Non-targeted identification of per- and polyfluoroalkyl substances at trace level in surface water using fragment ion flagging

Thijs J. Hensema, Bjorn J.A. Berendsen, Stefan P.J. van Leeuwen[*]

Wageningen University & Research, Wageningen Food Safety Research, Wageningen, the Netherlands

HIGHLIGHTS

- Novel PFASs were discovered with non-targeted screening using Fragment Ion Flagging.
- Fragment Ion Flagging allowed the identification of PFASs at trace levels.
- Screening techniques such as homologous series detection proved to be ineffective.
- Fragment based mass spectrometry was demonstrated to be an effective alternative.
- Novel PFASs were found in environmental samples, i.e. surface water.

ABSTRACT

The extent of unidentified Per- and Poly-fluoroalkyl substances (PFASs) found in environmental samples has led to the development of non-targeted screening methods. The study presented here reports the use of liquid chromatography hyphenated with high resolution mass spectrometry to detect and identify unknown and unexpected PFASs by fragment ion flagging (FIF). By exploring all ion fragmentation spectra for several characteristic fragments including $C_nF_{2n+1}^-$, $C_nF_{2n-1}^-$, $C_nF_{2n-3}^-$, $C_nF_{2n-7}^-$, $C_nF_{2n-11}^-$ and $C_nF_{2n+1}O^-$ the presence of widely different PFAS species can be anticipated without the need for targeted screening methods. These fragments are then related to their precursor ion by retention time matching and subsequently identified. With this methodology 40 PFASs were (tentatively) identified in four surface water samples sampled throughout the Netherlands. To the best of the authors' knowledge, four PFASs found through FIF are newly discovered species and have not been mentioned in any database or literature. This methodology eliminates the dependence on commonly reported full scan feature selection techniques such as mass defect filtering, homologous series detection and intensity threshold filtering, allowing the identification of PFASs at trace levels. Additionally, eight of the (tentatively) identified PFASs are not part of homologous series, stressing the shortcomings of commonly reported non-targeted PFASs screening methods and demonstrating the importance of more effective identification strategies such as FIF. Moreover, we like to emphasise that this approach is applicable to real-life environmental samples with PFASs at background concentration levels.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Per- and Poly-Fluoroalkyl Substances (PFASs) are man-made chemicals that have been used for a variety of applications (Buck et al., 2011). The unique physical and chemical properties of PFASs impart oil and water repellence, temperature resistance, and friction reduction to a wide range of products. PFASs have been detected in numerous environmental compartments across the globe (Houde et al., 2011). Due to their persistence and bio-accumulative and toxic (PBT) properties, several PFASs have been phased out by the industry and their production has been restricted (AnnextoH, 2020; −9/12: Listing of perf, 2019). Subsequently, a shift towards alternative PFASs with unknown toxicity and environmental fate is made, including shorter chain PFASs that do not bioaccumulate but are persistent and potentially toxic (OECD/UNEP, 2013; Wang et al., 2013).

The list of PFASs has severely expanded in recent years. In their inventory, the Organisation for Economic Co-operation and

T.J. Hensema, B.J.A. Berendsen and S.P.J. van Leeuwen

Development (OECD) has identified 4730 PFASs (2018). With current targeted methods, only a small selection of well-known PFASs are covered, ignoring unknown (or unreported) PFASs. The use of High Resolution Mass Spectrometry (HRMS) for non-targeted screening of novel or emerging PFASs is commonly reported in recent literature (Newton et al., 2017; Gebbink et al., 2017; Yu et al., 2018; Prevedouros et al., 2006). The majority of studies used full scan data for the selection of signals of interest. This results in an excessive amount of data. In an effort to reduce the number of potentially relevant signals, several feature selection techniques are used, such as homologous series detection (Newton et al., 2017; Gebbink et al., 2017; Yu et al., 2018; Prevedouros et al., 2006; Liu et al., 2015; Strynar et al., 2015), mass defect filtering (Gebbink et al., 2017; Prevedouros et al., 2006; McCord and Strynar, 2019; Hatton et al., 2018; D'Agostino and Mabury, 2014) and intensity threshold filtering (Yu et al., 2018; Strynar et al., 2015; McCord and Strynar, 2019; Barzen-Hanson et al., 2017). Although these approaches demonstrated their capabilities for identification of PFASs, they come along with drawbacks.

Homologous series detection is often used (Newton et al., 2017; Gebbink et al., 2017; Yu et al., 2018; Prevedouros et al., 2006; Liu et al., 2015; Strynar et al., 2015) as the two main synthesis routes of PFASs generally yield homologous series of varying numbers of CF_2 or C_2F_4 groups. However, PFASs not part of any homologous series would be missed when relying on this technique. Furthermore, it is known that substances belonging to a homologous series (i.e. long chain PFASs versus short chain PFASs) do not necessarily have the same fate and thus, will not end up in the same reservoir.

Mass defect filtering is a frequently applied feature selection approach in the detection of PFASs as well (Gebbink et al., 2017; Prevedouros et al., 2006; McCord and Strynar, 2019; Hatton et al., 2018; D'Agostino and Mabury, 2014). Highly fluorinated compounds can be differentiated from hydrogen substituted compounds by their relatively low mass defect. However, a negative mass defect could also indicate the presence of other atoms. Moreover, poly-fluorinated substances might, besides a fluorinated moiety, also contain large hydrogen substituted moieties leading to a (relatively large) positive mass defect, which hampers the identification potential based on mass defect.

Because homologous series detection and mass defect filtering are often not sufficient in effectively reducing the number of relevant signals, intensity threshold filtering is applied (Yu et al., 2018; Strynar et al., 2015; McCord and Strynar, 2019; Barzen-Hanson et al., 2017). The major drawback of this is the loss of trace level compounds and analytes with low ionization efficiency. Note that the intensity of a signal in non-targeted screening is in no way an indication of the potency or concentration of the related analyte.

The use of the above-mentioned feature selection techniques leads to an ambiguous non-targeted screening process yielding a relatively high percentage of false negative results. To more effectively anticipate the presence of PFASs in environmental samples, a more effective approach was developed.

An alternative approach that overcomes the mentioned drawbacks is Fragment Ion Flagging (FIF). By exploring characteristic fragments and/or neutral losses in MS, important features can be more quickly selected leading to a higher confidence of PFASs identification. Additionally, this removes data processing steps used in classical non-targeted screening (e.g. intensity threshold filtering). As many PFASs share the same molecular moieties i.e. C_nF_{2n+1}, these moieties can be observed in MS fragment experiments. However, these moieties fragment differently depending on the molecular structure of the analyte.

In recent literature FIF has been applied for the non-targeted identification of PFASs. In recent years, studies were published using this strategy on waste water from a fluorochemical manufacturing park in China (Liu et al., 2015), Aqueous Film Forming Foam (AFFF's) (Xiao et al., 2017) and fish (Liu et al., 2018), demonstrating the potential of this approach. These studies focussed on samples with relatively high PFASs concentrations (ppb range). In surface waters, PFASs are present at trace levels (ppt range) (Gebbink et al., 2017; Heydebreck et al., 2015) (even in industry rich areas) and such low levels are a huge challenge for non-targeted techniques. The aim of this project was to study the effectiveness of FIF for the detection and identification of PFASs in environmental samples at trace level concentrations, and to optimise the workflow. Additionally, suspect screening of full scan data is applied for comparison. As to not overcomplicate sample preparations and data processing, a focus is made on the detection of mainly anionic species. To the best of the authors' knowledge the work presented here is the first application of FIF for non-targeted screening of PFASs in surface waters (after sample enrichment) with background concentrations. Furthermore, based on the applied approach, four new PFASs are presented, not previously reported.

2. Experimental

2.1. Materials

The perfluoroalkyl carboxylic acids (PFCAs) (C_4–C_{13}), perfluoroalkyl sulfonic acids (PFSAs) (C_4, C_6, C_7, C_8 and C_{10}) and hexafluoropropylene oxide-dimer acid (HFPO-DA, a.k.a. GenX) were purchased from Wellington Laboratories, Ontario, Canada. Hexafluoropropylene oxide-trimer acid (HFPO-TA) was purchased from abcr GmbH, Karlsruhe, Germany. All H–PFCAs (C_3–C_5, C_7–C_9 and C_{11}) were acquired from Apollo Scientific, Manchester, United Kingdom. Acetonitrile and methanol were purchased from Actuall-Chemicals, Oss, The Netherlands. Ammonium solution (25%) and sodium acetate trihydrate were obtained from Merck, Darmstadt, Germany. Surface water samples were collected and stored in 1 L-high density polyethylene (HDPE) containers and prepared using weak anion exchange (WAX) Solid Phase Extraction (SPE) (Oasis WAX cartridge, 6 cc, 150 mg, 30 μm, Waters).

2.2. Sample collection

Surface water was sampled on four locations spread over the Netherlands, being the river Rhine in Tuindorp (at the Dutch-German border where the river Rhine enters The Netherlands), Noordzeekanaal in Amsterdam, Beneden Merwede in Dordrecht and the Scheldt in Zeeland. These places were selected because they are located in industrial areas, or industrial areas are located upstream. The location Beneden-Merwede is close to a fluorochemical production plant (Gebbink and van Leeuwen, 2020), whereas the other locations have no clear linkage to fluorochemical point sources. A metal bucket was used to transfer collected water to a 1 L-HDPE container. Preceding each collection, the bucket was rinsed twice with water from the sampling location. Samples were stored at 4 °C the next day.

2.3. Sample preparation

The sample enrichment procedure focussed on anionic PFASs species, using a previously validated WAX SPE method, developed for the analyses of PFASs in milk and eggs (Berendsen et al., 2020). The recovery for PFCAs and PFSAs (C5–C12) is typically above 70%. Before loading of the samples on to the SPE columns, the SPE columns were conditioned with 8 mL of methanol followed by 5 mL of Milli-Q water. Each water sample was prepared by loading 200 mL on two separate SPE cartridges, in total 400 mL per sample. After

washing of the cartridge with 5 mL sodium acetate buffer (pH 4.0) followed by 3 mL of methanol, elution was done by 5 mL of 2% NH_4OH in ACN. After elution, the two eluents obtained from a single sample were combined. Eluates were evaporated to dryness (40 °C, N_2). Hereafter, samples were dissolved in 100 μL of mobile phase B (20 mM in methanol/water, 95/5) and ultrasonicated for 5 min, followed by the addition 100 μL of mobile phase A (20 mM ammonium bicarbonate in H_2O) and again ultrasonication for 5 min. With every batch of samples, a method blank was included and prepared as described above using in-house MilliQ water as sample. Additionally, instrument blanks were also included in the injection sequence by injection of MilliQ to monitor for signals from the LC-MS system.

2.4. Instrumentation

Chromatographic separation was carried out on a Dionex Ultimate 3000 system using a Waters Acquity UPLC BEH C18 column (1.7 μm, 2.1 × 100 mm) at a column oven temperature of 50 °C. The LC system was coupled to a Thermo Fischer Q-Exactive orbitrap system using Electron Spray Ionization operating in negative ionization mode (ESI^-). The mobile phase consisted of eluent A (20 mM ammonium bicarbonate in H_2O) and eluent B (20 mM in methanol/water, 95/5, v/v). The gradient used at a flow of 0.3 mL min^{-1} was: 0–0.1 min, 0% mobile phase B, 0.1–15.1 min, linear increase to 100% B, with a final hold for 5 min. Hereafter the system returned to its initial conditions within 1 min with a final equilibration time of 9 min, resulting in a total runtime of 30.1 min. The injection volume was 5 μL.

2.5. Mass spectrometer analysis

FIF was carried out using All Ion Fragmentation (AIF) at a resolution of 70.000 Full Width Half Maximum (FWHM) with stepped collision energies of 30 and 60 eV. The Automatic Gain Control (AGC) target was set to 3e6 and a maximum injection time of 200 ms was used. The scan range for AIF spectra was between m/z 60 and m/z 900.

Full scan was used in parallel to AIF for assignment of the molecular ion mass and calculation of the unequivocal molecular formula. Full scan analysis was done at a resolution of 140.000 FWHM (at m/z 200) with a scan range of m/z 110 to m/z 1100; the AGC target was set to 1e6, and the maximum injection time was 200 ms. Common contaminants continuously measured as background noise served as internal lock masses. Internal lock masses are used to correct for mass drifts, allowing sub-ppm accuracy in full scan analysis. The lock masses used were mainly saturated fatty acids found as background noise. For each analysis, lock masses were re-established by injecting a solvent blank and extracting signals visible in the majority of the scans. Care was taken in selecting masses at different sections of the mobile phase gradient to cover the entire chromatogram. Additionally, lock masses were selected

to cover a mass range as wide as possible. Only m/z values where an unequivocal molecular formula was calculated were used as lock masses. Lock mass corrections are done automatically by the Q-Exactive system during data acquisition.

In combination with FIF, full scan data was also used for suspect screening and homologues series detection of PFASs. Full scan data was processed using Compound Discoverer 3.0, the settings used are listed in Table S1. Homologues series detection was done using a modified R package (Loos and Singer, 2017). Suspect lists were downloaded from the U.S. Environmental Protection Agency (EPA) CompTox Chemistry Dashboard (Williams et al., 2017) and the NORMAN Suspect List Exchange (Suspect List Excha, 2019).

Data Independent Analysis (DIA) was used for recording MS/MS spectra of masses of interest selected through FIF. The use of DIA allows isolation of ions regardless of their abundance and was preferred over Data Dependent Analysis (DDA) because it provides richer MS/MS spectra in contrary to DDA where only MS/MS spectra are recorded if a certain intensity threshold for a specified mass is reached. Hence DIA is preferred for mass isolation. ESI^- DIA analyses were done at a resolution of 70.000 FWHM and stepped collision energies of 10, 30 and 60 eV were used. The AGC target was set to 3e6, and the maximum injection time was 200 ms. Mass of interest were isolated within an isolation window of 1 m/z.

2.6. Workflow characteristic FIF

The process of FIF is illustrated in Fig. 1. Details on settings in each step are discussed later. AIF is applied (step 1) and subsequently characteristic PFASs fragments can be found through Extracted Ion Chromatograms (XICs, step 2). Based on reference standards of known PFASs combined with literature review the following relevant fragment ions were selected: $C_nF_{2n+1}^-$ (n = 2–10), $C_nF_{2n-1}^-$ (n = 2–10), $C_nF_{2n-3}^-$ (n = 3–11) $C_nF_{2n-7}^-$ (n = 5–11), $C_nF_{2n-11}^-$ (n = 7–12), $C_nF_{2n+1}O^-$ (n = 1–3), with a maximum allowed error of 5 ppm. The characteristic neutral loss of HF was investigated as well.

Through a second injection, full scan analysis was performed, the molecular ion that yielded the characteristic fragment ions in AIF was found through retention time matching (step 3). Subsequently targeted MS/MS spectra were recorded (step 4) to allow structure elucidation (step 5). Each PFAS found was assigned a confidence level as defined by Schymanski et al. (2014). However, identification based on FIF does not fit well in the original Schymanski criteria and therefore we propose an addition to this approach, as discussed in the Results and Discussion section. In cases a reference standard could be obtained from a supplier (e.g HFPO-TA), a confidence level up to 1 could be reached (see results and discussion).

2.7. Unequivocal molecular formula

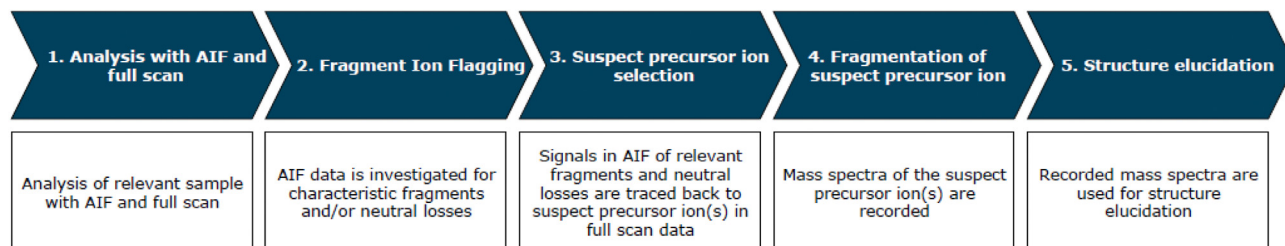Molecular formulas were calculated using software developed

Fig. 1. Workflow for fragment ion flagging for non-targeted screening.

by Lommen (2014) and is based on the Seven Golden Rules for heuristic filtering rules defined by Kind and Fiehn (2007). A maximum mass error of 1 ppm was permitted.

3. Results and discussion

The use of FIF for the analysis of four surface water samples led to the (tentative) identification of 31 PFASs with nine additional PFASs identified as background contaminants in solvent and method blanks. Notably, none of the analysed surface water samples contained PFAS homologous series other than PFCAs and PFSAs, homologous series of unexpected PFASs were not present. Applying homologues series detection of CF_2 and C_2F_4 on full scan data did indeed not result in the detection of homologues series apart from PFCAs and PFSAs (Fig. S1). Other studies often do find homologues series. This can possibly be attributed to the nature of these papers, focussing on PFASs contaminated sources, whereas the samples analysed here are not directly related to known contaminated sites, which may reduce the chances of finding homologue series. Additionally, due to policies restricting the production of long-chain PFASs (AnnextoH, 2020; −9/12: Listing of perf, 2019), synthesis of long chain PFASs is more strictly controlled likely resulting in singular homologous. This demonstrates it is extremely important not to rely on homologous series detection when using full scan feature selection. Additionally, to reduce data in full scan feature selection, an intensity threshold is often used, removing trace level compounds as well as compounds with low ionization efficiency. As such, FIF offers an advantage over full scan feature selection for effectively extracting relevant signals from the data without the need for intensity threshold filtering. Note that suspect screening of the surface waters in this study did not yield additional PFASs compared to the application of FIF. Moreover, suspect screening only identified 24 PFASs compared to the 40 PFASs identified with FIF (Table S2 and Table S3, respectively). Besides novel PFASs, suspect screening also missed several well-known PFASs included in the suspect screening list, this could be caused by problems with integration of peaks of low abundance ions in full scan analysis. This supports the assumption that FIF is more likely to detect unknown PFASs compared to full scan suspect screening. However, most likely, in some cases full scan will be more sensitive.

3.1. Confidence level of identification by FIF

The identification of unknown compounds is often supported by criteria defined by Schymanski et al. (2014) to indicate the confidence of the identification. According to the Schymanski scheme, unknown compounds are firstly identified by their exact mass (Fig. 2). By obtaining additional identification evidence (e.g. the unequivocal molecular formula) the level of confidence increases. The Schymanski scheme strictly speaking does not incorporate a starting point of an exact mass of interest that already intrinsically carries information on the identity of a substance (in this case a tentative candidate of the PFAS group). Therefore, with FIF, a higher level of identification is already obtained right from the start. The authors therefore propose to add levels to the Schymanski criteria to fit small molecule identification with FIF. These are illustrated in Fig. 2 as level 3F1, 3F2 and 3F3 in which F refers to FIF. Similar to the original Schymanski criteria, identification starts at the lowest level, with FIF this is 3F3. These additional levels are parallel to the original Schymanski criteria as these criteria are not directly applicable to small molecule identification using full scan MS and vice versa.

As done with FIF, by detection of characteristic fragments, information such as substituents, class and partial molecular formula is gained right from the start. Moreover, detected fragments could also potentially indicate a structural moiety. Therefore, it seems appropriate that FIF offers an increased initial identification confidence compared to classical identification with full scan MS, hence the adjusted criteria are added as sublevels of level 3. Each of the modified Schymanski criteria indicate the identification confidence level in the FIF process; 3F3 indicates the presence of relevant fragment signals but a (currently) unknown molecular ion. Because the molecular ion is unknown in this stage, it is important to note the retention time as an additional identifying factor as identical fragments may be present across the chromatogram. As more information is gained on the molecular ions, the identification confidence increases. Similar to the original Schymanski criteria, a distinction is made between unknown molecular formulas and known molecular formulas. Within each of the adjusted identification levels there are several degrees of confidence and are related to the size, weight and number of fragments detected. However, as noted by Schymanski et al., such sublevels reduce the generic applicability of the confidence scale and should be defined on a per-study basis.

3.2. Application of FIF

In Fig. 3 several XICs of characteristic fragments are shown after analysis of a surface water, demonstrating the second step in the FIF workflow (Fig. 1). Applying FIF clearly yields a practical overview of the signals that are relevant for further investigation. Subsequently retention time matching in full scan analyses is applied to all observed signals to establish the respective molecular ion for each individual signal (step 3 in Fig. 1). All PFASs that were eventually detected in surface water are listed in Table 1 accompanied with the
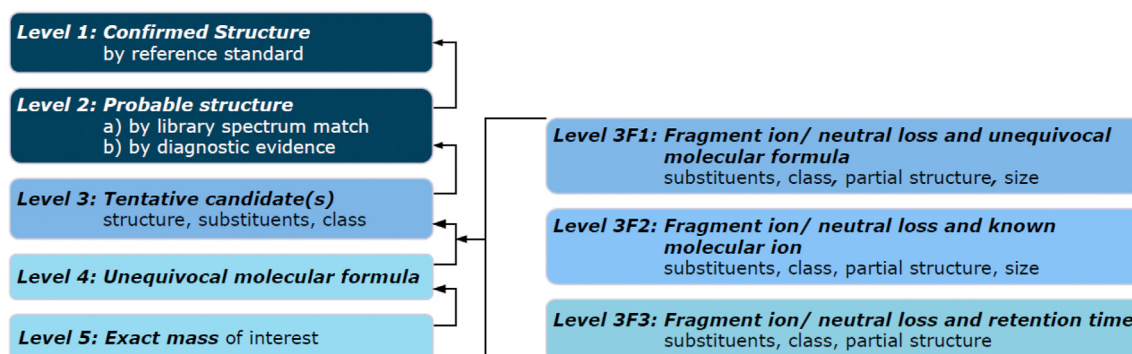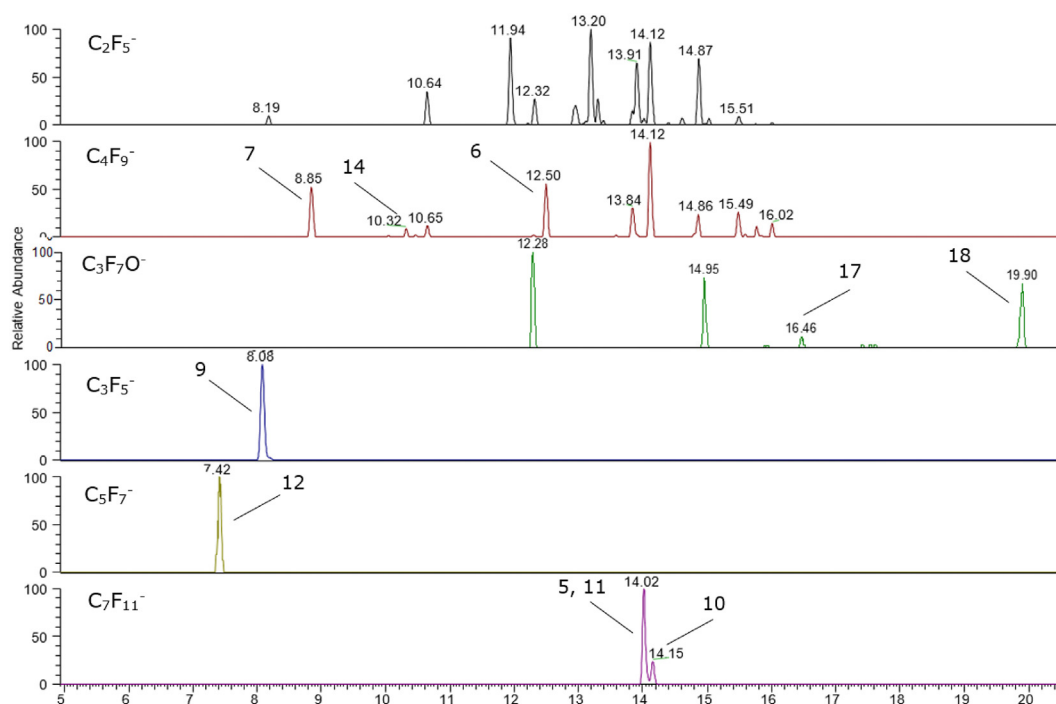
Fig. 2. Schematic showing the small molecule identification criteria as defined by Schymanski et al. (Schymanski et al., 2014) with the adapted criteria applicable for identification with FIF.

Fig. 3. XICs of frequently found fragments in surface water sampled in Zeeland. The signals above, found with FIF, have been annotated according to the numbered structures presented in Table 1. Several signals were only found in full scan and are not presented above. Most of the signals from the $C_2F_5^-$ and $C_4F_9^-$ XICs are largely yielded by PFCAs and PFSAs and have not been annotated to avoid overcomplication of the figure.

fragments that resulted in their detection with FIF. A complete list of PFASs and their masses and mass errors is included in the supporting information (Table S2 and Table S3).

Out of the 31 PFASs identified in surface waters across the Netherlands, 14 were identified as PFCAs (e.g. PFOA) and PFSAs (e.g. PFOS) and could be confirmed by reference standards (i.e. Schymanski level 1). Seven additional PFASs were identified as level 2, four as level 3 and four PFASs are currently identified as level 3F3, as discussed below.

PFCAs and PFSAs, structures 1 and 2 in Table 1, were observed with FIF by fragments in the series of $C_nF_{2n+1}^-$ and demonstrate the ease of PFAS FIF as seen in Fig. 3. The majority of PFCAs and PFSAs were found in all surface water samples and were confirmed as level 1 as discussed above. Four PFCAs and PFSAs, i.e. perfluoropropanoic acid, perfluoropropanesulfonic acid, perfluoropentanesulfonic acid and perfluorononanesulfonic acid were assigned a level 2a confidence as reference standards were not available in our lab at the time of study. GenX was detected in all surface waters as well (confirmed as level 1).
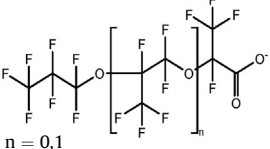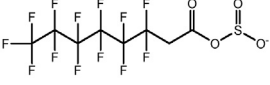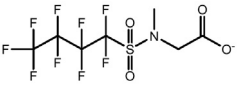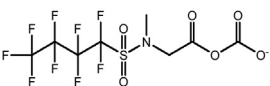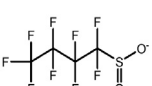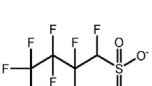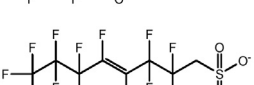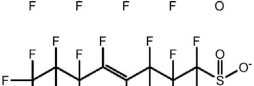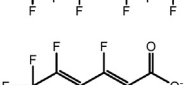
FIF scans revealed eight additional PFASs which were present in both the surface water samples and instrument blank measurements indicating these PFASs to be contaminants from the LC system. These included six H−PFCAs, which are PFCAs with a single F-atom substituted by a H-atom (Table 1, Structure 4). Five H−PFCAs, 4-H-perfluorobutanoic acid, 5-H-perfluoropentanoic acid, 7-H-perfluoroheptanoic acid, 8-H-perfluorooctanoic acid and 9-H-perfluorononanoic acid were confirmed by reference standards and were therefore assigned a level 1, mass spectra are given in the supporting information (Fig. S2 to Fig. S6). The H-PFCA 6-H-perfluorohexanoic acid was assigned a confidence of level 2b as a reference standard was not available. Although these H−PFCAs were found as contaminants from the LC system, Newton et al. (2017) found several H−PFCAs (C_4−C_7) in surface waters downstream of manufacturing plants. Hexafluoropropylene oxide trimer

acid (HFPO-TA, structure 3, n = 1), was also confirmed by a reference standard and therefore a level 1 was assigned. Two additional reoccurring peaks at 16.46 min and 19.90 min seen in the XIC of $C_3F_7O^-$ (Fig. 3) were assigned a level 3F3 confidence. The molecular ions of the latter two peaks could not be established, which led to the conclusion that these peaks are the product of complete insource fragmentation of compounds that were unstable at the instrumental settings in this study. This was further confirmed by applying AIF to masses with an m/z between 185 and 900: through this the mass of $C_3F_7O^-$ (m/z 184.9843) was not detected whereas isolation and fragmentation of masses below m/z 185 did result in the detection of the fragment. This indicates that the $C_3F_7O^-$ is a product of in-source fragmentation. As known from literature (Song et al., 2018) and our own experience with targeted methods, in-source fragmentation is a reoccurring problem seen in ether compounds including GenX and HFPO-TA. However, the retention times observed indicate these ether fragments are most likely yielded by compounds larger than GenX and HFPO-TA. One frequently mentioned fluoroether compound is HFPO-TeA, a compound belonging to the same homologous series. Fluoroether fragments were found at RT 16.46 and RT 19.90 in FIF (Song et al., 2018), of which one is speculated to be HFPO-TeA. Due to the lack of reference standards in our lab at the time of the study, these were not confirmed.

According to step 2 of the FIF workflow (Fig. 1), the masses for $C_3F_3^-$, $C_4F_5^-$, $C_6F_9^-$ and $C_7F_{11}^-$ were flagged in AIF eluting at 14.02 min as seen in Fig. 3. Through retention time matching (step 3 in Fig. 1), two suspect molecular ions, i.e. m/z 440.9478 and m/z 460.9340, were found in full scan in all surface water samples. Subsequently, both masses were isolated using DIA, and mass spectra were recorded (step 4 of the FIF workflow). Both suspect molecular ions were determined to be PFASs yielding similar characteristic fragments. The ion m/z 460.9340, structure 11, will be discussed later. With an unknown molecular formula but known precursor ion,

Table 1
PFASs (tentatively) identified in surface water samples sampled throughout the Netherlands as well as PFASs identified as background signals. Surface water samples are annotated by a single letter; T: Tuindorp, Z: Zeeland, A: Amsterdam, D: Dordrecht. Background signals are annotated with the letter 'B'.

Structure[a]	Molecular formula	mass	Confidence level	Ions detected in FIF	Found in
1[b]	Varies[c]		1 (n = 3 level 2)	$C_nF_{2n+1}^-$	Varies[c]
2[b]			1 (n = 3, 5, 9 level 2)	$C_nF_{2n+1}^-$	Varies[c]
3			1	$C_2F_5^-$, $C_3F_7O^-$ (for n = 0 $C_2F_5O^-$ as well)	Varies[c]
4[b]			1 (n = 6 level 2)	$C_nF_{2n-1}^-$	B
5	$C_8H_2F_{13}O_4S^-$	440.94717	2b	$C_3F_3^-$, $C_4F_5^-$, $C_6F_9^-$, $C_7F_{11}^-$	T, Z, A, D
6	$C_7H_5F_9NO_4S^-$	369.98011	2a	$C_4F_9^-$	T, Z, A, D
7	$C_8H_5F_9NO_6S^-$	413.96994	2b	$C_4F_9^-$	Z, A, D
8	$C_4F_9O_2S^-$	282.94808	2b	$C_4F_9^-$	Z, A
9	$C_4HF_8O_3S^-$	280.95241	3	$C_3F_5^-$, $C_4F_7^-$	T, Z, A, D
10	$C_8H_2F_{13}O_3S^-$	424.95226	3	$C_3F_3^-$, $C_4F_5^-$, $C_6F_9^-$, $C_7F_{11}^-$	T, Z, A, D
11	$C_8F_{15}O_3S^-$	460.93341	3	$C_3F_3^-$, $C_4F_5^-$, $C_6F_9^-$, $C_7F_{11}^-$	T, Z, A, D
12	$C_6F_7O_2^-$	236.97920	3	$C_5F_7^-$	Z
13	$C_4F_9^-$ at RT 9.49 min	N/A	3F3	$C_4F_9^-$	Z
14	$C_4F_9^-$ at RT 10.33 min		3F3	$C_4F_9^-$	Z
15	$C_3F_3^-$ at RT 11.86 min		3F3	$C_3F_3^-$	A
16	$C_4F_5^-$ at RT 7.53 min		3F3	$C_4F_5^-$	A
17	$C_2F_5^-$ and $C_3F_7O^-$ at RT 16.45 min		3F3	$C_2F_5^-$, $C_3F_7O^-$	B
18	$C_2F_5^-$ and $C_3F_7O_-$ at RT 19.91		3F3	$C_2F_5^-$, $C_3F_7O^-$	B

[a] Structures with confirmation level 1 were confirmed and other structures are proposed structures with different levels of confidence.
[b] A complete list of fragments detected in FIF is found in Table S2.
[c] Compounds of varying lengths were not observed in all samples. Complete list found in Table S3.

these signals can be assigned a 3F2 confidence. As noted in the mass spectrum for structure 5, *m/z* 440.9478, the difference between the molecular ion and the fragment ion $C_7F_{11}^-$ belonged to the consecutive neutral losses of SO_3HF and $COHF$ as seen in Fig. 4, resulting in the unequivocal molecular formula of $C_8H_2O_4F_{13}S^-$ for compound

5, at this point the compound is assigned a level 3F1 confidence.

The first neutral loss of SO_3HF and the presence of an SO_3H^- peak indicates the presence of a sulfonate group. Due to the lack of an SO_4^-/SO_4H^- peak and the presence of a neutral loss of $COHF$ it can be concluded the sulfonate group is not a sulphate group. With
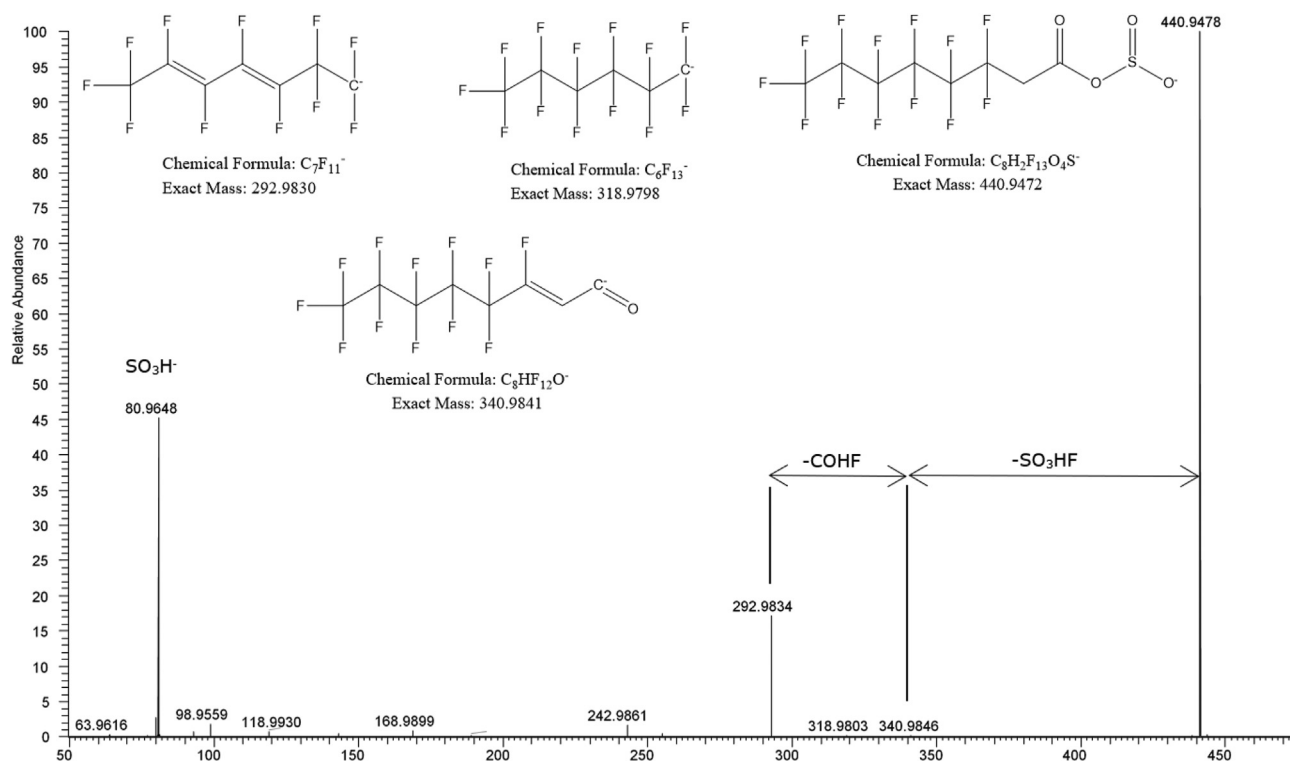
Fig. 4. Mass spectrum and assigned fragments of $C_8H_2F_{13}O_4S^-$, structure 5. The fragments essential to the identification of the compound have been noted in the figure, other fragments were detected but were at low abundance and/or not essential to the identification process, these fragments have been included in Fig. S7.

the molecular formula, the double bond equivalent can be calculated. Not including the SO_3 group, a singular double bond is calculated to be present. In the mass spectrum a saturated carbon chain, C_6F_{13}, is noted, leaving no room for a double bond. Consequently, the double bond is located between the fluorinated carbon chain and the SO_3 end group. Considering the lack of an SO_4^- fragment it can be concluded the fourth oxygen atom is double bonded to carbon atom neighbouring the sulfonate group. Lastly, the presence of an SO_2 (m/z 63.9616) peak leads to the orientation of the sulfonate group demonstrated in Table 1. The proposed structure derived from the above spectrum has not been reported earlier to the best of our knowledge and lacking a possibility for further confirmation (due to a lacking standard), we assigned this compound a confidence level 2b.

Structure 6 was found by flagging the fragment $C_4F_9^-$ eluting at RT 12.5 min and was found in all surface water samples. Following step 2 and 3 of the workflow (Fig. 1) the precursor ion of m/z 369.9800 was determined giving a level 3F2 confidence. Subsequently, a mass spectrum of this ion was recorded (Fig. S8) revealing the consecutive neutral losses of $C_2H_2O_2$, NCH_3 and SO_2 giving the unequivocal molecular formula $C_7H_5F_9NO_4S^-$. The structure as assigned to compound 6 in Table 1 was reported earlier by Newton et al. (2017) as MeFBSAA (N-methyl perfluorobutane sulfonamidoacetic acid) and they confirmed this as level 1. Based on the match between their and our spectra, but lacking a confirmation by reference standard in our case, this compound was assigned a level 2a confidence.

Eluting at RT 8.85 min the fragment $C_4F_9^-$ was found leading to the precursor ion m/z 413.9700 which was found in surface water from Zeeland, Dordrecht and Amsterdam. After isolation with DIA as described in the Experimental section (mass spectra included in Fig. S9), compound 7 demonstrated near identical fragmentation as compound 6, differentiated only by an additional CO_2 loss.

Fragmentation of structure 7 suggests the compound to be a linear PFAS, resulting in the neutral loss of CO_2, followed by the loss of $C_2H_2O_2$, NCH_3 and SO_2, as seen with structure 6 as well. The fragmentation pattern of structure 7 led to the structure as proposed in Table 1. To the best of our knowledge, this compound has not been previously mentioned in any literature or database and was given a level 2b confidence.

Concerning structure 8, to cover insource fragmentation during full scan, characteristic fragments were also flagged during full scan analyses. In doing so, an additional C_4F_9 peak was observed at RT 10.96 min in surface water from Zeeland and Amsterdam. This fragment being the product of insource fragmentation might also contribute to the fact that the fragment was not observed in AIF. Nevertheless, a loss of SO_2 was noted in the full scan data leading to the precursor ion m/z 282.9480. This ion was isolated with DIA, and an MS/MS spectrum was acquired and can be found in Fig. S10. The detection of SO_2 loss was accompanied by an SO_2F^- peak led to the unambiguous structure as given in Table 1 and was assigned a level 2b confidence accordingly.

At RT 8.08 min the m/z values for $C_3F_5^-$ and $C_4F_7^-$ were observed with FIF in all surface water samples. The m/z value 280.9528 was determined to be the molecular ion for structure 9, as from this ion a neutral loss of SO_3HF is seen, resulting in a $C_4F_7^-$ fragment ion. Thus, yielding the molecular formula $C_4HF_8SO_3^-$ for compound 9. After isolation three consecutive neutral losses were observed, HF, SO_2 and O. The loss of SO_2 would suggest an SO_2 group bonded to the fluorinated carbon chain with an ether bond. However, the loss of HF preceding the SO_2 loss creates a double bond in the vicinity of the sulphur containing moiety. Sulfonate groups are known to rearrange when a double bond in nearby (Wang et al., 2003; Binkley et al., 1993), resulting in a neutral loss of SO_2. The lack of an SO_2 ion peak supports this theory. The position of the hydrogen atom is unknown as mass spectral information does not provide

such information. Hence this structure was given a level 3 confidence. Mass spectral information can be found in Fig. S11.

Eluting at 14.15 min the ions $C_3F_3^-$, $C_4F_5^-$, $C_6F_9^-$ and $C_7F_{11}^-$ were flagged in all surface water samples. The molecular ion of m/z 424.9536 was suspected of being the precursor ion as a neutral loss of $SO_3CH_2F_2$ was recognized leaving the $C_7F_{11}^-$ ion. The neutral loss and the $C_7F_{11}^-$ fragment yield the formula $C_8H_2F_{13}O_3S^-$ for compound 10. With this, the double bond equivalent is calculated to be 1. Unfortunately, the location of the double bond and the two hydrogen atoms cannot be determined with mass spectrometry, leaving this structure at a level 3 confidence. The MS/MS spectrum and fragment assignment is included in Fig. S12.

The flagging of the ions $C_3F_3^-$, $C_4F_5^-$, $C_6F_9^-$ and $C_7F_{11}^-$ eluting at 14.02 min in all surface water samples led to two suspect precursor ions which were both found to be PFASs, m/z 440.9478 (structure 5, discussed above) and m/z 460.9340. Structure 11, m/z 460.9340, is near identical to PFOS, differentiated only by a singular double bond in the fluorinated carbon chain. This compound fragments predictably as no other functional groups or hydrogen atoms are present, hence yielding fluorinated carbon chain fragments and fragments such as SO_3^- and SO_3F^-. This compound was assigned a level 3 confidence as the position of the double bond could not be determined. However, chromatography suggests multiple isomers might be present, the spectra and chromatography are included in Fig. S13.

Structure 12 was tentatively identified after the m/z value of $C_5F_7^-$ was found in AIF eluting at 7.42 min in surface water sampled in Zeeland. In the full scan mass spectrum, a neutral loss of CO_2 was observed leading to the molecular ion 236.9792. Isolation of this mass provided a mass spectrum (Fig. S14) with one additional peak of interest, m/z 108.9911. The molecular formula of this fragment ion was calculated giving the unequivocal formula $C_3OF_3^-$, possibly yielded by rearrangement of the ions. The detection of a CO_2 neutral loss indicates this compound to be a carboxylic acid. Unfortunately, both double bond positions are currently unknown leaving compound 12 at a level 3 confidence.

Six additional peaks yielded by characteristic fragments were found. However, these PFASs have yet to be identified. As discussed earlier, two ether fragments were found that most likely originate from products of complete in-source fragmentation. Therefore, the determination of the molecular ion proved to be extremely difficult. A possible solution is to confirm by reference standard based on retention time match and in-source fragments match. And possibly, the in-source fragmentation can be minimised by softer ionization conditions or injection at another MS platform. However, this was not further explored in this study.

The fragments $C_3F_3^-$ and $C_4F_5^-$ were detected at RT 11.86 min and RT 7.43 min respectively in Amsterdam surface water. The m/z value of $C_4F_9^-$ yielded two additional peaks at RT 9.49 min in full scan and at RT 10.32 min in AIF in surface water from Zeeland. These PFASs were assigned a 3F3 confidence as the precursor ion could not be identified with sufficient certainty.

4. Conclusions

The use of FIF has proven to be a useful tool in the non-targeted identification of PFASs. Multiple PFASs not present in homologous series demonstrate the shortcomings of traditional full scan feature selection. Moreover, to reduce data in full scan feature selection, intensity threshold filtering is often used removing trace level compounds as well as compounds with low ionization efficiency. FIF circumvents the dependency on both intensity threshold filtering and homologous series detection which allowed the identification of PFASs which would not have been found otherwise. The application of suspect screening also yielded the

detection of fewer PFASs compared to FIF. In this study we demonstrate that with this approach identification of unknown PFASs is feasible at background concentrations in environmental samples, after SPE enrichment and clean-up. Although we focussed on anionic species, we assume that FIF is also applicable to cationic and zwitterionic PFASs using the appropriate sample pretreatment. We also expect that this approach will also be applicable to other sample types including fish, human milk and serum. The practicality of the FIF could be further increased with software capable of identifying neutral losses, allowing the identification of PFASs yielding non-characteristic fragments. Moreover, using FIF to investigate fragments with the formula $C_xH_yF_z^-$ could extend the reach of FIF even further.

Credit Author Statement

Thijs J Hensema: Conceptualization, Methodology, Formal analysis, Investigation, Writing — Original Draft, Visualization. **Bjorn JA Berendsen**: Conceptualization, Methodology, Validation, Writing — Review & Editing, Supervision, Visualization. **Stefan PJ van Leeuwen**: Conceptualization, Resources, Writing — Review & Editing, Visualization, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2020.128599.

References

PFASs, 2018. summary Report on Updating the OECD 2007 List of Per- and Polyfluoroalkyl Substances.

SC-9/12, 2019. Listing of Perfluorooctanoic Acid (PFOA), its Salts and PFOA-Related Compounds, pp. 2—4.

OECD/UNEP Global PFC Group, United Nations Environment Programme, 2013. Synthesis paper on per- and polyfluorinated chemicals (PFCs), environment, health and safety, environment directorate, OECD. In: IOMC Inter-organization Program. Sound Manag. Chem., pp. 1—58

Annex XVII to REACH, 2020. ECHA, pp. 4—6.

Barzen-Hanson, K.A., Roberts, S.C., Choyke, S., Oetjen, K., McAlees, A., Riddell, N., McCrindle, R., Ferguson, P.L., Higgins, C.P., Field, J.A., 2017. Discovery of 40 classes of per- and polyfluoroalkyl substances in historical aqueous film-forming foams (AFFFs) and AFFF-impacted groundwater. Environ. Sci. Technol. 51, 2047—2057. https://doi.org/10.1021/acs.est.6b05843.

Berendsen, B.J.A., Lakraoui, F., Leenders, L., van Leeuwen, S.P.J., 2020. The analysis of perfluoroalkyl substances at ppt level in milk and egg using UHPLC-MS/MS. Food Addit. Contam. Part A Chem. Anal. Control. Expo. Risk Assess. https://doi.org/10.1080/19440049.2020.1794053.

Binkley, R.W., Flechtner, T.W., Tevesz, M.J.S., Winnik, W., Zhong, B., 1993. Rearrangement of aromatic sulfonate anions in the gas phase. Org. Mass Spectrom. https://doi.org/10.1002/oms.1210280708.

Buck, R.C., Franklin, J., Berger, U., Conder, J.M., Cousins, I.T., De Voogt, P., Jensen, A.A., Kannan, K., Mabury, S.A., van Leeuwen, S.P.J., 2011. Perfluoroalkyl and polyfluoroalkyl substances in the environment: terminology, classification, and origins, Integr. Environ. Assess. Manag. 7, 513—541. https://doi.org/10.1002/ieam.258.

D'Agostino, L.A., Mabury, S.A., 2014. Identification of novel fluorinated surfactants in

aqueous film forming foams and commercial surfactant concentrates. Environ. Sci. Technol. 48, 121–129. https://doi.org/10.1021/es403729e.

Gebbink, W.A., van Leeuwen, S.P.J., 2020. Environmental contamination and human exposure to PFASs near a fluorochemical production plant: review of historic and current PFOA and GenX contamination in The Netherlands. Environ. Int. 137 https://doi.org/10.1016/j.envint.2020.105583.

Gebbink, W.A., Van Asseldonk, L., Van Leeuwen, S.P.J., 2017. Presence of emerging per- and polyfluoroalkyl substances (PFASs) in river and drinking water near a fluorochemical production plant in The Netherlands. Environ. Sci. Technol. 51, 11057–11065. https://doi.org/10.1021/acs.est.7b02488.

Hatton, J., Holton, C., DiGuiseppi, B., 2018. Occurrence and behavior of per- and polyfluoroalkyl substances from aqueous film-forming foam in groundwater systems. Remediation 28, 89–99. https://doi.org/10.1002/rem.21552.

Heydebreck, F., Tang, J., Xie, Z., Ebinghaus, R., 2015. Alternative and legacy per- fluoroalkyl substances: differences between European and Chinese river/estu- ary systems. Environ. Sci. Technol. 49, 8386–8395. https://doi.org/10.1021/acs.est.5b01648.

Houde, M., De Silva, A.O., Muir, D.C.G., Letcher, R.J., 2011. Monitoring of per- fluorinated compounds in aquatic biota: an updated review. Environ. Sci. Technol. 45, 7962–7973. https://doi.org/10.1021/es104326w.

Kind, T., Fiehn, O., 2007. Seven Golden Rules for Heuristic Filtering of Molecular Formulas Obtained by Accurate Mass Spectrometry, vol. 20, pp. 1–20. https://doi.org/10.1186/1471-2105-8-105.

Liu, Y., Pereira, A.D.S., Martin, J.W., 2015. Discovery of C5-C17 Poly- and per- fluoroalkyl substances in water by in-line Spe-HPLC-Orbitrap with in-source fragmentation flagging. Anal. Chem. 87, 4260–4268. https://doi.org/10.1021/acs.analchem.5b00039.

Liu, Y., Qian, M., Ma, X., Zhu, L., Martin, J.W., 2018. Nontarget mass spectrometry reveals new perfluoroalkyl substances in fish from the yangtze river and tangxun lake, China. Environ. Sci. Technol. 52, 5830–5840. https://doi.org/10.1021/acs.est.8b00779.

Lommen, A., 2014. Ultrafast PubChem searching combined with improved filtering rules for elemental composition analysis. Anal. Chem. 86, 5463–5469. https://doi.org/10.1021/ac500667h.

Loos, M., Singer, H., 2017. Nontargeted homologue series extraction from hyphen- ated high resolution mass spectrometry data. J. Cheminf. 9 https://doi.org/10.1186/s13321-017-0197-z.

McCord, J., Strynar, M., 2019. Identification of per- and polyfluoroalkyl substances in the cape fear river by high resolution mass spectrometry and nontargeted screening. Environ. Sci. Technol. 53, 4717–4727. https://doi.org/10.1021/acs.est.8b06017.

Newton, S., McMahen, R., Stoeckel, J.A., Chislock, M., Lindstrom, A., Strynar, M., 2017. Novel polyfluorinated compounds identified using high resolution mass spec- trometry downstream of manufacturing facilities near decatur, Alabama. En- viron. Sci. Technol. 51, 1544–1552. https://doi.org/10.1021/acs.est.6b05330.

Prevedouros, K., Cousins, I.T., Buck, R.C., Korzeniowski, S.H., 2006. Sources, fate and transport of perfluorocarboxylates. Environ. Sci. Technol. 40, 32–44. https://doi.org/10.1021/es0512475.

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ. Sci. Technol. 48, 2097–2098. https://doi.org/10.1021/es5002105.

Song, X., Vestergren, R., Shi, Y., Huang, J., Cai, Y., 2018. Emissions, transport, and fate of emerging per- and polyfluoroalkyl substances from one of the major fluo- ropolymer manufacturing facilities in China. Environ. Sci. Technol. 52, 9694–9703. https://doi.org/10.1021/acs.est.7b06657.

Strynar, M., Dagnino, S., McMahen, R., Liang, S., Lindstrom, A., Andersen, E., McMillan, L., Thurman, M., Ferrer, I., Ball, C., 2015. Identification of novel per- fluoroalkyl ether carboxylic acids (PFECAs) and sulfonic acids (PFESAs) in nat- ural waters using accurate mass time-of-flight mass spectrometry (TOFMS). Environ. Sci. Technol. 49, 11622–11630. https://doi.org/10.1021/acs.est.5b01215.

NORMAN suspect list exchange (n.d.). https://www.norman-network.com/?q=node/236. (Accessed 22 January 2019).

Wang, Z., Hop, C.E.C.A., Kim, M.S., Huskey, S.E.W., Baillie, T.A., Guan, Z., 2003. The unanticipated loss of SO2 from sulfonamides in collision-induced dissociation. Rapid Commun. Mass Spectrom. 17, 81–86. https://doi.org/10.1002/rcm.877.

Wang, Z., Cousins, I.T., Scheringer, M., Hungerbühler, K., 2013. Fluorinated alterna- tives to long-chain perfluoroalkyl carboxylic acids (PFCAs), perfluoroalkane sulfonic acids (PFSAs) and their potential precursors. Environ. Int. 60, 242–248. https://doi.org/10.1016/j.envint.2013.08.021.

Williams, A.J., Grulke, C.M., Edwards, J., McEachran, A.D., Mansouri, K., Baker, N.C., Patlewicz, G., Shah, I., Wambaugh, J.F., Judson, R.S., Richard, A.M., 2017. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. J. Cheminf. 9, 1–27. https://doi.org/10.1186/s13321-017-0247-6.

Xiao, F., Golovko, S.A., Golovko, M.Y., 2017. Identification of novel non-ionic, cationic, zwitterionic, and anionic polyfluoroalkyl substances using UPLC–TOF–MShigh-resolution parent ion search. Anal. Chim. Acta 988, 41–49. https://doi.org/10.1016/j.aca.2017.08.016.

Yu, N., Guo, H., Yang, J., Jin, L., Wang, X., Shi, W., Zhang, X., Yu, H., Wei, S., 2018. Non- target and suspect screening of per- and polyfluoroalkyl substances in airborne particulate matter in China. Environ. Sci. Technol. 52, 8205–8214. https://doi.org/10.1021/acs.est.8b02492.