Chemosphere 303 (2022) 135087

Contents lists available at ScienceDirect

Chemosphere

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

Non targeted screening of nitrogen containing disinfection by-products in formation potential tests of river water and subsequent monitoring in tap water samples

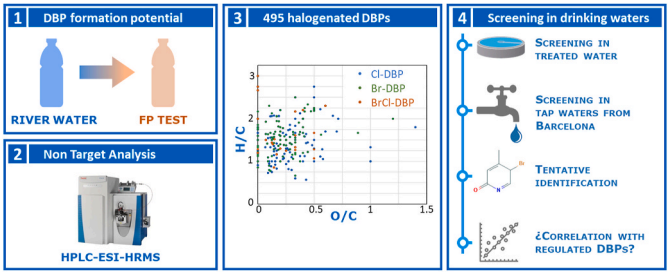
Josep Sanchís [a,b], Paula E. Redondo-Hasselerharm [c], Cristina M. Villanueva [c,d,e,f], Maria José Farré [a,b,*]

[a] Catalan Institute for Water Research (ICRA), Girona, Spain
[b] Universitat de Girona (UdG), Girona, Spain
[c] ISGlobal, Barcelona, Spain
[d] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[e] CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[f] IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain

HIGHLIGHTS

- A non-targeted approach with Orbitrap MS was used to tentatively identify 495 disinfection byproducts in formation potential tests of river water.
- Less DBP were found in DWTP and 40 tap water samples from Barcelona due to the removal of precursors at the plant.
- 15 halogenated DBPs were common in DWTP samples and 76 were detected in more than three tap waters.
- THMs, were poorly correlated with the concentrations of the newly detected DBPs..

GRAPHICAL ABSTRACT



ARTICLE INFO

Handling Editor: Xiangru Zhang

Keywords:
Halogenated DBPs
Non-target analysis
Orbitrap
Drinking water
Formation potential test

ABSTRACT

The generation of disinfection by-products during water chlorination is a major concern in water treatment, given the potential health risks that these substances may pose. In particular, nitrogen-containing DBPs are believed to have greater toxicological significance than carbon-based DBPs. Hence, high performance liquid chromatography coupled to high-resolution mass spectrometry (HPLC-HRMS) in positive mode was employed to identify new non-volatile nitrogen containing disinfection by-products (DBPs) and to assess their presence in potable water. Nine water samples were taken in the Llobregat river, in the context of a water reuse trial, near the catchment of a drinking water treatment plant (DWTP) in 2019. River samples were disinfected with chlorine under controlled formation potential tests conditions and analysed with a non-target approach. The peak lists of raw and chlorinated samples were compared exhaustively, resulting in an extensive list of 495 DBPs that include bromine and/or chlorine atoms. 172 of these species were found frequently, in three or more chlorinated samples. The empirical formulae of these DBPs were unambiguously annotated on the basis of accurate *m/z*

* Corresponding author. Catalan Institute for Water Research (ICRA), Girona, Spain.
E-mail address: mjfarre@icra.cat (M.J. Farré).

https://doi.org/10.1016/j.chemosphere.2022.135087
Received 23 February 2022; Received in revised form 20 May 2022; Accepted 21 May 2022
Available online 24 May 2022

measurements, isotopic patterns and common heuristic rules. Most of the annotated species (310) contained bromide, which is consistent with the relatively high bromide content of the Llobregat basin (>0.3 mg/l).

Drinking water samples were taken at the outlet of the DWTP during the same sampling period. According to their analysis, a large portion of the DBPs detected after the formation potential tests do not reach real-life drinking water, which suggests that the treatment train successfully removes a significant fraction of DBP precursors. However, 131 DBPs could still be detected in the final product water. A larger sampling was carried in the Barcelona water distribution network, during six consecutive weeks, and it revealed the presence of 78 halogenated DBPs in end-consumer water, most of which were nitrogen-containing. MS/MS fragmentation and retention times were employed to tentatively suggest molecular structure for these recalcitrant DBPs.

1. Introduction

Water disinfection is a cornerstone operation in drinking water management and public health protection, as it prevents the proliferation of infectious water-borne diseases. However, common disinfection methods, such as chlorination, chloramination and ozonation, trigger the formation of a myriad of undesired molecules, namely disinfection by-products (DBPs), some of which may produce negative effects to the human health after prolonged exposure and/or consumption (Rook, 1976; Evlampidou et al., 2020; Villanueva et al., 2015).

During the last decades, some halogenated DBPs have raised concern because of their toxicity and carcinogenicity (Evlampidou et al., 2020; Villanueva et al., 2004, 2015; Shi et al., 2020; Evans et al., 2020), hence some organic and inorganic halogenated species have been included in drinking water regulation world-wide, including the Guidelines for Canadian Drinking Water Quality (Government of Canada, 2021), the recent EU Directive 2020/2184 on the quality of water intended for human consumption (European Parliament, 2020), the US National Primary Drinking Water Regulation (stage 2) and the Australian Drinking Water Guideline guidelines (Natural Resource Management Ministerial Council Environment Protection and Heritage Council Australian Health Ministers' Conference, 2006). As can be seen in Table 1, these drinking water regulation and recommendations focus their attention on a limited number of low molecular weight DBPs, including THMs, inorganic oxyhalides and haloacetic acids (HAAs) assuming that they are good indicators for all DBPs formed. However, as global change triggers new challenges for water management and given the diversity of precursors, the total number of DBP species and its concentrations is expected to increase in the future. On the one hand, recent drought conditions and water shortages have led many drinking water utilities to consider or fully incorporate alternative water supplies, such as stormwaters, eutrophic waters, and advanced treated waste waters. Impaired waters used as drinking water supplies feature precursors polls fundamentally different from natural organic matter, which alters the array of formed DBPs. Also, changes in disinfectants away from chlorine to meet stringent regulations are generating a whole set of new species. For example, chloramines may generate iodinated-DBPs and *N*-nitrosamines, and ozone promotes bromate, tribromomethane and haloacetaldeydes (Sedlak and von Gunten, 2011), etc. Among this intricated spectrum of halogenated DBPs, regulated DBPs may not even be the primary toxicity drivers, according to recent toxicity-weighted studies (Chuang et al., 2019). In particular, nitrogen-containing DBPs are thought to be more toxic than carbon based DBPs, although none is regulated so far in drinking water guidelines (Plewa et al., 2008; Muellner et al., 2007; Komaki and Ibuki, 2022).

Up to date, more than 700 DBPs have been identified (Richardson and Plewa, 2020), in several disinfected water matrixes and applying different non-target-oriented methodologies, typically based on high-resolution mass spectrometry (HRMS). In Sanchís et al. (2020), reservoir water samples were chlorinated in batch-scale reactors and their dissolved organic matter (DOM) profile were characterised, resulting in the detection of >650 *m/z* signals with tentatively assigned as formulae containing bromine (Br_{1-2}) or chlorine (Cl_{1-2}) (Sanch í s et al., 2020a). Similarly, Wang et al. (2015) detected the presence of 181, 179, and 37 brominated DBPs in seawater chlorinated with sodium dichloroisocyanurate, trichloroisocyanuric acid and chlorine dioxide, respectively (Wang et al., 2015), and Lu et al. (2021) observed 189 chlorinated DBPs in chlorinated Suwannee River humic acids (Lu et al., 2021). Typical drinking water potabilization trains include DOM removal and filtration operations, all of which contribute to eliminate potential DBP precursors and mitigate the final concentration of DBPs (Williams et al., 2019). Yang et al. (2019) and co-authors recently

Table 1
Selection of regulated DBPs and their maximum allowable concentrations in drinking water.

DBPs	Maximum allowable concentration (μg/l)			
	EU Directive 2020/2184	Guidelines for Canadian Drinking Water Quality	US National Primary Drinking Water Regualation (stage 2)	Australian Drinking Water Guideline
TTHMs	100	100	80	250
HAA5	—[a]	80	60	–
Bromate	10	10	10	20
Chlorite	—[b]	1000	1000	800
Chlorate	—[b]	1000	–	–
Chloral hydrate	–	–	–	100
Chloroacetic acid	–	–	–	150
Dichloroacetic acid	–	–	–	100
Trichloroacetic acid	–	–	–	100
Cyanogen chloride	–	–	–	80
2-chlorophenol	–	–	–	300 (0.1)[c]
2,4-dichlorophenol	–	–	–	200 (0.3)[c]
2,4,6-trichlorophenol	–	–	–	20 (2)[c]
Formaldehyde	–	–	–	500
N-nitrosodimethylamine	–	0.04	–	0.1

[a] To be reduced to <60 μg/l by 2026.
[b] To be reduced to ≤0.25 mg/l (≤0.7 mg/l where chlorine dioxide is applied as disinfectant).
[c] The values inside parentheses were based on aesthetic criteria, instead of health criteria.

review the application of (LC/)MS/MS precursor ion scan for evaluating the occurrence, formation and control of polar halogenated DBPs in disinfected waters (Yang et al., 2019).

Still, the fingerprint of real potable water samples presents a plethora of signal that have been tentatively assigned as halogenated features. Sanchís et al. (2021) determined the presence of more than 100 halogenated ($Br_{0-2}Cl_{0-2}$) features in real drinking water samples, all of which had been generated during the potabilization process (Sanch í s et al., 2021). And, recently, Postigo et al. (2021) studied the chemodiversity of DBPs in four DWTPs and observed that only 19 features (of 86) were common in all the studied samples (Postigo et al., 2021a).

The elevated number of potential DBPs and their seemingly high variability can be justified because of the complex nature of DOM. DOM is formed by an intricated mixture of organic substances (Kellerman et al., 2014; Dittmar and Koch, 2006; Gros et al., 2021; Wagner et al., 2015), both biogenic and anthropogenic, and contain a rich diversity of functional groups and moieties that react with disinfection agents following a variety of pathways and with a significant influence of the physicochemical parameters of the medium (Doederer et al., 2014), the presence and intensity of light irradiance (Sanch í s et al., 2021) and the DOM profile itself (Farr é et al., 2019). In this situation, the convenience of adopting regulated volatile DBPs (and, particularly, of TTHMs) as general surrogates of DBP formation and exposure is debatable (Furst et al., 2021; Kolb et al., 2017). Also, because many relevant studies are essentially based on HRMS fingerprinting (Sanch í s et al., 2020a; Lavonen et al., 2013; Zhang and Yang, 2018; Phungsai et al., 2018) and other characterisation approaches (Lyon et al., 2014; Watson et al., 2018; Fabbricino et al., 2019), aiming to assess general trends in the DOM with a holistic view, there is a need to further understand the identity of the detected DBPs by employing analytical approaches that combine non-target screening and structural information elucidation (i.e. hyphenated techniques based on chromatography and high resolution mass spectrometric with hybrid analysers).

The present study aimed (1) to build an exhaustive list of semi-volatile/non-volatile DBPs generated after performing formation potential tests in river water; (2) to screen the detected nitrogen-containing DBPs in real potable water samples, determining its detection frequency and hence their relevance in drinking waters; and (3) to explore the adequacy of TTHMs and individual THMs as surrogates of the generation of DBPs occurring in tap water samples.

To achieve these objectives, batch-scale formation potential tests were performed in river water samples, taken during five different weeks near the catchment point of a DWTP in Barcelona. The DOM profiles, before and after disinfection, were characterised by liquid chromatography (LC) coupled to HRMS with an Orbitrap mass spectrometer using positive ionisation mode. The generated DBPs were identified and screened in real drinking water samples, which had been taken at the outlet of the DWTP. Finally, a screening was performed in 40 tap water samples taken from different districts of Barcelona, and their DBPs profiles. The adequacy of THMs as surrogates of halogenated DBP formation was tested.

2. Methods and materials

2.1. Sampling

Surface water samples were taken in the Llobregat River in June/July 2019, during five different weeks (see Table 2), in the context of a water reuse trial described in a previous study (Sanch í s et al., 2020b). Those samples from site *B* were taken at the catchment point of an important DWTP, which supplies 5.9×10^3 l/s water to the metropolitan area of Barcelona, while those samples taken in site *A* were taken ~8.7 km upstream. The river flow ranged between 3 and 5 m^3/s during the sampling and the water presented the physicochemical values that are presented in Table 2. Sampling point *B* was directly impacted by the nearby discharge of a wastewater treatment plant effluent, while sampling point *A* contains the diffuse anthropogenic impact from multiple wastewater effluents located along the higher course of the river. Exceptionally, as discussed in a previous study (Sanch í s et al., 2020b), the sample Llobregat_5 B received the discharge of tertiary effluent disinfected through chlorination, with a potential impact on the profile of DBPs that would be found in the real environment. Additionally, final potable water from the DWTP outlet were taken. The whole sampling was performed taking in consideration the hydraulic retention time in the river and inside the DWTP, in order to obtain comparable results in all the sites.

Grab samples were taken in amber glass bottles and transported with refrigeration to the laboratory, where they were immediately coarse-filtered with 0.7 μm mesh size glass-fiber (Whatman, UK) and analysed. All the sampling material employed had been previously rinsed thoroughly with nitric acid, ultrapure water, methanol and acetone to avoid sample contamination. Field blanks consisting of ultrapure water were transported, filtered and analysed weekly, in parallel with real samples, in order to prevent false positives.

A second sampling of potable waters was conducted between August 31st and October 16th of 2020. A tap water sample was collected in each of the 42 postal codes from the city of Barcelona (Redondo-Hasselerharm, Accepted for publication in Journal Of Exposure Science And Environmental Epidemiology) At each location, cold tap water was allowed to run for at least 2 min prior to the collection of the actual sample. Tap water was then slowly poured into a 2.5 l glass bottle containing ascorbic acid, filled to the top, and shaken. Bottles were kept at 4 °C until analysis. Forty-two samples were collected in this sampling and 40 were included in the present study. A summary of their characteristics can be found in Table S1.

Table 2
Summary of samples from the Llobregat river. TOC, stands for Total Organic Carbon. TN stands for Total Nitrogen. TON stands for Total Organic Nitrogen, estimated as TN minus the measured concentrations of nitrates, nitrites and ammonia.

#	Code	Sampling site	Sampling date	TOC (mg/l)	TN (mg/l)	TON (mg/l)	Cl^-(mg/l)	Br^-(mg/L)
1	Llobregat_1 A	Llobregat river (site A)	03-June-2019	6.1	2.5	0.92	246.9	0.41
2	Llobregat_1 B	Llobregat river (site B)	03-June-2019	8.7	4.6	1.4	311.5	0.47
3	Llobregat_2 A	Llobregat river (site A)	18-June-2019	4.2	2.1	0.71	222.1	0.36
4	Llobregat_2 B	Llobregat river (site B)	18-June-2019	6.5	5.1	1.3	284.4	0.38
5	Llobregat_3 A	Llobregat river (site A)	25-June-2019	3.5	1.6	0.76	344.9	0.76
6	Llobregat_3 B	Llobregat river (site B)	25-June-2019	5.1	3.8	1.27	342.2	0.64
7	Llobregat_4 A	Llobregat river (site A)	08-July-2019	–	2.2	0.77	236.3	0.42
8	Llobregat_5 A	Llobregat river (site A)	16-July-2019	5.7	2.1	0.74	197.2	0.30
9	Llobregat_5 B	Llobregat river (site B)	16-July-2019	6.3	3.2	1.18	263.1	0.37
10	DW_1	Drinking water outlet	04-June-2019	0.6	2.0	0.2	76.4	<0.01
11	DW_2	Drinking water outlet	19-June-2019	0.9	2.2	0.3	109.2	<0.01
12	DW_3	Drinking water outlet	26-June-2019	0.8	2.1	0.3	156.8	<0.01
13	DW_5	Drinking water outlet	17-July-2019	2.0	1.6	0.2	137.9	<0.01

2.2. DBP formation potential tests

Batch disinfection formation potential tests with coarse-filtered Llobregat surface water samples were performed in 250 ml sealed bottles, with no head-space, in an incubator (25 ± 1 °C) for 24 h. The chlorine dose was adjusted to obtain a free chlorine residual of 1–3 mg/l after the reaction. Free chlorine was individually checked for all the tests after 24 h, using a photometric test kit LCK 310 (Hach Lange GmbH, Germany). Then, samples were quenched with ~500 μM Na_2SO_3 (≥98.0%, BioUltra grade) and extracted for injection in the LC-HRMS.

A procedural blank consisting on ultrapure water was chlorinated, incubated, and analysed following the same experimental procedure.

2.3. Non-target analysis

The following methodology was employed to extract analyse all the water samples and it was based on a protocol for DOM characterisation, which has been thoroughly described in a previous study (Sanch í s et al., 2021).

Briefly, 250 ml of vacuum-filtered water were acidified with formic acid (ACS grade, Merck, Germany) and extracted by solid-phase extraction with styrene-divinylbenzene Bond Elut™ PPL cartridges (Agilent, Santa Clara, USA), previously conditioned with methanol and acidified ultrapure water (Optima®, "LC/MS grade", Fisher Chemical). After loading the sample, the cartridges were dried and eluted with 2 ml of methanol. The extracts were stored at −20 °C until their instrumental analysis.

Right before their LC-HRMS analysis, extracts were diluted with ultrapure water 1:1. Chromatographic separation was achieved in reverse phase with a ZORBAX Eclipse XDC18 column (150 × 4.6 mm, 5 μm particle size; Agilent Technologies, USA) and using a Waters Acquity UPLC System (Waters, Milford, MA, USA) pumping at 0.5 ml/min. Initial mobile phase conditions consisted of 5:95 acetonitrile:ammonium formate (0.01 M, pH 3.0) and were held for 1 min. Acetonitrile percentage increased linearly to 95%, at minute 10, and these conditions were held during 5 min. Initial conditions were recovered after 1 min and kept during 4 additional inter-run minutes, to ensure a proper column equilibration.

Ionisation was performed with an electrospray ionisation source (H-ESI II probe), in positive polarity and applying the following parameters: Voltage, 3.5 kV; sheath and auxiliary gas flows 40 and 20 a. u.; probe and heater temperatures, 350 and 300 °C; and S-Lens RF Level, 70%. The mass spectrometric analysis was performed with a Q Exactive™ (Thermo Fisher Scientific, USA) in data-dependant scan (DDS), an acquisition mode that combines a high-resolution full-scan with successive MS (Evlampidou et al., 2020) events, in which the most abundant ions are isolated and fragmented in order to obtain further structural information. The main scan range was *m/z* 70–1,000, with a resolution power of 100,000 FWHM (*full width at half maximum*), and data-dependent MS (Evlampidou et al., 2020) events were performed on the 5 most abundant ions, with a normalised collision energy of 30%.

2.4. Non-target analysis: data treatment

Raw LC-MS files were processed with Compound Discoverer (ThermoFischer Scientific), version 3.2. Relevant parameters concerning spectrum extraction, peak deconvolution, and chromatogram alignment are detailed in Table S2. This workflow was applied to raw and lab-chlorinated river samples, resulting in a preliminary peak-list that contained both halogenated and non-halogenated features.

An automatic filter was built in Compound Discoverer in order (1) to filter out halogenated DBPs candidates (automatically flagged by Compound Discoverer through isotopic pattern recognition), (2) to subtract blank contaminants, and (3) to exclude noisy peaks/artifacts). The filter was set with a peak area threshold of 10^6 a. u and selected those features that fulfilled simultaneously these conditions:

- a peak area smaller than 10^6 a. u. In the blank, in the chlorinated blank and in the raw river sample,
- a peak area higher or equal than 10^6 a. u. In the chlorinated river sample, and
- a satisfactorily resolved isotopic pattern for chlorinated/brominated molecules.

The application of this filter (see Figure S1) resulted in a list of DBPs candidates, which were curated by visual inspection with XCalibur 2.2 (ThermoFischer Scientific) in order (1) to ensure an acceptable peak shape in chlorinated samples, (2) to ensure that they were absent in raw river samples and blanks, and (3) to double check their isotopic pattern. Thirteen different halogenated isotopic patterns were considered, according to the mass difference of $^{37}Cl/^{35}Cl$ and $^{81}Br/^{79}Br$, and the intensity ratios detailed in Table S3. Then, their empirical formulae were assessed individually considering a restricted count of elements ($C_{1-30}H_{0-60}O_{0-15}N_{0-7}S_{0-1}Cl_{0-4}Br_{0-3}$), a maximum tolerable error of 2.5 ppm, a double bound equivalent (*DBE*, defined as in equation (1)), compressed between −0.5 and 15, and a DBE minus oxygen index (*DBE-O*, equation (2))compressed between −10 and 10^{39}.

$$DBE = C - \frac{H + Cl + Br}{2} + \frac{N + P}{2} + 1 \quad (1)$$

$$DBEO = DBE - number\ of\ O \quad (2)$$

In those cases when no empirical formula fulfilled these restrictions, an extended range of DBE-O values was tested, and the presence of one P and one additional S was considered.

An especial care was taken to prevent halogenated ESI-fragments to be treated as independent compounds. Those pairs of peaks that co-eluted and presented compatible elemental formulae (i.e., the peak with the highest *m/z* was annotated with a formula with more or the same number of atoms than the peak with the lowest *m/z*) were checked, in terms of peak shape and experimental behaviour (Pearson's $r > 0.9$) and discarded if necessary.

The identified features were represented in Van Krevelen diagrams according to their O/C and H/C ratios and were classified according to compositional spaces (based on previous fingerprint classification systems (Sanch í s et al., 2020a; Minor et al., 2014)). It should be highlighted that the labels of these van Krevelen regions do not reflect accurately the complex nature of the whole set of features contained therein and are employed merely for statistical purposes and work intercomparison.

The structures of some DBPs were tentatively identified using the Metfrag webtool (Ruttkies et al., 2016). Briefly, for each unknown feature, a search was performed in the Pubchem library, looking for all the substances that presented the specified empirical formula. The Metfrag algorithm fragmented, in-silico, each of the retrieved candidates. This process typically led to a relatively long list of Pubchem candidates, ranked according to the similarity between their in-silico fragmented spectrum and the real MS^n (n = 2) spectrum experimentally obtained by Orbitra-HRMS. When only one candidate was obtained with a Metfrag index >0.9, this was accepted as the preferred candidate. In those cases when several Metfrag candidates showed similarly good scores, the fragment ion search tool (FISh) of Compound Discoverer was employed to assess the coherency of their experimental MS (Evlampidou et al., 2020) fragments. In those cases were candidates included halogenated heterocycles, the intrinsic reactivity of heterocycles was taken into account to discard unlikely candidates (Grimmett, 1993a, 1993b, 1994).

Statistical analyses, including principal component analyses (PCA) and paired/unpaired *t*-tests were performed with custom R scripts.

2.5. Analysis of THMs

The analysis of THMs was conducted in the frame of another study

(under submission) (Redondo-Hasselerharm, Accepted for publication in Journal Of Exposure Science And Environmental Epidemiology). The analytical methodology and the detailed results can be found there, and the method is briefly presented in the Supporting Information (Text S1).

3. Results and discussion

3.1. Identification of halogenated DBPs in chlorinated river water

The non-target analysis of chlorinated and non-chlorinated samples after formation potential tests resulted in an extensive peak-list with 35,465 features. The PCA of these data revealed major differences between samples according to their chlorination/non-chlorination status. More precisely, the scores graph of PC1 vs. PC2 (52.6% and 19.9% of explained variance, respectively), displayed raw water samples clustering together at higher PC1 values, while DBP formation potential extracts appeared at lower PC1 values (see Fig. 1). Non chlorinated river samples from the sampling point A, in general, appeared at lower PC2 values than those from the sampling point B, although a certain degree of overlapping could be observed. Potable water samples from the DWTP outlet, which were included in the study, also clustered altogether at negative PC1 values and were clearly distinguished from raw river samples. Similarity among samples was also checked through the Cao dissimilarity index (Figure S2a) and the Spearman's rank correlation matrix (Figure S2b), both showing clustering of river waters, DBP formation potential tests and drinking water samples. None of the tested multivariant models identified the sample Llobregat_5 B (which received the emission of chlorinated tertiary effluent) as an outlier sample.

2084 peaks occurred exclusively in DBP formation potential tests. Since blanks had been subtracted (including extracts of ultrapure water, chlorinated ultrapure water and raw river water extracts), this subset of peaks can be unequivocally assigned as small molecules unintentionally generated during the reaction of active chlorine with organic matter from the dissolved and subcolloidal phases (i.e. DBPs). It should be highlighted that halogenated compounds represent a minor fraction of the whole set of DBPs in the peak-list, the study of which is not the scope of the present study.

In total, 495 halogenated DBPs were identified in at least one chlorinated river samples.

The complete list of DBPs is summarised in Table S4. The vast majority of substances could be annotated with only one empirical formula, since only one empirical formula fulfilled the restrictions detailed in **subsection 2.4**. In those few cases where two formulae fulfilled the restrictions, the one with an O/C ratio between 0.2 and 1.0 and a H/C ratio between 0.2 and 2.25 was prioritised.

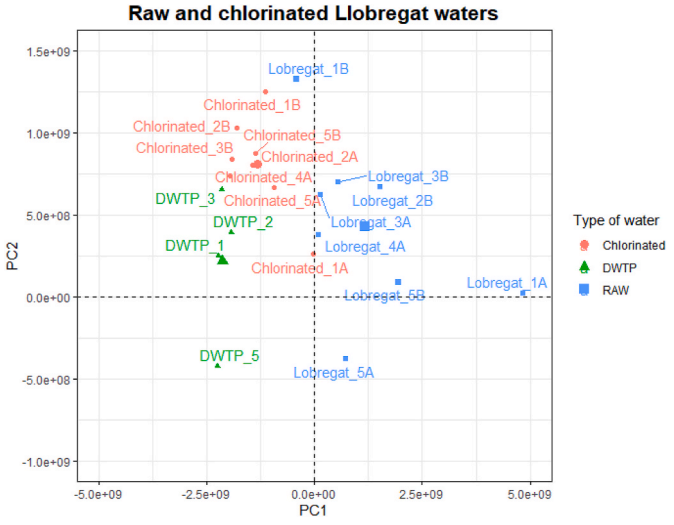


Fig. 1. Scores graph of chlorinated and non-chlorinated Llobregat river samples.

It should be highlighted that some halogenated substances were already in raw Llobregat river waters prior to their disinfection. Those peaks that did not experience a ≥10 fold increase in the DBP-FP tests, could not be unambiguously assumed to be DBPs and hence were discarded. Those that did experience a ≥10 fold increase were considered as halogenated DBPs and their areas in the river were subtracted. Raw river samples generally presented a small number of DBPs (from 12 to 18, as indicated in Table 3) with small peaks (average areas ranging from 5×10^3 to 6×10^4), in contrast to DBP formation potentials, where peaks exhibited average areas from 2.3×10^5 to 1.4×10^6. As an exception, the sample Llobregat_River_5 B, which had received the impact of a tertiary effluent disinfected by chlorination, exhibited a relevant number of halogenated DBPs, 75, significantly higher than the other samples that had received non-disinfected effluents. Similar results were obtained in a previous study that characterised its DOM profile with another analytical methodology (Sanch í s et al., 2021). The environmental impact of these substance is currently unknown and further tests should be carried out to unequivocally elucidate the molecular structure of these DBPs and understand their risk. While disinfecting wastewater effluent prior to its emission to the environment may be advantageous from a bacteriological perspective, the potential toxicological risks of these halogenated substances should not be overlooked.

3.2. Characterization of halogenated DBPs

Fig. 2 displays the Van Krevelen diagram containing the detected DBPs. Most of the DBPs, 272, were brominated, while 185 of them were chlorinated and 38 presented both halides. The predominance of brominated DBPs can be justified because of the characteristically high concentrations of dissolved bromide in the Llobregat River (>0.3 mg/l, see Table 2).

As can be observed in the same Van Krevelen diagram, most of the substances were relatively O-poor (95% of them presented a O/C ratio ≤0.5, with an average O/C = 0.22), with 95 annotated formulae (19.2%) with 0 oxygen atoms and only 2 substances with a O/C > 1.0; while the H/C ratios ranged from 0.29 to 2.50 (1st and 3rd quartiles: 1.00 and 1.47, average H/C = 1.43). These ratios are very similar to those obtained in previous reports with disinfected Llobregat samples (average O/C and H/C, 0.24 and 1,43, respectively) (Sanch í s et al., 2021), although those analyses were based on the direct infusion of organic matter, with no previous chromatographic separation of extracts, and therefore could be potentially interfered by ESI-fragments. In comparison with other works in the literature, Lavonen et al. (2013) observed slightly more oxidised and aromatic chlorinated DBPs (with average O/C values compressed between 0.59 and 0.62, and average H/C values ranging from 0.72 to 1.01, depending on the sample), although these results were obtained in negative ionisation mode, where other molecules may be more easily ionised.

Table 3
Number of halogenated DBPs detected in chlorinated river sample and DWTP from the Llobregat.

Sampling	Day	Number of halogenated DBPs				
		Sampling site A		Sampling site B		DWTP outlet
		Raw sample	DBP Formation potential	Raw sample	DBP Formation potential	
Week 1	1	15	203	18	275	33
Week 2	16	12	135	16	199	105
Week 3	23	17	113	17	126	40
Week 4	36	14	96	–	–	–
Week 5	44	13	58	75	130	75

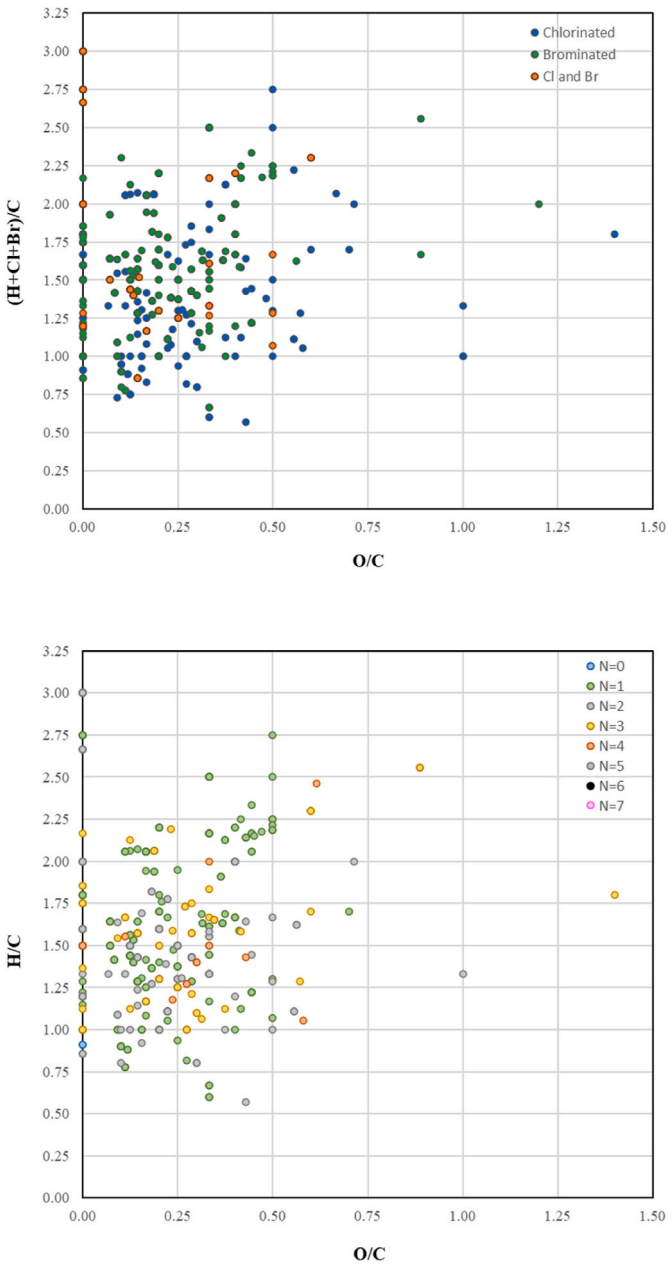


Fig. 2. Van Krevelen diagrams with all the halogenated DBPs detected in formation potential tests, classified according to their halogen content (up) and nitrogen content (down).

As can be seen in Table 4, according to their O/C and H/C ratios, lipids-like substances represented the most abundant compositional group (accounting for 147 DBPs, 80 of which brominated), followed by peptide-like (99 halogenated DBPs, 66 brominated) and condensed hydrocarbons-like substances (85 halogenated DBPs). Chlorinated and brominated DBPs were evenly distributed in compositional spaces below H/C = 1.5 (Br-DBPs/Cl-DBPs ≈ 1 in condensed hydrocarbons, lignins-like and tannins-like regions), while more aliphatic compositional spaces showed a prevalence of brominated features (Br-DBPs/Cl-DBPs ≈ 1.5 lipids-like, peptids-like and aminosugars-like).

Significant differences were observed among samples, despite of their geographical and temporary proximity. Table 3 compares the number of DBPs detected in each sample. As can be observed, the sample encoded as Llobregat_5 A presented the smallest number of positives, with 58 DBPs, while Llobregat_1 B presented as many as 275 individual DBPs. The sampling site B presented statistically more DBPs than sampling site A ($p = 0.030$ according to paired *t*-test). Also, the number of detected DBPs decreased monotonically through time in both sampling sites, and the decrease was statistically significant in site A (Spearman's $\rho = -1.0$, $p = 0.017$). The frequency of detection of DBPs was also highly variable throughout the sample set. Most DBPs were detected in only one, two or three samples (210, 113 and 39 DBPs, respectively) while only two compounds were ubiquitous in the nine analysed samples: DBP #1 ($m/z = 93.0213$ $t_R = 2.91$ min, tentatively annotated as $C_2H_6N_2Cl^+$) and #6 ($m/z = 119.0372$, $t_R = 4.03$ min, annotated as $C_4H_8N_2Cl^+$). Overall, since the experimental disinfection conditions were equivalent in all the disinfection tests, such a high diversity in the DBP profiles must be attributed to miscellaneous differences in the composition of the raw river DOM.

3.3. Screening of DBPs in DWTP outlet samples

Four samples were taken at the outlet of the DWTP during the same sampling campaign. These drinking waters presented significantly less DBPs and at lower concentrations than surface samples disinfected in the lab through formation potential tests. In total, 131 DBPs were detected in at least one DWTP sample and, as can be seen in Table 3, the number of detected DBPs varied considerable from sample to sample, ranging from 33 to 105 compounds per sample. Only 15 DBPs were detected in all four DWTP samples.

Table 4 display the *m/z* detected DBPs in drinking water samples and compares it with those detected in DBP formation potential tests. The pool of DBPs was composed by smaller molecules than the total pool of DBPs. This decrease was statistically significant ($p < 0.05$ for *t*-tests) for every testable compositional space (lipids, peptides, condensed hydrocarbons, lignins, and others) and ranged from −47 *m/z* in the lipids group until −71 *m/z* in the lignin region. The number of features also decreased in all the compositional spaces, but it was slightly more pronounced in the condensed compositional spaces (lignins and condensed hydrocarbons, −83% and −82%, respectively) than in the aliphatic ones (lipids and peptides, −65% and −73%, respectively).

These decreases of the number of DBPs and their smaller size, can be attributed to the successful removal of DBP precursors in the DWTP. The studied DWTP (Area Metropolitana de Barcelona, 1181), right after water catchment and sedimentation, performs a pre-chlorination step with chlorine dioxide, followed by flocculation, decantation, sand filtration, dilution with groundwater, and then, two parallel treatment processes: one consisting of ozonation plus carbon filtration; and another consisting of ultrafiltration (0.02 μm) plus UV radiation, reverse osmosis and remineralisation. After this treatment, both treated waters are combined and chlorinated in labyrinth channels, becoming the final treated water that we have analysed in the present work.

As result of these treatments, the TOC of the final potable waters ranged from 0.6 to 2.0 mg/l, as can be observed in Table 2, while in the river catchment (site B) this parameter ranged from 5.7 to 8.6 mg/l. Accordingly, DWTP samples presented less DBPs than site B (1.7–8.33 times less). Those DBPs that were detected in both types of samples presented significantly smaller peak areas in DWTPs (paired *t*-test: $p < 10^{-4}$), with median peak areas 5.8, 3.0, 5.8 and 3.8 times smaller.

3.4. Screening and tentative identification of DBPs in tap water

In a second phase, the list of 495 DBPs was also screened in 40 tap waters from Barcelona households and public fountains, which were sampled approximately one year after the Llobregat river experiments. Each tap water sample was collected from a different postal code, in order to fully evaluate the potential differences among different areas of the city.

In total, 76 DBPs were detected in these tap waters. Among them, 58 were present in ≥10% of the analysed samples. Most of these DBPs (all but one) were nitrogenated and most of them (65%) were brominated DBPs. The prevalence of brominated DBPs had been already observed in

Table 4
Classification of DBPs according to their compositional space, their halogens content and their *m/z*.

Compositional space		Number of DBPs, according to their halogen content				*m/z* of DBPs, according to their occurrence (as average ± s.d.)	
Tag	Definition	Whole pool of DBPs	Chlorinated	Brominated	Chlorinated and brominated	Whole pool of DBPs	DBPs detected in drinking waters
Lipids	O/C ≤ 0.2 (H + X)/C ≥ 1.5	**147**	56	80	11	243 ± 96 (n = 147)	196 ± 58 (n = 53)
Peptides	0.2 = O/C < 0.5 (H + X)/C ≥ 1.5	**99**	27	66	6	373 ± 143 (n = 99)	314 ± 100 (n = 27)
Aminosugars	0.5 = O/C < 0.7 (H + X)/C ≥ 1.5	**11**	4	6	1	425 ± 81 (n = 11)	— (n = 0)
Carbohydrates	O/C ≥ 0.7 (H + X)/C ≥ 1.5	**5**	2	3	0	367 ± 99 (n = 5)	487 (n = 1)
Condensed hydrocarbons	O/C ≤ 0.25 0.66 ≥ (H + X)/C < 1.2	**85**	39	40	6	236 ± 67 (n = 85)	227 ± 43 (n = 15)
Lignins	0.26 = O/C < 0.66 0.66 ≥ (H + X)/C < 1.5	**53**	27	22	4	320 ± 112 (n = 53)	249 ± 42 (n = 9)
Tannins	O/C ≥ 0.66 0.66 ≥ (H + X)/C < 1.5	**2**	2	0	0	226 ± 0 (n = 2)	226 ± 0 (n = 2)
Others		**93**	28	55	10	300 ± 121 (n = 93)	248 ± 38 (n = 24)
Total		**495**	**185**	**272**	**38**	292 ± 122 (n = 495)	237 ± 52 (n = 131)

the previous section, in disinfected Llobregat waters and in previous works in the literature, e.g. Tao et al., 2020 (Tao et al., 2020). A large number of nitrogenated DBPs was also observed in that work (>50% of the detected DBPs) although in the present study the number of nitrogenated DBPs was dominant, mostly because we only analysed the samples in positive mode looking for this specific family of compounds. This can be attributed both to the original composition of the Llobregat river DOM, which is known to commonly present a significantly abundant fraction of ionisable N-containing features (48.1 ± 7.2% in the "sampling site_A" (Wang et al., 2015)), and to the analytical approach that has been employed in the present study, with acidic acid and positive ionisation mode, which may have favoured the detection of N-containing functional groups over other types of DBPs.

Similarly than in DWTP samples, DBPs in tap water samples presented peak areas lower than in formation potential tests. The least abundant DBP presented an average peak area of 10^3 a. u, and the median average peak area was of merely 4×10^4 a. u. The DBP tentatively identified as 2-(2-bromoallylamino)acetonitrile (*m/z*174.9869 and $t_R = 8.86$) presented the largest peaks, with average area of 2.0×10^7 a. u.

Table S5 tabulates their formulae, frequencies of detection and presents their tentative structures, as identified by Metfrag and FISh scores (see subsection 2.4). Molecular structures could be tentatively assigned for 54 of the peaks. Figure S3 and Figure S4 display the chromatograms, spectra and Metfrag results that allowed the identification of one DBP as an example (#40).

Several pairs of DBPs presented the same structures, only distinguished by its halogen content (e.g., DBPs #21 and #22, tentatively annotated as 2-(2-chloroallylamino)acetonitrile and 2-(2-bromoallylamino)acetonitrile; or DBPs #16 and #17, annotated as 4-chloro- and 4-bromo-1,5-dimethyl-triazole.

The peaks with *m/z* = 93.0215, $t_R = 2.91$ (DBP #1, $[C_2H_6N_2Cl]^+$); *m/zm/z* = 136.9709 and $t_R = 2.98$ (DBP #2, $[C_2H_6N_2Br]^+$); and *m/z* = 142.0187 and $t_R = 3.01$ (DBP #12, $[C_4H_{10}NCl_2]^+$); were particularly ubiquitous, being spotted in 85%, 98% and 93% of the 40 tap water samples, respectively. According to PubChem library, 15 plausible chemical structures exist that can be assigned to DBPs #1 and #2 (listed in Table S6a), including halogenated acetamides, hydrazines, amines and *N*-haloamines, etc. However, this particular pair of DBPs didn't show satisfactory MS (Evlampidou et al., 2020) spectra because of their low abundance and because of their poor chromatographic retention in reverse phase LC, so MetFrag could not directly assist in DBPs #1 and #2 tentative assignation. Nevertheless, it can be hypothesised that their structures are likely closely related to those of their neighbour dihalogenated homologues, DBP #3 ($[C_2H_5N_2Cl_2]^+$) and DBP #5 ($[C_2H_5N_2Br_2]^+$), which eluted at very close retention times and presented clear MS (Evlampidou et al., 2020) spectra.

DBP #3 and its isomer DBP #4 were tentatively assigned as 2,2-dichloroethanimidamide (International Chemical Identifier key, InChIKey: IZRMJONVEVZERC-UHFFFAOYSA-N) and *N*'-(dichloromethyl) formamidine (InChIKey: CFEYYGWGIDXPKK-UHFFFAOYSA-N), respectively, on the basis of their best ranked Metfrag candidate and their logP values (as estimated by PubChem via XLogP3 3.0). Analogously, Metfrag ranked 2,2-dibromoethanimidamide (InChIKey: KWSQTQQYZJEWAM-UHFFFAOYSA-N) and *N*-(dibromomethyl)formamidine (InChIKey: HSEXSJBVOYAPPI-UHFFFAOYSA-N) as top candidates for DBP #5 and DBP #6. It should be highlighted that pairs DBPs #3/#4 and DBPs #5/#6 exhibited different behaviour among samples and presented no correlation, indicating that they were not chromatographic artifacts (e.g., peaks duplicated due to two prototropic tautomers in pH equilibrium) but truly different substances.

According to the tentative identification of DBPs #3 and #5, DBP #1 and #2 were annotated as 2-haloethanimidamides.

Regarding DBP #12, Metfrag ranks butylimino (chloro)-λ (Villanueva et al., 2015)-chlorane (InChIKey: QMSPZLZCQMEPLB-UHFFFAOYSA-N) as the best candidate, which is chemically implausible. Two candidates, *N*-(dichloromethyl)propan-1-amine and its isomer *N*-(dichloromethyl)propan-2-amine stood in second place, with a Metfrag score of 0.94 and 14 successfully annotated fragments. Its isomer DBP #13 was tentatively annotated as *N*,4-dichlorobutan-1-amine, with a Metfrag score of 0.89 and 29 annotated fragments (or as its branched isomers, *N*,1-dichloro-2-methylpropan-2-amine, with the same exact Metfrag qualifications).

DBP #16, with formula $C_4H_6ClN_3$, was tentatively identified as 4-chloro-1-ethyl-triazole or as 4-chloro-1,5-dimethyl-triazole (InChIKey: 1 S/C4H6BrN3/c1-2-8-3-4 (5)6-7-8/h3H, 2H2,1H3 and 1 S/C4H6BrN3/c1-3-4 (5)6-7-8 (3)2/h1-2H3, respectively). Other highly ranked candidates included 1,2,4-triazoles, such as 3-bromo-5-ethyl-1*H*-1,2,4-triazole, 3-bromo-4-ethyl-1,2,4-triazole, and 3-bromo-4,5-dimethyl-1,2,4-triazole. Triazoles rarely occur in natural substances, but they can be found in azole fungicides (e.g., tetraconazole, penconazole, metconazole, and ipconazole, all of which were recently included in the EU watchlist (European Commission, 2020)), in antibiotics (cefatrizine and tazobactam) and other pharmaceutical compounds (Wasilenko et al., 1996).

In addition, DBPs #32 and #34 were tentatively annotated as 2-chloro-5-methoxy-3*H*-indole and 3-(3-bromo-1*H*-indol-7-yl)propanal, respectively. Several biomolecules (i.e. melatonin, indole-3-carbinol, serotonin, etc.) and drugs (e.g., sumatriptan, vinblastine and diindolylmethane) present this heterocycle in their core structure. Postigo et al. (2021) recently stated that chlorine reacts with contaminants of emerging concern (e.g., through electrophilic attack, oxidation reactions and addition reactions on unsaturated bonds) (Postigo et al., 2021b), but considering the structures of DBPs #32 and #34 it is unclear whether

these DBPs were formed from singular precursors or if they were formed from the hydrolysis of miscellaneous natural-occurring macromolecules. In the future, more efforts should be devoted to identifying potential precursors of these DBPs (for instance, using the library developed in Postigo et al. (2021)) and to understand the environmental factors that affect their formation.

The DBP #36, with *m/z*262.0071 and t_R 6.59 min, was tentatively annotated as 3-(2-aminooxyethyl)-6-bromo-2-methoxyphenol (previous records of similar halogenated guaiacols generated after chlorination have been reported (Michałowicz et al., 2007)), and the DBP #42, with *m/z* 243.9737 and t_R 3.04 min, was tentatively identified as *N*-chloroisoleucine, a halogenated amino acid previously described in Nweke and Scully Jr. (1989) (Nweke and Scully, 1989). Also, the DBP#51 was tentatively annotated as a 6-chloro-melamine with an *N*′-ethoxyethanol group (InChIKey = JJRJHOBBHQZAAB-UHFFFAOYSA-N), and it was identified as a reasonable degradation product of melamine resin.

Significantly, some typical nitrogen-containing DBPs (i.e. haloacetonitrile or haloacetoamides) were not present among the tentatively identified compounds. Because the methodology was essentially non-target and analytical conditions were not optimised for these or any other DBPs, their absence in our list does not necessarily imply that they were not generated in the sample.

3.5. Multivariant analysis of tap water samples

In order to further understand the profile and the relative abundances of DBPs in tap water samples, they were studied by multivariant analysis. The PCA loading graph can be found in Figure S5. PCA showed that two principal components sufficed to explain >90% of the model variance. PC1 was mainly related to the DBP with *m/z* 174.9869 and t_R = 8.86 (tentatively identified as 2-(2-bromoallylamino)acetonitrile), while PC2 was related to the two DBPs with *m/z* 126.9826 (tentatively assigned as 1,2-diamino-1,2-dichloroethene and 2,2-dichloroethanimidamide).

PC1 obtained with either raw or normalised data, showed a moderate degree of monotonic association (Spearman's $\rho = -0.6$, $p = 2 \times 10^{-4}$) with the sample's latitude, expressed as Y UTM coordinate. In contrast, no significant relationship was observed regarding the sample's longitude. Therefore, it can be concluded that the DBP profile is influenced to a certain degree by the faucet location in Barcelona. It should be highlighted that the Llobregat river is one major source of potable water for the city, but it's not the only one. Typically, the potable water from the northernmost areas of Barcelona has a lesser contribution of Llobregat's water, and water is there preferably obtained from Ter River potabilization and groundwater sources. The salinity of the tap waters also supports this conclusion: Llobregat River has a remarkably high conductivity, due to geologic and anthropogenic factors, and thus water conductivity was found to moderately correlate with the latitude of the sampling point (Spearman's $\rho = -0.4$, $p < 1 \times 10^{-6}$).

3.6. Correlation of DBPs and THMs

The concentrations of individual and total THMs are displayed in Figure S6. Tap water samples presented TTHMs ranging from 17 to 83 μg/l, with predominance of tribromomethane (average concentration, 20 μg/l), followed by trichloromethane, dibromochloromethane and bromodichloromethane. These values were in compliance with drinking water regulations. A comprehensive discussion about THMs concentrations is out of the scope of the present study, and the THMs profile is discussed in deep in Redondo-Hasselerharm et al. (Redondo-Hasselerharm, Accepted for publication in Journal Of Exposure Science And Environmental Epidemiology).

The adequateness of THMs as surrogates of DBPs in potable waters was tested in tap water samples. According to Spearman correlation tests, none of the identified DBPs correlated with TTHMs ($\rho < 0.35$ in all the cases), and individual THMs showed a limited success in explaining the behaviour of 17 halogenated DBPs. The tentative identity of these DBPs and their Spearman's rank order coefficient are summarised in Table 5.

TCM and BDCM showed a moderate correlation with seven DBPs ($0.6 = \rho < 0.8$, $p < 0.05$), all of which chlorinated, which were tentatively identified as 2-chloroethanimidamide; 2,2-dichloroethanimidamide; *N*-(dichloromethyl)formamidine; *N*,4-dichlorobutan-1-amine or *N*,1-dichloro-2-methylpropan-2-amine; 2-chloro-5-methoxy-3*H*-indole; 5-amino-2-(chloroamino)-5-oxo-pentanoyl chloride; and an unidentified compound with formula $C_6H_{10}O_2N_4Cl_2S$.

DBCM and TBM showed a certain degree of correlation ($\rho \geq 0.6$, $p < 0.05$) with ten other substances, most of which brominated, which were tentatively identified as 2-(2-chloroallylamino)acetonitrile; 4-bromo-3-ethylidene-pyrazole; *N*-(bromochloromethyl)propan-2-amine or *N*-(bromochloromethyl)propan-1-amine; 2-(2-bromoallylamino) acetonitrile; 4-bromo-1,5-dimethyl-triazole; 7-bromo-4*H*-indazol-3-amine; 6-chloro-4,5-dimethyl-*N*-(pyridazin-3-ylmethyl)pyridazin-3-amine; 6-bromo-4,4-dimethyl-4a, 5,6,7,8,8a-hexahydro-1*H*-benzo[d][1,3]oxazin-2-one; 4-(3-bromo-2-hydroxy-5-isopropyl-phenyl)butanoic acid; and an unidentified compound with formula $C_6HO_2N_2ClS_2$.

Overall, only two compounds were correctly explained by a THM. Putative 4-bromo-1,5-dimethyl-triazole and 6-bromo-4,4-dimethyl-4a,5,6,7,8,8a-hexahydro-1*H*-benzo[d][1,3]oxazin-2-one were strongly correlated ($\rho \geq 0.8$, $p < 0.05$) with DBCM.

Table 5
Features that correlate strongly with THMs in tap water samples, classified according to their Spearman's rank correlation coefficient. Their tentative identification is summarised in the Supporting information (Table S5).

DBP	$0.6 = \rho < 0.7$ & $p < 0.05$	$0.7 = \rho < 0.8$ & $p < 0.05$	$\rho \geq 0.8$ & $p < 0.05$
TTHMs	–	–	–
TCM	*m/z* = 93.022; t_R = 2.91 min	*m/z* = 126.983; t_R = 3.03 min	–
	m/z = 142.019; t_R = 3.99 min	*m/z* = 126.983; t_R = 3.97 min	
	m/z = 180.021; t_R = 6.99 min	*m/z* = 272.998; t_R = 7.99 min	
	m/z = 199.004; t_R = 4.03 min		
BDCM	*m/z* = 126.983; t_R = 3.03 min	–	–
	m/z = 126.983; t_R = 3.97 min		
	m/z = 272.998; t_R = 7.99 min		
DBCM	*m/z* = 172.971; t_R = 7.21 min	*m/z* = 174.987; t_R = 8.86 min	*m/z* = 175.982; t_R = 7.60 min
	m/z = 185.968; t_R = 3.04 min		*m/z* = 262.044; t_R = 8.75 min
	m/z = 211.982; t_R = 9.49 min		
	m/z = 301.044; t_R = 11.9 min		
TBM	*m/z* = 131.037; t_R = 8.68 min	*m/z* = 174.987; t_R = 8.86 min	–
	m/z = 231.916; t_R = 4.01 min	*m/z* = 175.982; t_R = 7.60 min	
	m/z = 250.086; t_R = 10.5 min	*m/z* = 211.982; t_R = 9.49 min	
		m/z = 262.044; t_R = 8.75 min	

4. Conclusions

The present work successfully discerned hundreds of HPLC-HRMS peaks produced during water chlorination and, among them, we tentatively identified halogenated features as DBPs. The chlorination of coarse-filtered river samples resulted in a long list of halogenated species, 495, most of which were brominated (as a consequence of the significant bromide content of the studied river). A high sample-to-sample variability was observed, which suggests that this DBPs library is far for being complete. River samples were obtained and chlorinated in a scenario of water reuse that does not necessarily represent the typical Llobregat flow regime and DOM composition. In the future, the present approach should be applied to other waters in order to obtain a more complete picture of the potential chemodiversity of DBPs, that covers the extreme complexity of DOM precursors and their variations depending on the river basin or the sampling season, and it should also be extended to negative ionisation polarity (which was not employed in the present manuscript) and to other analytical conditions that can effectively widen the scope of detected DBPs.

Nevertheless, in practice, DWTP samples and domestic tap water showed a much more reduced profile of DBPs, which is justified because of the successful removal of precursors in the treatment plant. 15 halogenated DBPs were common in all DWTP samples and 76 were detected in more than three domestic tap waters. Our attention was focused in the (tentative) identification of nitrogen-containing halogenated DBPs found in tap waters as consumers are expected to be more frequently exposed to them. To the best of our knowledge, these DBPs had not been d in previous studies. While the presented structures were in good agreement with the annotated empirical formulae and with the experimental MS (Evlampidou et al., 2020) scans, the approach is limited by the number of molecules entered in the Pubchem library, by instrumental uncertainty (i.e., the eventual presence of interferents that may distort the MS (Evlampidou et al., 2020) scan) and by the non-target workflow uncertainty. Therefore, the unequivocal identification of the DBPs would require the analysis of analytical standards, most of which are not commercially available. In the future, special attention must be paid to reporting the identity of these DBPs with full confidence, in order to assess their risk for the consumers.

The profile of DBPs detected in the different postal codes from Barcelona city was, to some extent, impacted by their geographical location (and, more precisely, with the latitude). The regulated DBPs included in the present study, THMs, were poorly correlated with their concentrations: 59 halogenated DBPs compounds out of 76 were not correlated with THMs ($\rho < 0.6$), and 15 were merely "moderately correlated" ($0.6 = \rho < 0.8$). This indicates that THMs are not adequate descriptors of most non-volatile DBPs and challenges their role as model DBPs. The search of satisfactory predictors of DBP formation is an ongoing research line and, according to the present study, future research should take into account the generation of a large and diverse list of halogenated species.

Author statement

Josep Sanchís: Methodology, Software, Investigation, Visualization, Writing – original draft. **Paula E. Redondo-Hasselerharm**: Sampling, Investigation, Writing – review & editing. **Cristina M.** Villanueva: Sampling, Investigation, Writing – review & editing. **Maria José Farré**: Conceptualization, Methodology, Investigation, Resources, Writing – review & editing, Project administration, Funding acquisition, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Maria Jose Farre reports financial support was provided by Spanish Ministry of Science, Innovation and Universities. Maria Jose Farre reports financial support was provided by European Fund for Regional Development. Cristina Villanueva reports financial support was provided by Ajuntament de Barcelona (Institut de Cultura, Pla Barcelona Ciència, 2019). Maria Jose Farre reports financial support was provided by Generalitat de Catalunya through Consolidated Research Group ENV 2017 SGR 1124. Maria Jose Farre reports financial support was provided by CERCA program. Cristina Villanueva reports financial support was provided by Spanish Ministry of Science and Innovation. Cristina Villanueva reports financial support was provided by CERCA program.

Acknowledgements

This project was funded by the Spanish Ministry of Science, Innovation and Universities, AEI-MICIU, and the European Fund for Regional Development under the National Program for Research Aimed at the Challenges of Society, through the project NDMA_Predict (CTM 2017-85335-R), waterDOM (PID 2020-114065RB-C21), and Ajuntament de Barcelona (Institut de Cultura, Pla Barcelona Ciència 2019 #19S01446-006). The authors thank Generalitat de Catalunya through Consolidated Research Group ENV 2017 SGR 1124 and Tech 2017 SGR 1318. ICRA, Spain researchers thank funding from CERCA program. MJF acknowledges her Ramon y Cajal fellowship (RyC-2015-17108), from the AEI-MICI. CMV and PRH acknowledge the support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019–2023" Program (CEX 2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program. The authors would like to thank the staff of the Scientific and Technical Services of the Catalan Institute of Water Research (ICRA), Spain and Institute of Environmental Assessment and Water Research (IDAEA) for their assistance; the ACA (Agència Catalana de l'Aigua) head of water control and quality department Anoni Munné; the AGBAR (Aigües de Barcelona) technicians, especially Miquel Paraira (Water Quality Director & Laboratory Manager), for their help providing samples from the Drinking Water Treatment Plant. We thank Cintia Flores and Alexandra Paraian (IDAEA-CSIC), Natalia Lopo (ICRA), and Lourdes Arjona and Antonia Valentín (ISGlobal) for their technical assistance during the sampling, analytical measurements and data interpretation. We also thank Patricia González, Anna Gómez, Sònia Navarro and Laia Font-Ribera (Public Health Agency of Barcelona) for providing valuable information on Barcelona's drinking water supplies. We acknowledge all the volunteers that participated in the project providing tap water samples.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2022.135087.

References

Area Metropolitana de Barcelona. ETAP de Sant Joan Despí (in Catalan) : https://www.amb.cat/web/ecologia/aigua/instalacions-i-equipaments/detall/-/equipament/etap-de-sant-joan-despi/348966/11818 ww.amb.cat/web/ecologia/aigua/instalacions-i-equipaments/detall/-/equipament/etap-de-sant-joan-despi/348966/11818 (accessed Aug 27, 2021).

Chuang, Y.-H., Szczuka, A., Mitch, W.A., 2019. Comparison of toxicity-weighted disinfection byproduct concentrations in potable reuse waters and conventional drinking waters as a new approach to assessing the quality of advanced treatment train waters. Environ. Sci. Technol. 53 (7), 3729–3738.

Dittmar, T., Koch, B.P., 2006. Thermogenic organic matter dissolved in the abyssal ocean. Mar. Chem. 102 (3–4), 208–217.

Doederer, K., Gernjak, W., Weinberg, H.S., Farré, M.J., 2014. Factors affecting the formation of disinfection by-products during chlorination and chloramination of secondary effluent for the production of high quality recycled water. Water Res. 48, 218–228.

European Commission, 2020. Commission implementing decision (EU) 2020/1161 of 4 August 2020 establishing a watch list of substances for union-wide monitoring in the field of water policy pursuant to directive 2008/105/EC of the European parliament and of the Council. Off. J. Eur. Union 257, 32–35.

European Parliament, 2020. DIRECTIVE (EU) 2020/2184 OF the EUROPEAN parliament and OF the council of 16 december 2020 on the quality of water intended for human consumption (recast). Off. J. Eur. communities 435, 1–62.

Evans, S., Campbell, C., Naidenko, O.V., 2020. Analysis of cumulative cancer risk associated with disinfection byproducts in United States drinking water. Int. J. Environ. Res. Publ. Health 17 (6), 2149.
Evlampidou, I., Font-Ribera, L., Rojas-Rueda, D., Gracia-Lavedan, E., Costet, N., Pearce, N., Vineis, P., Jaakkola, J.J.K., Delloye, F., Makris, K.C., 2020. Trihalomethanes in drinking water and bladder cancer burden in the European union. Environ. Health Perspect. 128 (1), 17001.
Fabbricino, M., Yan, M., Korshin, G.V., 2019. Effects of chlorination on the fluorescence of seawater: pronounced changes of emission intensity and their relationships with the formation of disinfection byproducts. Chemosphere 218, 430–437.
Farré, M.J., Jaén-Gil, A., Hawkes, J., Petrovic, M., Catalán, N., 2019. Orbitrap molecular fingerprint of dissolved organic matter in natural waters and its relationship with NDMA formation potential. Sci. Total Environ. 670, 1019–1027.
Furst, K.E., Bolorinos, J., Mitch, W.A., 2021. Use of trihalomethanes as a surrogate for haloacetonitrile exposure introduces misclassification bias. Water Res. X 11, 100089.
Government of Canada. No Title https://www.canada.ca/en/health-canada/services/environmental-workplace-health/reports-publications/water-quality/guidelines-canadian-drinking-water-quality-summary-table.html (accessed Feb 3, 2021).
Grimmett, M.R., 1993a. Halogenation of heterocycles: I. Five-membered rings. Adv. Heterocycl. Chem. 57, 291–411.
Grimmett, M.R., 1993b. Halogenation of heterocycles: II. Six-and seven-membered rings. Adv. Heterocycl. Chem. 58, 271–345.
Grimmett, M.R., 1994. Halogenation of heterocycles: III. Heterocyeles fused to other aromatic or heteroaromatic rings. Adv. Heterocycl. Chem. 59, 245–369.
Gros, M., Catalán, N., Mas-Pla, J., Čelić, M., Petrović, M., Farré, M.J., 2021. Groundwater antibiotic pollution and its relationship with dissolved organic matter: identification and environmental implications. Environ. Pollut. 117927.
Kellerman, A.M., Dittmar, T., Kothawala, D.N., Tranvik, L.J., 2014. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. Nat. Commun. 5, 3804.
Kolb, C., Francis, R.A., VanBriesen, J.M., 2017. Disinfection byproduct regulatory compliance surrogates and bromide-associated risk. J. Environ. Sci. 58, 191–207.
Komaki, Y., Ibuki, Y., 2022. Inhibition of nucleotide excision repair and damage response signaling by dibromoacetonitrile: a novel genotoxicity mechanism of a water disinfection byproduct. J. Hazard Mater. 423, 127194.
Lavonen, E.E., Gonsior, M., Tranvik, L.J., Schmitt-Kopplin, P., Köhler, S.J., 2013. Selective chlorination of natural organic matter: identification of previously unknown disinfection byproducts. Environ. Sci. Technol. 47 (5), 2264–2271.
Lu, Y., Song, Z.-M., Wang, C., Liang, J.-K., Hu, Q., Wu, Q.-Y., 2021. Nontargeted identification of chlorinated disinfection byproducts formed from natural organic matter using Orbitrap mass spectrometry and a halogen extraction code. J. Hazard Mater. 126198.
Lyon, B.A., Cory, R.M., Weinberg, H.S., 2014. Changes in dissolved organic matter fluorescence and disinfection byproduct formation from UV and subsequent chlorination/chloramination. J. Hazard Mater. 264, 411–419.
Michałowicz, J., Duda, W., Stufka-Olczyk, J., 2007. Transformation of phenol, catechol, guaiacol and syringol exposed to sodium hypochlorite. Chemosphere 66 (4), 657–663.
Minor, E.C., Swenson, M.M., Mattson, B.M., Oyler, A.R., 2014. Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. Environ. Sci. Process. Impacts 16 (9), 2064–2079.
Muellner, M.G., Wagner, E.D., McCalla, K., Richardson, S.D., Woo, Y.-T., Plewa, M.J., 2007. Haloacetonitriles vs. Regulated haloacetic acids: are nitrogen-containing DBPs more toxic? Environ. Sci. Technol. 41 (2), 645–651.
Natural Resource Management Ministerial Council Environment Protection and Heritage Council Australian Health Ministers' Conference, 2006. Australian Guidelines for Water Recycling: Managing Health and Environmental Risks (Phase 1).
Nweke, A., Scully Jr., F.E., 1989. Stable N-chloroaldimines and other products of the chlorination of isoleucine in model solutions and in a wastewater. Environ. Sci. Technol. 23 (8), 989–994.
Phungsai, P., Kurisu, F., Kasuga, I., Furumai, H., 2018. Changes in dissolved organic matter composition and disinfection byproduct precursors in advanced drinking water treatment processes. Environ. Sci. Technol. 52 (6), 3392–3401.
Plewa, M.J., Wagner, E.D., Muellner, M.G., Hsu, K.-M., Richardson, S.D., 2008. Comparative Mammalian Cell Toxicity of N-DBPs and C-DBPs. ACS Publications.
Postigo, C., Andersson, A., Harir, M., Bastviken, D., Gonsior, M., Schmitt-Kopplin, P., Gago-Ferrero, P., Ahrens, L., Ahrens, L., Wiberg, K., 2021a. 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.
Postigo, C., Gil-Solsona, R., Herrera-Batista, M.F., Gago-Ferrero, P., Alygizakis, N., Ahrens, L., Wiberg, K., 2021b. A step forward IN the detection OF byproducts OF anthropogenic organic micropollutants IN chlorinated water. Trends Environ. Anal. Chem., e00148
Redondo-Hasselerharm, P. E.; Cserbik, D.; Flores, C.; Sanchís, J.; Farré, M. J.; Alcolea, J. A.; Planas, C.; Caixach, J.; Villanueva, C. M. Measuring and Modelling Human Exposure to Multiple Disinfection By-Products in Drinking Water. *(Under Submiss)*.
Richardson, S.D., Plewa, M.J., 2020. To regulate or not to regulate? What to do with more toxic disinfection by-products? J. Environ. Chem. Eng. 8 (4), 103939.
Rook, J.J., 1976. Haloforms in drinking water. Journal-American Water Work. Assoc 68 (3), 168–172.
Ruttkies, C., Schymanski, E.L., Wolf, S., Hollender, J., Neumann, S., 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J. Cheminf. 8 (1), 1–16.
Sanchís, J., Jaén-Gil, A., Gago-Ferrero, P., Munthali, E., Farré, M.J., 2020a. Characterization of organic matter by HRMS in surface waters: effects of chlorination on molecular fingerprints and correlation with DBP formation potential. Water Res. 176 https://doi.org/10.1016/j.watres.2020.115743.
Sanchís, J., Gernjak, W., Munné, A., Catalán, N., Petrovic, M., Farré, M.J., 2020b. Fate of N-nitrosodimethylamine and its precursors during a wastewater reuse trial in the Llobregat river (Spain). J. Hazard Mater. 124346.
Sanchís, J., Petrović, M., Farré, M.J., 2021. Emission of (chlorinated) reclaimed water into a mediterranean river and its related effects to the dissolved organic matter fingerprint. Sci. Total Environ. 760, 143881.
Sedlak, D.L., von Gunten, U., 2011. The chlorine dilemma. Science 331 (6013), 42–43.
Shi, Y., Ma, W., Han, F., Geng, Y., Yu, X., Wang, H., Kimura, S.Y., Wei, X., Kauffman, A., Xiao, S., 2020. Precise exposure assessment revealed the cancer risk and disease burden caused by trihalomethanes and haloacetic acids in shanghai indoor swimming pool water. J. Hazard Mater. 388, 121810.
Tao, D., Wang, R., Shi, S., Yun, L., Tong, R., Guo, W., Liu, Y., Hu, S., 2020. The identification of halogenated disinfection by-products in tap water using liquid chromatography–high resolution mass spectrometry. Sci. Total Environ. 740, 139888.
Villanueva, C.M., Cantor, K.P., Cordier, S., Jaakkola, J.J.K., King, W.D., Lynch, C.F., Porru, S., Kogevinas, M., 2004. Disinfection byproducts and bladder cancer: a pooled analysis. Epidemiology 15 (3), 357–367.
Villanueva, C.M., Cordier, S., Font-Ribera, L., Salas, L.A., Levallois, P., 2015. Overview of disinfection by-products and associated health effects. Curr. Environ. Heal. reports 2 (1), 107–115.
Wagner, S., Riedel, T., Niggemann, J., Vähätalo, A.V., Dittmar, T., Jaffé, R., 2015. Linking the molecular signature of heteroatomic dissolved organic matter to watershed characteristics in world rivers. Environ. Sci. Technol. 49 (23), 13798–13806.
Wang, X., Mao, Y., Tang, S., Yang, H., Xie, Y.F., 2015. Disinfection byproducts in drinking water and regulatory compliance: a critical review. Front. Environ. Sci. Eng. 9 (1), 3–15.
Wasilenko, W.J., Palad, A.J., Somers, K.D., Blackmore, P.F., Kohn, E.C., Rhim, J.S., Wright Jr., G.L., Schellhammer, P.F., 1996. Effects of the calcium influx inhibitor carboxyamido-triazole on the proliferation and invasiveness of human prostate tumor cell lines. Int. J. Cancer 68 (2), 259–264.
Watson, K., Farré, M.J., Leusch, F.D.L., Knight, N., 2018. Using fluorescence-parallel factor Analysis for assessing disinfection by-product formation and natural organic matter removal efficiency in secondary treated synthetic drinking waters. Sci. Total Environ. 640, 31–40.
Williams, C.J., Conrad, D., Kothawala, D.N., Baulch, H.M., 2019. Selective removal of dissolved organic matter affects the production and speciation of disinfection byproducts. Sci. Total Environ. 652, 75–84.
Yang, M., Zhang, X., Liang, Q., Yang, B., 2019. Application of (LC/) MS/MS precursor ion scan for evaluating the occurrence, formation and control of polar halogenated DBPs in disinfected waters: a review. Water Res. 158, 322–337.
Zhang, H., Yang, M., 2018. Characterization of brominated disinfection byproducts formed during chloramination of fulvic acid in the presence of bromide. Sci. Total Environ. 627, 118–124.