Journal of Hazardous Materials 448 (2023) 130906

Contents lists available at ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

Enhanced selectivity for acidic contaminants in drinking water: From suspect screening to toxicity prediction

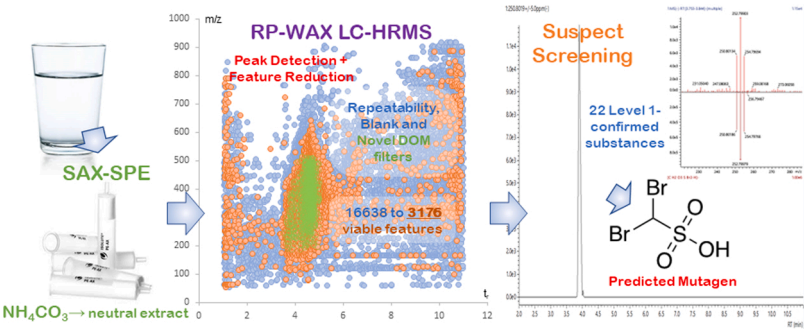
Davide Ciccarelli [a,b], D. Christopher Braddock [c], Andrew J. Surman [d], Blanca Ivonne Vergara Arenas [d], Tara Salal [d], Tim Marczylo [b,e], Paolo Vineis [b,f], Leon P. Barron [a,b,*]

[a] Environmental Research Group, MRC Centre for Environment and Health, School of Public Health, Imperial College London, 86 Wood Lane, London W12 0BZ, UK
[b] NIHR-HPRU Chemical and Radiation Threats and Hazards, NIHR-HPRU Environmental Exposures and Health, MRC Centre for Environment and Health, School of Public Health, Imperial College London, 86 Wood Lane, London W12 0BZ, UK
[c] Department of Chemistry, Imperial College London, 82 Wood Lane, London W12 7TA, UK
[d] Department of Chemistry, King's College London, Britannia House, 7 Trinity Street, London SE1 1DB, UK
[e] UK Health Security Agency, Harwell Science Campus, Femi Avenue, Harwell, Didcot OX11 0GD, UK
[f] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG, UK

HIGHLIGHTS

- Suspect screening LC-HRMS workflow tailored for trace organic acid contaminants.
- NH_4HCO_3 extends applicability of strong anion exchange solid-phase extraction.
- Creation of a novel filter for reduction of dissolved organic matter features.
- Dibromomethanesulfonic acid confirmed and quantified in UK drinking water.
- New formation pathways for tribromo- and trichloro-hydroxycyclopentenedione proposed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Dr. B. Lee

Keywords:
High-resolution mass spectrometry
Strong anion exchange
Feature reduction
Dissolved organic matter
Disinfectant by-products
Perfluorinated acids

ABSTRACT

A novel analytical workflow for suspect screening of organic acidic contaminants in drinking water is presented, featuring selective extraction by silica-based strong anion-exchange solid-phase extraction, mixed-mode liquid chromatography-high resolution accurate mass spectrometry (LC-HRMS), peak detection, feature reduction and compound identification. The novel use of an ammonium bicarbonate-based elution solvent extended strong anion-exchange solid-phase extraction applicability to LC-HRMS of strong acids. This approach performed with consistently higher recovery and repeatability (88 ± 7 % at 500 ng L^{-1}), improved selectivity and lower matrix interference (mean = 12 %) over a generic mixed-mode weak anion exchange SPE method. In addition, a novel filter for reducing full-scan features from fulvic and humic acids was successfully introduced, reducing workload and potential for false positives. The workflow was then applied to 10 London municipal drinking water samples, revealing the presence of 22 confirmed and 37 tentatively identified substances. Several poorly investigated and

* Corresponding author at: Environmental Research Group, MRC Centre for Environment and Health, School of Public Health, Imperial College London, 86 Wood Lane, London W12 0BZ, UK.
E-mail address: leon.barron@imperial.ac.uk (L.P. Barron).

https://doi.org/10.1016/j.jhazmat.2023.130906
Received 20 December 2022; Received in revised form 27 January 2023; Accepted 28 January 2023
Available online 31 January 2023

potentially harmful compounds were found which included halogenated hydroxy-cyclopentene-diones and dibromomethanesulfonic acid. Some of these compounds have been reported as mutagenic in test systems and thus their presence here requires further investigation. Overall, this approach demonstrated that employing selective extraction improved detection and helped shortlist suspects and potentially toxic chemical contaminants with higher confidence.

1. Introduction

A vast array of natural and anthropogenically-derived chemical substances exists in our environment. When present in water, some also survive treatment and can contaminate municipal drinking water supplies [1-3]. Currently, as of 2023, > 204 million substances exist on the Chemical Abstracts Service (CAS) Registry and, of these, it is estimated that 350,000 chemical products have been licenced for manufacture and sale on a global level [4]. As a consequence, identification and monitoring of large numbers of chemical compounds across multiple environmental compartments presents an enormous challenge. Recently, analytical approaches for non-target analysis (NTA) and suspect screening (SS) using liquid chromatography and high-resolution accurate mass spectrometry (LC-HRMS) have emerged and detect much larger numbers of compounds in a sample. However, NTA still presents multiple challenges, from ensuring detection of physico-chemically diverse compounds to reliably translating spectra of unknown substances into defined structures. Gaining comprehensive and confident knowledge of all sample components remains an aim still far from being fully achieved.

Current sample preparation methods for NTA of water are designed to concentrate as wide an array of sample components as possible into a single solution for injection. These have employed mostly single-sorbent solid phase extraction (SPE) [5-7], multi-layer SPE [8-11], or evaporative concentration [10,12,11]. Despite this obvious logic, extraction of such a wide array of compound chemistries for instrumental analysis can present new analytical problems (or exacerbate existing problems) which could affect performance and reliability including, but not limited to: (a) insufficient chromatographic resolution which can in turn affect separation accuracy and precision when using retention time (t_R) as a metric for compound identification, especially in liquid chromatography (LC); (b) increased droplet surface ion competition in electrospray ionisation (ESI) reducing or causing variable detector sensitivity and noise; and (c) much increased spectral complexity, which in turn increases the chances of incorrect spectral interpretation due to interferences in MS and MS/MS. Therefore, opting for some selectivity in the extraction step based on relatively simple, but still broad-reaching commonality in chemical properties for molecules present in water may improve the value and confidence in the overall SS/NTA approach. The obvious disadvantage of such an approach is an increased workload with multiple different extraction modes required for chemicals with different properties being targeted. However, recognising that scientific knowledge of drinking water composition remains partial, and that even wide-scope extraction methods are highly unlikely to cover the entire chemical space anyway, multiple orthogonal and complementary methods are likely more effective at reducing this uncertainty and are ultimately more practical, reliable, and arguably easier to automate as a result. Selective extraction of acidic, basic and neutral substances in separate fractions can be achieved by employing serial or parallel SPE with sorbents that are highly specific for each selected chemical space. The focus of this work lies in anionic components in drinking water as many such substances have recently emerged as very toxic in humans including a large array of perfluoroalkyl substances (PFAS) and disinfectant by-products (DBPs), such as haloacetic acids (HAAs) or oxyhalides, to name just a few. Sample preparation methods employing separate modes selective for acidic or basic analytes have been adopted for the analysis of persistent and mobile organic contaminants (PMOCs) for example [13,14], employing mixed-mode ion exchangers which are co-polymerised with an organic polymer backbone conferring wider extraction selectivity. Silica-based strong-anion exchange (SAX) resins have not been used in this field to date, due to the ionic strength required to elute strong acids effectively, but also finding such solutions that are compatible with chromatography and/or mass spectrometry detection.

In addition to extraction, and keeping anionic selectivity as the focus, the instrumental analysis requires similar levels of consideration. Consensus regarding the limitations of reversed-phased liquid chromatography (RPLC) for SS/NTA of very polar and ionic compounds has resulted in new methods employing hydrophilic interaction liquid chromatography [15,11,14], mixed-mode RP-anion exchange-cation exchange [13,16] and supercritical fluid extraction [5,11,14].

Over the last few years, SS and NTA analysis of drinking water has led to the detection of an increasing number of contaminants, including DBPs [17-20], PFAS [21,22], pharmaceutical and personal care products [3,23], industrial chemicals [14,24], leachables from bottles [25] and pipes [26], and pesticides [27]. However, most of the detected contaminants have intermediate polarities due to the vastly prevalent use of RPLC-based methods, and to the best of our knowledge a method optimised specifically for simultaneous detection and identification of relatively hydrophobic and hydrophilic acidic contaminants is still missing.

For SS and NTA, a carefully optimised data-processing workflow is also critical. Several studies have highlighted the marked difference in output of different algorithms for peak detection [28-30]. Consensus is also growing in the scientific community about the importance of feature reduction [31-33,23,28] to eliminate unnecessary workload and reduce false positives. In particular, blank and repeatability filters are increasingly becoming integrated in data-processing NTA workflows. Additionally, several studies have characterised natural dissolved organic matter (DOM) by LC-HRMS [34-36] including relatively harmless fulvic and humic acids. The concentration of DOM in UK tap water is generally in the range of a few parts per million [37]. By applying a concentration factor during sample preparation, a large number of DOM mass spectral features will be generated. Therefore, a means to reduce or remove such features in silico would reduce the SS/NTA workload and likelihood of generating false positives (while retaining the option to reassess the data for such substances post-hoc if required).

The aim of this work was to develop and apply a comprehensive workflow for SS of organic acids and to help identify potentially new toxic substances from this class in municipal drinking water with high confidence. The objectives included design, optimisation, evaluation and application of the sample preparation, non-targeted instrumental analysis, peak detection, feature reduction and suspect screening stages using initially a selection of well-known acidic contaminants, followed by wider application to SS of drinking water samples from London, UK. Specifically, we present the novel use of SAX SPE employing a high ionic strength elution solvent (ammonium bicarbonate) which can be eliminated conveniently before LC-HRMS for detection of both weak and strong organic acids. Additionally, a new RP-weak-anion exchange (WAX) LC-HRMS method was developed to separate organic acids with very different polarities and expected to be present at concentrations spanning several orders of magnitude. For this purpose, a list of 23 substances was compiled with the aim of covering a broad polarity range (LogD at pH 7.5 ranging from 3.6 to −3.9, Table S1, predicted with ACD/Labs Percepta), and of representing classes of compounds expected to be found in drinking water: hydrophobic and hydrophilic PFAS [22, 38], DBPs [39], pesticides [40], metabolites of industrial chemicals

[41], and sweeteners [42].

Lastly and importantly for SS/NTA of acids in drinking water, a list of known elemental compositions present in drinking and surface water DOM was generated and to our knowledge, this is the first time such a list has been employed for feature reduction in SS. Overall, this new approach can be used to improve confidence of detection and identification of new acidic and potentially toxic agents in drinking water.

2. Materials and methods

2.1. Reagents

LC-MS grade ammonium bicarbonate, as well as ammonium hydroxide puriss. (≥25 % NH_3 basis) and diethylamine puriss. (≥99.5 %) were purchased from Sigma-Aldrich (Steinheim, Germany). LC-MS grade acetonitrile, formic and acetic acid were acquired from Fisher Scientific (Waltham, USA). LC-MS grade ultrapure water and methanol were supplied respectively by Supelco (Bellefonte, USA) and VWR (Radnor, USA). A total of 23 hazardous and/or physicochemically diverse acidic substances (see Table S1 for details and physicochemical properties) were purchased from Sigma-Aldrich, Accustandard (New Haven, CT, USA), Supelco (Bellefonte, PA, USA) and Manchester Organics (Runcorn, UK). A Suwannee River Fulvic Acid Standard (SRFAS) was obtained from the International Humic Substances Society (St. Paul, USA). For confirmation of shortlisted suspect screening candidates, 15 commercially available reference materials were purchased from a variety of suppliers (see Table S2).

2.2. Water sampling and sample preparation

Ten municipal drinking water samples were taken from households across London, UK, from the 23rd until the 31st of March 2022 (GPS coordinates in Table S3). Cold water was collected in 30 mL Nalgene polypropylene bottles (Sigma-Aldrich) after having been rinsed three times, then was subsequently frozen at − 20 °C and kept in the dark until analysis. The same procedure was used to collect and store ultrapure water to be used as process blanks. For method performance evaluation and comparison, a single tap water sample was collected in a 1 L Schott bottle from the laboratory supply (measured pH = 7.5) and employed for all preparations. For SPE, Isolute PE-AX in bulk and polypropylene single fritted SPE cartridges (10 mL reservoir, polyethylene frits) were purchased from Biotage (Uppsala, Sweden). A 500 mg mass of this silica-based strong anion exchanger in the acetate form was packed into polypropylene cartridges and conditioned with 4 mL of methanol followed by 4 mL of ultrapure water. Subsequently, 10 mL of sample was loaded. The cartridges were then washed with 8 mL of an ultrapure water:methanol solution (9:1 v/v), dried, and eluted with 8 mL of a 0.2 M ammonium bicarbonate solution in ultrapure water:methanol (1:9 v/v). The eluates were then evaporated to dryness at 38 °C, and reconstituted with 100 μL of 1:1 v/v ultrapure water:methanol (injection solvent). All extracts were then transferred to 9 mm screw top clear glass Verex vials from Phenomenex (San Juan, USA) with 9 mm Solid Top Polyethylene caps from Restek (Bellefonte, USA) and 0.2 mL glass inserts from VWR for injection onto the LC-HRMS.

2.3. Instrumentation

All analysis was performed on a Shimadzu LCMS9030 liquid chromatography-quadrupole-time-of-flight mass spectrometer (LC-QTOF-MS) equipped with a Nexera XR LC system and a standard ESI source (Shimadzu Corporation, Kyoto, Japan). Detailed chromatographic conditions and acquisition settings are reported in Table S4. For chromatography, a Waters Atlantis Premier BEH C_{18} AX (2.5 μm, 2.1 × 100 mm) column with pre-column (VanGuard FIT Cartridge, 2.5 μm, 2.1 × 5 mm, containing the same stationary phase) for LC analysis was acquired from Waters (Tampa, FL, USA). Mobile phase A was a 5 mM ammonium bicarbonate in water (pH 6.9, adjusted with acetic acid) and mobile phase B was a 5 mM ammonium bicarbonate in acetonitrile:ultrapure water (9:1 v/v with pH of the water adjusted to 8.9 with diethylamine before adding the organic solvent). The flow rate and injection volume were 0.4 mL/min and 5 μL respectively.

2.4. Method performance evaluation and comparison

Working range, linearity, repeatability, recoveries and matrix effects were evaluated according to Eurachem Fitness for Purpose of Analytical Methods guidelines [43]. To determine working range and linearity, matrix-matched calibrants were used and analysed with triplicate injection (N = 13 calibrants tested; range: ~1–2000 ng L^{-1}). Chromatographic t_R imprecision was determined with seven consecutive injections of a sample spiked at ~500 ng L^{-1}. For recovery and repeatability, n = 7 replicate matrix-matched mixture standards (~500 ng L^{-1}) were analysed after SPE. Background subtraction was also performed using unspiked matrix (n = 3) after SPE. The recovered signal was then compared to additional unspiked samples (n = 3) which were processed using SPE but reconstituted with a solution of ~50 μg L^{-1} of all analytes in injection solvent (theoretical 100 % recovery concentration). The latter solutions were also used to evaluate matrix effect by comparing them to a standard of equivalent concentration prepared directly in injection solvent (n = 3), with no SPE. The equations used for recovery and matrix effect are in Table S5. In this work, all PFAS having linear and branched isomers were quantified as a sum of all isomers. Method recovery was also compared to that of an existing SPE procedure for acids [16], modified to match the concentration factor of the proposed method. Developed for PMOCs analysis, this procedure closely follows the manufacturer's instruction for generic acid compounds analysis without including a washing step. Oasis WAX SPE cartridges for PFAS analysis (150 mg, 6 mL reservoir) were purchased from Waters. The cartridges were conditioned with 5 mL of 2 % formic acid in methanol followed by 5 mL of ultrapure water. A sample volume of 10 mL was loaded and the cartridges eluted with 5 % ammonia in methanol. The eluates were then evaporated to dryness at 38 °C, reconstituted with 100 μL of ultrapure water:methanol 1:1, centrifuged at 2000 rpm for 5 min, and transferred to vial for injection. For this second method, n = 7 preparations in matrix were used in the same way, each again spiked at ~500 ng L^{-1} of all analytes and also prepared in parallel with two sets of triplicates of unspiked samples which were later reconstituted after SPE either in injection solvent (for background subtraction) or with a ~50 ng L^{-1} solution of all analytes (as theoretical 100% recovery), as above. The latter extracts were also used to evaluate matrix effect as above (see Table S5). The same unspiked laboratory water sample prepared in triplicate with both methods was used for a comparison of the number of features generated pre- and post-reduction, as well as the number of substances detected. Suspect screening was separately performed for the features removed by the DOM filter, and library matches were manually reviewed to ensure no false negative results were generated by DOM features reduction (n = 6).

2.5. Creation of a DOM filter for features reduction

The laboratory tap water sample (also used for method performance evaluation and comparison), as well as a 50 mg/L SRFAS in methanol: ultrapure water 1:1, were employed. In order to take into account seasonal and geographical variations of fulvic acid content [44], the DOM composition of SRFAS was considered. Both solutions were run with the proposed LC-QTOF-MS method, and full scan data were analysed. The average of spectra acquired from 3.2 to 5.0 min was used to compile a list of 1955 DOM elemental compositions with correspondent calculated *m/z* values for the $[M\text{-}H]^-$ or $[M+1\text{-}H]^-$. Each elemental composition included in the list was generated manually in Microsoft® Excel (WA, USA) on the basis of well-known spectral patterns [34,45] and verified to be present in the spectrum of either a tap water or fulvic acid

standard. Only elemental compositions containing exclusively carbon, hydrogen and oxygen were included. Manually generated elemental compositions found in the sample with $[M-H]^-$ having a spectral intensity of less than 100 in both spectra were not included in the list, generating a *m/z* interval ranging from 187 to 611 Da (the list is available in Table S6). In this way, elemental compositions unlikely to generate features were excluded to avoid unnecessary computational burden and increased risk of false negatives.

2.6. Data processing workflow for NTA

For quantitative analysis and data-independent analysis (DIA) deconvolution, LabSolutions Insight Explore v3.8 SP1 was employed (Shimadzu Corporation, Kyoto, Japan). Peak detection, alignment, as well as repeatability and blank features reduction filters were performed with MS-DIAL v4.8 freeware (Riken, Kanagawa, Japan) [46] with further details in Table S7. Peak detection parameters were optimised to achieve detection of selected analytes to the lowest calibrant concentration. Sample extracts were run in triplicate injections with the corresponding extracted ultrapure water blank. Features not present in all triplicates or present in the blank with areas above 20% of that in the aligned sample data were removed. The MS-DIAL output was exported to Microsoft Excel, where features with t_R included between 3.2 and 5.0 min and a mass difference with elemental compositions included in the DOM filter equal or lower than 5 ppm were removed. A flowchart representing the whole suspect screening workflow is reported in Fig. S1.

2.7. Suspect screening of hazardous and potentially toxic organic acids

An MS1 database for SS was compiled from various sources: 2153 chemicals with probability of negative ESI ionisation calculated as more than 0.3 in the NORMAN Suspect List Exchange database (SusDat) [47]; 640 PFAS contained in PFASNTREV19 [48] and 541 chlorinated transformation products and disinfection by-products (DBPs) from ChlorineTPS [49], all reported as suitable for negative polarity ESI-MS. The mass accuracy threshold for suspect shortlisting was set to 5 ppm. All library matches were reviewed manually to assess alignment, peak shape, isotopic composition (M, M+1, M+2 etc.) and possible alternative structures in PubChem for the same elemental composition. Retention time shifts and mass accuracy were verified for known compounds in municipal water samples at alignment and suspect screening stages. Samples for qualitative analysis were prepared and analysed together with standards and spiked samples for quality control. For candidate peaks with intense signals, DIA deconvolution was investigated for diagnostic MS/MS fragments. For the remaining candidates, a dedicated product ion scan was performed during additional chromatographic runs (details in Table S4). Tentatively identified substances were confirmed where a standard was commercially available. Toxicity data for substances identified with confidence level 1, 2a and 2b [50] were either found in literature or predicted with the US EPA Toxicity Estimation Software Tool (EPA TEST). The endpoint considered for predicted toxicity were oral rat LD50, developmental toxicity and mutagenicity with nearest neighbour method.

3. Results and discussion

3.1. Analytical method performance for target analytes and comparison with WAX-SPE

The selection of analytes for method performances evaluation was made to cover a broad chemical space of organic acids. To our knowledge, it is the first time that ammonium bicarbonate was used as additive for elution of both strong and weak organic acids from SAX-SPE. Ammonium bicarbonate degrades into ammonia and carbon dioxide at or above 35 °C. This also offers the additional new benefit of a nearly neutral pH extract which significantly reduces the risk of hydrolysis and acid-catalysed reactions. This is especially important for NTA, where the reactivity of unknown substances cannot be easily predicted. The selectivity of bicarbonate for SAX quaternary ammonium is intermediate between that of formate and chloride. Improved selectivity was evident from the full-scan acquisition total ion chromatograms (TICs, Fig. S2): by subtracting the total area of the TIC chromatogram generated by extracted ultrapure water from the average total area generated by samples (n = 3), the proposed method returned 24% of the signal intensity generated by the WAX-SPE comparison method. Furthermore, the WAX-SPE comparison method resulted in significant amounts of brown precipitate absent from the method proposed here, which had to be removed by centrifugation.

Quantitative analytical performance of the full method in matrix-matched standards showed that the lowest calibrated concentration across analytes ranged from 1 to 100 ng L^{-1}, whilst the highest calibrated concentration ranged from 449 to 2148 ng L^{-1}. Coefficients of determination (R^2) were $\geq$ 0.995 for all analytes with at least six calibration points (details in Table S8). Percentage relative standard deviation (% RSD) of t_R for all analytes were $\leq$ 1 %. Percentage recoveries (Fig. 1, Tables S9a and S9b) obtained with the method were consistently high (88 $\pm$ 7 %, n = 7, expressed as mean $\pm$ standard deviation) across all analytes. On average, %RSD for analyte peak area was 7 %, with the poorest precision for 3,5-dichlorobenzoic acid at 15 %. Peak area variations with injection of standard in solvent at the spiking concentration (Table S10) show how the imprecision of the proposed method was mainly driven by analytes that exhibited low ionisation efficiency and consequently low signal in full-scan QTOF-MS (for example, %RSD for dichlorprop was 12 % with the extracted solution, 14 % with the standard in solvent). By comparison, the modified WAX-SPE methodology percent recoveries were lower and less consistent at 63 $\pm$ 9 % (n = 7), with an average %RSD of peak areas of 15 %. Acifluorfen could not be quantified due to co-eluting interferences from the matrix. All analytes affected by poor repeatability with the comparison method showed

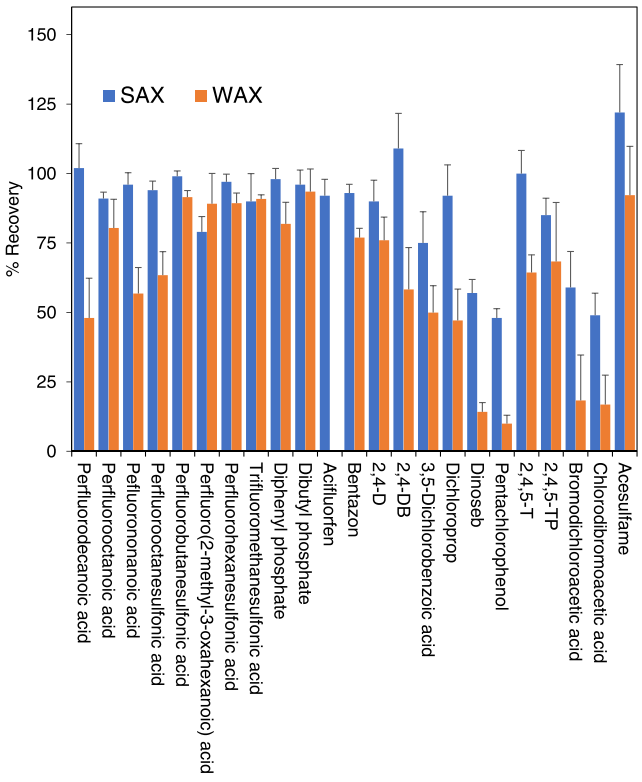


Fig. 1. Average recoveries for 23 selected organic acid contaminants with the proposed (SAX, n = 7) and comparison (WAX, n = 7) SPE methods. Error bars represent the standard deviation.

instrumental repeatability ≤ 14 % RSD in injection solvent at the same concentration in vial. Therefore, with a median instrumental imprecision of 3 % RSD across analytes, the additional variance of the SAX-SPE method was by comparison much lower. The larger imprecision associated with the WAX-SPE approach is due to low and compound-specific ionisation efficiency and signal suppression from co-extracted matrix. However, by employing a more sensitive targeted multiple reaction monitoring approach, the overall imprecision of both methods would be expected to decrease significantly. Matrix effects (mainly as ion suppression) for the proposed method herein were 7 % lower and more consistent on the SAX-based method (12 ± 9 % versus 19 ± 20 % for WAX, expressed as mean ± standard deviation, n = 3, details in Table S11a and S11b). Within the limitations of the experimental design and choice of analytes, silica-based SAX still showed promising performance overall when compared to mixed-mode SPE with a WAX-based resin.

3.2. Effect of feature reduction for NTA of drinking water

After alignment, the total number of unfiltered features generated by analysing the same laboratory sample in triplicate with the proposed SAX-SPE method was, on average, less than a quarter of that generated with the WAX-based SPE method (16,064 versus 67,603 features, respectively, n = 3). Features generated by sample preparation, mobile phases and instrument were successfully removed using both the blank and repeatability filters. Fig. 2 shows the difference in number, *m/z* and intensity of features in the laboratory sample using both extraction approaches after application of blank and repeatability filters. Consequently, this feature reduction approach was applied to all 10 water samples collected from across London using the SAX-SPE method. On average across all samples, 16,638 features were generated in the raw dataset (RSD = 7 %), 7371 were left after the alignment and repeatability filters were applied (RSD = 5 %), and 3930 left after the blank filter (RSD = 9 %). An additional 19 % of the remaining features were removed on average by applying the DOM filter, down to 3176 on average (RSD = 10%). By comparison, the WAX-SPE method generated on average 7543 features in the sample tested in triplicate after all filters were applied. For the SAX-SPE method, subsequent library matches removed with the application of DOM filter were, on average, 29 per sample (n = 6). Isotopic composition, peak shape and DIA MS/MS spectra of all removed features returning a library match showed no potential false negative results. Fig. 3 shows the ion plot of both features removed by the DOM filter and those remaining, generated by the sample used for method comparison extracted with SAX-SPE (n = 3). To our knowledge, this is first time such a DOM feature reduction filter has been successfully implemented and applied for NTA of water. Though peak detection and alignment parameters were set to ensure detection of substances at low concentrations, library matches generated by features with low signal and no detectable isotopic compositions were discarded but via manual data-processing. The development of an automated workflow capable of performing all stages of the workflow for SS was beyond the scope of this work, but should include this DOM feature reduction step, together with customised databases and automatic selection of features with isotopic composition compatible with candidate elemental formula to considerably reduce the time required for data-processing.

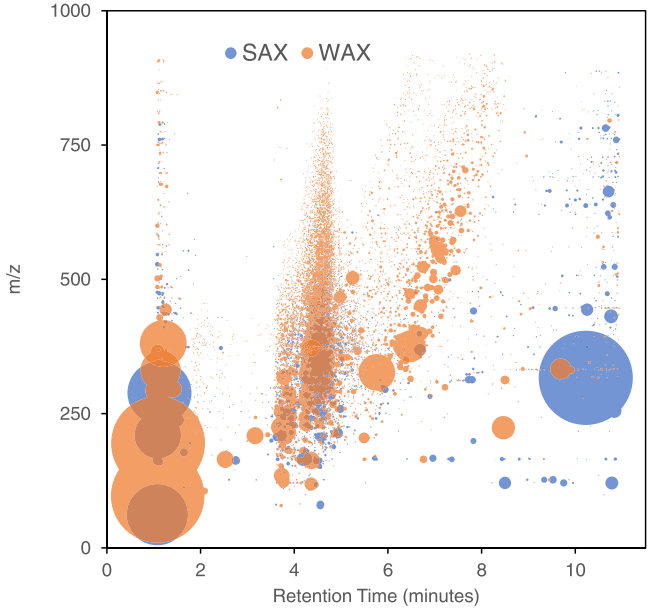


Fig. 2. Bubble chart representing relative intensities of all features generated in a tap water sample (n = 3), after alignment and application of repeatability and blank filters.

3.3. Suspect screening of organic anions

A total of 22 substances were identified with confidence level 1 in the 10 municipal water samples, as well as 4 at level 2a, 1 at level 2b, 11 at level 3 and 21 at level 4 (details in Table S12). Structural elucidation for MS/MS product ions of compounds identified with level of confidence 2a, 2b and 3 was then attempted; their elemental composition is reported in Table S13. For compounds with fragmentation patterns not compatible with tentative structures or compounds with chemistries not compatible with extraction selectivity, molecular structural elucidation was attempted manually. Table 1 reports the final list of substances identified with higher confidence levels 1, 2a and 2b along with their experimental or predicted toxicity data and other details including t_R, CAS number, molecular formula and *m/z*. Detection of well-known and expected compounds showed that the method was effective for reliable targeted analysis. For example, widely reported compounds such as four additional HAAs (of which three were US EPA-regulated) and six perfluorinated acids were found in these samples. However, several new or poorly investigated compounds were also identified showing its added value in NTA. To the best of our knowledge, this is first time dibromethanesulfonic acid was confirmed at confidence Level 1 in drinking water. It was previously reported together with other halogenated sulphonic acids in drinking water at confidence level 2b [51,52]. It was present in all samples, but no experimentally-derived or predicted

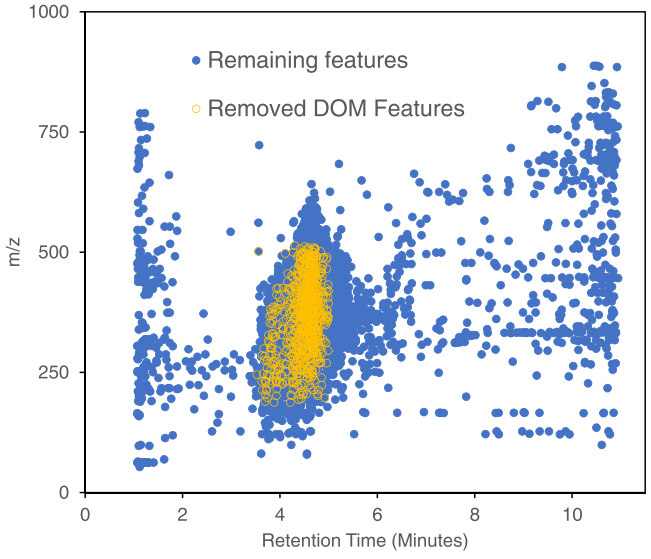


Fig. 3. Ion plot of features removed by DOM filter (yellow) and remaining after DOM filter application (blue) in the same tap water sample (n = 3).

Table 1
Example substances identified with level of confidence 1, 2a and 2b, with experimental or predicted toxicity data, t_r, *m/z* of $[M-H]^-$, elemental composition and CAS when available. (Note: LD50= median lethal dose; NOAEL= no observed adverse effect level; LOAEL= lowest observed effect level; BMDL= benchmark dose level; NOEC= no observed effect concentration).

Analyte (confidence level)	Toxicity - Experimental	Toxicity - Predicted	t_R (min)	*m/z* [M-H]-	Elemental Composition	CAS
Perfluoropentanoic acid (1)		Oral rat LD50 1507 mg/kg; Mutagenicity negative[a]	5.6	262.9760	$C_5HF_9O_2$	2706–90–3
Perfluoropentanesulfonic acid (1)		Oral rat LD50 305 mg/kg; Mutagenicity negative[a]	6.3	348.9398	$C_5HF_{11}O_3S$	2706–91–4
5-Bromosalicylic acid (1)	LD50 100 mg/kg (intravenous, mouse, acute)[b]	Mutagenicity Negative; Developmental NON-toxicant[a]	5.2	214.9349	$C_7H_5BrO_3$	96–97–9
3-Bromo-5-chloro-2-hydroxybenzoic acid (1)		Oral rat LD50 990 mg/kg; Mutagenicity Negative; Developmental NON-toxicant[a]	5.8	248.8960	$C_7H_4BrClO_3$	4068–58–0
Dibromomethanesulfonic acid (1)		Oral rat LD50 141 mg/kg; Mutagenicity Positive (1.0); Developmental toxicant (0.67)[a]	4.0	250.8019	$CH_2Br_2O_3S$	859073–88–4 (sodium salt)
3,5-Dibromo-2-hydroxybenzoic acid (1)		Oral rat LD50 277 mg/kg; Mutagenicity Negative[a]	5.9	292.8454	$C_7H_4Br_2O$	3147–55–5
2,6-Dibromo-4-nitrophenol (1)	LD50 (intravenous, mouse, acute): 56 mg/kg[b]	Mutagenicity Negative[a]	5.6	293.8407	$C_6H_3Br_2NO_3$	99–28–5
6:2 fluorotelomer sulphonic acid (6:2 FTSA) (1)	NOAEL 15 mg/kg-day (oral, rat, reproduction developmental, chronic and subchronic); LOAEL 45 mg/kg-day (oral, rat, reproduction developmental, and chronic); LD50: > 300 < 2 000 mg/kg bw (oral, rat, acute)[c]		6.4	426.9679	$C_8H_5F_{13}O_3S$	27619–97–2
Trichloroacetic acid (1)	LOAEL= 6 mg/kg-day (oral, mouse, chronic); LEL = 4.1 mg/kg-day (oral, rat, subchronic); NOAEL < 4.10 mg/kg-day (oral, rat, subchronic)[d]; LD50 400 mg/kg bw (oral, mouse, acute)[e]		4.5	160.8969	$C_2HCl_3O_2$	76–03–9
Bromochloroacetic acid (1)	LOAEL > 572 mg/kg-day (oral, frog, chronic)[f]	Oral rat LD50 200 mg/kg; Mutagenicity Positive (1.0); Developmental toxicant (0.67)[a]	3.7	170.8853	$C_2H_2BrClO_2$	5589–96–8
Dibromoacetic acid (1)		Oral rat LD50 122 mg/kg; Mutagenicity Positive (1.0); Developmental toxicant (0.67)[a]	3.9	214.8349	$C_2H_2Br_2O_2$	631–64–1
Tribromoacetic acid (1)		Oral rat LD50 75 mg/kg; Mutagenicity Positive (1.0)[a]	4.7	292.7454	$C_2HBr_3O_2$	75–96–7
MCPA (2-methyl-4-chlorophenoxyacetic acid) (1)	NOAEL 0.15 mg/kg-day (oral, dog, chronic)[g]; NOAEL 0.21 mg/kg-day (oral, dog, chronic); LOAEL 1.02 mg/kg-day (oral, dog, chronic); LOAEL 0.2 mg/kg-day (oral, dog, chronic)[d]; LD50 950 mg/kg bw (oral, rat, acute)[e]		4.9	199.0167	$C_9H_9ClO_3$	94–74–6
Mecoprop 2-(4-Chloro-2-methylphenoxy)propanoic acid (1)	LOAEL 0.5 μg/kg-day (topical, house mouse, development); NOAEL 0.5 μg/kg-day (topical, house mouse, development)[h] NOAEL 1 mg/kg-day (oral, rat, chronic)[i]; LOAEL 9 mg/kg-day (oral, rat, subchronic)[d]; LD50 650 mg/kg bw (oral, rat, acute)[c]		5.3	213.0324	$C_{10}H_{11}ClO_3$	93–65–2
Acesulfame (1)	LD50 7431 mg/kg bw (oral, rat, acute)[j]	Developmental NON-toxicant (0.33)[a]	4.2	161.9867	$C_4H_5NO_4S$	33665–90–6
Chlorodibromoacetic acid (1)		Oral rat LD50 122 mg/kg; Mutagenicity Positive (1.0); Developmental toxicant (0.67)[a]	3.8	248.7954	$C_2HBr_2ClO_2$	5278–95–5
Bromodichloroacetic acid (1)		Oral rat LD50 1217 mg/kg; Mutagenicity Positive (0.67); Developmental toxicant (0.67)[a]	3.6	204.8459	$C_2HBrCl_2O_2$	71133–14–7
Perfluorooctanesulfonic acid (1)	NOAEL 0,6 μg/kg-day (oral, human, subchronic)[k]; NOAEL 5 μg/kg-day (oral, mouse, subchronic)[i]; LOAEL 1 mg/kg-day (oral, rabbit, developmental); LOAEL > 2.5 mg/kg-day (oral, rabbit, developmental)[d]		6.659	498.9297	$C_8HF_{17}O_3S$	1763–23–1
Perfluorohexanesulfonic acid (1)	NOAEL 6 μg/kg-day (oral, human, subchronic)[k]; NOAEL 1 mg/kg-day (oral, rat, subchronic)[o]		6.115	398.9361	$C_6HF_{13}O_3S$	355–46–4
Perfluorobutanesulfonic acid (1)	BMDL 18.9 mg/kg-day (oral, rat, chronic and subchronic)[i]; NOAEL 60 mg/kg-day (oral, rat,		5.237	298.9425	$C_4HF_9O_3S$	375–73–5

(*continued on next page*)

Table 1 (*continued*)

Analyte (confidence level)	Toxicity - Experimental	Toxicity - Predicted	t_R (min)	*m/z* [M-H]-	Elemental Composition	CAS
	subchronic)[i]; LD50 2000 mg/kg bw (oral, rat, acute)[c]					
Trifluoromethane sulphonic acid (1)	NOAEL 40 mg/kg-day (oral, rat, subacute); LD50 400 mg/kg bw (oral, rat, acute)[c]		1.94	148.9526	CF_3SO_3H	1493–13–6
Perfluoroheptanoic acid (1)	NOEC 5 mg/kg (oral, rainbow trout, acute morphology); LOAEL 100 mg/kg-soil (environmental, earthworm, acute mortality); LC50 > 2.8 mM (static, water flea, acute)[h]		6.3	362.9696	$C_7HF_{13}O_2$	375–85–9
Dichlorometane sulphonic acid (2a)		Oral rat LD50 807 mg/kg; Mutagenicity Negative; Developmental toxicant (0.67)[a]	3.3	162.9029	$CH_2Cl_2O_3S$	not found
Bromochloromethanesulfonic acid (2a)		Oral rat LD50 1024 mg/kg; Mutagenicity Negative; Developmental toxicant (0.67)[a]	3.7	206.8524	CH_2BrClO_3S	not found
Ricinoleic acid (2a)		Oral rat LD50 12,931 mg/kg; Mutagenicity Negative; Developmental toxicant (0.67)[a]	7.5	297.2435	$C_{18}H_{34}O_3$	141–22–0
2,2,4-tribromo-5-hydroxycyclopent-4-ene-1,3-dione (2a)		Oral rat LD50 113 mg/kg; Mutagenicity Positive (1.0);	5	344.7403	$C_5HBr_3O_3$	not found
2,2,4-trichloro-5-hydroxycyclopent-4-ene-1,3-dione (2b)		Oral rat LD50 627 mg/kg; Mutagenicity Positive (0,67); Developmental NON-toxicant[a]	4.9	212.8919	$C_5HCl_3O_3$	not found

[a] U.S. Environmental Protection Agency Toxicity Estimation Tool (TEST).
[b] U.S. Army Armament Research & Development Command.
[c] European Chemicals Agency eChemPortal 2020.
[d] U.S. Environmental Protection Agency ToxRefDB.
[e] World Health Organisation International Programme on Chemical Safety IPCS.
[f] Naomi Weber, Ty Higuchi, John Tessari & D. N. Rao Veeramachaneni (2004) Evaluation of the Effects of Water Disinfection By-products, Bromochloroacetic and Dibromoacetic Acids, on Frog Embryogenesis, Journal of Toxicology and Environmental Health, Part A,67:12, 929–939.
[g] U.S. Environmental Protection Agency Health Effects Assessment Summary Tables.
[h] U.S. Environmental Protection Agency ECOTOXicology Database.
[i] European Food Safety Authority.
[j] O'Neil, M.J. (ed.). The Merck Index - An Encyclopaedia of Chemicals, Drugs, and Biologicals. Cambridge, UK: Royal Society of Chemistry, 2013., p. 8.
[k] U.S. Agency for Toxic Substances and Disease Registry Minimal Risk Levels 2020.
[l] Provisional Peer-Reviewed Toxicity Values (Oak Ridge National Laboratory).

toxicity data could be found in the literature. However, dibromomethanesulfonic acid was predicted here with EPA TEST as both mutagenic and a developmental toxicant, while dichloromethanesulfonic and bromochloromethanesulfonic acid were predicted as developmental toxicant only. Two metabolites of branched alkylbenzene sulfonates, 5-methyl-6-(4-sulfophenyl)heptanoic acid and 3,3,4-trimethyl-5-(4-sulfophenyl)hex-4-enoic acid were tentatively identified respectively in eight and 10 out of 10 samples. MS/MS spectra showed a high number of characteristic product ions, highly informative about their alkyl chains structures (details in Table S13). Despite this richness of MS/MS information and the high probability of para-position substitution, no certain structure based solely on MS data could be assigned and they were reported at Level 3.

The importance of increased method selectivity for acids for high-quality structural elucidation was experienced multiple times whilst processing SS results. Overall, the method performance evaluation showed that the number of substances detected with the SAX-SPE method was consistently higher than that of the WAX-SPE method and regardless of the confidence level, with a total of 45 substances detected with SAX, versus 38 with WAX (n = 3, details in Table S14). Only one substance was detected exclusively with the comparison method (2-bromo-4-nitrophenol, at Level 4). The extracted ion chromatogram and spectra of 2,2,4-tribromo-5-hydroxycyclopent-4-ene-1,3-dione, dibromomethanesulfonic acid, and 6:2 fluorotelomer sulphonic acid (6:2 FTSA) in a London tap water sample are shown in Fig. 4 as examples of the quality of data generated by this method. An example of how selective extraction could facilitate qualitative analysis was in the case of a peak detected in one sample and returning library matches for three isomers of monochloro-cresols (with a measured $[M-H]^-$ *m/z* of 141.0110 and a matching elemental composition of $C_7H_6ClO^-$). However, these structures are not expected to be acidic and consequently are not expected to be retained by this silica-based SAX-SPE method. Confidence was further enhanced as a washing step could be included during SPE, which likely further enhanced anion selectivity through elimination of weak hydrogen bonding or dipole interactions with the silica polymer backbone. Such washing steps are generally not included in any other NTA sample preparation approaches to our knowledge. Research into other relevant substances revealed 4-chloro-2-methylphenolate to be an in-source fragmentation product of 2-methyl-4-chlorophenoxyacetic acid (MCPA), a well-known acidic herbicide. The extracted ion chromatogram of the $[M-H]^-$ of both herbicides showed matching peaks. The identity was further confirmed by injection of a reference material. An erroneous identification was consequently averted using this approach.

3.4. Structural elucidation of halogenated hydroxycyclopentenediones (HCDs)

Features compatible with halogenated HCDs were detected at high frequency: eight of these acidic substances were observed in on average in nine out of 10 samples. Tentative identifications of these features have previously been reported in drinking water, but with uncertainty about

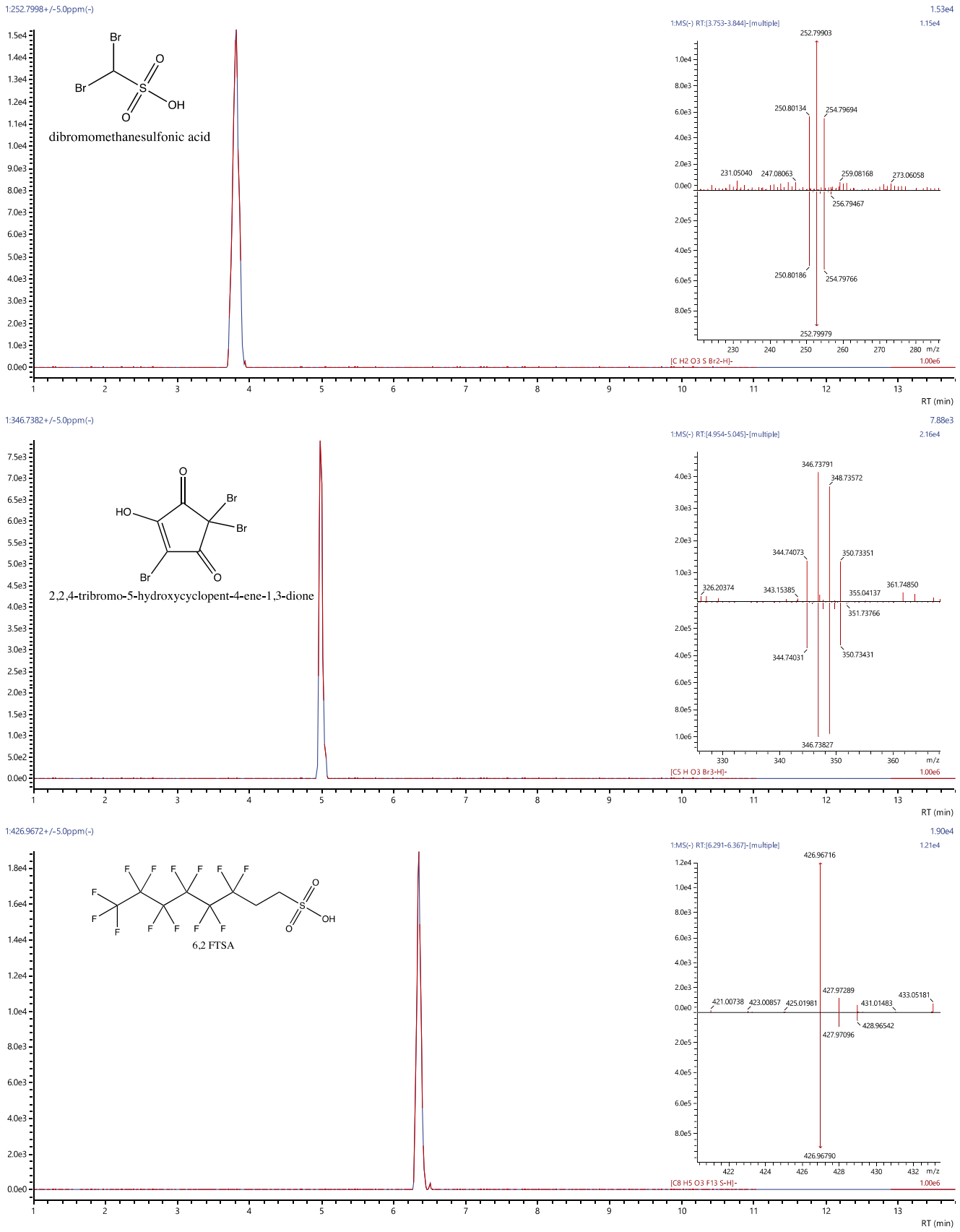


Fig. 4. Extracted ion chromatogram, spectra and structures of three substances of interest identified in a London tap water sample.

their structures [19,53]. In particular, halogenated furoic acids were proposed as alternative structures for the same elemental compositions. No evidence of CO_2 neutral loss from deprotonated molecular ion (expected with a furoic acid structure) was observed in any of the MS/MS spectra of these species. The MS/MS spectra of trihalogenated-HCDs (Fig. S3) showed an unusual loss of a radical halogen, followed by neutral loss of CO: this distinctive fragmentation pattern suggests the formation of a radical on a sp^3 carbon (when considering the original molecule), substituted with another halogen and further stabilised by resonance. The only potential structures of relevant elemental composition carrying a sp^3 carbon are HCDs and methylfurandiones. However, the latter are not acidic and consequently unlikely to be retained during sample preparation or efficiently ionised in negative ESI (predicted pK_a and logD at pH 7.5 for 2,2,4-tribromo-5-hydroxycyclopent-4-ene-1, 3-dione were respectively 4.5 and −1.32, source ACD Labs Percepta). Other common and characteristic MS/MS fragments were C_4ClO^- (dichloro, bromodichloro and dibromochloro HCDs) and C_4BrO^- (tribromo, dibromo and dibromochloro HCDs). These MS fragmentation patterns match those found by Pan et al. [54], who previously reported synthesis, isolation and characterisation of tribromo-HCD. However, Pan et al. have identified the compound by combining exclusively MS and infra-red (IR) spectroscopy data, and employing the IR spectrum to define the functional groups present on the molecule. Structures like hydroxypyranones, which have been excluded by Pan et al. solely on the basis of an IR spectrum, require additional evidence to be ruled out. Furthermore, no nuclear magnetic resonance (NMR) spectrum was published, leaving uncertainty about the position of functional groups in the molecule. The molecular structure was presented in a figure as one of the possible isomers (2,2,4-tribromo-5-hydroxycyclopent-4-ene-1, 3-dione), without specifying the structure with a name.

Attempts to synthesise and isolate enough tribromo-HCD to record a ^{13}C NMR spectrum failed (details in the SI), due to poor recovery from the aqueous and severe product losses during evaporation. However, by spiking drinking water samples, it was possible to confirm that the synthesised substance was very likely the same as previously observed. Liquid/liquid extraction of purified fractions (containing 85 % water and 15 % acetonitrile) with dichloromethane gave a 1:9 partition between the organic and the aqueous solvents. Evaporation of a purified fraction at 38 °C under a 0.5 mL/min stream of nitrogen employing Turbovap gave an average recovery of 10 % (n = 3). Freeze-drying of the same fractions after preliminary freezing [55] gave on average 13 % recovery (Labconco, Kansas City, US). It is unclear whether losses during evaporation are due to compound volatility or to a possible equilibrium with volatile species (in analogy with HAAs).

Criquet et al. investigated the reactivity of phenolic compounds with bromine [56], and found that compounds with three hydroxyl groups in *meta* position to each other (like the one selected by Pan et al. as precursor) react with bromine both by electron transfer with oxidation of hydroxyls to carbonyls, and by electrophilic aromatic substitutions of hydrogen with bromine. Chlorination of resorcinol, a compound with two hydroxyl groups in *meta* instead of three, generates a compound tentatively identified as trichloro-cyclopentenedione [57], which is missing a hydroxyl group when compared to trichloro-HCD. Formation pathways for tribromo-HCD have been proposed by Pan et al. and Huang et al. *Huang et al.,* [58]. However, an alternative pathway where the reaction mechanisms governing ring-opening and closure are clearly defined is presented in Fig. 5. The irreversible retro-aldol ring-opening reaction in an acidic environment leads to decarboxylation and subsequent intra-molecular aldol ring-closure, followed by oxidation. Substances with elemental compositions compatible with intermediates I, II and III were detected. Despite the pathway proposed here representing 2,4,6-trixhydroxybenzaldehyde as a precursor, intermediate II should be considered a "universal" precursor of tribromo-HCD, a structure that any potential polyphenolic precursor would have to generate for the reaction to proceed further. Consequently, the reactions occurring from intermediate II onwards can be adopted for all potential precursors. Two key elements allowed definition of an individual structure for tribromo-HCD: (a) electrophilic aromatic substitution of bromine occurs exclusively with bromine atoms in the *meta* position (determining bromines in beta position in the HCD); (b) the regioselectivity of hypobromous acid electrophilic addition to double bonds prevents additional hydroxyls from creating a bond with carbons already bound to bromine (determining the position of bromines and hydroxyl in the HCD). The generation of trichloro-HCD with hypochlorous acid should follow an identical formation pathway. Consequently, all considerations made above for structural elucidation can be extended to trichloro-HCD.

Given that tribromo-HCD has previously been isolated, that a MS/MS spectrum was made available, and that an unambiguous spectrum-structure match could be hypothesised, this substance was reported at Level 2a here ("probable structure"). Its analogue, 2,2,4-trichloro-5-

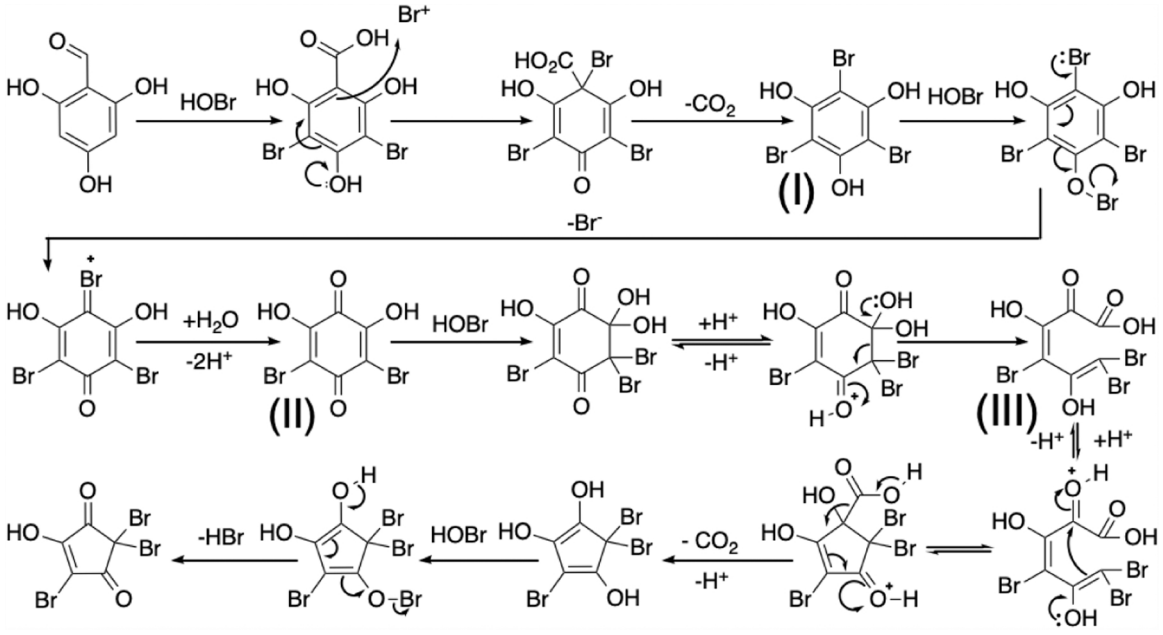


Fig. 5. Proposed formation pathway for tribromo-HCD.

hydroxycyclopent-4-ene-1,3-dione, was reported at Level 2b here on the basis of MS/MS fragmentation pattern presenting all key events that, in analogy with 2,2,4-tribromo-5-hydroxycyclopent-4-ene-1,3-dione, lead to the identification of a single structure. However, it must be noted that final confirmation with a third analytical technique is still required to achieve unequivocal structural elucidation, and that arguably it is difficult to classify such cases employing confidence levels as defined by Schymanski et al. The other five halogenated HCDs were reported at Level 3 and one at Level 4 because either multiple structural isomers and enantiomers were possible for each one of them, or an MS/MS spectrum could not be obtained due to low signal. No toxicity data could be found for these substances either, but all halogenated HCDs structures reported in this paper at all levels of confidence were predicted to be mutagenic by the EPA TEST software. Commercial reference materials were not available, indicating that further investigations are needed to assess their toxicity and exposure via drinking water.

3.5. Quantification of confirmed substances

Substances identified at confidence level 1 were quantified in a seven-day composite sample of laboratory tap water (50 mL aliquots sampled daily on workdays from 31st October until 8th November, 2022). Standard addition regression included a minimum of six spiked standards for each compound. Unspiked samples and ultrapure water extracted blanks were prepared in triplicate. Method detection and quantification limits were estimated with extracted reagent blanks (10 measurements), following strictly sections 6.2.3 and 6.2.4 of the Eurachem guidelines [43]. Limits of detection (LOD) and quantification (LOQ), average calculated concentrations and standard deviations for 12 compounds found in this particular sample (out of 22 in total identified in all samples analysed with qualitative analysis) are reported in Table S15. Coefficients of determination were ≥ 0.995 for all analytes with at least six calibration points. Recoveries of compounds not included in 3.1 were not tested. However, in consideration of linearity data and results reported in 3.1 and Table S15, the authors were confident in the quantitative results. Dibromomethanesulfonic acid was the compound found at the highest concentration (934 ng L^{-1}), in line with the concentration ranges of other halogenated methanesulfonicacids previously found in drinking water [52]. Trichloroacetic, bromodichloro and chlorodibromo acetic acids were found at concentrations of 725, 256 and 189 ng L^{-1} respectively (4.4, 1.2 and 0.7 nM). These DBPs are cytotoxic, genotoxic, mutagenic and teratogenic though at higher concentrations (>1000 nM) than were detected here [59-62]. The US EPA and the EU set a maximum contaminant level (MCL) for five specific HAAs at 60 μg L^{-1}, but this list does not include two of the HAA substances detected here (chlorodibromo- and bromodichloroacetic acid). Five PFAS were also quantified at concentrations < 10 ng L^{-1}, and more evidence is emerging showing their effects in humans. In comparison to other developed countries, very little occurrence data exists for PFAS in municipal drinking water in the UK and no NTA methods are currently employed for shortlisting of relevant compounds. Lastly, acesulfame is an artificial sweetener and has been shown to survive treatment to result in drinking water concentrations at the ng-μg/L range [63] and was quantified here, at 144 ng L^{-1}.

4. Conclusions

A new and improved SS workflow has been successfully developed and applied to selective identification of acidic substances in drinking water using SAX-SPE and mixed-mode RP-AX LC-HRMS for the first time. In comparison to the current standard method for analysis of acidic substances, this new approach offered broadly better recovery, repeatability, fewer matrix effects and better selectivity for a selection of 23 analytes, including some containing only one carbon (trifluoromethanesulfonic acid). This could find broad application also for quantitative analysis, where much decreased matrix interference is highly desirable to improve accuracy and sensitivity. A comprehensive SS data-processing workflow was also tailored for acidic components including peak detection, comprehensive feature reduction, a novel filter to eliminate features generated by DOM, and a database for SS of acidic contaminants. The application of this new DOM filter significantly reduced the number of incorrect library hits without generating false negatives. To our knowledge, for the first time selective SPE extraction was employed not just to enhance detection, but also to support substance identification by excluding potential structures not presenting affinity for the sorbent. We believe this novel concept constitutes a step forward towards enabling more reliable structural confirmation and elucidation in SS (and indeed also NTA). Analysis of 10 municipal water samples from across London confirmed the presence of 22 substances including several poorly investigated contaminants. Notable novel findings prompted by SS were: (a) dibromomethanesulfonic acid was confirmed to be present in drinking water for the first time with a standard, then quantified as the most concentrated analyte, and finally predicted as mutagenic and developmental toxicant; (b) novel reaction mechanisms for the formation of tribromo- and trichloro- HCDs were proposed together with in-depth analysis of MS/MS fragmentation patterns, allowing for previously unsupported unequivocal structural elucidation; furthermore, halogenated HCDs were predicted to be mutagenic. Occurrence of dibromomethanesulfonic acid and halogenated HCDs indicate the need for further investigation to rule out potential threats to public health from these substances. Although further automation is needed, overall, this method constitutes an optimal platform for improved detection and identification of acidic contaminants and pollutants in drinking water.

Environmental Implication

A novel suspect screening and non-target analysis workflow, which improves detection and facilitates better identification of organic acidic contaminants in drinking water is presented. A novel and selective extraction, mixed-mode chromatography and a fully optimised data-processing workflow are introduced. Classes of pollutants like per- and polyfluoroalkyl substances or disinfection by-products include thousands of molecules, the majority of which have not been identified and/or routinely monitored. This gap can be addressed with this method, and is suitable for chemically diverse acidic species, as demonstrated by the occurrence of several new, poorly investigated and/or potentially toxic compounds in municipal drinking water.

Authorship Statement

Davide Ciccarelli: Conceptualisation, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Visualization**; D. Christopher Braddock:** Conceptualisation, Formal analysis, Writing – review and editing; **Andrew J. Surman:** Formal analysis, Investigation, Writing – review and editing, Supervision; **Blanca Ivonne Vergara Arenas:** Investigation, Supervision; **Tara Salal:** Investigation; **Tim Marczylo:** Writing – review and editing, Supervision; **Paolo Vineis:** Writing – review and editing, Supervision; **Leon Barron:** Conceptualisation, Methodology, Validation, Writing - original draft, Writing – review and editing, Supervision, Funding acquisition, Visualisation, Project administration.

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.

Data availability

Data will be made available on request.

Acknowledgements

This study was funded by the National Institute for Health and Care Research (NIHR) Health Protection Research Units in Chemical and Radiation Threats and Hazards and Environmental Exposures and Health, both partnerships between the UK Health Security Agency and Imperial College London. The views expressed are those of the authors and not necessarily those of the NIHR, UK Health Security Agency or the Department of Health and Social Care. Blanca Ivonne Vergara Arenas was financially supported by the Council of Science and Technology of Mexico (CONACYT, Postdoctoral Fellowship Grant). We are grateful to Christine Hinz for her support with Shimadzu software and data processing with MS-DIAL, to Carmelo Ferraro for the support with Excel, to Kieran Nolan for advice on compound synthesis, and to Timothy Gant for feedback. This work was pre-printed on ChemRxiv (DOI: 10.26434/chemrxiv-2022-df8gb-v3).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jhazmat.2023.130906.

References

[1] Newton, S.R., McMahen, R.L., Sobus, J.R., Mansouri, K., Williams, A.J., McEachran, A.D., et al., 2018. Suspect screening and non-targeted analysis of drinking water using point-of-use filters. Environ Pollut 234, 297–306.
[2] Tröger, R., Ren, H., Yin, D., Postigo, C., Nguyen, P.D., Baduel, C., et al., 2021. What's in the water?–Target and suspect screening of contaminants of emerging concern in raw water and drinking water from Europe and Asia. Water Res 198, 117099.
[3] Wang, Y.-Q., Hu, L.-X., Zhao, J.-H., Han, Y., Liu, Y.-S., Zhao, J.-L., et al., 2022. Suspect, non-target and target screening of pharmaceuticals and personal care products (PPCPs) in a drinking water system. Sci Total Environ 808, 151866.
[4] Wang, Z., Walker, G.W., Muir, D.C., Nagatani-Yoshida, K., 2020. Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. Environ Sci Technol 54 (5), 2575–2584.
[5] Bieber, S., Greco, G., Grosse, S., Letzel, T., 2017. RPLC-HILIC and SFC with mass spectrometry: polarity-extended organic molecule screening in environmental (water) samples. Anal Chem 89 (15), 7907–7914.
[6] Hernández, F., Ibáñez, M., Portolés, T., Cervera, M.I., Sancho, J.V., López, F.J., 2015. Advancing towards universal screening for organic pollutants in waters. J Hazard Mater 282, 86–95.
[7] Lara-Martín, P.A., Chiaia-Hernández, A.C., Biel-Maeso, M., Baena-Nogueras, R.M., Hollender, J., 2020. Tracing urban wastewater contaminants into the Atlantic ocean by nontarget screening. Environ Sci Technol 54 (7), 3996–4005.
[8] Čelić, M., Jaén-Gil, A., Briceño-Guevara, S., Rodríguez-Mozaz, S., Gros, M., Petrović, M., 2021. Extended suspect screening to identify contaminants of emerging concern in riverine and coastal ecosystems and assessment of environmental risks. J Hazard Mater 404, 124102.
[9] Gago-Ferrero, P., Bletsou, A.A., Damalas, D.E., Aalizadeh, R., Alygizakis, N.A., Singer, H.P., et al., 2020. Wide-scope target screening of> 2000 emerging contaminants in wastewater samples with UPLC-Q-ToF-HRMS/MS and smart evaluation of its performance through the validation of 195 selected representative analytes. J Hazard Mater 387, 121712.
[10] Köke, N., Zahn, D., Knepper, T.P., Frömel, T., 2018. Multi-layer solid-phase extraction and evaporation—enrichment methods for polar organic chemicals from aqueous matrices. Anal Bioanal Chem 410 (9), 2403–2411.
[11] Neuwald, I., Muschket, M., Zahn, D., Berger, U., Seiwert, B., Meier, T., et al., 2021. Filling the knowledge gap: a suspect screening study for 1310 potentially persistent and mobile chemicals with SFC-and HILIC-HRMS in two German river systems. Water Res 204, 117645.
[12] Mechelke, J., Longrée, P., Singer, H., Hollender, J., 2019. Vacuum-assisted evaporative concentration combined with LC-HRMS/MS for ultra-trace-level screening of organic micropollutants in environmental water samples. Anal Bioanal Chem 411 (12), 2555–2567.
[13] Montes, R., Aguirre, J., Vidal, X., Rodil, R., Cela, R., Quintana, J.B., 2017. Screening for polar chemicals in water by trifunctional mixed-mode liquid chromatography–high resolution mass spectrometry. Environ Sci Technol 51 (11), 6250–6259.
[14] Schulze, S., Zahn, D., Montes, R., Rodil, R., Quintana, J.B., Knepper, T.P., et al., 2019. Occurrence of emerging persistent and mobile organic contaminants in European water samples. Water Res 153, 80–90.
[15] Gago-Ferrero, P., Schymanski, E.L., Bletsou, A.A., Aalizadeh, R., Hollender, J., Thomaidis, N.S., 2015. Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS. Environ Sci Technol 49 (20), 12333–12341.
[16] Montes, R., Rodil, R., Cela, R., Quintana, J.B., 2019. Determination of persistent and mobile organic contaminants (PMOCs) in water by mixed-mode liquid chromatography–tandem mass spectrometry. Anal Chem 91 (8), 5176–5183.
[17] Kimura, S.Y., Cuthbertson, A.A., Byer, J.D., Richardson, S.D., 2019. The DBP exposome: development of a new method to simultaneously quantify priority disinfection by-products and comprehensively identify unknowns. Water Res 148, 324–333.
[18] Postigo, C., Andersson, A., Harir, M., Bastviken, D., Gonsior, M., Schmitt-Kopplin, P., et al., 2021. Unraveling the chemodiversity of halogenated disinfection by-products formed during drinking water treatment using target and non-target screening tools. J Hazard Mater 401, 123681.
[19] Tao, D., Wang, R., Shi, S., Yun, L., Tong, R., Guo, W., et al., 2020. The identification of halogenated disinfection by-products in tap water using liquid chromatography–high resolution mass spectrometry. Sci Total Environ 740, 139888.
[20] Wang, Y., Luo, C., Yang, M., Ren, J., Wang, W., Yong, L., et al., 2020. Target quantification and semi-target screening of halogenated carboxylic acids in drinking water using ultra-high performance liquid chromatography-quadrupole orbitrap high-resolution mass spectrometry. J Chromatogr A 1614, 460710.
[21] Guardian, M.G.E., Boongaling, E.G., Bernardo-Boongaling, V.R.R., Gamonchuang, J., Boontongto, T., Burakham, R., et al., 2020. Prevalence of per- and polyfluoroalkyl substances (PFASs) in drinking and source water from two Asian countries. Chemosphere 256, 127115.
[22] Wang, Y.-Q., Hu, L.-X., Liu, T., Zhao, J.-H., Yang, Y.-Y., Liu, Y.-S., et al., 2022. Per- and polyfluoralkyl substances (PFAS) in drinking water system: target and non-target screening and removal assessment. Environ Int 163, 107219.
[23] Hinnenkamp, V., Balsaa, P., Schmidt, T.C., 2022. Target, suspect and non-target screening analysis from wastewater treatment plant effluents to drinking water using collision cross section values as additional identification criterion. Anal Bioanal Chem 414 (1), 425–438.
[24] Kolkman, A., Vughs, D., Sjerps, R., Kooij, P.J., van der Kooi, M., Baken, K., et al., 2021. Assessment of highly polar chemicals in Dutch and Flemish drinking water and its sources: presence and potential risks. ACS EST Water 1 (4), 928–937.
[25] Tisler, S., Christensen, J.H., 2022. Non-target screening for the identification of migrating compounds from reusable plastic bottles into drinking water. J Hazard Mater 429, 128331.
[26] Diera, T., Thomsen, A.H., Tisler, S., Karlby, L.T., Christensen, P., Rosshaug, P.S., et al., 2022. A non-target screening study of high-density polyethylene pipes revealed rubber compounds as main contaminant in a drinking water distribution system. Water Res, 119480.
[27] Taylor, A.C., Mills, G.A., Gravell, A., Kerwick, M., Fones, G.R., 2022. Pesticide fate during drinking water treatment determined through passive sampling combined with suspect screening and multivariate statistical analysis. Water Res 222, 118865.
[28] Hohrenk, L.L., Itzel, F., Baetz, N., Tuerk, J., Vosough, M., Schmidt, T.C., 2019. Comparison of software tools for liquid chromatography–high-resolution mass spectrometry data processing in nontarget screening of environmental samples. Anal Chem 92 (2), 1898–1907.
[29] Jacob, P., Wang, R., Ching, C., Helbling, D.E., 2021. Evaluation, optimization, and application of three independent suspect screening workflows for the characterization of PFASs in water. Environ Sci: Process Impacts 23 (10), 1554–1565.
[30] Myers, O.D., Sumner, S.J., Li, S., Barnes, S., Du, X., 2017. Detailed investigation and comparison of the XCMS and MZmine 2 chromatogram construction and chromatographic peak detection methods for preprocessing mass spectrometry metabolomics data. Anal Chem 89 (17), 8689–8695.
[31] Bader, T., Schulz, W., Kümmerer, K., Winzenbacher, R., 2016. General strategies to increase the repeatability in non-target screening by liquid chromatography-high resolution mass spectrometry. Anal Chim Acta 935, 173–186.
[32] Fraisier-Vannier, Ol, Chervin, J., Cabanac, G., Puech, V., Fournier, S., Durand, V., et al., 2020. MS-CleanR: a feature-filtering workflow for untargeted LC–MS based metabolomics. Anal Chem 92 (14), 9971–9981.
[33] Helmus, R., Ter Laak, T.L., van Wezel, A.P., de Voogt, P., Schymanski, E.L., 2021. patRoon: open source software platform for environmental mass spectrometry based non-target screening. J Cheminform 13 (1), 1–25.
[34] Li, L., Zhuo, X., Fang, Z., Wang, X., 2021. Investigation of the molecular structure complexity of dissolved organic matter by UPLC-orbitrap MS/MS. Talanta 230, 122320.
[35] Sanchís, J., Jaén-Gil, A., Gago-Ferrero, P., Munthali, E., Farré, M.J., 2020. Characterization of organic matter by HRMS in surface waters: effects of chlorination on molecular fingerprints and correlation with DBP formation potential. Water Res 176, 115743.
[36] Verkh, Y., Rozman, M., Petrovic, M., 2018. A non-targeted high-resolution mass spectrometry data analysis of dissolved organic matter in wastewater treatment. Chemosphere 200, 397–404.
[37] Gough, R., Holliman, P.J., Willis, N., Freeman, C., 2014. Dissolved organic carbon and trihalomethane precursor removal at a UK upland water treatment works. Sci Total Environ 468, 228–239.
[38] Björnsdotter, M.K., Yeung, L.W., Kärrman, A., Jogsten, I.E., 2019. Ultra-short-chain perfluoroalkyl acids including trifluoromethane sulfonic acid in water connected to known and suspected point sources in Sweden. Environ Sci Technol 53 (19), 11093–11101.
[39] Gallidabino, M.D., Hamdan, L., Murphy, B., Barron, L.P., 2018. Suspect screening of halogenated carboxylic acids in drinking water using ion exchange chromatography–high resolution (Orbitrap) mass spectrometry (IC-HRMS). Talanta 178, 57–68.

[40] Cosgrove, S., Jefferson, B., Jarvis, P., 2019. Pesticide removal from drinking water sources by adsorption: a review. Environ Technol Rev 8 (1), 1–24.
[41] Funk, S.P., Duffin, L., He, Y., McMullen, C., Sun, C., Utting, N., et al., 2019. Assessment of impacts of diphenyl phosphate on groundwater and near-surface environments: Sorption and toxicity. J Contam Hydrol 221, 50–57.
[42] Buchner, E.-M., Happel, O., Schmidt, C.K., Scheurer, M., Schmutz, B., Kramer, M., et al., 2019. Approach for analytical characterization and toxicological assessment of ozonation products in drinking water on the example of acesulfame. Water Res 153, 357–368.
[43] B. Magnusson and U. Ornemark, Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0. Available from www.eurachem.org, 2014.
[44] Sharp, E.L., Parsons, S.A., Jefferson, B., 2006. Seasonal variations in natural organic matter and its impact on coagulation in water treatment. Sci Total Environ 363 (1–3), 183–194.
[45] D'Andrilli, J., Foreman, C.M., Marshall, A.G., McKnight, D.M., 2013. Characterization of IHSS Pony Lake fulvic acid dissolved organic matter by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and fluorescence spectroscopy. Org Geochem 65, 19–28.
[46] Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., et al., 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods 12 (6), 523–526.
[47] Dulio, V., van Bavel, B., Brorström-Lundén, E., Harmsen, J., Hollender, J., Schlabach, M., et al., 2018. Emerging pollutants in the EU: 10 years of NORMAN in support of environmental policies and regulations. Environ Sci Eur 30 (1), 1–13.
[48] Liu, Y., D'Agostino, L.A., Qu, G., Jiang, G., Martin, J.W., 2019. High-resolution mass spectrometry (HRMS) methods for nontarget discovery and characterization of poly-and per-fluoroalkyl substances (PFASs) in environmental and human samples. Trends Anal Chem 121, 115420.
[49] Postigo, C., Gil-Solsona, R., Herrera-Batista, M.F., Gago-Ferrero, P., Alygizakis, N., Ahrens, L., et al., 2021. A step forward in the detection of byproducts of anthropogenic organic micropollutants in chlorinated water. Trends Environ Anal Chem 32, e00148.
[50] Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., et al., 2014. Identifying Small Molecules Via High Resolution Mass Spectrometry: Communicating Confidence. ACS Publications.
[51] Zahn, D., Frömel, T., Knepper, T.P., 2016. Halogenated methanesulfonic acids: a new class of organic micropollutants in the water cycle. Water Res 101, 292–299.
[52] Zahn, D., Meusinger, R., Frömel, T., Knepper, T.P., 2019. Halomethanesulfonic acids – a new class of polar disinfection byproducts: standard synthesis, occurrence, and indirect assessment of mitigation options. Environ Sci Technol 53 (15), 8994–9002.
[53] Gonsior, M., Schmitt-Kopplin, P., Stavklint, H., Richardson, S.D., Hertkorn, N., Bastviken, D., 2014. Changes in dissolved organic matter during the treatment processes of a drinking water plant in Sweden and formation of previously unknown disinfection byproducts. Environ Sci Technol 48 (21), 12714–12722.
[54] Pan, Y., Li, W., Li, A., Zhou, Q., Shi, P., Wang, Y., 2016. A new group of disinfection byproducts in drinking water: trihalo-hydroxy-cyclopentene-diones. Environ Sci Technol 50 (14), 7344–7352.
[55] Petersen, E., Lorentzen, J., Flink, J., 1973. Influence of freeze-drying parameters on the retention of flavor compounds of coffee. J Food Sci 38 (1), 119–122.
[56] Criquet, J., Rodriguez, E.M., Allard, S., Wellauer, S., Salhi, E., Joll, C.A., et al., 2015. Reaction of bromine and chlorine with phenolic compounds and natural organic matter extracts–Electrophilic aromatic substitution and oxidation. Water Res 85, 476–486.
[57] Boyce, S.D., Hornig, J.F., 1983. Reaction pathways of trihalomethane formation from the halogenation of dihydroxyaromatic model compounds for humic acid. Environ Sci Technol 17 (4), 202–211.
[58] Huang, Y., Li, H., Zhou, Q., Li, A., Shuang, C., Xian, Q., et al., 2018. New phenolic halogenated disinfection byproducts in simulated chlorinated drinking water: identification, decomposition, and control by ozone-activated carbon treatment. Water Res 146, 298–306.
[59] Pals, J.A., Ang, J.K., Wagner, E.D., Plewa, M.J., 2011. Biological mechanism for the toxicity of haloacetic acid drinking water disinfection byproducts. Environ Sci Technol 45 (13), 5791–5797.
[60] Rahman, M.B., Driscoll, T., Cowie, C., Armstrong, B.K., 2010. Disinfection by-products in drinking water and colorectal cancer: a meta-analysis. Int J Epidemiol 39 (3), 733–745.
[61] Richardson, S.D., Plewa, M.J., Wagner, E.D., Schoeny, R., DeMarini, D.M., 2007. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. Mutat Res/Rev Mutat Res 636 (1–3), 178–242.
[62] Villanueva, C.M., Cantor, K.P., Cordier, S., Jaakkola, J.J., King, W.D., Lynch, C.F., et al., 2004. Disinfection byproducts and bladder cancer: a pooled analysis. Epidemiology 357–367.
[63] Dietrich, A.M., Pang, Z., Zheng, H., Ma, X., 2021. Mini Review: will artificial sweeteners discharged to the aqueous environment unintentionally "sweeten" the taste of tap water? Chem Eng J Adv 6, 100100.