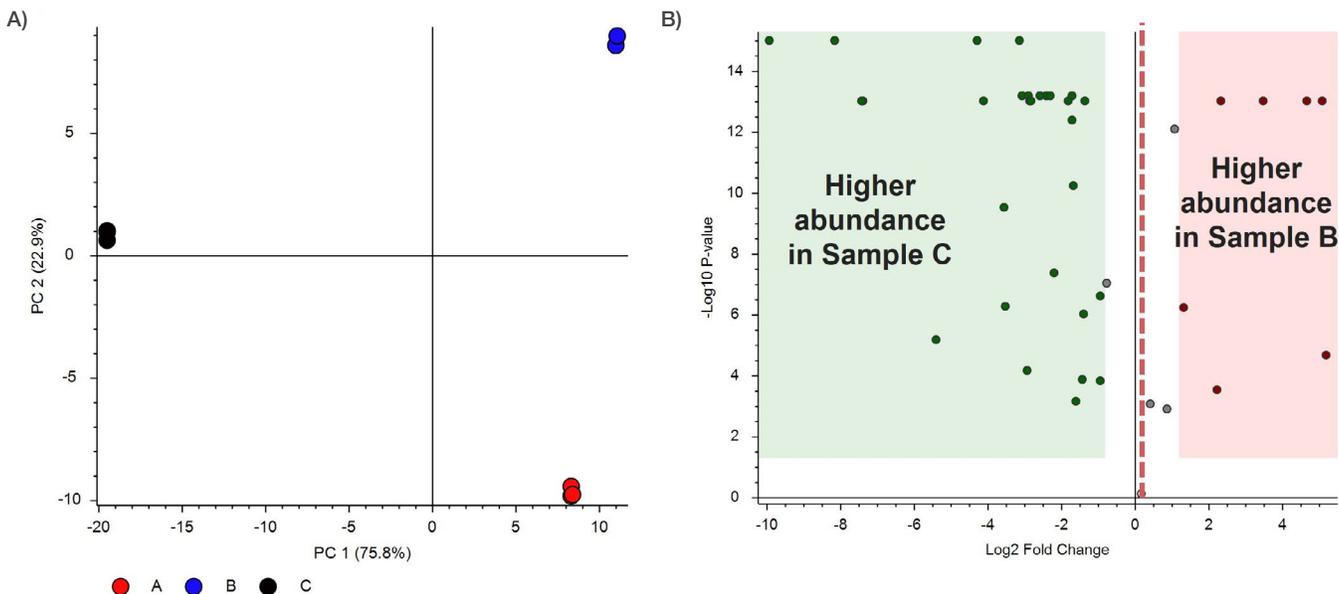
**Figure 8. (A) Kendrick Mass Defect plot of study data created using the Compound Discoverer software mass defect plotting tool with the Kendrick formula set to “CF2”. Data points belonging to a specific homologous series of bis-perfluoro-N-alkyl sulfonamides (bis-FAS) are highlighted in red. (B) Overlaid extracted chromatograms of the series in the AFFF-impacted soil sample (Sample C).**



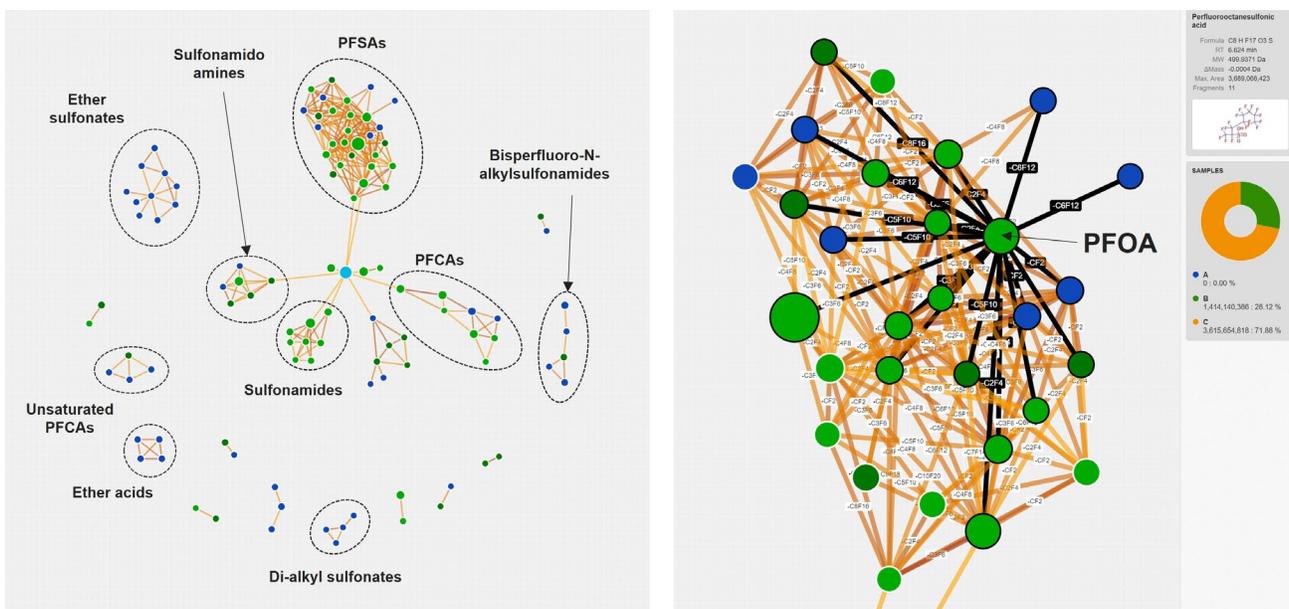
**Figure 9. (A) Standard Mass Defect plot and (B) Orthogonal MS<sup>1</sup> plot created in Compound Discoverer software using the Result Chart tool. The Standard Mass Defect is calculated using the Calculate Mass Defect node and the “Calc. MW” is the calculated molecular weight for the neutral compound. The “md/C” and “m/C” values (B) were calculated using the PFAS.R script in the Scripting Node step of the processing workflow and are published in the Compounds table of the Results file. All PFAS annotations with confidence level between 1–2 are highlighted in blue, while those with confidence levels between 3–4 are highlighted in green. All other non-PFAS compounds are shown as gray triangles.**

Compound Discoverer 3.4 software also provides statistical comparison tools to help users understand the differences between samples. As shown in Figure 10, the software’s principal component analysis (PCA) shows a statistically significant difference in the PFAS composition of all three samples. The differential analysis plot was applied to the AFFF mix and AFFF-impacted soil sample, with each data feature plotted by log fold change versus probability. Investigation of each data feature in the region of the plot with Log<sub>2</sub> Fold change > 1 and -Log p-value > 1.3 revealed that the majority of the PFAS

that were more abundant in Sample C compared to Sample B are a series of C<sub>4</sub>–C<sub>8</sub> perfluorocarboxylic acids, C<sub>6</sub>–C<sub>10</sub> perfluorosulfonic acids, and perfluoroalkyl sulfonamides. Previously, these compounds had been observed in electrochemical fluorination AFFF products and found to persist in AFFF-impacted soils.<sup>10</sup> More notably, the AFFF-impacted soil contained 125 PFAS that were not detected in the AFFF mixture, leading to the significant difference in overall composition between the two samples shown in the PCA plot (Figure 10A).



**Figure 10.** Statistical analysis plots were used to evaluate the PFAS composition of the three samples analyzed. The PCA plot (A) shows significant differences between the PFAS compositions of the samples. Differential analysis of Samples B and C (B) revealed more than 20 PFAS compounds at higher levels in Sample C. Further investigation determined that most of those compounds are part of a series of perfluorocarboxylic acids and perfluorosulfonic acids.



**Figure 11.** Comprehensive Molecular Network diagram created in Compound Discoverer software using the FluoroMatch PFAS fragment database and the general PFAS fragment compound class as fragment seeds to generate clusters of structurally related compounds. The connections between nodes were based on the thresholds: “Score” = 50, “Coverage” = 70, and “Matched Fragments” = 5. The general names of the various homologous series in the individual clusters are shown in the dashed circles. (Right) close-up view of the PFASs cluster, highlighting PFOA and its link to multiple PFASs along with the differences in the linked compounds’ elemental formulas. The size of each node depicts the maximum observed peak areas across all samples. The network diagram is interactive for dynamic data visualization.

Identifying homologous series is an important step in fingerprinting the sources of PFAS in samples and determining the appropriate remediation strategies. The Molecular Networks node (Figure 1) enables the user to view structurally related PFAS by associating features that meet user-defined thresholds of matching fragments and spectral match score. The Molecular Network diagram (Figure 11) used the FluoroMatch PFAS

fragment database to cluster structurally related compounds using fragment seeds. The clustered homologous series are shown in Figure 11 by dashed circles. The close-up view of the “PFASs” cluster (Figure 11, right side) highlights perfluorooctanesulfonic acid and its links to multiple PFASs, along with the differences in the linked compounds’ elemental formulas.

## Conclusion

By accessing a comprehensive set of libraries, databases, and other PFAS-specific tools, Compound Discoverer 3.4 software provides a powerful framework that leverages the ultra-high mass accuracy and resolution of Orbitrap mass analyzer technology to consistently annotate over 40,000 PFAS with the highest confidence possible.

This application applied the PFAS NTA workflow to “unknown” samples from the NIST PFAS NTAILS.<sup>5</sup> The Schymanski<sup>1</sup> annotation confidence scale was easily implemented in Compound Discoverer software using a feature added to revision 3.4 that allows custom prioritization of spectral libraries and databases along with data filtering and compound tagging. Built-in statistical analysis and visualization tools, including PCA, differential analysis, mass defects, and molecular networks plots, facilitate comparison of PFAS sample composition.

While this application note described a workflow for environmental samples, it is as easily applied to other application areas, including metabolomics, lipidomics, clinical, pharmaceutical, and food safety analyses to name a few. For example, the PFAS NTA workflow (using Compound Discoverer revision 3.3 SP3 software) has been applied to identify unknown PFAS in food contact materials.<sup>11</sup>

## Resources

The resources used in the Compound Discoverer 3.4 software non-targeted PFAS workflow, including a step-by-step instruction guide, filter sets, tag set, PFAS.R script installation and user guide, and other relevant information, can be downloaded from [mycompounddiscoverer.com](https://mycompounddiscoverer.com/pfas-analysis/) at: <https://mycompounddiscoverer.com/pfas-analysis/>

## References

1. Schymanski, E.L.; et al. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48*, 2097–2098.
2. Wang, Z.; et al. A new OECD Definition of Per- and Polyfluoroalkyl Substances. *Environ. Sci. Technol.* **2021**, *55* (23), 15575–15578. DOI: 10.1021/acs.est.1c06896
3. BP4NTA Best Practices for Non-targeted Analysis. (Accessed March 21, 2025). <https://nontargetedanalysis.org>
4. Charbonnet, J. A.; et al. Communicating Confidence of Per- and Polyfluoroalkyl Substance Identification via High-Resolution Mass Spectrometry. *Environ. Sci. Technol. Lett.* **2022**, *9*, 473–481.
5. Place, B.J.; et al. Per- and Polyfluoroalkyl Substances – Non-Targeted Analysis Interlaboratory Study Final Report. NIST Internal Report, (2024) NIST IR 8544. **2024**. <https://doi.org/10.6028/NIST.IR.8544>
6. Kaufman, A.; et al. Simplifying Nontargeted Analysis of PFAS in Complex Food Matrixes. *J. AOAC Int.* **2022**, *105*(5), 1280–1287 <https://doi.org/10.1093/jaoacint/qsac071>
7. Getzinger, G.J.; et al. Structure Database and *In Silico* Spectral Library for Comprehensive Suspect Screening of Per- and Polyfluoroalkyl Substances (PFASs) in Environmental Media by High-resolution Mass Spectrometry. *Anal. Chem.* **2021** Feb 9;93(5):2820–2827. <https://doi.org/10.1021/acs.analchem.0c04109>
8. Innovative Omics. FluoroMatch – Covers entire non-targeted PFAS workflow. (Accessed March 24, 2025). <https://innovativeomics.com/software/fluoromatch-flow-covers-entire-pfas-workflow/>
9. Zweigle, J.; et al. Efficient PFAS prioritization in non-target HRMS data: systematic evaluation of the novel MD/C-m/C approach. *Anal. Bioanal. Chem.* **2023**, *415*(10):1791–1801. <https://doi.org/10.1007/s00216-023-04601-1>
10. Houtz, E.F.; et al. Persistence of perfluoroalkyl acid precursors in AFFF-impacted groundwater and soil. *Environ. Sci. Technol.* **2013**, *47*, 8187–8195.
11. Hu, J.; et al. Comprehensive Screening of Per- and Polyfluoroalkyl Substances (PFAS) in Food Contact Materials: Utilizing Combustion Ion Chromatography for Total Organic Fluorine (TOF) Analysis. *J. AOAC Int.* **2025**. <https://doi.org/10.1093/jaoacint/qsaf003>

Learn more at [thermofisher.com/orbitrap](https://thermofisher.com/orbitrap)  
and [thermofisher.com/compounddiscoverer](https://thermofisher.com/compounddiscoverer)