Updated Guidance for Communicating PFAS Identification Confidence with Ion Mobility Spectrometry

Anna K. Boatman[1], Jessie R. Chappel[2-4], Kaylie I. Kirkwood-Donelson [5], Jonathon F. Fleming[6], David M. Reif[6], Emma L. Schymanski[6*], Julia E. Rager[2-4*], and Erin S. Baker[1*]

1. Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, USA
2. Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, USA
3. The Institute for Environmental Health Solutions, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, USA
4. Curriculum in Toxicology and Environmental Medicine, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27514, USA
5. Immunity, Inflammation, and Disease Laboratory, National Institute of Environmental Health Sciences, Durham, North Carolina 27709, USA
6. Predictive Toxicology Branch, Division of Translational Toxicology, National Institute of Environmental Health Sciences, Durham, North Carolina 27713, USA
7. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 Avenue du Swing, L-4367, Belvaux, Luxembourg

* Correspondence: Emma L. Schymanski: emma.schymanski@uni.lu, Julia E. Rager: jrager@unc.edu, Erin S. Baker: erinmsb@unc.edu

Abstract

Over the last decade, global contamination from per- and polyfluoroalkyl substances (PFAS) has become apparent due to their detection in countless matrices worldwide, from consumer products to human blood to drinking water. As researchers implement non-targeted analyses (NTA) to more fully understand the PFAS present in the environment and human bodies, clear guidance is needed for consistent and objective reporting of the identified molecules. While confidence levels for small molecules analyzed and identified with high-resolution mass spectrometry (HRMS) have existed since 2014, unification and automation of these levels is needed due to inconsistencies in reporting and continuing innovations in analytical methods. Here, we (i) investigate current practices for confidence level reporting of PFAS identified with liquid chromatography (LC), gas chromatography (GC), and/or ion mobility spectrometry (IMS) coupled with high resolution mass spectrometry (HRMS) and (ii) propose a simple, unified confidence level guidance that incorporates both PFAS-specific attributes and IMS collision cross section (CCS) values.

Keywords

ion mobility spectrometry (IMS), non-targeted analysis (NTA), high-resolution mass spectrometry (HRMS), identification confidence, per- and polyfluoroalkyl substances (PFAS)

Synopsis

Unified and simplified requirements guide confidence level assignment in non-targeted PFAS identification efforts with ion mobility spectrometry.

Introduction

Per- and polyfluoroalkyl substances (PFAS) are a large and growing class of anthropogenic fluorinated chemicals[4-6], with estimates ranging from 4,700 chemicals on the widely cited National Institute of Standards and Technology (NIST) PFAS Suspect List[7] to over 7 million PFAS in PubChem.[8] As PFAS production, use, and monitoring expands, these numbers are likely to increase. To date, researchers have typically focused on a limited number of PFAS due to the low availability of analytical standards, but the increased use of non-targeted analyses (NTA) with high resolution mass spectrometry (HRMS) has enabled broader monitoring and novel chemical discovery.[4-6] Similarly, the use of analytical techniques beyond gas or liquid chromatography (GC or LC) coupled to HRMS is expanding. For example, numerous studies cite ion mobility spectrometry (IMS) as a promising technology for PFAS research due to its ability to separate PFAS from biomolecules,[4, 5, 9, 10] and its adoption is already growing.

When reporting NTA results, it is critical to assign a confidence level to each identified molecule.[2] Confidence levels provide context for chemical identifications and are important for transparent reporting of results,[11] interpretation of exposure and toxicity data,[12] and appropriate response in the event of an environmental release of unknown chemicals.[13] Schymanski *et al.*[2] first defined confidence levels for small molecule identifications made using HRMS in 2014 and others have subsequently adapted these levels. For example, Celma *et al*.[14] incorporated CCS values into the 2014 levels, and Charbonnet *et al.*[3] defined a PFAS-specific system, adding additional sublevels to the 2014 levels using PFAS-specific attributes. However, manual level assignments following any of these schemes are often subjective and prone to unintended user bias (often in the form of overestimation), because lower confidence identifications are often perceived as “worse” than higher confidence ones. Additionally, little guidance is available for what to do when some parameters do not meet the defined requirements.[14-16] Even our own extensive experience with PFAS NTA using LC-IMS-HRMS has frequently generated situations

where confidence levels are not defined by any existing guidance.[17, 18] Further, there are vast and growing numbers of PFAS candidates, and a rapidly increasing number of researchers performing PFAS NTA on different analytical platforms including GC-HRMS, LC-HRMS, LC-IMS-HRMS[19-23], strengthening the need for simple and objective confidence level assignment to support PFAS NTA. Here, we investigate the current reporting of PFAS NTA identifications, address identified gaps by clarifying and unifying existing guidance, and introduce a simplified checklist to help standardize confidence level assignments for PFAS identified using LC-IMS-HRMS.

Materials and Methods

To investigate current practices for confidence level reporting of PFAS identified in NTA using HRMS, information was aggregated across our team's extensive experience in this field as well as a dedicated literature review examining recently published papers worldwide. Our team's experience includes the analysis of hundreds of analytical PFAS standards and numerous environmental matrices (water, serum, plasma, whole blood, tissue, etc.).[10, 18, 24-27] For the literature review, a search was performed in PubMed (RRID: SCR_004846) on 30 September 2024 using the following search query: ((PFAS[Title/Abstract]) OR (polyfluor*[Title/Abstract]) OR (perfluor*[Title/Abstract])) AND ((non-target*[Title/Abstract]) OR (nontarget*[Title/Abstract]) OR (untarget*[Title/Abstract])). A total of 426 publications from as early as 1997 were identified through this query, with the 94 studies published in 2023 selected for manual review to assess research trends in more detail. Inclusion in the in-depth analysis required articles specifically use HRMS for PFAS analysis, yielding 38 studies in total from 2023. Information from these articles was then summarized, including platform type, chemical identification reporting approach, matrix analyzed, number of PFAS identified, and confidence level scheme cited. This information was then used to help create the guidance criteria for assigning confidence levels to PFAS identified using LC-IMS-HRMS by understanding gaps in literature and practices which need to be considered in the guidance.

Results and Discussion

Assessing Current Practices for Confidence Levels in PFAS NTA

Our literature search identified 38 papers published in 2023 using HRMS for PFAS NTA worldwide. The most common platforms were LC-Orbitrap and LC-QTOF (see **Table S1** for full breakdown of instruments and experiment types). One study used Fourier transform ion cyclotron resonance (FTICR) and no chromatography, and one study did not specify the type of HRMS instrument used. This illustrates the wide range of platforms available for PFAS NTA and the lack of transparency in reporting. Some studies analyzed PFAS exclusively, while others evaluated a broader range of pollutants that also included PFAS. The types of experiments conducted were diverse, including environmental monitoring (searching for unknown molecules in samples collected from subjects or nature), remediation studies (evaluating PFAS breakdown via intentional degradation technology treatment), industrial analysis (investigating composition of consumer products), and more. A variety of matrices were analyzed, including water (most common), blood, soil, air and analytical standards. Studies identified a wide range of PFAS (0 to 102 PFAS per study; **Figure S1**), and interestingly, more than a quarter of these studies (12) did not report confidence levels (**Figure 1**). Among the 26 studies that reported confidence levels, the 2014 levels are the most cited for PFAS identifications. Additionally, we observed a wide distribution of confidence level assignments, with Level 2 being the most frequently assigned

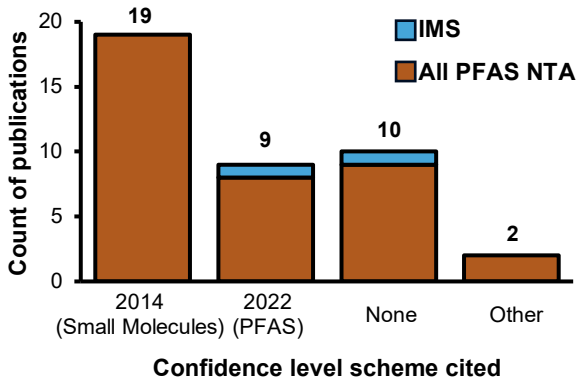

Figure 1. Confidence level scheme cited in the 38 papers from 2023. 2014 (Small Molecules): cited Schymanski *et al*. scale for small molecules. 2022 (PFAS): cited Charbonnet *et al*. scale for PFAS. None: no confidence levels were reported. Other: cited different confidence levels that do not use a 1-5 scale.

(**Figure 2**). Previous guidance by Schymanski *et al*.[2] and Celma *et al*.[14] stated that Level 2b identifications are usually rare;[2, 14] however, we found that was not the case with PFAS NTA. In fact, 8 of the 10 studies from the literature search that used sublevels for Level 2 reported Level 2b and/or 2c identifications (range: 5 to 87 PFAS at Level 2b/c PFAS). An in-depth evaluation of the methods showed that the confidence level guidelines were applied inconsistently (three such examples are shown in **Figure S2**). Examples include reported mass errors exceeding author-defined thresholds or assigned levels that contradict established guidance, *e.g*., Level 2b for structures with explicitly noted potential isomers (one exact structure is required for Level 2b; isomers possible points to Level 3). These inconsistencies complicate cross-study comparisons even across identical analytical platforms and call into question the accuracy of reported confidence levels. These inconsistencies not only hinder the interpretation of published data but also impede comparison of IMS to non-IMS results, and underscore the need for more objective, standardized confidence level assignments. Clearly, there is an opportunity to unify existing confidence level frameworks in a way that enables clear, simple and more comparable reporting as IMS is increasingly adopted in PFAS NTA.

To understand potential improvement areas for confidence levels, the 38 NTA studies were evaluated further. A noted lack of universal standards for acceptable NTA performance[15, 16, 28] was evident in the literature. For instance, mass tolerances were sometimes not defined or set as high as 20 ppm for HRMS data. Automation is one potential solution for these inconsistencies,

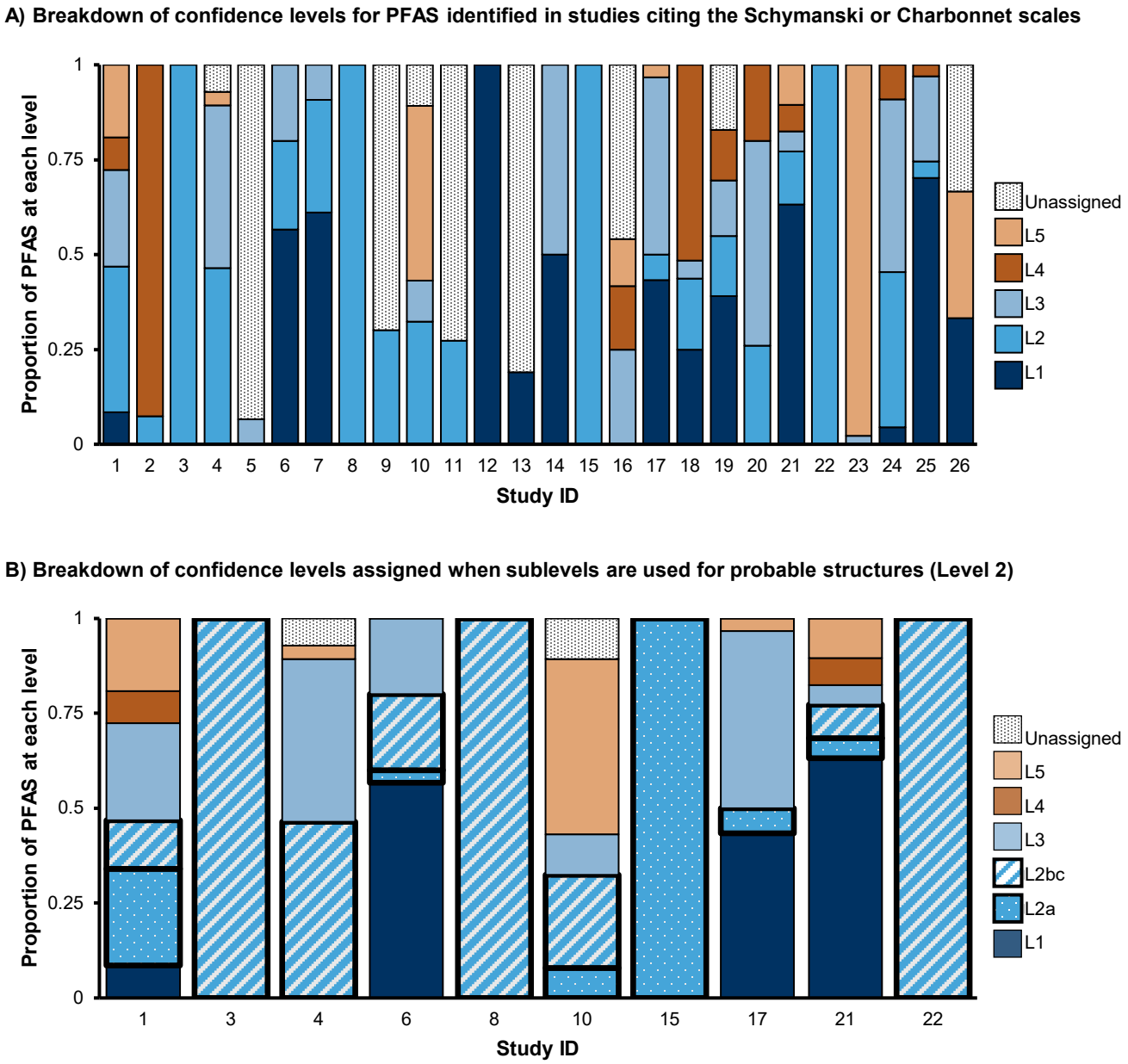

Figure 2. Distribution of confidence levels among the 26 papers that identified > 0 PFAS and cited Schymanski[2] or Charbonnet[3] scales. **A)** Percent of PFAS identified at each level by study. In several instances not all PFAS identified were assigned a confidence level, e.g. Study 9 reported 93 PFAS identified, assigned Level 2 to 28 of them, and did not assign confidence levels to the remainder. **B)** Breakdown of the 10 papers that assigned sublevels to Level 2 probable structures. 9 papers did not assign sublevels, and 8 papers did not report any Level 2 detections. Level 2 identifications are shown with thicker outlines and pattern fill (dots for 2a, diagonal stripes for 2b and 2c). The Charbonnet scale defines 2b and 2c as equivalent confidence, so these levels are combined to facilitate comparison to the Schymanski scale.

where confidence levels would be assigned objectively based on predefined criteria for mass accuracy and other dimensions. For automation to function, discrete criteria must be provided. Existing scales have attempted to define criteria for mass, CCS, and RT accuracy. However, given the wide variety of analytical techniques and modeling tools used in PFAS NTA, universally suitable acceptance tolerances cannot be defined. For example, while experimental CCS values are highly reproducible with reported literature reproducibility of ≤ 0.30% in some cases,[29] predicted CCS values are more variable. In other literature beyond the 38 studies evaluated here, one study noted machine learning tools with errors up to 8%,[25] one calculated trendlines for CCS vs *m/z* trendlines within a PFAS class with no calculated error[7] and one used confidence intervals rather than an accuracy percentage.[10, 21] All methods have utility depending on the experimental context, and defining an explicit tolerance to apply universally would be detrimental. Thus, we faced a challenge in unifying objective tolerance criteria.

Addressing Gaps in Existing Confidence Level Assignments

To address the current challenges in communicating confidence for PFAS identified using LC-IMS-HRMS our first step was to unify and simplify the existing confidence level guidance for small molecules[2], IMS[14], and PFAS[3]. An overarching observation of our literature search was an inconsistent application of current guidance, leading us to conclude that simplification and clarification of existing guidance was necessary. The updated requirements are summarized in a checklist[1] (**Figure 3**), described in detail in the following section, and provided as an extended printable document in the **Supplemental Checklist**. One key change is the removal of sublevels. Existing scales distinguish between a probable structure based on a library match (Level 2a) and one based on diagnostic evidence if no library entry exists (Level 2b). Charbonnet *et al.* also defined a Level 2c, defined as equivalent to Level 2b for probable structures elucidated using PFAS homologous series. However, in our experience analyzing PFAS using LC-IMS-HRMS, we have detected PFAS that exist in spectral libraries but not CCS libraries. We have also detected

PFAS that match our in-house CCS library, fall exactly on CCS vs *m/z* and RT vs *m/z* trendlines for a homologous series, and do not have observable or diagnostic mobility-aligned fragments. Both scenarios would not qualify as a Level 2a under current guidance because a CCS or spectral library match was absent; however, more evidence is available than required for a 2b/c identification. Additionally, IMS libraries are growing, but most spectral libraries do not include CCS values and the vast majority of PFAS in PubChem do not yet have experimental CCS values.

Of the 8,099 unique CIDs in PubChem with experimental CCS values, only 208 are PFAS according to the OECD definition implemented in PubChem[8] (12 Nov. 2024). As spectral and

LC- or GC-IMS-HRMS Confidence Level Requirement Checklist for PFAS
All boxes must be checked for PFAS identifications at each level

Confirmed Structure (Level 1)
Exactly one structure confirmed by reference standard analyzed under identical conditions

- ☐ Exactly one structure
- ☐ Reference standard analyzed in house using same analytical methods
- ☐ Exact *m/z* matches reference standard
- ☐ Isotopic pattern matches reference standard
- ☐ CCS matches reference value
- ☐ RT or RI matches reference value
- ☐ One or both of:
  - Mobility-aligned fragments match reference spectrum
  - Diagnostic evidence *e.g.* PFAS synthesis, homologues, parent-TPs

Probable Structure (Level 2)
Exactly one structure with sufficient evidence to rule out all other structural possibilities

- ☐ Exactly one structure
- ☐ Exact *m/z* matches structure
- ☐ Isotopic pattern matches structure
- ☐ CCS matches library or predicted value
- ☐ RT or RI matches library or predicted value
- ☐ One or both of:
  - Mobility-aligned fragments match library or predicted spectrum
  - Diagnostic evidence *e.g.* PFAS synthesis, homologues, parent-TPs

Tentative Candidate Structure(s) (Level 3)
One or more candidate structures satisfy all experimental evidence

- ☐ One or more structures
- ☐ Exact *m/z* matches all candidates
- ☐ Isotopic pattern supports all candidates
- ☐ CCS library or predicted value supports all candidates
- ☐ RT or RI library or predicted value supports all candidates
- ☐ One or both support all candidates:
  - Mobility-aligned fragments
  - Diagnostic evidence *e.g.* PFAS synthesis, homologues, parent-TPs

Unequivocal Formula (Level 4)
Exactly one molecular formula is possible based on all available evidence

- ☐ Exactly one possible formula
- ☐ Exact *m/z* supports formula
- ☐ Isotopic pattern supports formula
- ☐ One or both support formula:
  - Mobility-aligned fragments
  - Diagnostic evidence *e.g.* PFAS synthesis, homologues, parent-TPs

Feature of Interest (Level 5)
At least one supporting dimension of data indicating the feature is of interest

- ☐ One or more of:
  - Exact *m/z* matches a suspect list PFAS
  - Mass defect is indicative of fluorine atoms
  - CCS vs *m/z* falls in fluorinated space
  - Other justifiable attributes not listed here

Figure 3. Requirement checklist for PFAS identified using LC- or GC-IMS-HRMS. All listed criteria must be met to assign each confidence level. An extended printable version of the checklist is available in the supplemental information and online.[1]

CCS libraries grow and as analytical instrumentation advances, the gray area between a library-based identification and a diagnostic evidence-based identification will likely blur further. Therefore, to align with our goal of reducing ambiguity, we are removing the sublevel distinction in the updated confidence levels. Other key changes from current guidance are described below and summarized in **Table S2.**

Updated PFAS LC- or GC-IMS-HRMS Confidence Levels

The proposed LC- or GC-IMS-HRMS confidence levels for PFAS unify existing guidance from three sources: small molecules confidence levels introduced in 2014,[2] levels incorporating CCS values from 2020,[14] and levels specific to PFAS defined in 2022.[3] The updated levels for PFAS identified using IMS range from 1 to 5 and have no sublevels. These levels also require that acceptable tolerances for all measured dimensions in each experiment must be defined based on the capabilities of their analytical platform. To avoid selecting confidence levels based on desired results, these tolerances should be defined in advance and widely accepted as appropriate tolerances for the analytical platform in question. Platform-specific error tolerances will reduce ambiguity and enable uniform confidence level assignment as technology advances. **Table 1** provides suggested tolerances for the Agilent 6560 IMS-QTOF mass spectrometer based on our own experience with PFAS NTA and reference material analysis (available as a printable worksheet in **Table S3**).

Table 1. Example tolerance definitions for each measured dimension. These values are recommended default settings for an Agilent 6560 LC-IMS-HRMS platform and have been used for analysis of PFAS in a variety of environmental matrices.

Parameter	Match type	Example tolerance definition
Exact *m/z*	All	Measured *m/z* is ± 10 ppm of the exact *m/z*; same adducts observed as with the reference standard (where applicable)
Isotopic pattern	All	Isotopic pattern visually matches the predicted isotopic distribution, with exceptions made for low-abundance features with no M+1 or M+2 detected
CCS	Reference standard	Measured CCS is ± 0.2 % of the reference standard value
	Library value	Measured CCS is ± 2.0 % of external library value

	Predicted value	Measured CCS value is ± 5.0 % of expected value based on homologous series trend lines
RT	Reference standard	Measured RT is ± 0.5 min of in-house library values (to account for matrix effects), and must elute in same order
	Predicted value	Measured RT increases with *m/z* within homologous series

Broadly, at ***Level 1*,** the identified PFAS is confirmed by matching experimentally validated values from a reference material in all dimensions. At ***Level 2***, sufficient evidence is present that exactly one structure is probable for the identified PFAS, with all other structural possibilities ruled out. At ***Level 3***, at least one candidate structure must fulfill all available evidence, and multiple candidate structures are acceptable. At ***Level 4***, a structure cannot be elucidated but an unequivocal molecular formula is determined, while at ***Level 5*** the feature cannot be identified beyond being a molecule of interest. Definitions and additional requirements are detailed below.

Definitions

- Match: The measured value for a feature is within the user-defined acceptable error tolerance of the comparison value (*i.e.* the reference standard, library, or predicted value)
- Diagnostic evidence: Experimental observations that corroborate compound identifications. Examples for PFAS could include homologues, compound synthesis information, MS/MS fragments, ionization behavior, parent-transformation product relationships, etc.[2]
- PFAS homologues: Structurally similar compounds differing by one or more repeating structural unit *e.g.* CF_2, CH_2CF_2, or CF_2O[30]
- Library: A reputable repository containing validated mass spectra and/or CCS values
- Prediction: A calculated spectrum or value obtained from a computational tool or extrapolated from experimental data. Example for PFAS include CCS and RT and values derived from homologous series trendlines, subclass-aligned fragments, and more[3, 25, 30]
- Support: The comparison value aligns with the measured value for a feature and may be outside of the user-defined acceptable error (only applicable at Level 3 and lower).

Confirmed Structure (Level 1)

All Level 1 identifications must meet the minimum requirements of exact *m/z*, isotopic envelope, CCS, and RT (in the case of LC) or retention index (RI; in the case of GC) match to a reference standard (**Figure 3**). Reference values from a matching analytical standard must be collected on the same analytical platform as the unknown. These spectra may come from separate injections of the analytical standards or from spiked internal native or mass labeled standards, either analyzed at the same time as the samples of interest, or analyzed previously in-house using identical analytical methods.[2, 14] The high reproducibility of intra-laboratory CCS values and the ability of IMS to separate PFAS and matrix interferences makes it possible for analysts to use in-house libraries containing RT values, CCS values, and mass spectra without having to run new reference standards alongside experimental samples within the same analytical run. The number of possible Level 1 identifications is thus limited by the number of standards available to the analyst. Existing scales require MS2 spectra for all detected features; however, fragmentation spectra are not always obtainable for low abundance features, which is particularly problematic for environmental analysis of complex biological matrices. Given the high reproducibility of CCS values, the ability of IMS to separate PFAS from interfering biomolecules, and the increased confidence afforded by adding a CCS dimension, mobility-aligned fragments are not required for Level 1 PFAS identifications if diagnostic evidence is available such as the presence of other homologues in the series. Analysts should, however, consider all empirical data available and if anything does not align with the identification, the identification should be downgraded.

Probable Structure (Level 2)

At Level 2, all evidence in all available dimensions must point to exactly one structure and eliminate all others from contention.[2] This is the highest confidence level possible when standards cannot be obtained for the molecule in question. All Level 2 identifications must match either a library value or a predicted value (if a library value is not available) for exact *m/z*, isotopic

envelope, CCS, and RT (for LC) or RI (for GC). The library value must come from a reputable spectral or CCS library or literature appropriate to the experiment.[31, 32] Predicted values may come from a variety of sources, for example CCS vs *m/z* and RT vs *m/z* trendlines for homologous series.[25, 27] Fragmentation spectra should be used to determine a structure when available; however, it is possible that a combination of MS1 data, CCS values, RT/RI, and diagnostic evidence (such as the presence of homologues at other masses) may be sufficient to justify a Level 2 assignment for some probable structures when no other possibility satisfies all available evidence.

Tentative Candidates (Level 3)

With Level 3 identifications, at least one candidate structure is supported by all available evidence.[2, 6, 14] As with Level 2, Level 3 requires either a library or predicted value for exact *m/z*, isotopic envelope, CCS, and RT or RI for all candidate structures. However, at Level 3, multiple tentative candidate structures are permitted, *i.e*., the requirement for exactly one structure only is relaxed. Fragmentation spectra are typically required to determine a structure and should be used when available; however, as with Level 2, diagnostic evidence (see ***Definitions***) may be sufficient to propose a tentative candidate structure in the absence of fragmentation data. When multiple candidates are proposed, it is recommended to report a “most likely” candidate if possible. Level 3 is also an appropriate assignment when a reference standard or library was used, but the calculated errors are outside of the given tolerance criteria. While Metz *et al*.[33] recently proposed a “probability” approach to rank cases with multiple “n” candidates, this is not yet integrated in this level system while we await community feedback on their proposal.

Unequivocal Molecular Formula (Level 4)

The definition of a Level 4 identification is an unequivocal formula for the detected feature.[2, 6, 14] The exact *m/z* and isotopic pattern must match exactly one possible molecular formula. Molecular formula determination typically requires an isotopic envelope and rich

fragmentation information, but these are not fixed requirements. The highest scoring formula of a list of possible formulas does not qualify as an unequivocal formula without additional diagnostic evidence, as these are often wrong, especially with the presence of fluorine in the candidate formulas – a fact that is unavoidable in PFAS NTA.

Feature of Interest (Level 5)

Level 5 is the catchall for any feature in the dataset that cannot be annotated with an unequivocal formula or candidate structure but is still of potential interest to the experimental question.[2, 14] This does not mean that every remaining feature is annotated with Level 5. Often in environmental NTA there will be thousands of features that do not even qualify for Level 5; for example, matrix lipids when evaluating PFAS, or statistically insignificant features when comparing between two groups of samples (*e.g.,* statistical comparisons of features detected in upstream vs downstream industrial effluent). The requirements for a Level 5 identification include a measured monoisotopic *m/z*, a measured CCS, and at least one piece of evidence that supports the feature being a PFAS of interest. In the move towards automation, "of interest" must be explicitly defined. The additional evidence could be database driven, such as an exact mass match to a suspect list PFAS (equivalent to a Level 5a per the 2022 PFAS levels[3]) with insufficient data to bump up the identification to a Level 4. The evidence could also be experimental (equivalent to a Level 5b per the 2022 PFAS levels[3]), such as a negative mass defect, feature presence in fluorinated space when plotting its CCS vs *m/z*, presence of CF_2 homologues, presence of CF_2-containing fragments, or any number of other justifications.

Discussion and Examples of PFAS Identified with IMS

In typical LC- and GC-HRMS workflows, chromatographic separation and confirmation with diagnostic MS2 fragments is required for compounds with very close *m/z* values. **Figure 4A** shows two PFAS, Hydro-EVE and 6:2 FTS, that have *m/z* values within 10 ppm and require LC separation on most platforms. However, their CCS values are distinct enough that IMS fully

resolves these two molecules.[10] An evaluation of our in-house library[26] containing experimentally measured RT and CCS values from authenticated standards for 175 PFAS revealed only 3 pairs

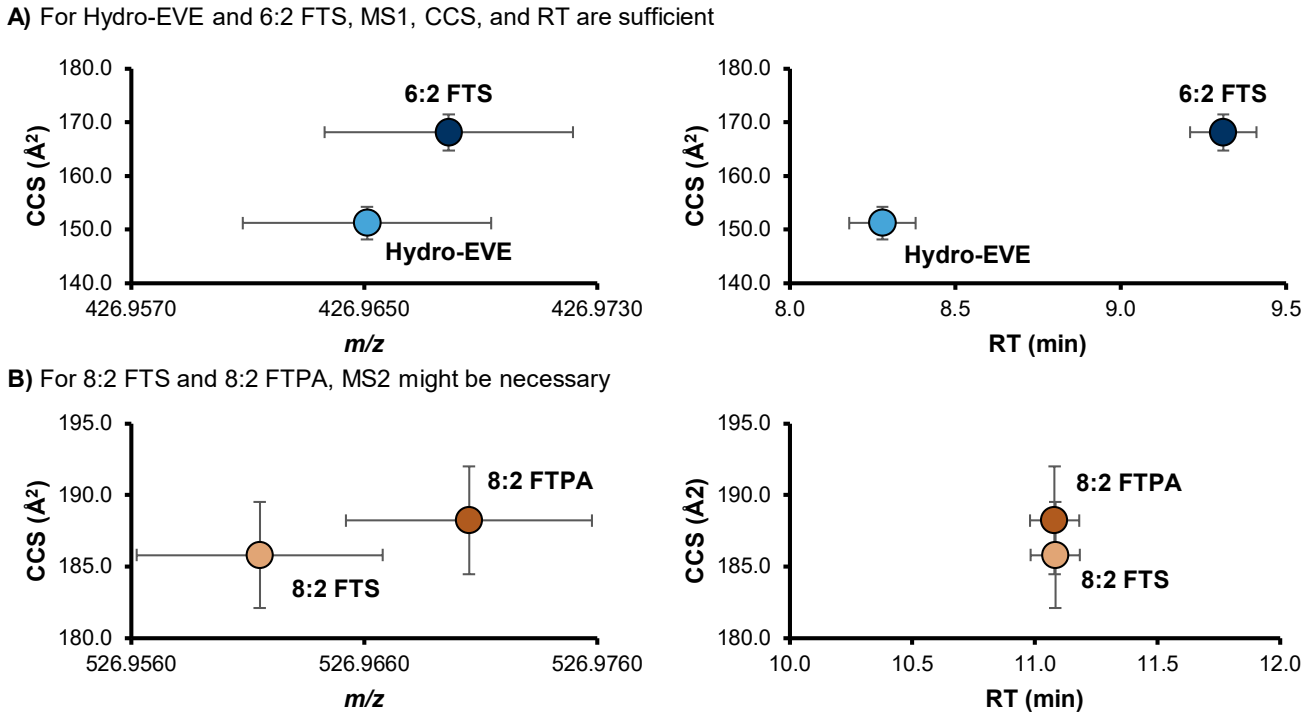

Figure 4. Two examples of PFAS measured using LC-IMS-HRMS. The vertical error bars show ±2.0% CCS, and the horizonal error bars show ± 10 ppm (for *m/z*) and ± 0.1 min (for RT). **A)** The exact *m/z* of deprotonated Hydro-EVE and 6:2 FTS are quite close (426.9651 and 426.9679 respectively, a difference of 6.57 ppm). However, their CCS values (151.2 and 168.1 A^2) diverge enough (11.1%) to fully resolve these two molecules even if they were to coelute. **B)** On the other hand, 8:2 FTS and 8:2 FTPA have a larger difference in *m/z* (526.9615 and 526.9705, 17 ppm) that are resolved by a tolerance of ± 5 ppm but have a chance of overlapping with a tolerance of ± 10 ppm. Because the CCS values are also close and these molecules coelute in this method, MS2 is recommended to confirm which structures are present in the sample. However, depending on the LC gradient and the resolving power of the IMS platform used, MS2 may not be necessary.

of PFAS with the potential to overlap with another PFAS in the library when LC, IMS, and HRMS dimensions are all considered with the tolerances shown in **Table 1**, not including branched isomers (*e.g.*, linear PFOA separates from branched PFOA, but individual branched isomers do not always separate from each other).[10] **Figure 4B** shows one of these pairs, 8:2 FTS and 8:2 FTPA, that may not fully separate using MS1, RT, and CCS depending on the resolving power of the IMS system (remaining pairs shown in **Figure S3**).

Existing confidence level guidance does not currently consider subjective parameters like geographical knowledge and whether certain PFAS may be coming from manufacturing products in the area. However, such context is often important to the PFAS identification workflow and may

sometimes factor into confidence level assignments when assigning tentative Level 3 structures, selecting a most likely candidate at Level 3, bumping a tentative Level 3 candidate to a probable Level 2 structure, or confirming a Level 1 identification in the absence of observable fragments. Our updated guidance permits the analyst to use such knowledge as diagnostic evidence so long as it is reported transparently. For example, if a researcher is studying PFAS in water upstream and downstream of a fluorochemical manufacturer and they have existing knowledge that only one of the candidates is produced by that manufacturer, the user could change the confidence from Level 3 to Level 2. All minimum requirements for mass spectra, CCS, and RT must still be met in this case. Thus, in the example illustrated in **Figure 4**, confidence levels would be assigned as follows, assuming a mass tolerance of ± 10 ppm and a CCS tolerance of ± 0.5%:

- 6:2 FTS and Hydro-EVE: Level 1
- 8:2 FTS and 8:2 FTPA: Level 1 (with diagnostic or fragmentation evidence), Level 3 otherwise (report feature as “8:2 FTS or 8:2 FTPA”)

Perspective

Effective NTA workflows are essential for monitoring and discovering PFAS and other emerging compounds, and transparent communication of results is critical. Multidimensional techniques like LC-IMS-HRMS are invaluable for analyzing complex environmental and biological samples; however, current methods suffer from inconsistent data quality and confidence reporting. Here, we adapted existing confidence level frameworks to PFAS identified using LC- and GC-IMS-HRMS to improve data quality and reporting of identification confidence in this rapidly growing area of research. Our research experiences amongst study coauthors combined with a dedicated literature search revealed the inconsistent use of existing confidence level guidance for PFAS across a wide range of NTA platforms. To address this gap, we unified and simplified existing confidence level guidance by creating a minimum requirements checklist for PFAS identified at each level. We additionally proposed improvements including user-defined,

instrument-specific tolerances for quality control and updated requirements for CCS and MS2 fragmentation across all levels. These updates provide easier methods for analysts to communicate confidence of PFAS identifications and increase the transparency of NTA results. While this guidance is currently intended for PFAS identified using LC- or GC-IMS-HRMS, researchers using non-IMS platforms focused on molecules other than PFAS may also benefit from these updates. Guidance should be updated as new analytical technologies appear and predictive tools improve, ultimately growing towards automation. To support this, we are developing a web-based application (Small Molecule Identification Scoring Made Easy, or "SMISE") for automated confidence level assignment as well as additional functionality to aid NTA researchers in proposing and eliminating candidate structures. Release of this tool will enable automated and objective confidence level assignment and unified reporting of evidence used in NTA PFAS identifications.

Declarations

Data Availability Statement

The Baker lab PFAS LC-IMS-HRMS library is freely available on Zenodo (DOI: https://doi.org/10.5281/zenodo.14341320). A standalone checklist is also available on Zenodo (DOI: https://doi.org/10.5281/zenodo.14743503).

Acknowledgements

Funding for this work was made possible through grants from the NIH National Institute of Environmental Health Sciences (P42 ES027704, P42ES031007, and T32ES007126), a cooperative agreement with the Environmental Protection Agency (STAR RD 84003201), and support was additionally provided through the Institute for Environmental Health Solutions at the University of North Carolina Gillings School of Global Public Health. ELS acknowledges funding support from the Luxembourg National Research Fund (FNR) for project A18/BM/12341006.

Conflicts of Interest

The authors have no competing financial interests to declare.

Supporting Information

The Supplemental Document (PDF) contains supplemental tables (including **S1**: Breakdown of experiment types. **S2**: Summary of proposed changes to existing guidance. **S3**: Worksheet for defining tolerances.) and figures (including **S1**: Number of PFAS identified in NTA studies from 2023. **S2**: Examples of inconsistent confidence levels. **S3**: LC-IMS-HRMS library evaluation of overlapping PFAS.).

The Supplemental Checklist (PDF) is a printable version of the checklist shown in **Figure 3**, with definitions added, to serve as a standalone reference

References

(1) Boatman, A., Chappel, J., Kirkwood-Donelson, K., Fleming, J., Reif, D., Schymanski, E., Rager, J., & Baker, E. LC- or GC-IMS-HRMS Confidence Level Requirement Checklist for PFAS. 1 ed.; Zenodo, 2025.

(2) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology* **2014**, *48* (4), 2097-2098. DOI: 10.1021/es5002105.

(3) Charbonnet, J. A.; McDonough, C. A.; Xiao, F.; Schwichtenberg, T.; Cao, D.; Kaserzon, S.; Thomas, K. V.; Dewapriya, P.; Place, B. J.; Schymanski, E. L.; et al. Communicating Confidence of Per- and Polyfluoroalkyl Substance Identification via High-Resolution Mass Spectrometry. *Environmental Science & Technology Letters* **2022**, *9* (6), 473-481. DOI: 10.1021/acs.estlett.2c00206.

(4) Perera, D.; Scott, W.; Smolinski, R.; Mukhopadhyay, L.; McDonough, C. A. Techniques to characterize PFAS burden in biological samples: Recent insights and remaining challenges. *Trends in Environmental Analytical Chemistry* **2024**, *41*, e00224. DOI: https://doi.org/10.1016/j.teac.2023.e00224.

(5) Strynar, M.; McCord, J.; Newton, S.; Washington, J.; Barzen-Hanson, K.; Trier, X.; Liu, Y.; Dimzon, I. K.; Bugsel, B.; Zwiener, C.; et al. Practical application guide for the discovery of novel PFAS in environmental samples using high resolution mass spectrometry. *Journal of Exposure Science & Environmental Epidemiology* **2023**, *33* (4), 575-588. DOI: 10.1038/s41370-023-00578-2.

(6) Charbonnet, J. A.; McDonough, C. A.; Xiao, F.; Schwichtenberg, T.; Cao, D.; Kaserzon, S.; Thomas, K. V.; Dewapriya, P.; Place, B. J.; Schymanski, E. L.; et al. Communicating Confidence of Per- and Polyfluoroalkyl Substance Identification via High-Resolution Mass Spectrometry. *Environmental Science & Technology Letters* **2022**, *9* (6), 473-481. DOI: 10.1021/acs.estlett.2c00206.

(7) PFAS Suspect List. National Institute of Standards and Technology Public Data Repository. 2023-01-05. https://data.nist.gov/od/id/mds2-2387 (accessed 2024-10-22).

(8) Schymanski, E. L.; Zhang, J.; Thiessen, P. A.; Chirsir, P.; Kondic, T.; Bolton, E. E. Per- and Polyfluoroalkyl Substances (PFAS) in PubChem: 7 Million and Growing. *Environmental Science & Technology* **2023**, *57* (44), 16918-16928. DOI: 10.1021/acs.est.3c04855.

(9) Díaz-Galiano, F. J.; Murcia-Morales, M.; Monteau, F.; Le Bizec, B.; Dervilly, G. Collision cross-section as a universal molecular descriptor in the analysis of PFAS and use of ion mobility spectrum filtering for improved analytical sensitivities. *Analytica Chimica Acta* **2023**, *1251*, 341026. DOI: https://doi.org/10.1016/j.aca.2023.341026.

(10) Dodds, J. N.; Hopkins, Z. R.; Knappe, D. R. U.; Baker, E. S. Rapid Characterization of Per- and Polyfluoroalkyl Substances (PFAS) by Ion Mobility Spectrometry–Mass Spectrometry (IMS-MS). *Analytical Chemistry* **2020**, *92* (6), 4427-4435. DOI: 10.1021/acs.analchem.9b05364.

(11) Peter, K. T.; Phillips, A. L.; Knolhoff, A. M.; Gardinali, P. R.; Manzano, C. A.; Miller, K. E.; Pristner, M.; Sabourin, L.; Sumarah, M. W.; Warth, B.; et al. Nontargeted Analysis Study Reporting Tool: A Framework to Improve Research Transparency and Reproducibility. *Analytical Chemistry* **2021**, *93* (41), 13870-13879. DOI: 10.1021/acs.analchem.1c02621.

(12) Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; McLean, J. A. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *Journal of the American Society for Mass Spectrometry* **2016**, *27* (12), 1897-1905. DOI: 10.1007/s13361-016-1469-y.

(13) Phillips, A. L.; Williams, A. J.; Sobus, J. R.; Ulrich, E. M.; Gundersen, J.; Langlois-Miller, C.; Newton, S. R. A framework for utilizing high-resolution mass spectrometry and nontargeted analysis in rapid response and emergency situations. *Environmental toxicology and chemistry* **2022**, *41* (5), 1117-1130.

(14) Celma, A.; Sancho, J. V.; Schymanski, E. L.; Fabregat-Safont, D.; Ibáñez, M.; Goshawk, J.; Barknowitz, G.; Hernández, F.; Bijlsma, L. Improving Target and Suspect Screening High-Resolution Mass Spectrometry Workflows in Environmental Analysis by Ion Mobility Separation. *Environmental Science & Technology* **2020**, *54* (23), 15120-15131. DOI: 10.1021/acs.est.0c05713.
(15) Schulze, B.; Jeon, Y.; Kaserzon, S.; Heffernan, A. L.; Dewapriya, P.; O'Brien, J.; Gomez Ramos, M. J.; Ghorbani Gorji, S.; Mueller, J. F.; Thomas, K. V.; et al. An assessment of quality assurance/quality control efforts in high resolution mass spectrometry non-target workflows for analysis of environmental samples. *TrAC Trends in Analytical Chemistry* **2020**, *133*, 116063. DOI: https://doi.org/10.1016/j.trac.2020.116063.
(16) Fisher, C. M.; Peter, K. T.; Newton, S. R.; Schaub, A. J.; Sobus, J. R. Approaches for assessing performance of high-resolution mass spectrometry–based non-targeted analysis methods. *Analytical and bioanalytical chemistry* **2022**, *414* (22), 6455-6471.
(17) Kirkwood-Donelson, K. I.; Dodds, J. N.; Schnetzer, A.; Hall, N.; Baker, E. S. Uncovering per- and polyfluoroalkyl substances (PFAS) with nontargeted ion mobility spectrometry–mass spectrometry analyses. *Science Advances 9* (43), eadj7048. DOI: 10.1126/sciadv.adj7048 (acccessed 2024/10/22).
(18) Boatman, A. K.; Chappel, J. R.; Polera, M. E.; Dodds, J. N.; Belcher, S. M.; Baker, E. S. Assessing Per- and Polyfluoroalkyl Substances in Fish Fillet Using Non-Targeted Analyses. *Environmental Science & Technology* **2024**, *58* (32), 14486-14495. DOI: 10.1021/acs.est.4c04299.
(19) Dodds, J. N.; Baker, E. S. Ion mobility spectrometry: fundamental concepts, instrumentation, applications, and the road ahead. *Journal of the American Society for Mass Spectrometry* **2019**, *30* (11), 2185-2195.
(20) Zheng, X.; Aly, N. A.; Zhou, Y.; Dupuis, K. T.; Bilbao, A.; Paurus, V. L.; Orton, D. J.; Wilson, R.; Payne, S. H.; Smith, R. D. A structural examination and collision cross section database for over 500 metabolites and xenobiotics using drift tube ion mobility spectrometry. *Chemical science* **2017**, *8* (11), 7724-7736.
(21) Picache, J. A.; Rose, B. S.; Balinski, A.; Leaptrot, Katrina L.; Sherrod, S. D.; May, J. C.; McLean, J. A. Collision cross section compendium to annotate and predict multi-omic compound identities. *Chemical Science* **2019**, *10* (4), 983-993, 10.1039/C8SC04396E. DOI: 10.1039/C8SC04396E.
(22) Izquierdo-Sandoval, D.; Fabregat-Safont, D.; Lacalle-Bergeron, L.; Sancho, J. V.; Hernández, F.; Portoles, T. Benefits of Ion Mobility Separation in GC-APCI-HRMS Screening: From the Construction of a CCS Library to the Application to Real-World Samples. *Analytical Chemistry* **2022**, *94* (25), 9040-9047. DOI: 10.1021/acs.analchem.2c01118.
(23) Celma, A.; Alygizakis, N.; Belova, L.; Bijlsma, L.; Fabregat-Safont, D.; Menger, F.; Gil-Solsona, R. Ion mobility separation coupled to high-resolution mass spectrometry in environmental analysis – Current state and future potential. *Trends in Environmental Analytical Chemistry* **2024**, *43*, e00239. DOI: https://doi.org/10.1016/j.teac.2024.e00239.
(24) Dodds, J. N.; Kirkwood-Donelson, K. I.; Boatman, A. K.; Knappe, D. R.; Hall, N. S.; Schnetzer, A.; Baker, E. S. Evaluating Solid Phase Adsorption Toxin Tracking (SPATT) for passive monitoring of per-and polyfluoroalkyl substances (PFAS) with Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS). *Science of The Total Environment* **2024**, *947*, 174574.
(25) Foster, M.; Rainey, M.; Watson, C.; Dodds, J. N.; Kirkwood, K. I.; Fernández, F. M.; Baker, E. S. Uncovering PFAS and Other Xenobiotics in the Dark Metabolome Using Ion Mobility Spectrometry, Mass Defect Analysis, and Machine Learning. *Environmental Science & Technology* **2022**, *56* (12), 9133-9143. DOI: 10.1021/acs.est.2c00201.
(26) Dataset for "Multidimensional library for the improved identification of per- and polyfluoroalkyl substances (PFAS)". (accessed 2024-12-20).

(27) Kirkwood-Donelson, K. I.; Dodds, J. N.; Schnetzer, A.; Hall, N.; Baker, E. S. Uncovering per- and polyfluoroalkyl substances (PFAS) with nontargeted ion mobility spectrometry-mass spectrometry analyses. *Sci Adv* **2023**, *9* (43), eadj7048. DOI: 10.1126/sciadv.adj7048 From NLM.
(28) BP4NTA. *QA/QC Metrics*. 2024. https://nontargetedanalysis.org/reference-content/results/qa-qc-metrics/ (accessed 2024 10/4/2024).
(29) Stow, S. M.; Causon, T. J.; Zheng, X.; Kurulugama, R. T.; Mairinger, T.; May, J. C.; Rennie, E. E.; Baker, E. S.; Smith, R. D.; McLean, J. A.; et al. An Interlaboratory Evaluation of Drift Tube Ion Mobility–Mass Spectrometry Collision Cross Section Measurements. *Analytical Chemistry* **2017**, *89* (17), 9048-9055. DOI: 10.1021/acs.analchem.7b01729.
(30) Koelmel, J. P.; Paige, M. K.; Aristizabal-Henao, J. J.; Robey, N. M.; Nason, S. L.; Stelben, P. J.; Li, Y.; Kroeger, N. M.; Napolitano, M. P.; Savvaides, T.; et al. Toward Comprehensive Per- and Polyfluoroalkyl Substances Annotation Using FluoroMatch Software and Intelligent High-Resolution Tandem Mass Spectrometry Acquisition. *Analytical Chemistry* **2020**, *92* (16), 11186-11194. DOI: 10.1021/acs.analchem.0c01591.
(31) BP4NTA. *Data Processing and Analysis*. 2024. https://nontargetedanalysis.org/reference-content/methods/data-processing-and-analysis/ (accessed 2025 1/20/2025).
(32) Elapavalore, A.; Kondić, T.; Singh, R. R.; Shoemaker, B. A.; Thiessen, P. A.; Zhang, J.; Bolton, E. E.; Schymanski, E. L. Adding open spectral data to MassBank and PubChem using open source tools to support non-targeted exposomics of mixtures. *Environmental Science: Processes & Impacts* **2023**, *25* (11), 1788-1801, 10.1039/D3EM00181D. DOI: 10.1039/D3EM00181D.
(33) Metz, T. O.; Chang, C. H.; Gautam, V.; Anjum, A.; Tian, S.; Wang, F.; Colby, S. M.; Nunez, J. R.; Blumer, M. R.; Edison, A. S.; et al. Introducing “Identification Probability” for Automated and Transferable Assessment of Metabolite Identification Confidence in Metabolomics and Related Studies. *Analytical Chemistry* **2025**, *97* (1), 1-11. DOI: 10.1021/acs.analchem.4c04060.