Project ID N°: **101036449**

Call: **H2020-LC-GD-2020-3**
Topic: **LC-GD-8-1-2020** - Innovative, systemic zero-pollution solutions to protect health, environment, and natural resources from persistent and mobile chemicals

Preventing **R**ecalcitrant **O**rganic **M**obile **I**ndustrial chemical**S** for **C**ircular **E**conomy in the soil-sediment-water **S**ystem

Start date of the project: **1st November 2021**

Duration: **36 months**

D1.4 - GC- and LC-HRMS methods and data treatments workflows for suspect screening and retrospective identification of relevant CECs in the water cycle and for identifying transformation products of PFAS during remediation process

Main authors: **Christian Dietrich (BAFG), Alicia Hartmann (BAFG), Alessandro Frugis (ACEA), Marco Mancini (ACEA), Valentina Gioia (ACEA), Marco Lazzazzara (ACEA), Dana Pierina Orlando-Veliz (IDAEA-CSIC), Manuel García Vara (IDAEA-CSIC), Rocío Inés Bonansea (IDAEA-CSIC), Marta Llorca (IDAEA-CSIC), Marinella Farré (IDAEA-CSIC), Miren López de Alda (IDAEA-CSIC), Frederik Zietschmann (BWB), Tobias Hensel (BWB)**
Lead Beneficiary: **BAFG**
Type of delivery: **R**
Dissemination Level: **PU**
Filename and version: **PROMISCES_D1-4_SuspectScreening (version 1)**
Website**: https://promisces.eu/**

Due date: **31 October 2024**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°101036449

Disclaimer

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Document History

This document has been through the following revisions:

Version	date	Author/Reviewer	Description
0.1	14/10/2024	C. Dietrich	First draft for review
0.2	07/11/2024	C. Dietrich	Second draft for review
0.3	14/11/2024	C. Dietrich	Implementation of review comments
0.4	28/11/2024	F. Sermondadaz	Version after QC
1.0	28/11/2024	C. Dietrich	Final Version for distribution

Authorisation

Authorisation	Name	Status	Date
Review	L. Amalric	Reviewer BRGM head of laboratory	08/11/2024
Validation	A. Togola	WP1 leader	25/11/2024
Quality Control	F. Sermondadaz	Administrative and financial manager	28/11/2024
Approval	J. Lions	Project coordinator	29/11/2024

Distribution

This document has been distributed to:

Name	Title	Version issued	Date of issue
All WP1 partners	WP1 Partners	Version 1	29/11/2024

Executive Summary

A large number of anthropogenic substances including PFAS and iPM(T)s are present in the water cycle. For most of these compounds, information about occurrence, sources and fate are scarce as targeted methods alone cannot cope with the large variety of these substances.

This document reports on the results regarding development of suspect screening workflows for PFAS and other iPM(T)s realized within the framework of the task dealing with "Suspect screening and database-assisted analysis".

Five different workflows have been developed using different analytical setups for analyses of different matrices ranging from ground- and surface water over wastewater and landfill leachates to lettuce and sediment and. Two workflows are based on liquid chromatography coupled to high-resolution mass spectrometry and are specifically designed for the screening of PFAS compounds in landfill leachates and their transformation products in ground- and surface water, wastewater treatment plant effluents, lettuce and sediment. Two other LC-HRMS-based workflows are dedicated to the identification of iPM(T)s in surface- and wastewater. These workflows are complemented by a GC-MS-based method enabling a screening for more unpolar compounds not amenable to most LC-MS methods and is dedicated to industrial wastewater.

Subsequent data treatment is performed using commercial as well as open-source software. Matching of the spectral information of the analytical data was performed with both commercial and public compound databases (massbank.eu and mzcloud.org). An overview about the main aspects of the five workflows is given in Table 1.

The developed workflows have been applied to various samples from the PROMISCES case studies CS#1, CS#2, CS#3, CS#4 and CS#7.

Table 1: Comparative summary of the five methods.

	ACEA	CSIC 1	BAFG	CSIC 2	BWB
Chromatography	HPLC	HPLC	HPLC	HPLC	GC
Analyser	QToF	QExactive	QToF	QToF	GC-MS
Analytes	PFAS	PFAS TPs	iPM(T)s	iPM(T)s	iPM(T)s
Matrix	landfill leachates	Surface water, groundwater, WWTP effluent, lettuce, sediments	Surface water, WWTP effluent	Waste water	WWTP effluent, Industrial wastewater
Sample preparation	Dilution and addition of internal standards	Centrifugation and addition of mix internal standards	Filtration	Addition of internal standards, lyophilization, and sample extract reconstitution	Addition of internal standards, SPE-Enrichment
Software	SCIEX OS, SCIEX LibraryView	Thermo Scientific Compound Discoverer 3.3 with public and homemade libraries	Homemade R scripts	Bruker MetaboScape® 2022b	Agilent Masshunter, Excel-script
Database	SCIEX Fluorochemical HR-MS/MS library 2.0	MassBank (massbank.eu) and NIST mass spectral libraries	In-House spectral database (also available on massbank.eu))	mzCloud (mzcloud.org)	NIST Mass Spectral Library, in-house database

Table of contents

List of abbreviations

bbCID	Broadband Collision Induced Dissociation
CE	Collision Energy
CS	Case Study
Da	Dalton
ESI	Electrospray Ionisation
GC	Gas Chromatography
HPLC	High-Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
IDA	information dependent acquisition
iPM(T)s	Industrial, Persistent, Mobile and potentially Toxic substances
IS	Internal Standard
LC	Liquid Chromatography
LOD	Limit of Detection
LOQ	Limit of Quantification
MeOH	Methanol
MS	Mass Spectrometer
MS/MS	Triple Quadrupole Mass Spectrometer
m/z	mass-to-charge ratio
NTS	Non-Target Screening
QToF	Quadrupole – Time-of-Flight
PFAS	Per- and polyfluoroalkyl substances
PTFE	Polytetrafluoroethylene
ppm	parts per million
RPM	Rounds Per Minute
RT	Retention Time
S/N	Signal-to-Noise Ratio
SPE	Solid Phase Extraction
TIC	Total Ion Chromatogram
ToF	Time-of-Flight
UPLC	Ultra-high Performance Liquid Chromatography
WP	Work Package
WWTP	Wastewater Treatment Plant
XIC	eXtracted Ion Chromatogram

List of Figures

List of Tables

1 Introduction

A large number of anthropogenic substances are present within the urban water cycle. Some of them may be hazardous to biota and biosystems or may affect human health and are therefore under regular observation. Considering the large variety of different PFAS and the overall high number of industrial chemicals, targeted methods for a limited number of well-known chemicals are not sufficient to assess the highly diverse occurrences in the water cycle.

One way to overcome this issue is screening for a large number of substances via database-assisted suspect screening. Depending on the analytical method and the corresponding database in use one could perform a screening for far more than 1000 substances with a single analytical method. As a further advantage, instrumental analyses of suspect screening approaches are intrinsically independent of the subsequent assignment of suspected compounds. This enables a retrospective screening of hitherto unregarded substances with an updated database even years after generation of the raw data. Nevertheless, such untargeted approaches cannot replace validated target methods as suspect screening data usually comes without internal standards and calibration for the substances in question. Data acquired from suspect screening is therefore in most cases limited to qualitative information. In order to unambiguously identify a compound with a reference standard and to obtain robust quantitative data a subsequent analysis using targeted methods, as provided in Deliverable D1.2 *"Targeted methods for relevant iPM(T) substances in waters ",* is necessary.

This deliverable reports on five different workflows for a database-assisted suspect screening of PFAS compounds, their transformation products and other iPM(T)s in various matrices using different analytical approaches.

2 Analytical methods and data treatment

2.1 Large suspect screening of PFAS in landfill leachates (ACEA)

2.1.1 General

In Italy, three full-scale landfill leachate treatment plants (LTP) were investigated in three different regions of the country (i.e., Marche, Veneto, Liguria) to perform analysis of target and suspect PFAS compounds. In two of the investigated landfill leachate plants (LTP1 and LTP2) conventional leachate treatments (i.e., clari-flocculation, biological process, ultrafiltration) are implemented, whereas in a third plant (LTP3) advanced technologies (i.e., reverse osmosis) are utilized.

In order to search the suspected PFAS, it was used a workflow involving: the preparation and pre-treatment of leachate samples, the addition of IS in the injectable liquid to the UPLC-Q-TOF, the acquisition of the chromatogram and the application of a peak recognition procedure for peaks above a certain threshold intensity using commercial software and a commercial library that contains information for the recognition of 261 PFASs.

The use of IS in the test sample improved the quality of PFAS suspect detection and provided a tool to improve the reliability of identification and for the semi-quantitative estimation of identified PFAS concentrations without the need for a subsequent analysis via targeted methods.

2.1.2 Chemicals and reagents

- Solvents (Acetonitrile, Water, Methanol), LC-MS grade, brand Biosolve.
- Formic acid ≥99%, LC-MS grade, brand VWR.
- Ammonium acetate, LC-MS grade >99%, brand VWR.
- 1 M ammonium acetate:
  - weigh 77.08 g of ammonium acetate, transfer to a 1000 mL flask and make up to volume with water, LC-MS grade;
  - filter with cellulose acetate, porosity of 0.2-0.5 µm.

  The solution, stored in a refrigerator at 2-8 °C, is considered stable for six months.

- **Mobile phase A**, Ammonium acetate 5 mM in Water:
  - pour into a 1000 mL cylinder, 100-200 mL of LC-MS water;
  - add 5 ml of 1 molar ammonium acetate (M);
  - make up to volume (1 liter) with water, LC-MS grade;
  - transfer to 1 liter bottle and shake vigorously.

  The solution, stored at room temperature in amber container, is stable for 14 days.

- **Mobile Phase B**:
  - Acetonitrile, LC-MS grade.

  The solution, stored at room temperature in amber container, is stable for one month.

As internal standard it was used a labelled mixture which contains the following compounds:

Table 2: List of Internal Standards used in the ACEA workflow.

MPFAC-24ES (Internal Standard)
M4PFBA
M5PFPeA
M5PFHxA
M4PFHpA
M8PFOA
M9PFNA
M6PFDA
M7PFUnDA
M2PFDoDA
M2PFTeDA
M8FOSA
d3-N-MeFOSAA
d5-N-EtFOSAA
M3PFBS
M3PFHxS
M8PFOS
M2-4:2 FTS
M2-6:2 FTS
M2-8:2 FTS
MPFAC-24ES (Internal Standard)
M4PFBA
M5PFPeA
M5PFHxA
M4PFHpA
M8PFOA
M9PFNA
M6PFDA
M7PFUnDA

2.1.3 Sample preparation and pre-treatment

Leachate samples were diluted by 1:100 ratio and fortified with 20 IS (labelled PFAS) at the concentration of 200 ng/L in the final liquid extract.

2.1.4 Chromatography

Mobile phase A: ammonium acetate 5 mmol in water

Mobile phase B: Acetonitrile 100%

Table 3: Chromatographic parameters used in the ACEA LC-HRMS method.

Elution gradient:

Time (min)	Flow (mL/min)	A (%)	B (%)
0.0	0.3	98.0	2.0
0.50	"	98.0	2.0
14.0	"	5.0	95.0
16.0	"	5.0	95.0
16.10	"	98.0	2.0
22.00	"	98.0	2.0

Column Oven: 40 °C

Autosampler: 15 µL/s

2.1.5 Mass spectrometry

The Electrospray Ionization (ESI) mass spectrometer operates in negative mode. It always acquires a full scan and a MSMS spectra. For the MSMS acquisition it can record data in two modalities: namely SWATH and IDA.

2.1.5.1 Criteria for SWATH acquisition (Sequential Window Acquisition of all Theoretical Mass Spectra)

TOF MS
TOF Start Mass: 200 Da
TOF stop mass: 850 Da
Accumulation time: 0.05s
Time bins to sum: 4

TOF MSMS
TOF Start Mass: 50 Da
TOF stop mass: 600 Da
Accumulation time: 0.04s
Method duration: 22min
Total scan time: 1.273 s
Estimated cycle: 1037

NB: the spectrometer always acquires both a Full Scan and a MSMS spectrum of the SWATH experiment.

For the so called SWATH mode (Sequential Window Aquisition of all Theoretical Mass Spectra), the acquisition windows are setted as found in Table 5.

Table 4: MS-Instrument parameters for suspect screening of PFAS in the ACEA LC-HRMS method.

Window	Precursor ion start mass (Da)	Precursor ion stop mass (Da)	Declustering potential (V)	DP Spread (V)	Collision energy (V)	CE Spread (V)	Time bins to sum*
1	200.0000	215.0000	-35	0	-30	10	6
2	214.0000	240.0000	-35	0	-30	10	6
3	239.0000	252.0000	-35	0	-30	10	6
4	251.0000	275.0000	-50	0	-35	30	6
5	274.0000	300.0000	-50	0	-35	30	6
6	299.0000	325.0000	-50	0	-35	30	6
7	324.0000	350.0000	-50	0	-35	30	6
8	349.0000	375.0000	-50	0	-35	30	6
9	374.0000	400.0000	-50	0	-35	30	6
10	399.0000	425.0000	-50	0	-35	30	6
11	424.0000	450.0000	-50	0	-35	30	6
12	449.0000	475.0000	-50	0	-35	30	6
13	474.0000	500.0000	-50	0	-35	30	6
14	499.0000	525.0000	-50	0	-35	30	6
15	524.0000	550.0000	-50	0	-35	30	6
16	549.0000	575.0000	-50	0	-35	30	6
17	574.0000	595.0000	-50	0	-35	30	6
18	594.0000	625.0000	-50	0	-35	30	6
19	624.0000	650.0000	-50	0	-35	30	6
20	649.0000	675.0000	-50	0	-35	30	6
21	674.0000	700.0000	-50	0	-35	30	6
22	699.0000	725.0000	-50	0	-35	30	6
23	724.0000	750.0000	-50	0	-35	30	6
24	749.0000	775.0000	-50	0	-35	30	6
25	774.0000	800.0000	-50	0	-35	30	6
26	799.0000	825.0000	-50	0	-35	30	6
27	824.0000	850.0000	-50	0	-35	30	6

*The bins to sum indicates how many ions the detector collects before returning the signal. Thus, the higher will be the bins to sum, the better the sensitivity at the expense of resolution and accuracy.

2.1.5.2 Criteria for IDA acquisition (Information Dependent Acquisition)

TOF MS
TOF Start Mass: 200 Da
TOF stop mass: 850 Da
Accumulation time: 0.15s
Time bins to sum: 4

IDA criteria
Maximum candidate ions: 17
Intensity threshold exceeds: 1000 cps
Exclude former candidate ions:
- for: 7s
- After: 3 occurences

Mass tolerance: +/- 50 mDa

TOF MSMS
TOF Start Mass: 50 Da
TOF stop mass: 850 Da
Accumulation time: 0.06

2.1.6 Data treatment

The SCIEX OS software processes the acquired chromatograms and performs the identification procedures by making use of the Library View Software application. Library View Software is a container of commercial and open source libraries. Library View Software has been implemented with the library fluorochemical_HR-MS/MS_2.0 which contains information for the recognition of 261 PFAS.

There are two modes to process the data: Analytic Mode and Explorer Mode.

In Analytics mode match is made of the collected mass spectra with what is in the library according to the criteria given in Table 6:

Table 5: Parameters for peak picking within the ACEA suspect screening workflow

Qualitative rule	Acceptable difference	Marginal difference	Inacceptable difference	Combined score weight (%)
Mass error (ppm)	< 5	< 10	≥ 10	35
% difference isotope ratio	< 5	< 20	≥ 20	30
Library hit score	> 70	> 50	≤ 50	35

If the integrated peaks meet the criteria of “acceptable differences” then it will go to Explorer mode processing for final confirmation.

In Explore mode, in the case of a “suspect” analysis, the chromatographic peak (XIC) corresponding to the exact mass of the compound previously found in Analytics mode is extracted from the TIC. At this point we will go to search, in the mass spectrum that generated the peak, for the precursor ion and any other characteristic fragments to have a confirmation of the analyte.

The entire workflow is depicted in the Figure 1.

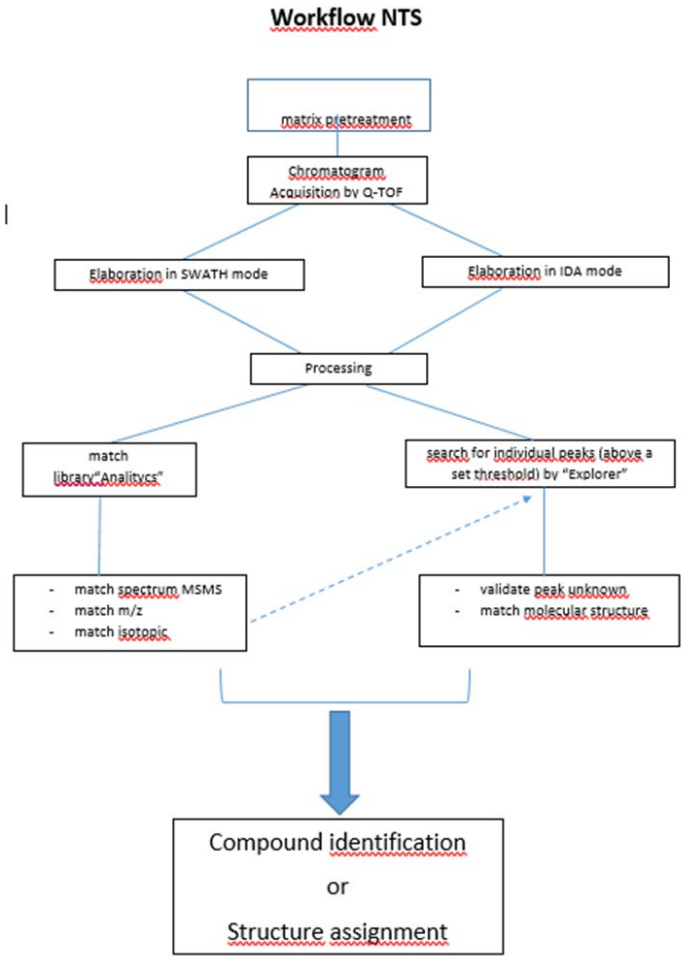

Figure 1: Illustration of the workflow used for suspect screening of PFASs in the ACEA LC-HRMS method

In a real case, the acquired data will undergo our suspect analysis. In the so called “Analytics mode” we’ve settled down some flagging rules useful to know if the acquired spectra meets the library data. The next screenshot replicates the conditions in Table 6.

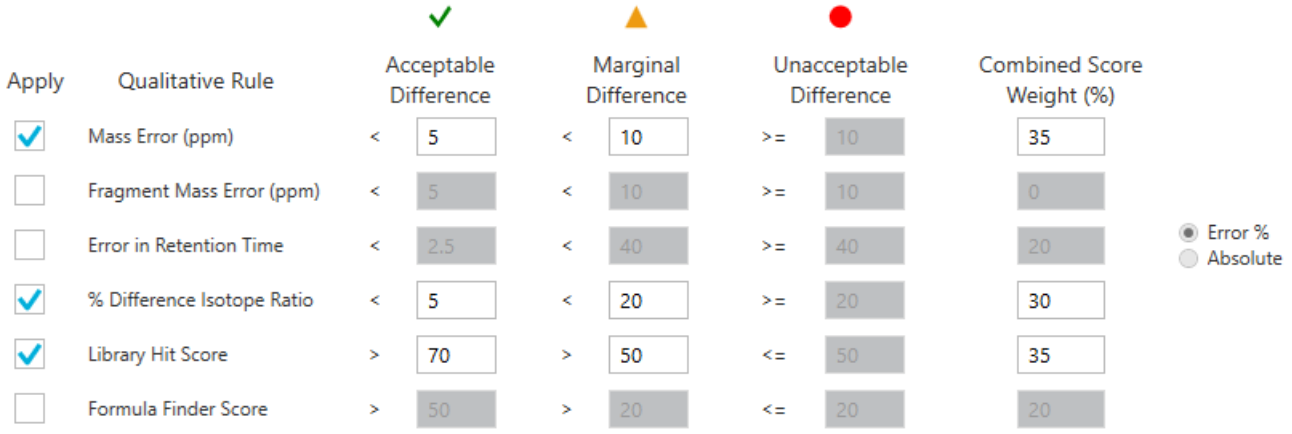

Figure 2: Screenshot of parameter setup in ACEA workflow

In the following example (see Figure 3) there is a partial match for the presence of PFHxS in our real sample (a leachate sample). Specifically, what is found is a positive match for *mass error* and *library confidence* with a marginal difference for the *isotope ratio confidence.* As we can see, in the reported spectra (namely XIC, MS and MSMS) we can visualize the chromatographic peak, the main ion and (in the right part) the fragmented ions matched with the library spectrum (reported as grey negative comparative spectrum).

Figure 3: Example of suspect presence of PFHxS in an analyzed sample.

Once we've noticed these matches, the samples which are suspected to contain compounds of interest will be investigated in the "Explorer mode" in order to obtain a confirmation of our hypothesis. For this purpose, we use the data collected with IDA experiment.

Indeed, in IDA mode (information dependent acquisition), the spectrometer can skim between the possible candidates which will undergo the fragmentation by following the criteria listed in 2.1.5.2.

Considering the setting in criteria b), we can support that in IDA experiment, skimming between the possible candidates which will undergo the fragmentation, we will record a cleaner spectrum.

In the reported example (Figure 4) we extract from the TIC chromatogram the XIC of the hypothesized molecule (PFHxS). By zooming and selecting a portion of the peak in the XIC, we extract the MS spectrum of the suspected compound (Figure 5). As we have obtained the XIC of the suspected compound using the brute formula of the molecule, we will look for its calculated mass to confirm his presence (see the zoomed MS spectrum in Figure 6 and the circled confirming ion).

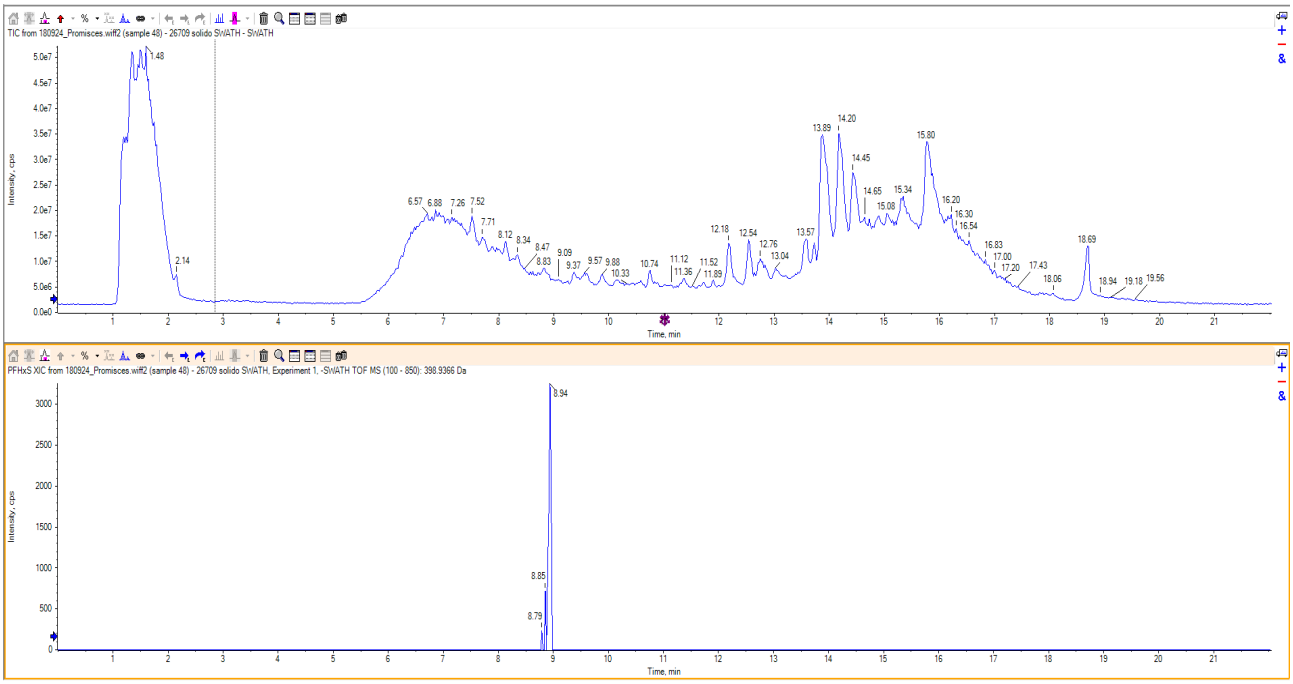

Figure 4: Extraction of PFHxS-XIC from the TIC chromatogram

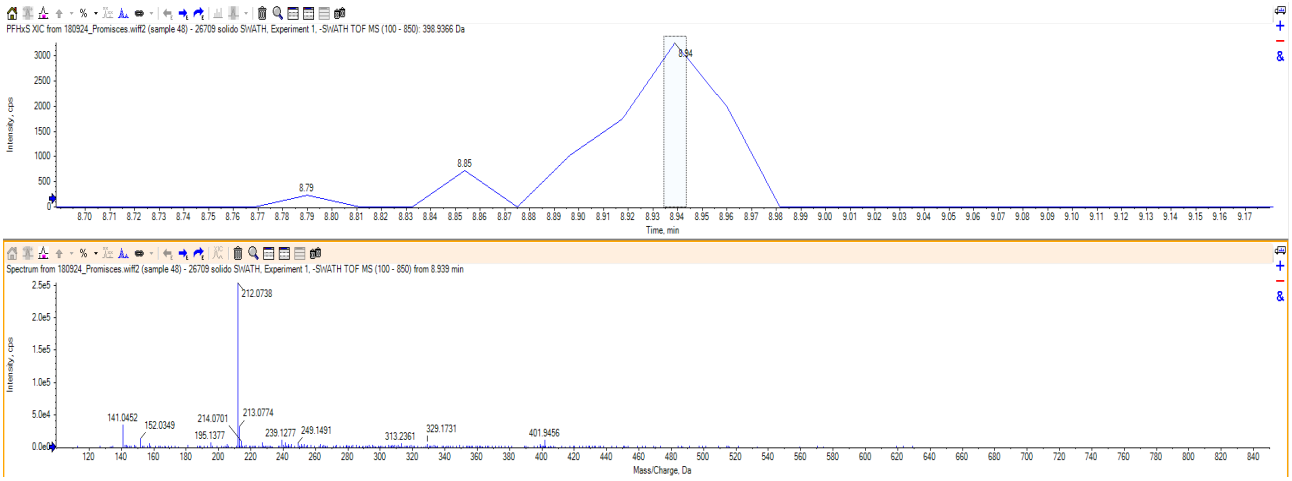

Figure 5: XIC peak peaking and MS spectrum extraction

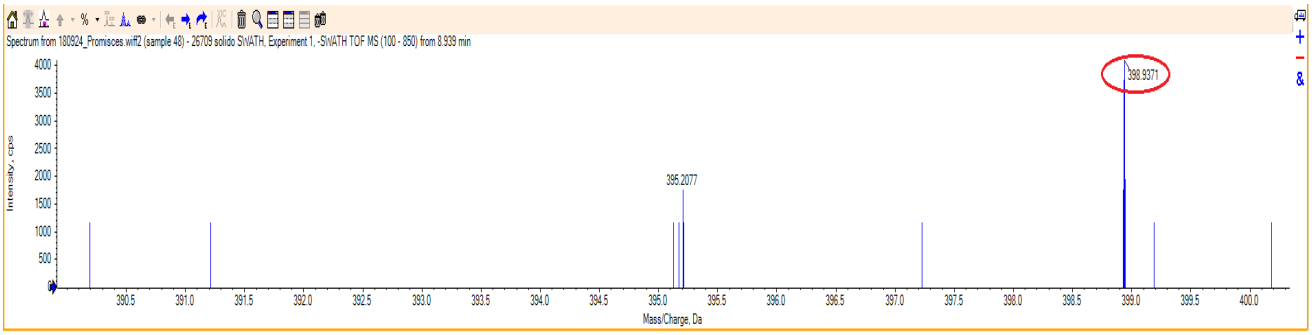

Figure 6: MS spectrum of PFHxS with its exact mass confirming ion.

As seen above the **"Analytics mode"** is a real and proper suspect analysis (rather than a NTS) where the comparison is done with a fluorinated library. Nevertheless if the focus must be on other compounds classes, there is still the possibility to compare our collected data and spectra with other libraries in order to satisfy the subject of investigation.

Instead, at the state of the art, the same cannot be said for the **"Explorer mode".** Indeed in this case, even if we manually explore the spectra, we start from the assumption that the compound is present or may be present. If conversely we extract the XIC and MS spectra without any previous assumption, the research of compounds would request much more effort and time indicating **that this procedure is not suitable and recommendable as routine analysis screening**.

2.1.7 Performance

Performance of the LC-HRMS measurement in terms of LOD (limit of detection) cannot be determined for this workflow, as the non-target method applied here works without calibration. Furthermore, it has to be noted that LODs in LC-MS measurements are highly substance-specific.

Of those compounds identified by this workflow that have been additionally quantified via targeted methods (cf. Deliverable D1.2 and CS#4), concentrations less than 1 µg/L could be determined.

The LODs estimates by the use of IS, were in the range 0.5-5 µg/L.

2.1.8 Application to real samples

The ACEA workflow has been applied to more than 10 leachate samples from landfill leachate treatment plants (LTP) in CS#4.
Out of 10 samples analysed, all PFAS previously identified by target analysis were correctly confirmed. Only PFBA was not automatically confirmed as the concentrations were close to the LOD.

The range of PFAS identified is between 1-40 µg/L. No other PFAS besides those analysed as targets were identified.
All internal standards were correctly identified

Other applications, outside the scope of this project, were carried out on wastewater samples in routine monitoring at the inlet and outlet of sewage treatment plants.

2.2 Suspect screening of PFAS transformation products (CSIC 1)

2.2.1 General

The study of PFAS and related products by means of suspect screening has been based on previous group experience with other groups of compounds. The workflow allows us to tentatively identify up to level 2 of confidence PFAS & related compounds according to Schymanski scale (Schymanski 2014) and eventually level 1 for those PFAS for which CSIC has the standards.

2.2.2 Chemicals and reagents

The chemicals used for the development of the workflow are shown in Table 7**.**

Table 6: List of PFASs used in CSIC 1 suspect screening method

Carboxylic acids	**Sulfonic acids & sulfonamides**	**New PFASs**
PFBA	PFBS	6:2 diPAP
PFPeA	PFPeS	8:2 diPAP
PFHxA	PFHxS	ADONA
PFHpA	PFHpS	EtFOSA
PFOA	PFOS	EtFOSAA
PFNA	PFNS	FOSAA
PFDA	PFDS	MeFOSAA
PFUnA	PFDoS	MeFOSA
PFDoA	FOSA	HFPO-DA (Gen-X)
PFTrDA	**Fluorotelomer sulfonic acids**	PFMOAA
PFTeDA	4:2 FTSA	PFMOPrA
PFHxDA	6:2 FTSA	PFMOBA
PFODA	8:2 FTSA	PFO_2HxA
	10:2 FTSA	PFO_3OA

Other chemicals and reagents used for the analysis of PFAS include ultrapure water, methanol, ammonium acetate and ammonium hydroxide.

Furthermore, the pre-concentration of waters is done by solid phase extraction with Oasis WAX 3cc cartridges.

2.2.3 Sample preparation and pre-treatment

Water samples are centrifuged before extraction at 2000 rpm at room temperature for 10 min. Afterwards, 200 mL of water is transferred to a PET (Polyethylene) container and spiked with a mixture of surrogate internal standards in methanol for a final concentration of 10 pg/mL in sample. Then, the sample is processed following a method adapted from another one previously developed by IDAEA-CSIC (Barbosa 2023). Briefly, solid phase extraction (SPE) cartridges are successively conditioned with 2 mL of methanol and 2 mL of ultrapure water (gravity conditions). Sample loading of 200 mL of surface water, and 100 mL of wastewater (WW) is done under vacuum conditions using PEEK capillary tubes, at a flow rate of 1 mL/min. The cartridges are then dried under vacuum for 15 min and PFAS eluted with 8 mL of methanol (0.1% NH_4OH) in Polypropylene (PP) tubes and evaporated under a gentle stream of nitrogen near to dryness. The final extracts are transferred to LC-vials with 250 µL inserts, dried, and reconstituted in 100 µL of ultrapure water/methanol (90:10).

A SPE blank sample is always carried out in parallel to real samples in order to monitor any cross-contamination.

2.2.4 Chromatography

The chromatographic separation is achieved using an Acquity LC (Waters, Milford, MA, USA) system, equipped with a C18 analytical column Hypersil GOLD PFP LC (50x3 µm) (Thermo Fisher Scientific, San Jose, CA). The mobile phase used consists of (A) aqueous ammonium acetate 20 mM and (B) methanol ammonium acetate 20 mM. Very briefly, at the starting point, the elution gradient is 20% B and within 5 min rise to 80% B, then in 5 min increases to 90% B and is maintained for 2 min more. Finally, initial conditions are achieved within 1 min and maintained for 1 min more. Therefore, the total run time is 12 min for each injection using a flow rate of 0.2 mL/min. The optimal injection volume is 10 µL.

2.2.5 Mass spectrometry

The chromatographic system is coupled to a Q-Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with an electrospray ionization (ESI) source, working in negative ionization conditions. Data is acquired in full scan (FS) (90-1500 Da) with a resolution of 70,000 FWHM and data-dependent scan (ddS) of the most intense ions at resolution of 15,000 FWHM. The entire system is controlled by Xcalibur 4.1 software.

2.2.6 Data treatment

The total ion chromatograms (TIC) obtained by FS acquisition are processed using the Xcalibur software (Thermo Fisher Scientific) for quantification purposes of the standards that are available in CSIC (level 1 of confidence).

Suspect screening of the samples for tentative identification of new PFASs is carried out by processing the data by Compound Discoverer 3.3 SP2 (Thermo Fisher Scientific). A homemade list of PFASs was assembled by PFAS IDAEA-CSIC team, with data from the literature and other databases containing the exact mass of 1,280 compounds. Structural information of the listed compounds, monoisotopic mass and properties such as LogP and logD have been calculated using Chemicalize platform of

ChemAxon (chemicalize.com) and included in the homemade database. Furthermore, the software has been provided with the information in different databases like the EFS HRAM Compounds database (Thermo Fisher Scientific), PFAS NIST database (data.nist.gov), ChemSpider (chemspider.com) for structural information, and MzCloud (mzcloud.org) as a mass spectra database.

An example of workflow used can be seen in Figure 7.

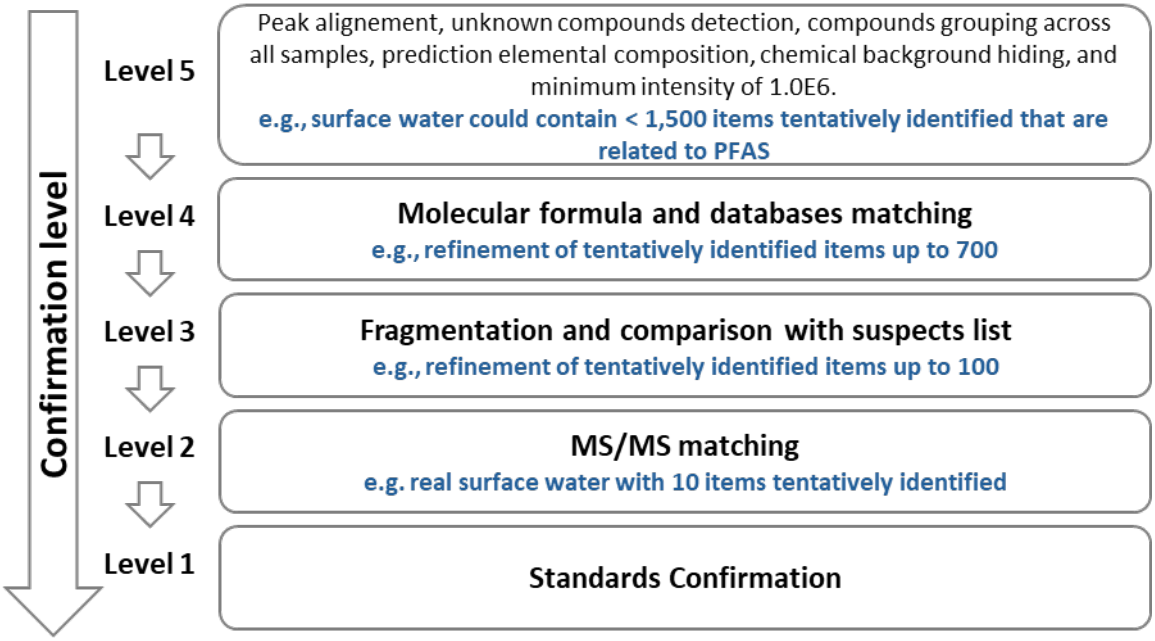

Figure 7: Example of workflow used for suspect screening of PFAS in IDAEA-CSIC.

2.2.7 Application to real samples

The IDAEA-CSIC workflow has been applied to ground coming from WP3 and CS#7 and wastewaters from WP4 and CS#3 in Spain. Furthermore, the workflow will be applied to samples coming from the experiments with lettuce irrigated with treated wastewater in collaboration with other partners from IDAEA-CSIC. Up to now, no new PFAS compounds in addition to those covered by targeted methods (cf. Deliverable D1.1) have been found, underlining the relevance and completeness of the previously chosen target analytes.

2.3 LC-MS Suspect screening of iPM(T)s in surface - and wastewater (BAFG)

2.3.1 General

The goal of this method is to screen for a large amount of substances present in surface water or treated wastewater. Analysis is performed via LC-HRMS in a database-assisted suspect-screening approach. Compounds tentatively identified by this method can be easily included into an existing LC-MS/MS targeted analysis method (cf. Deliverable D 2.1).

2.3.2 Chemicals and reagents

Eluents for chromatography: acetonitrile and formic acid (both LC-MS grade) were received from Merck (Darmstadt, Germany) and Sigma-Aldrich (Seelze, Germany), respectively. Ultrapure water was prepared with a Milli-Q water purification system (Merck Millipore).

2.3.3 Sample preparation and pre-treatment

Samples from monitoring studies (case studies CS#1 and CS#2) were send to BAFG by project partners and were stored in the refrigerator at 5°C until analysis.

Prior to analysis samples were filtered through 0.45 µL PTFE syringe filters. No further pre-treatment was performed.

2.3.4 Chromatography

Chromatographic separation is achieved with a Zorbax Eclipse Plus C18 column (Agilent Technologies) using an Agilent 1260 infinity (Agilent Technologies). Aliquots of the sample under investigation (80 µL) are directly injected into the system without further pre-treatment. Ultrapure water with 0.1% formic acid (eluent A) and acetonitrile with 0.1% formic acid (eluent B) are used as eluent at a flow rate of 0.3 ml/min and a column oven temperature of 40°C with the gradient shown in Table 8:

Table 7: Chromatographic gradient used for suspect screening of iPM(T)s in the BAFG LC-HRMS method.

Time	Eluent A	Eluent B
0 min	98%	2%
1 min	98%	2%
2 min	80%	20%
16.5 min	0%	100%
22 min	0%	100%
22.1 min	98%	2%
25 min	98%	2%

2.3.5 Mass spectrometry

Mass spectrometric analysis is performed with a TripleToF 6600 hybrid quadrupole time-of-flight mass spectrometer (Q-ToF-MS/MS) (SCIEX, Darmstadt, Germany) equipped with an ESI source and operated in positive and negative ionization mode in two separate runs. Data acquisition was done by means of full scan experiments ranged from 100 to 1200 Da. Subsequently, MS2 spectra of the eight most intense peaks were recorded via information dependent acquisition (IDA) with a collision energy (CE) set with potential 40 V and a collision energy spread (CES) set with potential 15 V. Acquisition time of the full scan and the IDA experiments were 150 ms and 30 ms, respectively.

2.3.6 Data treatment

For acquisition of raw data the instrument control software (Analyst TF 1.7, SCIEX) was used. Subsequently, acquired raw data files were transferred to the open data format mzXML using ProteoWizard 3.0 (proteowizard.sourceforge.io) The overall data treatment workflow is depicted in Figure 8.

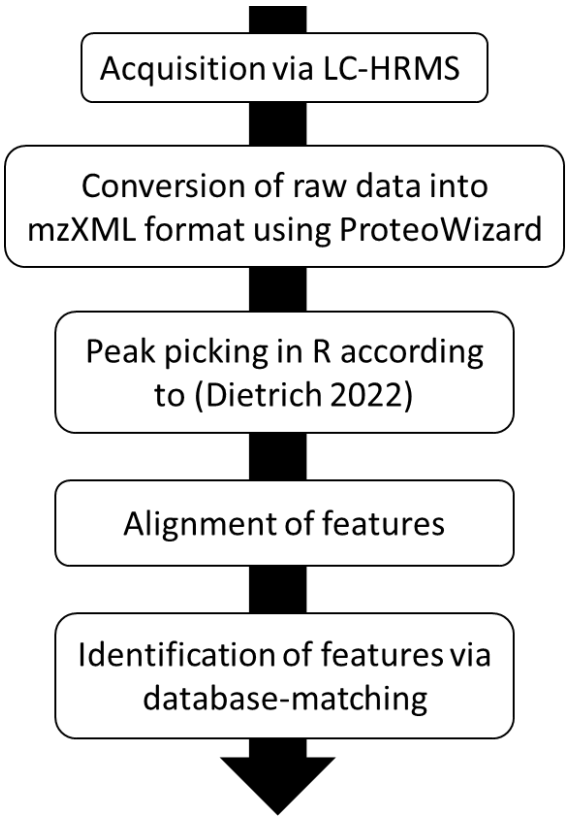

Figure 8: Data treatment workflow for database-assisted suspect screening of iPM(T)s using BAFG method

Peak picking is performed by an algorithm written in R, which is described in detail in (Dietrich 2022). The algorithm extracts chromatograms (XIC – eXtracted Ion Chromatogram) of all previously recorded full scan data within a certain mass range (so-called « bin ») using the following parameters (Table 9). As a result, a list of features is obtained for each sample (feature = signal with defined m/z and retention time).

Table 8: Parameters for peak picking within the BAFG suspect screening workflow

Parameter	Value
m/z range	100 – 1200 (whole range)
m/z bin size	0.02 Da
Retention time range	120 – 1200 s
Min. peak intensity	1
Signal/Noise-ratio	3
Peak width	5 - 60 s

After completion of peak picking in all samples, an alignment of all features is performed: features detected in different samples having the same mass and retention time within an m/z-tolerance of 20 ppm and a retention time tolerance of 20 s are regarded as the same feature and are grouped accordingly. M/z values and retention times of aligned features are registered as the respective mean values of all features included.

Identification of suspects is achieved by matching high resolution mass, MS2 spectra and retention time of the aligned features with an in-house database, which was fed with the respective information from analyses of authentic reference standards on the same machine. High resolution mass spectra as well as MS2 have been uploaded to Massbank for further external use.[1]

2.3.7 Performance

Performance of the LC-HRMS measurement in terms of LOD (limit of detection) cannot be determined for this workflow, as the non-target method applied here works without calibration. Furthermore, it has to be noted that LODs in LC-MS measurements are highly substance-specific. Of those compounds identified by this workflow that have been additionally quantified via targeted methods (cf. Deliverable D1.2 and Milestone MS2), concentrations less than 1 µg/L could be determined.

Performance of the peak picking process was evaluated in terms of False Positive detections. While picking peaks close to the background noise, about 20% of the results in the feature list were actually noise that were wrongly identified as features. Checking for signals with a signal-noise-ratio (S/N) of at least 10 improved the false positive rate to about 5%.

Assignment of features by database matching is prone to errors if signal quality is low (especially the MS2 spectrum) or if multiple peaks can be assigned to various database entries. Manual inspection of the results from automated database screening revealed that about 30% of the assignments were

[1]https://massbank.eu/MassBank/Result.jsp?type=rcdidx&idxtype=site&srchkey=BAFG&sortKey=name&sortAction=1&pageNo=1&exec=

incorrect. Main reasons for this were multiple assignment of one substance to multiple features and assignment of false positive peaks from peak picking (see above) which by chance matched with one of the substances in the database. While the first issue can be easily coped with by simply filtering multiple assignments, the latter issue only concerns one-time assignments within one single sample.

2.3.8 Application to real samples

The BAFG workflow has been applied to 58 urban effluent samples from CS#1 which furnished 49 tentatively identified substances including 31 industrial compounds. Database assisted suspect screening of 144 surface- and groundwater samples from CS#2 indicated the presence of 280 tentatively identified substances including 75 industrial compounds. A complete list of all features identified in samples from CS#2 can be found in ZENODO (https://zenodo.org/records/14011987).

2.4 LC-MS Suspect screening of iPM(T)s in wastewater (CSIC 2)

2.4.1 General

The goal of this suspect screening analysis is to evaluate three different sample preparation methods to obtain a broad identification of iPM(T) substances in effluent wastewater samples from CS#3. Subsequently, the identified compounds were prioritized based on their persistent, bioaccumulation, mobility and toxicity.

2.4.2 Chemicals and reagents

Solvents for chromatography (acetonitrile, methanol and HPLC water) were purchased from Thermo Fisher Scientific Inc. (Waltham, MA, USA). Formic acid LC-MS grade was obtained from Sigma-Aldrich (Seelze, Germany).

2.4.3 Sample preparation and pre-treatment

Three different sample preparation methods were tested to evaluate their effect on non-selective extraction of analytes from treated wastewater effluent samples.

Firstly, 400 mL of each effluent water sample (n=3) were spiked with a mixture of isotopically labelled standard compounds to monitor analytical performance and instrumental sensitivity.

Sample preparation method 1 (SPM-1) consisted of the lyophilization of 200 mL of the previously spiked each sample spiked with the isotopically labelled standards using a freeze dryer LyoAlfa (Telstar) with a final condenser temperature of -64 °C and a vacuum pressure of 0.031 mBar. The residue is then redissolved in 15 mL of methanol followed by 15 mL of ethyl acetate. The resulting extract is transferred to a glass centrifuge tube and centrifuged at 4,000 rpm and 20 °C for 10 min (Eppendorf® Centrifuge 5810R). Subsequently, the supernatant is evaporated under a stream of N_2 at 10 psi to 0.2 mL and reconstituted with HPLC grade water to 1mL. Finally, the extract is centrifuged at 10,000 rpm and 20 °C for 10 min before transferring it to an HPLC vial for analysis.

For the other two sample treatments, 2 mL of the sample were centrifuged under the same conditions described above. Then, sample preparation method 2 (SPMT-2) consisted of direct injection of the sample in the HPLC system, and sample preparation method 3 (SPMT-3) consisted of the online solid-phase extraction (SPE) of the sample using a polymeric column.

2.4.4 Chromatography

Chromatographic separation of the analytes was carried out with a reversed-phase Luna Omega Polar C18 column (3µm 50 x 2.1 mm, Phenomenex, Torrance, CA, USA). The column oven was set at 30 °C. A flow-rate of 0.3 mL min-1 and an injection volume of 5 µL were employed for the analysis. The mobile phase composition in positive ionization mode consisted of (A) H_2O and (B) acetonitrile both with 0.1 % formic acid and in negative ionization mode were (A) H_2O with 5 mM ammonium acetate and (B) acetonitrile.

Table 9: Chromatographic gradient used for suspect screening of iPM(T)s in the CSIC LC-HRMS method.

Time	Eluent A	Eluent B
0 min	95%	5%
1 min	95%	5%
18.5 min	0%	100%
20.5 min	0%	100%
21.5 min	95%	5%
25 min	95%	5%

2.4.5 Mass spectrometry

Mass spectrometric analysis was performed with an Impact II Q-TOF system (Bruker Daltonics, Billerica, MA, USA) with Vacuum Insulated Probe Heated Electrospray Ionization (VIP-HESI). Full-scan spectra were recorded between a mass range of 70 m/z and 1,000 m/z at an ionization energy of 6.0 eV in positive and negative ionization. MS/MS data were obtained in two acquisition modes with a nominal collision energy of 25 eV (ramped between 25 – 60 eV): broadband Collision Induced Dissociation (bbCID), a data independent mode (DIA), and AutoMSMS, a data dependant mode (DDA).

2.4.6 Data treatment

The retrospective analysis of the samples consisted of the contrast of the data obtained with a suspect list, mzCloud database (about 8,000 CECs from different applications) that was available at NORMAN Substance Database.

Data-dependent acquisition was processed with the suspect list selected through MetaboScape® 2022b (Bruker Daltonics, Billerica, MA, USA), a software for compound identification. Compounds were identified based on exact mass (m/z) with a mass error tolerance of < 5 ppm and scoring of the difference between the measured isotopic pattern and the theoretical pattern of the ion (mSigma, narrow score: 50 and wide score: 250). Following data processing, the matched compounds were manually evaluated considering their MS/MS spectra and the fragmentation pattern available in mzCloud under similar instrumental conditions. Additionally, the peak areas of identified compounds were obtained from full-scan data in DIA raw files using the extracted ion chromatogram function in Compass Data Analysis software version 5.3 (Bruker Daltonics, Billerica, MA, USA).

2.4.7 Performance

Performance of the LC-HRMS measurement in terms of LODs (limit of detections) cannot be determined for this workflow, as the nontarget method applied here works without calibration.

However, many of the identified compounds were included in a subsequently developed target analytical method which demonstrated good performance in terms of matrix effects, recoveries, reproducibility and LOD/LOQ (in most cases in the low n/L or pg/L range).

Furthermore, for every matching compound with the nontarget workflow, different identification levels were assigned according to Schymanski, where level 1 corresponds to confirmation with a

reference standard, level 2 represents probably structure by matching with spectrum data library, and the rest of levels (3 to 5) describe substance class, formula or mass of interest.

As regards the comparison of different sample treatment methods, lyophilization was the method providing the largest number of identified compounds at an acceptable level of confidence (all above level 2a): 117, versus 37 and 49 when using direct injection and on-line preconcentration, receptively.

2.4.8 Application to real samples

The CSIC nontarget workflow has been applied to 3 wastewater treatment plant effluent samples from CS#3, collected in three different days of the week (Wednesday, Friday and Sunday). A total of 119 compounds were found, and 22 of them were confirmed with reference standards. Compounds from industrial applications were the second most commonly detected category of contaminants (in total 33 compounds/metabolites). The list of compounds prioritized for subsequent development of target analytical methods for determination of relevant iPM(T) substances in waters from the case study CS#3 can be found in ZENODO (https://zenodo.org/records/13844325).

2.5 GC-MS Suspect screening of iPM(T)s in surface water (BWB)

2.5.1 General

The GC screening method is intended to provide a rapid qualitative and semi-quantitative assessment of medium-volatile trace compounds in water samples. Through a solid-phase extraction (SPE) enrichment step, the samples are both concentrated and purified. Considering the distribution of concentrations within the sample material and the specific properties of individual compounds, a minimum concentration of approximately 10 ng/L per substance is required. Data comparison is performed using both the NIST and an internal database, and the categories of reliability of the results varies depending on the degree of concordance between the sample and the databases. The establishment of a sample library for wastewater samples enables the detection of changes in sample composition. Additionally, the comparison of trace compound compositions from indirect discharger samples (industrial wastewater) with other samples is possible.

2.5.2 Chemicals and reagents

All chemicals and materials used for the method are summarized in Table 11.

Table 10: Chemicals and materials used for suspect screening of iPM(T)s in the BWB GC-MS method.

Chemicals/Material	**Manufacurer**
SPE cartridges, BondElut-ENV, 500 mg/6 mL	Agilent Technologies, USA
Auto Trace	Thermo Fisher Scientific, USA
Turbo Vap II	Biotage, SWE
Nitrogen (5.0)	Linde plc, IR
Helium (5.0)	Linde plc, IR
Dichlormethane (HPLC grade)	Merck KGaA, GER
Methanol (HPLC grade)	Merck KGaA, GER
Etylacetate (HPLC grade)	Merck KGaA, GER
Sodium sulfate (H_2O-free, p.a.)	Merck KGaA, GER
Diphenyl d10	Sigma-Aldrich, USA

2.5.3 Sample preparation and pre-treatment

The sample is filled into a 1L glass bottle with a ground glass stopper, transported promptly and under cool conditions. The sample can be stored at 4°C and needs to be extracted within 48 hours. Two bottles per sample are filled, with one bottle retained as a backup sample. A 1L sample water is filled in a 1L graduated cylinder, and 50 µL of the internal standard is added. Solid-phase extraction is performed using an Auto Trace automatic SPE extractor, with Bond Elut-ENV cartridges. The cartridges are conditioned with 5 mL ethyl acetate and 10 mL methanol consecutively prior loading 1000 mL sample. After drying the cartridge over a nitrogen flow for 45 minutes, the analytes are extracted by 2 mL of ethyl acetate. The sample extract is dried with sodium sulfate, transferred into a glass evaporation tube, concentrated to 500 µL at 30°C, transferred into a vial, sealed airtight, and subsequently analyzed by GC-MS.

2.5.4 Chromatography

The samples are analyzed using an Agilent GC-MS, with the instrument parameters summarized in Table 12.

Table 11: Chromatographic parameters used for suspect screening of iPM(T)s in the BWB GC-MS method.

Parameter	
Gas chromatograph	Agilent 7890A
Injection volume	1 µL
Injector parameters	65 °C – 300 °C (ramp at 720 °C/min)
Analytical column	HP-5 MS UI, 30m x 0.250mm x 0.25 ID
Carrier gas	Helium

If a high concentration of trace substances is expected, the extract can be diluted with ethyl acetate. The samples are injected fully automatically. The temperature program of the column oven is shown in Table 13, and after each measurement, the oven cools to ambient temperature for approximately 5 minutes before a new sample run starts.

Table 12: Temperature gradient used for suspect screening of iPM(T)s in the BWB GC-MS method.

Time	**Temperature**
0 min	40 °C
2 min	40 °C
7 min	70 °C
30 min	300 °C
35 min	300 °C

To determine the retention time index (RI) a standard solution of multiple alkanes (chain length: C7-C33) is analyzed with each measurement sequence along with an ethyl acetate blank.

2.5.5 Mass spectrometry

Mass spectrometry is performed using an Agilent 7000 Triple-Quad equipped with a (electron impact) EI source. The run parameters of the device are summarized in Table 14.

Table 13: MS-Instrument parameters for suspect screening of iPM(T)s in the BWB GC-MS method.

Parameter	
Mass spectrometer	Agilent 7000 GC-MS-Triple-Quad
Source temp.	300 °C
Transferline temp.	300 °C
Detector temp.	280 °C
Mass range	50-500 m/z
Ionisation mode	Electron impact (EI)
Ionisation energy	70 eV

2.5.6 Data treatment

The obtained spectra were evaluated through library comparison to identify the chemical compounds present in the samples. A semi-quantitative analysis can be performed by comparing the peak areas with the ones from the internal standard diphenyl D10 (0.5 µg/L, ~$2*10^6$ area units), based on its fragment ion at 164 m/z. For the measurement to be considered reliable, a peak area ratio of at least 1000:1 between the internal standard diphenyl D10 and the baseline is required. This criterion ensures that the signal is distinguishable from noise and that the results are of sufficient quality.

Additionally, the measured spectra were subjected to a Retention Index (RI) analysis using Agilents MassHunter Unknowns Analysis (UA) software. The individual RI is determined automatically by the UA software after subtracting the blank measurements. This step reduces background noise and enhances the accuracy of the spectral identification. The software then compares the features with the stored libraries (NIST and in-house) and outputs a list with compound suggestions and various matching factors. Adapted from Schymanski et al. (Schymanski 2014), 6 categories of confidence are defined, see Table 15.

Table 14: Categories of confidence for suspect screening of iPM(T)s in the BWB GC-MS method.

Category	Details	Peak area >10,000	Delta RI <100	NIST library match factor >80 %	Inhouse library match factor >80 %	Reference standard
1	Confirmed structure	✓	✓	✓	✓	✓
2	Probable structure (very high confidence)	✓	✓	✓	✓	✗
3	Tentative structure (high confidence)	✓	✓	✓	✗	✗
4	Structure suggestion, low confidence	✓	✗	✓	✗	✗
5	No Structure suggestion	✓	✗	✗	✗	✗
6	No Structure suggestion	✗	✓/✗	✓/✗	✓/✗	✓/✗

For a feature to be included in the compound list, a minimum peak area must be reached, the deviation of the defined RI must not exceed 100 and there must be a match with at least one library.

The spectral comparison uses the NIST 2020 database, which contains over 31,000 compounds, along with 1.3 million spectra. The in-house BWB-GCMS library, which includes over 1,000 compounds, is updated regularly and is based on reference standards analyzed on the GC-MS. This complements the identification process, expanding the scope of potential matches and enhancing the robustness of the analysis.

2.5.7 Performance

A limit of quantification cannot be determined for this method, as no quantification is performed. As a quality assurance criterion, the peaks are compared with the internal standard, as described in Section 2.4.6. “Data treatment“. Quality assurance is ensured through a control standard with known concentration.

2.5.8 Application to real samples

This method was applied to more than 72 wastewater samples and, for comparison, to several industrial indirect discharger samples from CS#1. On average, more than 80 peaks can be identified in wastewater samples. Example data provided in Table 16 and Figure 9; as established in additional measurements (not expensed within PROMISCES).

Table 15: Example data of four wastewater samples analyzed by the BWBs GC-MS suspect screening method.

Proposed substances	**CAS**	**Sample**	**3495**	**4153**	**4493**	**4856**
		Category	**Area % of IS-Area**			
diisobutyl phthalate	84-69-5	2	2415	12852	9716	1741
dibutyl phthalate	84-74-2	1	-	-	2507	3240
acetic acid, butyl ester	123-86-4	3	-	-	1981	-
7,9-di-tert-butyl-1-oxaspiro(4,5)deca-6,9-diene-2,8-dione	82304-66-3	2	689	590	626	669
benzoic acid, 4-ethoxy-, ethyl ester	23676-09-7	2	330	158	434	406
diethyl phthalate	84-66-2	1	313	243	201	336
acetophenone	98-86-2	1	-	47	199	137
diethyl carbitol	112-36-7	3	268	-	193	-
dibutyl adipate	105-99-7	2	55	209	173	310
di(propylene glycol) methyl ether	13588-28-8	3	457	23	154	-

The 10 most abundant features of four different wastewater samples after data treatment are presented in Table 16. The categories given are explained in Table 15, the values for the area are relative to the area of the IS. For reference the chromatograms of those four samples are shown in Figure 9.

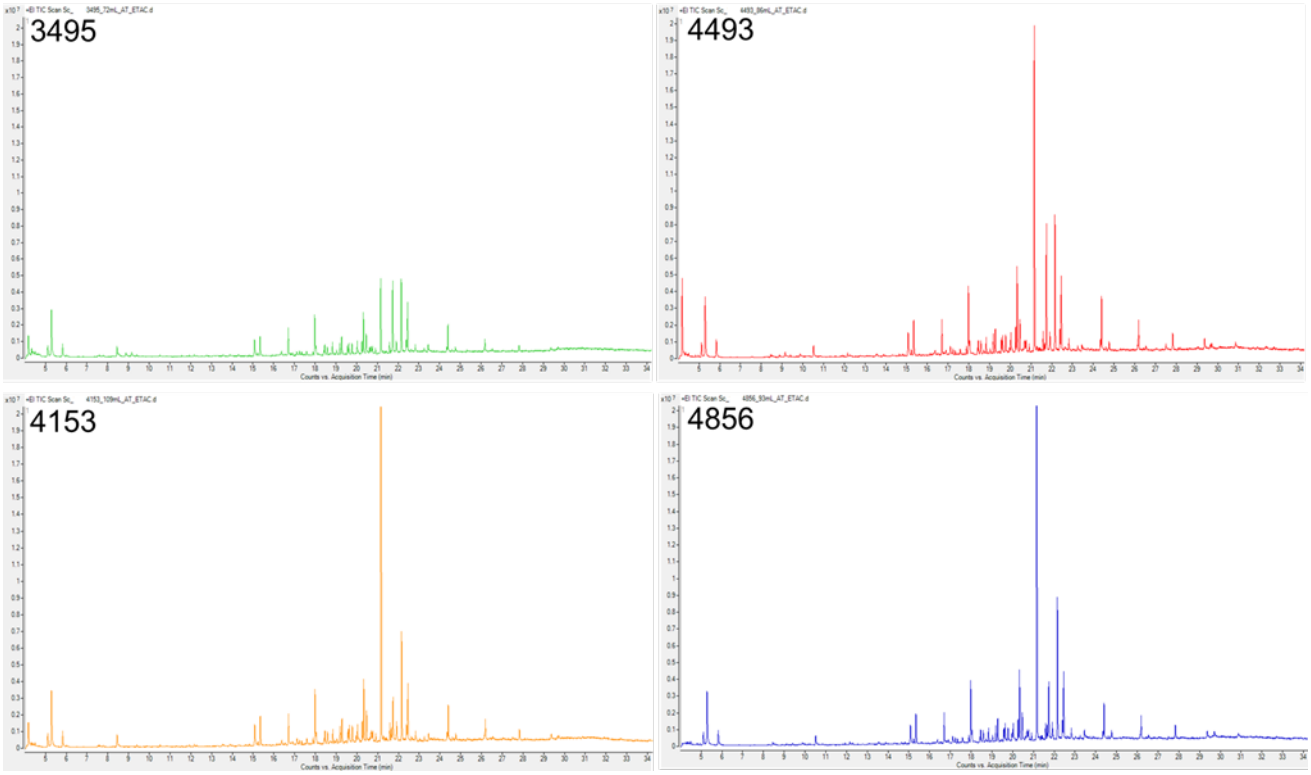

Figure 9: Chromatograms of the four example samples.

3 Conclusions

This document reports on five different approaches for a database-assisted suspect screening of CECs in various matrices. Two methods using liquid chromatography coupled to high-resolution mass spectrometry are specifically designed for the screening of PFAS compounds and their transformation products, respectively. Two other LC-HRMS-based methods are dedicated to the identification of iPM(T)s in surface- and wastewater. The latter two methods are complemented by a GC-MS-based method enabling a screening for more unpolar compounds not amenable to most LC-MS methods.

Each method has been designed to be applied in a specific case study and thus for analysis in different matrices. Sample preparation therefore ranges from simple filtration or dilution over centrifugation to lyophilization. Except for the GC-MS method, which requires a solid phase extraction and subsequent elution with a solvent suitable for GC measurements, all other mentioned sample preparation techniques aim to preserve the original CEC composition of the sample as far as possible.

While signal detection is performed with a high-resolution MS for the four LC-MS-based methods, the GC-MS method works with a triple quadrupole machine. Nevertheless, in all five workflows raw data is acquired via an untargeted approach without pre-defining compounds of interest.

In all cases this raw data is subsequently subjected to a separate data treatment step including comparison of the acquired spectral data to a database. For this, commercial as well as open-source software is used in combination with commercial and public compound databases.

Being based on an untargeted approach, the suspect screening workflows presented here commonly suffer from reduced sensitivity compared to targeted analysis. Furthermore, quantitative information is limited due to missing calibration. To address these issues, still targeted analysis is the method of choice. The big benefit of the presented suspect screening methods is the possibility to comparatively fast perform a screening for a large number of compounds without the need to develop a new target method. Furthermore, the sequential approach of data acquisition and separate data treatment allows for retrospective screening of substances even after many years. Information about new compounds identified via suspect screening can then be used to systematically adapt targeted methods for quantification of these substances.

Application of the workflows dedicated to screening of PFAS and PFAS TPs in cases studies CS#3, CS#4 and CS#7 confirmed the relevance and completeness of the previously chosen target analytes as already reported in Deliverable D1.1 *" Methods for PFAS in waters and complex matrices "*. Suspect screening of iPM(T)s revealed the occurrence of more than 100 compounds in samples from CS#1, CS#2 and CS#7. After further prioritization of these findings, a first selection of substances has already been included into the compound lists of the respective targeted methods (Deliverable D1.2 *" Targeted methods for relevant iPM(T) substances in waters "*).

4 References

Barbosa, M. O., Ratola, N., Homem, V., Pereira, M. F. R., Silva, A. M., Ribeiro, A. R., Llorca, M. & Farré, M. (2023). Per-and poly-fluoroalkyl substances in Portuguese rivers: spatial-temporal monitoring. Molecules, 28(3), 1209.

Dietrich C, Wick A, Ternes TA (2022). Open-source feature detection for non-target LC–MS analytics. *Rapid Commun Mass Spectrom*., 36(2):e9206. doi: 10.1002/rcm.9206.

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, 4, 2097-2098. doi: 10.1021/es5002105.