Impact of method parameters on the performance of suspect screening for the identification of trace organic contaminants in surface waters

Pedro A. Segura*, Mathieu Racine, Alexia Gravel, Emmanuel Eysseric, Anne-Marie Grégoire, Diane Rawach, François-Xavier Teysseire

* Tel: 1-(819) 821-7922. Fax: 1-(819) 821-8019. E-mail: pa.segura@usherbrooke.ca

Department of Chemistry, Université de Sherbrooke, Sherbrooke, QC J1K 2R1

Keywords: Contaminants of emerging concern; pesticides; suspect screening; high-resolution mass spectrometry; solid-phase extraction.

The final version of this paper was published in the *Canadian Journal of Chemistry:*

Segura P. A., Racine M., Gravel A., Eysseric E., Grégoire A.-M., Rawach D., Teysseire F.-X. (2019) Impact of method parameters on the performance of suspect screening for the identification of trace organic contaminants in surface waters. *Canadian Journal of Chemistry* 97:197–211. https://doi.org/10.1139/cjc-2018-0298

Abstract

The performance of a suspect screening method to detect diverse small-molecule trace organic contaminants (TOCs) was systematically evaluated using a set of 39 model compounds. The suspect screening method used a commercial algorithm for peak identification based on liquid chromatography-high resolution mass spectrometry data. Experiments showed that ionization efficiency, ion transfer parameters and chromatography could affect the detection of TOCs. As expected, compounds with low ionization yields and poorly retained compounds in chromatographic columns are more difficult to identify in the samples at environmental concentrations. Similarly, TOCs with large deviations from the average mass of the compounds screened were not transmitted efficiently in the mass spectrometer thus negatively affecting their detection. The suspect screening method was validated in terms of recovery and limits of identification of the model compounds using three different types of solid-phase extraction cartridges (reversed-phase with polar groups, mixed-mode anion exchange and mixed mode cation exchange). Experiments showed that more than two thirds of the model compounds had recoveries > 75% with each of the three cartridges and comparison of limits of identification showed that more than half of the model compounds could be identified at concentrations between 6 and 100 ng L^{-1}. However, it was observed that the amount of co-extracted compounds was higher in mixed-mode ion exchangers compared to the reversed-phase cartridge. Application of the suspect screening method using the three different cartridges to surface water samples showed that between 0 to 3% of the positive matches found by the peak identification algorithm were classified as probable structures. Solutions to improve suspect screening of TOCs are proposed and discussed.

1. Introduction

Despite the progress in environmental analysis since the 20th century, many trace organic contaminants (TOCs) have gone unnoticed in the environment for years because of inadequate methodology [1]. Pharmaceuticals and personal care products (PPCPs) are perhaps the most recent example. These compounds have been present in surface waters for many years but only until the late 1990's they became a topic of concern among scientists. Today more than 200 PPCPs have been detected in environmental waters around the world [2]. While the development, of liquid chromatography-tandem mass spectrometry has been key component for the discovery of PPCPs in the environment [3], many of those PPCPs could have been detected with techniques available since the 1980s, *e.g.* solid-phase extraction, liquid chromatography-uv/fluorescence detection, derivatization and gas chromatography-mass spectrometry, as some published methods suggest it [4,5].

The reason why traditional methods of environmental analysis have failed to detect new contaminants is mainly because they are target methods, *i.e.* they only are able to detect specific compounds for which they were developed. Therefore, the rest of the contaminants potentially present in the samples are ignored. In many cases, targeted approaches used to determine the occurrence of contaminants in the environment could lead to a gradual phenomenon known as the "Matthew Effect": researchers target only the compounds detected by previous studies and compounds that are not detected end up being completely ignored with time [6].

Without a doubt, target methods are the most efficient way to quantify analytes since these methods can be specifically tailored and optimized to extract, separate and quantify a limited array of compounds. However, the resulting data, especially in the case of occurrence studies, are potentially biased if not used with a previous knowledge of the presence of all relevant compounds in a sample. Therefore, when studying the occurrence of TOCs, a quantitative targeted approach can be more comprehensive after a qualitative knowledge of the composition of the sample is obtained.

Since most research on the fate and occurrence of TOCs such as PPCPs has focused mainly on target studies, our information about their occurrence and fate is still fragmentary. Therefore, there is a pressing need to improve the characterization of environmental samples. Also, a better understanding of the presence of TOCs in natural waters will make possible to concentrate our efforts on the most concerning compounds, *e.g.* those having a greater toxicity or that degrade in the most persistent transformation products, and find solutions to reduce their presence in the environment.

In order to solve this dilemma, one of the most successful approaches that have been applied to solve the limitation of target methods of analysis, is suspect screening methods, which aim to

detect relevant compounds in a wide mass-to-charge ratio range (*e.g. m/z* 100 to 1000) in a sample. In suspect screening methods, samples are treated using generic extraction techniques, such as reversed-phase solid-phase extraction (SPE) or liquid-solid extraction, and analysis is usually carried out with liquid chromatography-high resolution mass spectrometry (LC-HRMS). Compounds present in the samples are identified by matching accurate masses, isotopic patterns and tandem mass spectra to compounds in databases of suspect contaminants, *i.e.* compounds potentially present in the samples such as known water contaminants or toxic compounds [7].

The interest in suspect screening methods has grown in the last decade and these methods have been successfully applied to the analysis of contaminants in wastewaters [8], ground waters[9] and surface and drinking waters [10]. The importance of suspect screening methods is proven by a recent study on the toxicity of environmental water samples which indicated that more than 99% of the response of two *in vitro* bioassays was due to unidentified compounds [11]. The potential of suspect screening methods to identify previously disregarded compounds has been demonstrated by multiple studies. For example, Moschet *et al.* [12] analysed surface water samples using LC-HRMS and processed their data using a commercial software and a home-made database to screen for 185 pesticides and their transformation products. Their approach was successful to unambiguously identify 13 pesticides and 5 transformation products. Schymanski et al. [13] used suspect screening to identify for the first time a transformation product of an industrial compound in wastewater treatment plant (WWTP) effluents. Later, Sjerps et al. [14] applied suspect screening to 151 samples comprising WWTP effluents, surface waters, groundwaters and drinking waters. They used a suspect list of 5219 chemicals and identified 174 compounds. Only 20% of them were mentioned in lists of potentially relevant chemicals, which showed the complementarity between suspect screening and target methods.

Nevertheless, suspect screening methods still suffer of many technical issues that limit their application. For example, in a suspect screening validation study [10], low sensitivity was observed in real samples: only 1 out of 15 compounds spiked at a concentration equivalent to 5 ng L^{-1} were identified while in a pure solvent spiked at the same concentration,13 out of 15 compounds were identified in the samples. Furthermore, despite the fact that the molecular formula of selected peaks can be assigned using the numerous databases and libraries available online containing TOCs (PubChem, Chemspider, NIST Library, Merck Index, etc.), their correlation with a specific compound remains problematic since there can be multiple structures corresponding to the same molecular formula [15]. Many studies also stressed the fact that there is a lack of open spectral libraries containing organic contaminants that are independent of instrumentation and settings compared to the exhaustive ones available for proteomics and metabolomics [13,16]. Those studies wished for databases to be extended with contaminants, which could possibly fix another problem: the need for standards to unmistakably identify an organic micropollutant. Indeed, the use of analytical standards may be useful in many cases, but they can also be expensive or even unavailable for many compounds and their stable transformation products.

The objective of this paper is to carry out a systematic study to determine the effect of three key parameters on the results obtained by suspect screening methods: chromatographic separation, ion transmission and solid-phase extraction. While suspect screening workflows may be affected by a diverse range of parameters and experimental conditions, we decided to focus on crucial parameters and conditions that are often used in suspect screening experimental setups. To the authors' knowledge, only one study has systematically evaluated the parameters affecting suspect screening [17] but it did not include all steps of the method. Therefore, a comprehensive study of method parameters on suspect screening performance has not yet been published and recent studies have focused mostly in the optimization of compound identification workflows [8,12,18,19]. In order to achieve the proposed systematic study, a set of 39 structurally diverse organic compounds representative of suspect contaminants such as PPCPs, pesticides and consumer product additives was selected. These compounds were carefully chosen to study the capacity of a suspect screening method to identify a list of 278 diverse TOCs (for more details consult the Excel file: SupplementaryMaterial(Databases-Results).xlsx) in extracts of surface water obtained after SPE using three different type of cartridges: reversed-phase with polar groups, mixed-mode weak anion exchange and mixed-mode weak cation exchange.

2. Materials and methods

2.1. Reagents and chemicals

The following organic compounds were used as model TOCs and were purchased from Sigma-Aldrich Canada : acetaminophen (acronym: ACT, purity 99.0 %), atrazine (ATZ, 98.1%), avobenzone (AVB, 99.5%), benzyl butyl phthalate (BBP, 98%), carbamazepine (CBZ, ≥ 98%), carbaryl (CAR, , ≥ 98%), chlorpyrifos (CPF, 99.8%), cyclophosphamide (CYC, 99.2%), sodium diatrizoate hydrate (DIA, ≥ 98%), dibutyl phthalate (DBP, 99%), diclofenac (DCF, ≥ 98.5%); 2,4-dichlorophenoxyacetic acid (2,4-D, 99.9%), dimethylaminophenazone (DAP, ≥ 98%), 17α-ethinylestradiol (EE2, 99.4%), ensulizole (ESZ, 98.0%), fluoxetine hydrochloride (FLX, 99.8%), gemfibrozil (GEM ≥98.5%), glyphosate (GLY, 99.7%), ibuprofen (IBU, 98%), linuron (LIN, 99.7%), metoprolol tartrate (MET, ≥98%), ofloxacin (OFX, 99.8%), propyl paraben (PPB, 99%), salicylic acid (SCA, 99.0%), roxithromycin (ROX, ≥ 90%), sucralose (SUC, 98%), sulfamethoxazole (SMX, 99.9%), tributyl O-acetylcitrate (ATBC, 98%), 2,2,4-trimethyl-1,3-pentanediol diisobutyrate (TXIB, 98.5%), o-toluenesulfonamide (OTSA, 99%), oxybenzone (OXB, ≥ 98%), tris(1,3-dichloro-2-propyl)phosphate (TDCPP, 95.7%), tributyl phosphate (TBP, ≥ 99%), tris(2-butoxyethyl) phosphate (TBEP, 94%) and thiabendazole (TBZ, 98.0%). Naproxen (NAP, 99.9%), methotrexate (MTX, ≥ 98%), sulisobenzone (SLB, ≥ 97%) and tris(1-chloro-2-propyl)phosphate (TCPP, 97.5%) were purchased from Santa Cruz Biotechnologies.

Solvent and additives such as methanol (MeOH, optima LC-MS grade), acetonitrile (ACN, Optima LC-MS), water (Optima LC-MS), formic acid (FA, Optima LC-MS) and ethylenediaminetetraacetic acid disodium salt (Na_2EDTA, ACS grade) were purchased from

Thermo Fisher Scientific. Ammonium acetate (NH_4Ac, HPLC grade) was obtained from EMD Millipore. Ammonium hydroxide was bought from Sigma-Aldrich Canada.

2.2 . Choice of model compounds

Since there are thousands of TOCs that could be found in each sample, it is necessary to develop a simple approach to identify key steps in suspect screening methods. The families of compounds that are of interest for this study are: pharmaceuticals, personal care products, organic additives of consumer products and pesticides. These families were selected because of their frequent detection in the environment and their potential harmful effects on aquatic species. Therefore, the compounds that will serve as models are representative of those families and could be found in surface waters. They were were used for optimization tests and overall method performance evaluation. The selection of model organic contaminants was done according to key criteria such as concentration in wastewaters and surface waters, production volumes and reports on usage, hydrophobicity and molecular structure. Structural diversity, *i.e.* different chemical functions, was a main factor in the final selection of the set of model compounds in order develop a method able to identify a wide range of organic contaminants and not only those that have been previously reported. In total, 39 compounds were selected: 17 pharmaceuticals, 4 active ingredients in personal care products, 11 consumer product additives and 7 pesticides. A detailed explanation of the procedure used in the choice of model compounds is presented in the Supplementary Material.

2.2. Sampling

Samples of surface water from the Magog river (45^o 16'14.4"N, 72^o 07'15.3"W) in the province of Québec, Canada were collected in clean amber HDPE bottles on May 7th, 2015 to evaluate SPE recoveries. A second sampling camping, on September 26th, 2016, was carried out to collect water from the St-François river (45° 26'42.3"N, 71° 55'26.1"W) to estimate method limits of identification and apply the optimized suspect screening method. Samples were conserved at -20 °C until analysis. Field blanks were prepared by filling HDPE bottles with deionized water (18 MΩ) and opening them during sample collection. These samples were subjected to all the sample preparation steps as the surface water samples.

2.2 Solid-phase extraction (SPE)

Three different types of styrene-divinylbenzene (PS-DVB) based SPE cartridges manufactured by Phenomenex, Strata-X, Strata-X-AW and Strata-X-CW, were tested in order to study the effect of

sorbent chemistry on the extraction recoveries for the model compounds. Cartridges had 6 mL of volume and contained 200 mg of sorbent. Strata-X is a reversed-phase sorbent with polar groups (PS-DVB modified with *N*-vinylpyrrolidone), Strata X-AW is a mixed-mode weak anion exchanger (Strata-X-AW, PS-DVB functionalized with a diamine groups) and Strata-X-CW is a mixed-mode weak cation exchanger (Strata-X-CW, PS-DVB functionalized with carboxylate groups). These sorbents were also used to evaluate differences in the number of compounds identified in surface water samples by the tested suspect screening method.

Spiked surface water samples were used to evaluate the performance of the suspect screening method. Before extraction, model compounds were spiked in a flask containing 250 mL of surface water for a final concentration of 200 ng L^{-1}. To each flask 50 mg of Na_2EDTA were added to improve recovery of compounds capable of forming chelates with dissolved metals. Extraction conditions for each cartridge are shown in Table 1. Extraction recoveries were determined by comparing the areas of three replicate surface water samples spiked before SPE to those of three replicate surface water samples spiked after SPE. Surface water samples used for the identification of suspect screening contaminants were prepared using this same procedure except that they were not spiked with model compounds.

2.3 Reversed-phase liquid chromatography

A Shimadzu Nexera UHPLC system was used for the separation of the analytes on a C_{18} Acquity UHPLC HSS T3 column (2.1 × 50 mm, 1.8 µm) from Waters Co. Column temperature was set to 25 ^{o} C and the flow rate of mobile phase was 500 µL min^{-1}. Several mobile phase gradients were tested to obtain optimal retention of all the compounds and minimize potential signal suppression caused by coelution. Since compounds could ionize in positive or negative electrospray ionization (ESI), different solvents and additives were tested. In ESI+, two organic solvents were evaluated for the mobile phase: ACN and MeOH both with 0.1 % (v/v) FA. For ESI-, MeOH and ACN were tested as solvent B both with 1 mM NH_4Ac. Final chromatographic conditions for ESI+ were: solvent A was 0.1% FA in H_2O and solvent B was 0.1% FA in MeOH. Elution gradient, as %B in the mobile phase, was: 0 min (2%), 13min (65.7%), 19 min (71%), 20 min (100%), 25 min (100%), 26 min (2%), 30 min (2%). For ESI-, final chromatographic conditions were: solvent A was 1 mM NH_4AC in H_2O and solvent B was 1 mM NH_4Ac in MeOH. The same gradient as in ESI+ was used for ESI-. The sample volume injected was 3 µL for ESI+ and 5 µL for ESI-. In ESI-, higher signal-to-noise ratios are observed compared to ESI+ [20] and for many compounds ESI- allows better sensitivity [21]. However, it has been demonstrated that buffers such as ammonium formate and ammonium bicarbonate can severely supress the signal in ESI- compared to 0.1% formic acid in ESI+. Therefore, in order to compensate for such effects and perform a fairer comparison it was decided to use a higher volume of sample for ESI- than for ESI+.

2.4 Electrospray ionization quadrupole-time-of-flight mass spectrometry

A liquid chromatograph described earlier was coupled to a Maxis quadrupole-time-of-flight mass spectrometer (QqTOFMS) manufactured by Bruker Daltonics and equipped with an electrospray ionization source. Source parameters for positive and negative ionization were: capillary, 2000 V (ESI+) or 3000V (ESI-), nebulizer gas, 4 bar, dry gas, 10 L min^{-1} and dry temperature, 200 ^{o}C. In a first stage of the experiments only MS^1 data was acquired and the scan range was *m/z* 50 to 1200. Scan ranges for the ion cooler radiofrequency voltages and the transfer time in both modes were from 55 to 330 Vpp and from 30 to 60 µs, respectively. For both parameters, the timing (proportion of the scan time that the respective value is applied) was 50%-50%. Pre-pulse storage was 5 µs. Mass calibration was done with sodium formate before each analysis. With these conditions, mass resolution (R_{FWHM}) for *m/z* 403 was, in average, 43000. In a second stage, MS^2 experiments were performed only with the tentative candidates (level 3 according to the identification level scheme proposed by Schymanski *et al.* [22]). In this way the amount of data acquired is reduced and only MS^2 data of interest for the suspect contaminants are recorded. Quadrupole isolation width of 3 Da, a collision energy of 30 eV and nitrogen as collision gas were used to generate product ion spectra.

2.5 Data analysis

A previously used home-made library containing 278 (264 unique exact masses and 14 isomers) of suspect surface water TOCs [10] was transferred to Compass Library Editor (Bruker Daltonics). This library contains TOCs such as pesticides, pharmaceuticals, personal care products, and consumer product additives such as plasticizers and flame retardants as well as reported transformation products (available as an Excel file, Supplementary Material). For each compound, ions $[M+H]^+$and $[M+Na]^+$ were added to the library for the ESI+ mode and ions $[M-H]^-$ and $[M+CH_3COOH-H]^-$ for the ESI- mode. The algorithm Molecular Feature (MF) of the DataAnalysis software (version 4.2) from Bruker was used to identify peaks of the suspect TOCs. The algorithm considers that a series of signals belong to a given compound if they have a high correlation in retention time and isotopic pattern. Thus, the optimum correlation coefficient (0.7) was selected to obtain the highest percentage of detection. Since the identification of a given compound by the algorithm is highly dependent on the signal to noise ratio (S/N) of each chromatographic peak, a value of 3 for that parameter was used to allow the maximum number of analyzed compounds differing from the noise. Subsequently, an optimum mass tolerance between experimental and calculated mass of the model compound of 7 mDa was chosen to obtain a minimum of both false negatives and false positives.

In order to evaluate quantitatively the quality of a match with the library, we used two match scores calculated by DataAnalysis: Fit and Reverse Fit. Both scores indicate how similar are the library and the acquired mass spectra, in terms of mass and relative intensities. However, they differ in the way the peaks are used for the calculation of the score: in the Fit score algorithm, peaks in the

acquired spectrum that are not present in the library spectrum are disregarded, while in the Reverse Fit, peaks in the library spectrum that are absent in the acquired spectrum are ignored. In both cases the maximum score is 1000. Using these two scores allows to eliminate incorrect matches since a high value (*e.g.* > 900) for both scores suggests a strong match. An optimal value of Fit and Reverse Fit of 940 was found by injecting in the LC-QqTOFMS solutions of the model compounds at a concentration of 100 μg L^{-1} in H_2O: MeOH 1:1 and applying the MF algorithm to identify and compare acquired and library mass spectra. Therefore, the MF algorithm is not merely a peak picking algorithm, it also evaluates the quality of a match by comparing experimental and theoretical isotopic patterns for a given suspect contaminant. In this way the number of false positives is reduced since an exact mass match is not sufficient to classify an accurate mass as an identified compound.

2.6 Evaluation of method limits of identification of model compounds in surface water samples

Method performance was evaluated using a mixture of the 39 model compounds at 4 different concentrations of 5, 25, 210 and 500 μg L^{-1} spiked in SPE extracts of the surface water from St-François River, considering the preconcentration factor of 833, those concentrations were equivalent to about 6, 30, 252 and 600 ng L^{-1}. Non-spiked water samples were also analyzed to measure the possible occurrence of the model compounds in the samples to the signal observed. This procedure allowed to evaluate the capacity of the suspect screening method to identify the presence of model compounds in the samples that could be extrapolated to other suspect contaminants present in the home-made database. This limit in the performance of the method was called the “limit of identification” (LOI), which depends on several factors such as analyte response and the presence of co-eluting interferences that may prevent the identification of suspect contaminants by the suspect screening method and is different than the limit of detection. Since some of the model compounds are know as common laboratory contaminants (e.g. dibutyl phthalate), compounds were considered to have been identified when the signal obtained in the spiked extracts was higher than the signal plus its standard deviation in the unspiked extracts from Saint-François River. A final correction was added to consider extraction recoveries of the individual model compounds, therefore a final value the “Corrected limit of identification” was calculated to determine the capacity of the method to correctly identify low concentrations of TOCs in the samples.

2.7 Approaches used to improve the identification confidence levels of suspect contaminants

The identification level scheme proposed by Schymanski *et al.* [22] was used in order to classify the different compounds identified by the screening method according to their identification confidence level (Figure 1). To obtain a higher identification confidence level, several filters or techniques were applied: elimination of candidates present at higher abundance in the field blanks

than in the samples, compounds with low signals (e.g. <1000 counts), estimation of retention times using ChromGenius software (ACD Labs), comparison of experimental and library tandem mass spectra (mzCloud) as well as injection of pure standards when they were available.

When the MF algorithm could not differentiate between 2 molecular formulas, we used ranking by spectral accuracy as means to identify the correct molecular formula. Spectral accuracy uses the whole experimental isotopic pattern to rank possible molecular formulas according to similarity to theoretical isotopic patterns [23]. This technique has been used to reduce the number of possible molecular formulas that can be assigned to a given accurate mass in environmental samples [24].

Retention times of TOCs potentially present in the samples were predicted by building a knowledge base of retention times of small organic molecules in ChromGenius version 2017.1.3. Briefly, 215 compounds (mostly pesticides and pharmaceuticals) were injected in the LC-QqTOFMS system using the same chromatographic method described earlier. Retention time and molecular structure of each compound were entered in the software in order to develop a model of prediction of retention time based on physicochemical properties and molecular similarity [25]. Software settings were: similarity (Tanimoto coefficient using the best 35 records) and equation (correlate retention time). The following options were also checked: LogD, LogP, polar surface area, molecular volume, molecular weight, molar refractivity, H donors and H acceptors. This method was validated using a subset of ten compounds to determine its accuracy (for more details consult Supplementary Material). It was observed that, the absolute difference between experimental and predicted retention times for those ten compounds was in average, 1.8 ± 1.6 min. In order to reduce the number of compounds eliminated due to uncertainty of the model of prediction of retention time, the upper 95% confidence limit, 2.9 min, was selected as threshold value for acceptable difference between predicted and experimental retention times. Therefore, it was assumed that absolute retention time differences larger than that threshold value suggest that the observed peak in the chromatogram is not the suspect TOC interest. While the chosen threshold value is not free of false negatives, it was the best compromise to ensure good confidence in the proposed tentative structures (identification confidence level 3 according to Schymanski et al.[22]).

For comparison of experimental and reference library tandem mass spectra, mzCloud (https://www.mzcloud.org/) database was employed. As of October 2018, mzCloud had more than 2.6 million spectra corresponding to more than 7400 unique compounds [26]. Since abundance of product ions may be different depending on the instrument and the collision induced dissociation parameters, Total Composite Spectrum in mzCloud of each potential match was used. A match was considered acceptable if the mass difference between experimental and library product ions was less than 5 mDa. Only product ions having a signal-to-noise ratio > 3 and abundance > 10% were used for comparisons.

3. Results and discussion

3.1 Optimisation of liquid chromatography and ion transmission

Two compositions were tested for solvent B (organic solvent) for ESI+ experiments: 0.1 % FA in MeOH and 0.1 % FA in ACN. All the compounds detected in 0.1 % FA in ACN were also detected 0.1 % FA in MeOH, however, in the former, nine compounds (SCA, EE2, IBU, GEM, 2,4-D, CPF, TCPP, PPB) were not detected (Figure 2). For that reason, further tests with ACN as solvent B were not carried out and it was decided to use 0.1% FA in MeOH as solvent B for ESI+ experiments. Lower response using ACN in ESI+ for most of the model compounds can be partly explained by its higher solution and gas phase basicity compared to MeOH. As reported previously, maximum response for analytes in ACN-H_2O mixtures is observed at lower proportions of organic solvent compared to MeOH-H_2O mixtures [27]. Therefore, the signal of strongly retained compounds should be lower in mobile phases containing ACN compared to those containing MeOH as it was observed for many of the model compounds in the present experiments.

For example, for ACT retention times (t_R) in ACN and MeOH were 1.9 and 2.6 min, respectively and the average ACN to MeOH peak area ratio was 1.3, thus indicating a higher signal for ACT with ACN as B solvent than with MeOH. The same was observed for DIA (t_R=2.1 min with ACN, t_R=3 min with MeOH) which had an intensity ratio of 1.2. In the case of late eluters such as DCF (t_R =10.3 min with ACN, t_R =14.2 min with MeOH) the intensity ratio was 0.27, the same with BBP (t_R=13.1 min with ACN, t_R =16.7 with MeOH) with an intensity ratio of 0.31. However, such effect was not generalized, for early eluters such as carbaryl (t_R =2.8 min with ACN, t_R =4.7 min with MeOH) the intensity ratio was 0.41. Therefore, while mobile phase composition may have had an impact on the ESI signal, analyte-dependent behaviour was also observed, as previously reported [28].

For ESI-, two solvent B compositions were tested: 1 mM NH_4Ac in MeOH and 1 mM NH_4Ac in ACN (Figure 3). As observed with ESI+, in ESI- 10 compounds less (SCA, ACT, 2,4-D, TDB, DCF, PPB, LIN, GEM, TDCPP and ROX) were detected in ACN compared to MeOH (same compounds plus MTX and IBU). Therefore, 1 mM NH_4Ac in MeOH was selected as solvent B in ESI-. Higher ionization efficiency in ESI- of small organic molecules using MeOH vs ACN is due mainly to the protic nature of MeOH that has a stabilizing effect on the deprotonated form of compounds having acid functional groups [29]. Previous reports have shown that the composition of the mobile phase can have important effects on the analyte signal in LC-MS [30,31]. However, the objective here was not to test an array of conditions but to evaluate the relative impact of the conditions most often used. Therefore, MeOH and ACN were the solvents tested regarding mobile phase composition with formic acid and ammonium acetate as additives.

Chromatographic methods were optimized by changing the slope of the gradient in order to obtain adequate separation of the model compounds (results not shown). Using the gradient described in the *Materials and methods* section and the selected mobile phase for ESI+, it was observed that 38 out of 39 model compounds eluted between 2 and 20.5 min. Only 1 compound, glyphosate (GLY), could not be retained sufficiently by the column and it eluted too close to the solvent front to be detected adequately. In the LC separation using the selected mobile phase for ESI- experiments, all the 12 compounds detected eluted between 2 and 20.5 min. These results showed that the chromatographic method used allowed the retention of a wide range of organic compounds. However, hydrophilic TOCs or their transformation products or metabolites could not be properly separated which adds a negative bias in suspect screening towards very polar compounds. Chromatographic methods based on hydrophilic interaction chromatography (HILIC) are more appropriated for the retention of such compounds [18]. HILIC has been used previously in a suspect screening study[18], however a separate HILIC method was not developed since one of the objectives of the present work was to test the performance of a single method.

In order to simplify the experiments, it was decided to run all subsequent experiments in ESI+ since all analytes tested could be detected in that mode and few advantages were observed using ESI-. Limiting ionization to a single type of source or polarity could bias the results in non-targeted screening workflows since many contaminants of interest may ionize better in ESI- than ESI+ or in atmospheric pressure chemical ionization compared to ESI. Nevertheless, the impact of using only ESI+ on the results is minor since the list of suspects that was used contained mostly compounds that can be ionized by ESI+.

Following mobile phase composition optimization tests, two key parameters for the optimal transmission of ions in the Bruker axis QqTOFMS instrument, ion cooler voltages and transfer time, were studied. Two functioning modes were compared: scan mode (set values vary between a minimum and maximum value) and fixed mode (static value). The ion cooler or cooling cell is an interface between the collision cell and the TOF mass analyser and is composed of an hexapole ion guide. The ion cooler improves ion focusing and loss of kinetic energy and reduces pressure in the orthogonal acceleration stage. It also allows the accumulation of product ions before transferring them to the TOF [32]. The transfer time limits the transferred mass range with longer transfer times allowing the transfer of ions of higher *m/z* to the TOF mass analyser. Since those parameters affect the transmission of ions as a function of their *m/z*, nine model compounds representing a wide range of *m/z* values were selected to optimize ion transmission: SCA (*m/z* 139), ACT (*m/z* 152), OTSA (*m/z* 172), GEM (*m/z* 251), SMX (*m/z* 254), FLX (*m/z* 310), MTX (*m/z* 455), DIA (*m/z* 614) and ROX (*m/z* 837).

With fixed ion cooler radiofrequency voltages of 192.0 Vpp and transfer time of 45.0 μs, only compounds with *m/z* values between 250 and 455 (GEM, SMX, AVB and MTX), were detected (Figure 4). Compounds with higher or lower *m/z* were not observed. When scanning ion cooler values between 55.0 and 330.0 Vpp and transfer times between 30.0 to 60.0 μs, the intensity of

most compounds decreased compared to the fixed mode, however all nine compounds were detected (Figure 4). Therefore, the scan mode was used for both ion cooler and transfer time to allow the detection of a larger number of compounds.

The bias caused by ion transmission settings according to *m/z* of sample components has been reported previously and in some cases, signal suppression caused by ion cooler settings was as high as 95% [33]. While ion cooler parameters may be unique for Bruker QqTOFMS, settings affecting ion transmission can be set in other mass spectrometers. These results demonstrate an inherent limitation of suspect screening methods using HRMS for the analysis of a wide range of water TOCs: compromises in terms of sensitivity must be done to detect a wide range of compounds. For that reason, sample preparation is a critical step in suspect screening methods since it could be used to enhance the signal of compounds with low ionization efficiency or poor ion transmission.

3.2 Solid-phase extraction recovery of model compounds

Previous suspect screening studies of TOCs have used different types of commercial SPE cartridges for sample extraction such as reversed-phase type Oasis HLB [poly(divinylbenzene-co-*N*-vinylpyrrolidone)] [9,10] and Chromabond HR-X (polystyrene-divinylbenzene co-polymer) [8] or multilayered cartridges using a mixture of reversed-phase sorbents such as Oasis HLB and Isolute ENV+ (PS-DVB functionalized with phenolic groups) and mixed-mode weak ion exchangers such as Strata-X-AW (PS-DVB functionalized with an ethylene diamine group) and Strata-X-CW (PS-DVB functionalized with a carboxylic group) [12,18]. Those types of polymeric sorbents show good retention of polar and nonpolar compounds compared to silica C_{18} sorbents used in the past [34].

Results of recovery experiments of the model compounds using the three tested SPE cartridges, reversed-phase with polar groups (RP), mixed-mode weak cation exchange (WCX) and mixed-mode weak anion exchange (WAX), are shown in Figure 5. These data show that all cartridges had acceptable recoveries (>75%) for about three quarters of the compounds tested.

Mixed-mode polymeric sorbents such as those used in this study posses both hydrophobic and ionic regions that are able to retain a multitude of compounds based on van der Waals or ionic interactions, respectively. For example, naproxen (NAP) is an organic acid that cannot be retained by ionic interactions with the carboxylic acid groups bonded to the PS-DVB particles of the solid phase in the WCX cartridge. Therefore, as expected, the recovery with the 5% FA/MeOH elution was nil and NAP was completely recovered with the ACN-MeOH 1:1 elution (103.0 % ±7.8). Therefore, this high recovery was due to van der Waals and π-π interactions between the compound and PS-DVB. In the case of the RP cartridge, retention of medium to low polar

compounds is possible because of H-bond and van der Waals interactions with vinylpyrrolidone groups bonded to the polymer.

Out of the 38 compounds that can be separated by the developed LC method only 2 had nil recoveries: EE2 and 2,4-D. This was probably due to matrix effects and/or low recovery. EE2 could be observed in MeOH:H_2O (1:1) spiked at 100 µg L^{-1} (Figure 2) but at a higher concentration (166 µg L^{-1}) in the presence of the surface water matrix, it could not be detected which suggests a strong signal suppression for this compound. In the case of 2,4-D, the compound was only observed during extraction recovery experiments in the river water samples spiked after SPE with the RP cartridges. It could not be detected in any of the WAX of WCX extracts. Recovery of this compound with the RP cartridges was probably low due to its low distribution constant (logD=-0.8) at the extraction pH. Previous reports have shown good recoveries in Strata-X cartridges for both EE2 and 2,4-D but in different experimental conditions, e.g. instrument or extraction pH [35,36].

While the recovery results demonstrate that it was possible to extract a large portion of the test compounds using at least one of the tested cartridges, for some compounds spiked in the samples detection was impossible. This shows again another limitation of suspect screening analysis: generic sample extraction methods using a single sorbent chemistry are not able to recover adequately many TOCs of interest and compounds with low ionization efficiencies or poorly transmitted to the mass analyzer are the most affected since the presence of signal suppression induced by the matrix can lower their signal to the point that they are not detectable.

3.3 Evaluation of the performance of the suspect screening method according to SPE cartridge

Evaluation of the overall performance of suspect screening methods is important to determine how and to which extent results can be biased by experimental parameters. Therefore, it was decided to spike the model compounds in river water SPE extracts (i.e. after extraction) at different concentrations in order to estimate the “limits of identification” of the method (LOI), that is the lowest concentration of a given model compound that can be found in the samples by the MF algorithm. Such limits not only depend on ionization efficiency and instrumental parameters like ion transfer as discussed earlier but also on the matrix. For example, extraction of coeluting interferences of close *m/z* values as those as of the suspect TOCs may impair the proper identification of isotopic patterns that are necessary to match experimental data to theoretical isotopic distribution of the TOCs present in the database. Also, the matrix may cause signal suppression or enhancement effects that could mask or improve the detection of peaks of low intensity in the isotopic pattern.

Figure 6 shows the limits of identification, for the 38 compounds that could be separated by the LC method, corrected according to the SPE preconcentration factor and extraction recoveries discussed earlier. In general, the cartridge that allowed the largest percentage of identified compounds was RP (84%) followed by WAX (79%) and WCX (74%). Cartridges also differed in the concentrations at which those compounds could be identified: 58% of compounds could be detected at concentrations between about 6 and 100 ng L^{-1} in both the RP and WCX cartridges while WAX the percentage was slightly lower, 53%.

Only six model compounds could not be identified using any of the cartridges, four of them (DIA, ESZ, OXB and TDCPP) were false negatives, *i.e.* they could be observed in the extracted ion chromatograms obtained manually but could not be identified automatically by the MF algorithm; and two compounds (EE2 and 2,4-D) were not detected or extracted in the spiked river water samples. DIA, ESZ, OXB and TDCPP were probably not detected by the MF algorithm due to the combination of factors such as low signal intensity and matrix interferences that reduced the Fit and Reverse fit values below the threshold value of 940.

These results showed that suspect screening methods using RP, WCX or WAX cartridges are sensitive enough to detect a wide array of organic contaminants at environmental concentrations, but a significant number of false negatives is usually observed. Therefore, experimental data shows that a single generic method will bias the results in suspect screening methods towards compounds that can be ionized, separated and recovered adequately with the chosen method parameters.

While the samples extracted with the RP cartridge (Strata-X) posed no problem through the analysis, some issues were encountered with both ion exchangers. Indeed, during the analysis of WAX cartridge extracts of surface water samples, chromatograms showed detector saturation caused by unidentified compounds eluting at retention times between 2.5 and 2.9 minutes. These compounds are most likely hydrophilic humic substances that were extracted by anion exchange and are eluted early in reversed-phase liquid chromatography [37]. To preserve the instrument's detector lifetime, that part of the chromatogram was sent to the waste rather than to the QqTOFMS using a divert valve between the column and the ESI source. Unfortunately, since the retention time of ACT is 2.6 minutes, this compound could no longer detected using the WAX cartridge. Another gradient was not applied to solve this issue since the objective here was to compare the performance of the tested cartridges with the same method. Also, after the reconstitution step with the WCX cartridge, the sample showed turbidity, since some compounds could not be dissolved in the solution. Therefore, an additional filtering step was added to insure column longevity. While the identification of the compounds causing turbidity of the extracts was outside the scope of the present work, it is possible that humic substances with cationic groups are responsible for such result. It is known that surface water can contain from 1 to 5 mg L^{-1} of dissolved humic substances [37], therefore they are present at concentrations about 10^5 higher than median concentrations of most PPCPs [38]. These humic substances besides being composed of compounds having acid functions that participated in anion exchange with the WAX cartridges, are also composed of

aromatic amines and peptides [39] that could have been retained by cation exchange in the WCX cartridge. However, more studies are necessary to validate this hypothesis. Survey view chromatograms (Figure 7), *i,e.* plots showing *m/z* as a function of retention time and a color scale to indicate signal intensity, illustrate the effect of cartridge type on the detection of sample components. As it can be seen in Figure 7, WAX cartridges show a series of peaks between *m/z* 200 and 800 and retention times between 4 and 15 min that are much more intense and numerous than in RP and WCX cartridges (peaks inside the red ellipse in Figure 7). Those peaks had a difference of mass of 44 Da and most likely correspond to protonated or cationized molecules of ethylene glycol oligomers ($H[OCH_2CH_2]_nOH$, where n=5 to 17). The presence of such compounds in wastewater effluents has been reported previously [18] and is explained by their high-volume production worldwide and multiple uses in consumer products [40]. Low molecular weight PEGs are also biodegradation products of polyethoxylated surfactants [41,42]. Other closely related compounds (inside black rectangle on Figure 7) having mass differences of 22 and 15 Da were identified in the samples extracted with WAX, however they could not be identified. A detailed analysis of the chromatograms (Figure SM-1 and Table SM-1, Supplementary material) showed that average mass differences of 22.0130 Da were observed between double charged ions and when the single charged ions were observed the average difference was 44.0256 Da. Theses results indicate the presence of another type of ethylene glycol oligomers and supports the hypothesis of the closely related nature of the peaks highlighted in Figure 7.

Also, by the end of the chromatographic separation, at around 20 min, a high number of signals are observed. Their high retention time suggests that they have a highly hydrophobic nature. Such series of peaks are less intense in both WCX and RP cartridges. Therefore, these results indicate that while WAX cartridges can offer interesting selectivity for the extraction of organic acids in surface waters, in the experimental conditions used in the present study they can also co-extract many natural and synthetic compounds from the samples that may not be relevant, could saturate detector signal and also may interfere with the identification of suspect TOCs.

3.4 Application of the developed suspect screening method to the analysis of surface water samples

Non-spiked samples of the St-François river near Sherbrooke, QC (Canada) were extracted and analyzed using the three different cartridges. These samples were extracted with the three tested sorbents in order to compare the effect of the sorbent type on the number of suspect contaminants that could be successfully identified. The workflow of identification used in this study, based on the Schymanski levels of confidence[22] is shown in Figure 1.

Results showed that between 68 to 100 peaks with accurate masses that matched the formula of suspect compounds in the home-made database were identified by the MF algorithm in the samples within 7 mDa of mass error (Table 2, for more details consult the Excel file:

SupplementaryMaterial(Databases-Results).xlsx). However, some of these matches had to be disregarded since, in some cases the same exact masses (within a mass error of 7 mDa) were present in the field blanks at a higher signal intensity than in the samples or were present in only one of the two replicate samples. Therefore, a series of filters and diagnostic techniques were applied to assign the Schymanski identification confidence levels.

In order to assign level 5 (accurate mass) to the positive matches we used the following filters: i) *Peak present in all replicate samples and ii) Peak in samples must be at least 3× higher than in the field blanks.* Application of those two filters reduced the data to between 17 and 29 level 5 compounds.

Since the MF algorithm compares the experimental isotope pattern to the theoretical isotopic patterns of the suspect contaminants in the home-made database to assign a match to an observed ion, the identification level assigned to those compounds passed the requirements to be accepted as identification confidence level 4 (unequivocal formula) as well. However, since the MF algorithm tolerates differences of at least 5% between the theoretical and experimental isotopic pattern, in one case the formulas were not completely unequivocal. An ion of *m/z* 267.1727 eluted at 20.3 min an it was identified by the MF algorithm as atenolol ($C_{14}H_{22}N_2O_3$, 266.1630 Da) and tri-isobutyl phosphate/tributyl phosphate ($C_{12}H_{27}O_4P$, 266.1647 Da) in the samples extracted with the WAX cartridges. Such result can be explained by the close monoisotopic mass values (< 2 mDa) and also similar isotopic patterns (about 3% difference between the relative abundance of the M+1 peaks and less than 0.2% for the M+2 peaks).

In order to assign an unambiguous molecular formula to *m/z* 267.1727, a more powerful technique for the unambiguous assignment of molecular formulas to accurate masses, spectral accuracy [23,24], was applied. Analysis of the experimental isotopic pattern showed that the neutral formula $C_{14}H_{22}N_2O_3$, corresponding to atenolol, was ranked first with an spectral accuracy of 29%. While this value is low compared to the threshold value of 98% spectral accuracy, commonly accepted as high similarity between experimental and theoretical isotopic pattern [24], no other molecular formula had a higher similarity within the spectral accuracy determination parameters (for details, consult Supplementary Material).

Additionally, it was observed that several compounds had multiple peaks identified in the chromatograms such as methylparaben which was identified at different retention times. Therefore, while it is possible that other formulas are possible for the suspect contaminants identified by the MF algorithm, the peaks that were assigned identification level 4 remain potential candidates that must be confirmed using other diagnostic techniques.

One of those diagnostic techniques is the comparison of experimental and predicted retention times. Since the objective of the identification workflow was to attain level 2b (probable structure) it was also decided to only estimate the retention time for compounds with signals higher than 1000 counts. This decision was based on the average minimum signal that is necessary to obtain meaningful tandem mass spectra in the LC-QqTOFMS system employed. Experiments showed that about half of the peaks with unequivocal molecular formulas could be eliminated by using both retention time estimation and a signal threshold. Therefore, between 9 and 16 tentative candidates were identified in the samples. Estimation of retention time was very helpful to eliminate false positives since, in some cases (e.g. butylparaben, ethylparaben, triethyl phosphate) the same compound had up to three different identified peaks in the chromatograms. In a few cases, the estimation of retention time could not eliminate isomers with level 4 identification level. For example given the wide window of acceptable difference in retention time (<2.9 min), tramadol (difference between observed and predicted retention time=0.96 min) and O-desmethylvenlafaxine (2.46 min) could be both tentative compounds. Nevertheless, estimation of retention time is a valuable tool for suspect screening and the prediction power of the approach can be improved by building larger and more diverse knowledge base of structures and retention times [25]. For this study a knowledge base of only 215 compounds was built. Better performance was observed with knowledge bases of more than 400 compounds [25].

The following step was to compare experimental and library tandem mass spectra in order to assign probable structures to the peaks identified as tentative candidates. Unfortunately for some of the tentative candidates, MS^2 experiments were not possible; thus, a higher level of identification could not be assigned.

Results of the compounds with level 2a identification level are shown in Table 3. In total, five compounds were identified as probable structures. Figure 8 shows the MS^2 spectra of some of these compounds. As the level 2a compounds were available in the laboratory, a final step of final structural confirmation was performed. The five level 2a compounds (gabapentin, caffeine, N,N-diethyl-m-toluamide, metoprolol and venlafaxine) were assigned a level 1 identification since retention times and MS^2 spectra matched. For more details consult the Excel file: SupplementaryMaterial(Databases-Results).xlsx) and the Supplementary material (Figures SM-6 to SM-10).

4. Conclusion

In the last few years important efforts have been done to improve identification workflows of HRMS data to reduce false positives and enhance the confidence of the results reported. However, false negatives remain an important pitfall of suspect screening analysis that must be tackled to allow proper evaluation of the presence of TOCs of interest in environmental samples.

A suspect screening method for TOCs in surface waters was developed and its performance was systematically evaluated. A set of 39 diverse model compounds was used to identify method steps responsible for possible bias in the identification of TOCs. The results demonstrated that sample preparation, chromatographic separation, ionization and ion transmission reduce the amount of TOCs than can be detected by suspect screening analysis and generate an important number of false negatives.

Sample preparation and chromatography remain critical steps in the methodology that must be bettered in order identify a larger number of suspect compounds. The results obtained in the present study suggest that the use of SPE cartridges combining different sorbent chemistries [12,18] are, at this moment, the best compromise in terms of cleaner extracts and higher number of extracts compounds. Nevertheless, to ensure proper detection of thousands of compounds at low concentrations, improvements in chromatography are urgently needed. As it has been observed in metabolomics and other fields based on the analysis of complex mixtures, traditional techniques cannot identify all the analytes of interest in a sample. According to Stoll et al. [43] The percentage of compounds observed in a chromatogram with a minimum resolution of 1 using modern LC columns and instruments is < 1% for samples containing more than 500 compounds. This means that in complex environmental samples most peaks coelute with many other compounds. Enhanced resolution of thousands of components present in complex samples can be achieved by two-dimensional liquid chromatography (2D-LC) because of its higher peak capacity, *i.e.* the maximum number of peaks that can be adequately separated during an experiment, compared to one-dimensional LC. Application of 2D-LC to the analysis of natural products [43] or food [44] have shown that this technique can improve the detection of sample components by reducing signal suppression caused by coelution and it also allows to resolve isomers. An specially interesting approach to solve the limitations of reversed-phase liquid chromatography for suspect screening is to use HILIC and reversed-phase columns in a 2D-LC setup, which has been applied to the separation of natural oxidants [45].

Regarding ionization, it is known that chemical ionization (CI) or electron ionization (EI) sources used in gas chromatography are generally able to ionize compounds of low polarity or low mass more efficiently than electrospray [46]. Gas chromatography coupled to HRMS (GC-HRMS) is an orthogonal technique to LC-HRMS that is increasingly commercially available [47]. The complementarity between LC-HRMS and GC-HRMS has been demonstrated by recent studies that showed an increased overall number of contaminants detected compared to a single technique [47,48]. For example, Fernandez *et al.* used a GC/LC-QqTOFMS instrument to detect both apolar and polar TOCs such as PAHs, musks, pesticides and pharmaceuticals in surface waters [47]. Thus, addition of GC-HRMS to suspect screening workflows widens the range of compounds that can be successfully identified in a given sample and compensates for bias caused by electrospray ionization.

Finally, bias in the detection of ions due to their *m/z* is an intrinsic limitation of design of the QqTOFMS used in the experiments and in general of TOF mass analysers with orthogonal acceleration [46]. It has been reported that Orbitrap mass analyzers do not suffer from such discrimination of low and high mass ions [49] and offer higher mass resolution and mass accuracy compared to TOF mass analysers. Nevertheless, issues with relative ion abundances and discrimination of low intensity mass signals in Orbitraps have been observed [50] and could be problematic for the identification of suspect TOCs based in their isotopic pattern.

In summary, progress in analytical instrumentation in the last years such as two-dimensional liquid chromatography (2D-LC) [43], addition of GC-MS to suspect screening workflows and improvements in ion transmission might in mass spectrometers might make possible the application of only a few methods to unravel the presence of TOCs in the environment in the near future.

5. Acknowledgements

This study was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery grant awarded to P.A.S. We would like to thank René Gagnon and Philippe Venne for their support with the LC-QqTOFMS. Funding sources did not have any involvement in the design, experiments, data interpretation, writing, revision, or submission of this study.

6. Author Contributions

P A. S. conceived and designed the experiments; M. R. carried out most of the experiments and data analysis; A. G. performed experiments to determine limits of identification; A.-M. G. developed and optimized the method for the estimation of retention times; D. R. and A. G. built the home-base database. P. A. S, M. R., and A. G. wrote the paper.

7. References

(1) Keith, L. H. In *Advances in the Identification and Analysis of Organic Pollutants in Water*; Keith, L. H., Ed.; Ann Arbor Science Publishers, Inc.: Ann Arbor, MI, 1981; Vol. 1, p 1190.
(2) Hughes, S. R.; Kay, P.; Brown, L. E. *Environ. Sci. Technol.* **2012**, *47*, 661.
(3) Barceló, D.; Petrović, M. *TrAC, Trends Anal. Chem.* **2007**, *26*, 2.

(4) Lajeunesse, A.; Gagnon, C. *Int. J. Environ. Anal. Chem.* **2007**, *87*, 565.
(5) Patrolecco, L.; Ademollo, N.; Grenni, P.; Tolomei, A.; Caracciolo, A. B.; Capri, S. *Microchem J* **2013**, *107*, 165.
(6) Daughton, C. G. *Sci. Total Environ.* **2014**, *466*, 315.
(7) Krauss, M.; Singer, H.; Hollender, J. *Anal. Bioanal. Chem.* **2010**, *397*, 943.
(8) Hug, C.; Ulrich, N.; Schulze, T.; Brack, W.; Krauss, M. *Environ. Poll.* **2014**, *184*, 25.
(9) ter Laak, T. L.; Puijker, L. M.; van Leerdam, J. A.; Raat, K. J.; Kolkman, A.; de Voogt, P.; van Wezel, A. P. *Sci. Total Environ.* **2012**, *427*, 308.
(10) Segura, P. A.; MacLeod, S. L.; Lemoine, P.; Sauvé, S.; Gagnon, C. *Chemosphere* **2011**, *84*, 1085.
(11) Escher, B. I.; van Daele, C.; Dutt, M.; Tang, J. Y.; Altenburger, R. *Environ. Sci. Technol.* **2013**, *47*, 7002.
(12) Moschet, C.; Piazzoli, A.; Singer, H.; Hollender, J. *Anal. Chem.* **2013**, *85*, 10312.
(13) Schymanski, E. L.; Singer, H. P.; Longrée, P.; Loos, M.; Ruff, M.; Stravs, M. A.; Ripollés Vidal, C.; Hollender, J. *Environ. Sci. Technol.* **2014**, *48*, 1811.
(14) Sjerps, R. M.; Vughs, D.; van Leerdam, J. A.; ter Laak, T. L.; van Wezel, A. P. *Water Res.* **2016**, *93*, 254.
(15) Wode, F.; van Baar, P.; Dünnbier, U.; Hecht, F.; Taute, T.; Jekel, M.; Reemtsma, T. *Water Res.* **2015**, *69*, 274.
(16) Zedda, M.; Zwiener, C. *Anal. Bioanal. Chem.* **2012**, *403*, 2493.
(17) Díaz, R.; Ibáñez, M.; Sancho, J. V.; Hernandez, F. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 355.
(18) Gago-Ferrero, P.; Schymanski, E. L.; Bletsou, A. A.; Aalizadeh, R.; Hollender, J.; Thomaidis, N. S. *Environ. Sci. Technol.* **2015**, *49*, 12333.
(19) Bade, R.; Bijlsma, L.; Miller, T. H.; Barron, L. P.; Sancho, J. V.; Hernández, F. *Sci. Total Environ.* **2015**, *538*, 934.
(20) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2001**, *20*, 362.
(21) Liigand, P.; Kaupmees, K.; Haav, K.; Liigand, J.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. *Anal. Chem.* **2017**, *89*, 5665.
(22) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. *Environ. Sci. Technol.* **2014**, *48*, 2097.
(23) Wang, Y.; Gu, M. *Anal. Chem.* **2010**, *82*, 7055.
(24) Eysseric, E.; Barry, K.; Beaudry, F.; Houde, M.; Gagnon, C.; Segura, P. A. *Anal. Chem.* **2017**, *89*, 9805.
(25) Tyrkkö, E.; Pelander, A.; Ojanperä, I. *Anal. Chim. Acta* **2012**, *720*, 142.
(26) mzCloud 2018.
(27) Zhou, S.; Hamburger, M. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1516.
(28) Raffaelli, A.; Bruins, A. P. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 269.
(29) Henriksen, T.; Juhler, R. K.; Svensmark, B.; Cech, N. B. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 446.
(30) Mallet, C. R.; Lu, Z.; Mazzeo, J. R. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 49.
(31) Gao, S.; Zhang, Z.-P.; Karnes, H. *J. Chromatogr. B* **2005**, *825*, 98.
(32) Bruker Daltonics Bremen, Germany, 2008; Vol. Bruker Daltonik GmbH.
(33) Klitgaard, A.; Iversen, A.; Andersen, M. R.; Larsen, T. O.; Frisvad, J. C.; Nielsen, K. F. *Anal. Bioanal. Chem.* **2014**, *406*, 1933.
(34) Pichon, V. *J. Chromatogr. A* **2000**, *885*, 195.

(35) Vulliet, E.; Wiest, L.; Baudot, R.; Grenier-Loustalot, M.-F. *J. Chromatogr. A* **2008**, *1210*, 84.
(36) Biesaga, M.; Jankowska, A.; Pyrzyńska, K. *Microchimica Acta* **2005**, *150*, 317.
(37) Ferrer, I.; Barcelo, D.; Thurman, E. *Anal. Chem.* **1999**, *71*, 1009.
(38) Hughes, S. R.; Kay, P.; Brown, L. E. *Environ. Sci. Technol.* **2013**, *47*, 661.
(39) Leenheer, J.; Croué, J. *Environ. Sci. Technol.* **2003**, *37*, 18A.
(40) Bernhard, M.; Eubeler, J. P.; Zok, S.; Knepper, T. P. *Water Res.* **2008**, *42*, 4791.
(41) Marcomini, A.; Zanette, M.; Pojana, G.; Suter, M. J. F. *Environ. Toxicol. Chem.* **2000**, *19*, 549.
(42) Eubeler, J. P.; Bernhard, M.; Knepper, T. P. *TrAC Trends in Analytical Chemistry* **2010**, *29*, 84.
(43) Stoll, D. R.; Carr, P. W. *Anal. Chem.* **2017**, *89*, 519.
(44) Cacciola, F.; Dugo, P.; Mondello, L. *TrAC Trends in Analytical Chemistry* **2017**, *96*, 116.
(45) Jandera, P.; Hájek, T.; Staňková, M.; Vyňuchalová, K.; Česla, P. *J. Chromatogr. A* **2012**, *1268*, 91.
(46) Gross, J. H. *Mass Spectrometry: A Textbook*; 2nd ed.; Springer: Berlin, Germany, 2011.
(47) Hernández, F.; Ibáñez, M.; Portolés, T.; Cervera, M. I.; Sancho, J. V.; López, F. J. *J. Hazard. Mater.* **2015**, *282*, 86.
(48) Moschet, C.; Lew, B. M.; Hasenbein, S.; Anumol, T.; Young, T. M. *Environ. Sci. Technol.* **2017**, *51*, 1553.
(49) Perry, R. H.; Cooks, R. G.; Noll, R. J. *Mass Spectrom. Rev.* **2008**, *27*, 661.
(50) Kaufmann, A.; Walker, S. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 1081.

Figures

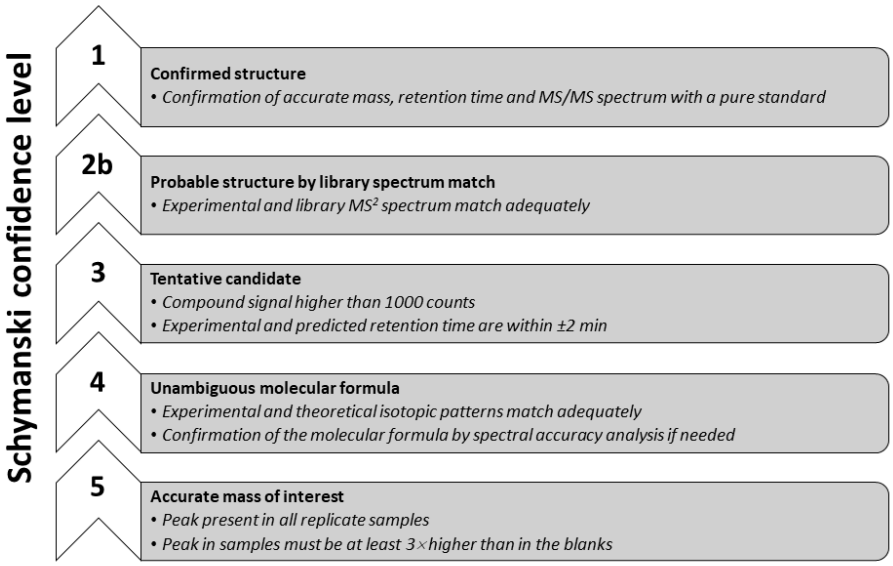

Figure 1. Suspect contaminant identification workflow based on the Schymanski diagram.

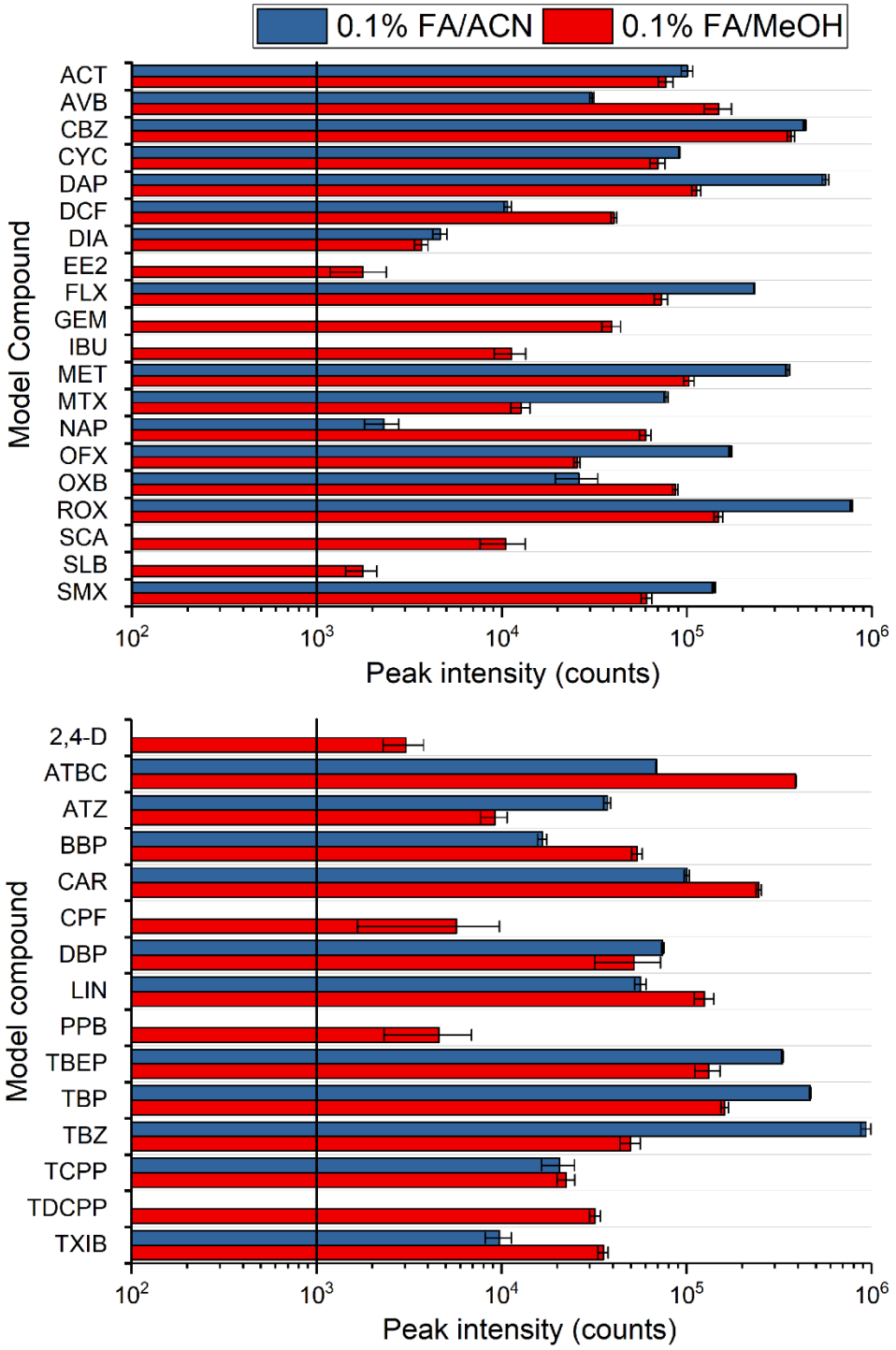

Figure 2. Average intensity of model compounds (top: PPCPs, bottom: pesticides and additives) spiked at 100 µg L^{-1} in H_2O: MeOH 1:1 and analyzed by LC-QqQTMS using ESI+ and a mobile phase composed of an aqueous solvent (solvent A: 0.1% FA in H_2O) and an organic solvent (solvent B: 0.1% FA in ACN or MeOH). Length of error bars indicate ± 1 standard deviation of three consecutive injections. Vertical bar at 10^3 counts indicates threshold intensity for adequate MS^2 spectra. ESZ, GLY, OTSA and SUC were not available when these experiments were performed.

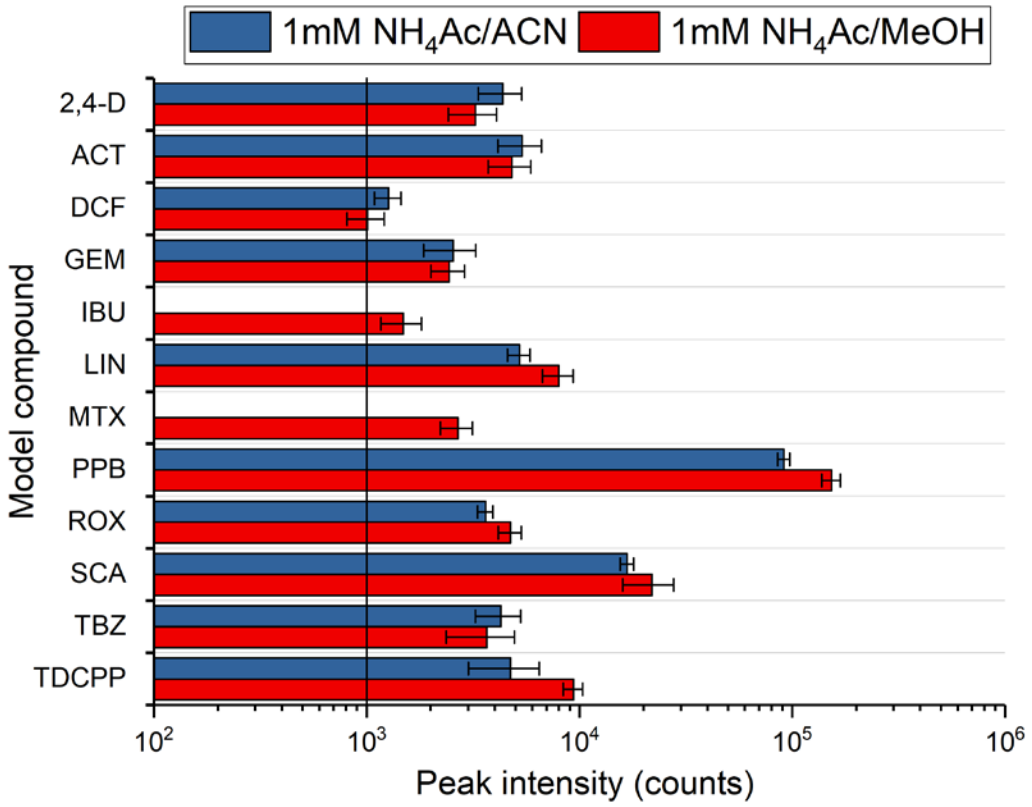

Figure 3. Average intensity of model compounds spiked at 100 µg/L in H2O: MeOH 1:1 and analyzed by LC-QqQTMS using ESI- and a mobile phase composed of an aqueous solvent (solvent A: 0.1% FA in H2O) and an organic solvent (solvent B: 1 mM NH4Ac in ACN or MeOH). Length of error bars indicate ± 1 standard deviation of three consecutive injections. Vertical bar at 103 counts indicates threshold intensity for adequate MS/MS spectra. ESZ, GLY, OTSA and SUC were not available when these experiments were performed.

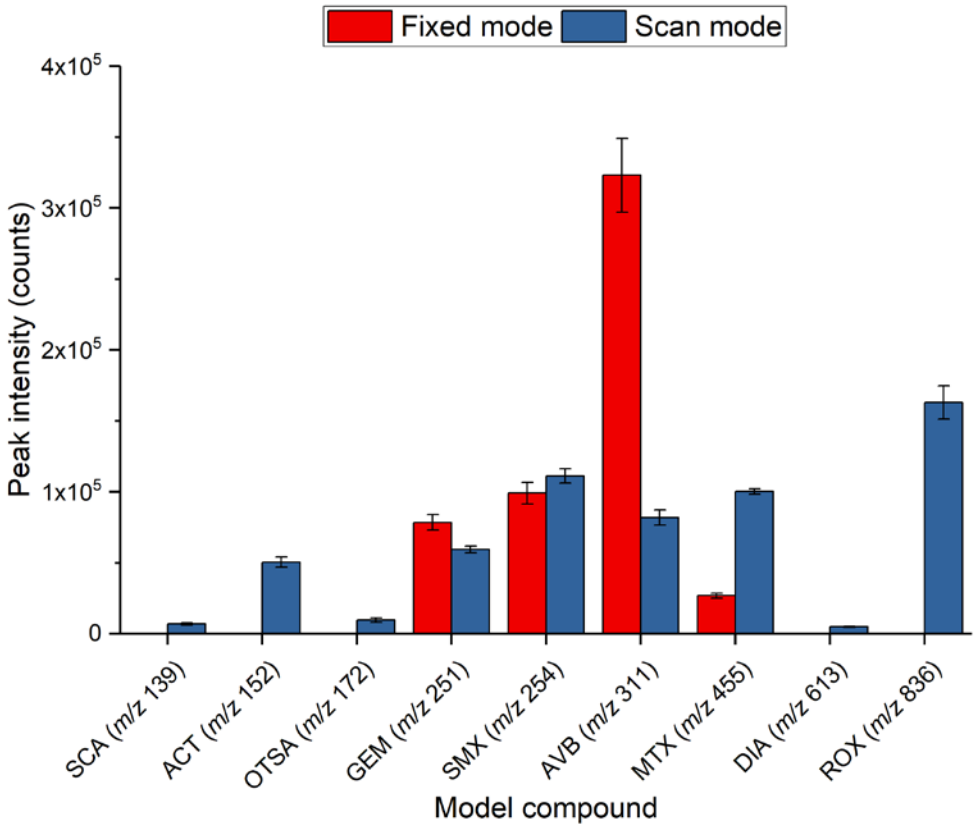

Figure 4. Peak intensity of model compounds in the fixed and scan modes of ion cooler and transfer time of the QqTOFMS.

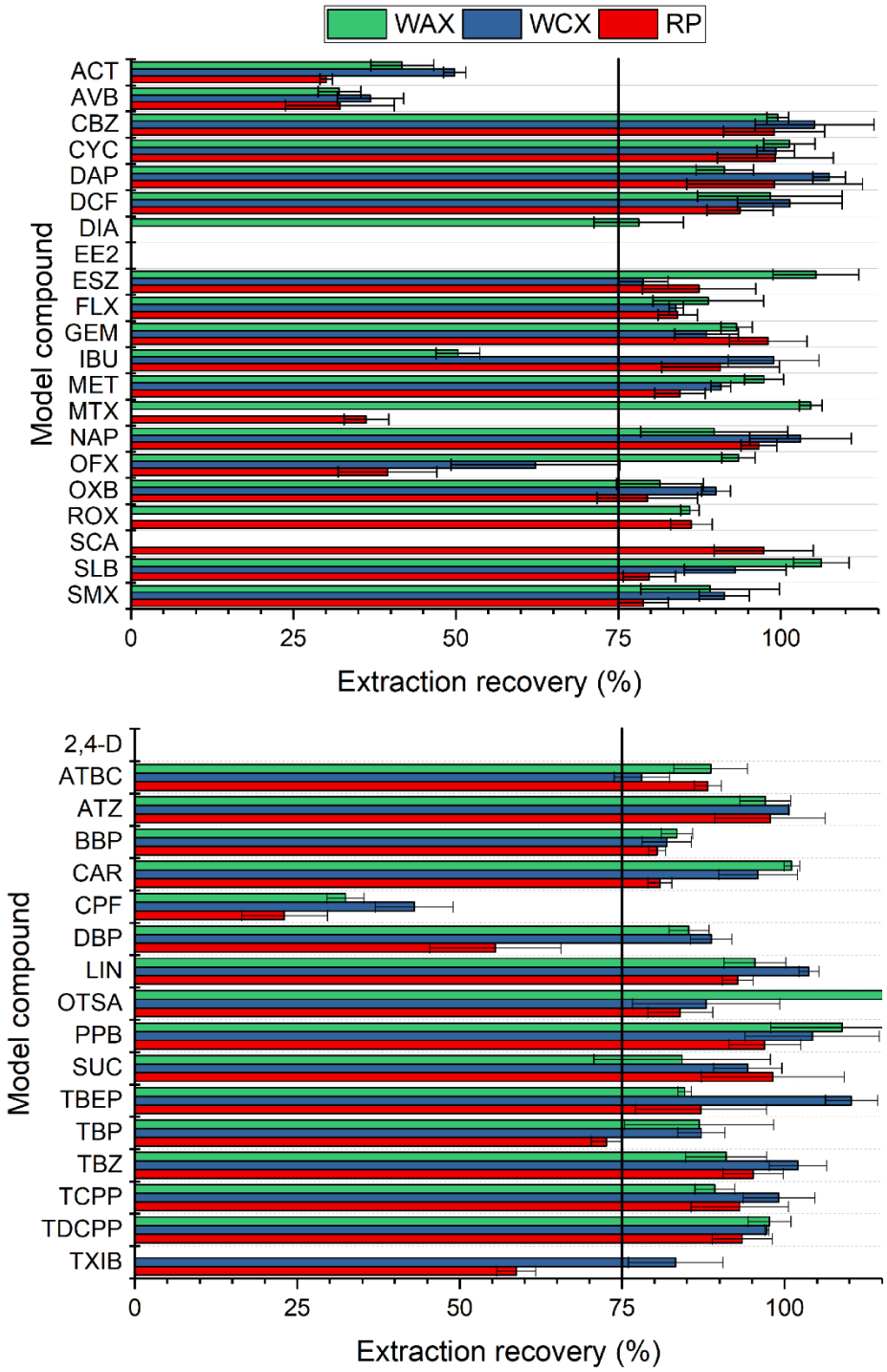

Figure 5. Extraction recoveries of the PPCPs (top), pesticides and consumer product additives (bottom) used as model compounds. RP: reversed-phase with polar groups SPE cartridge. WCX: mixed-mode weak cation exchange cartridge, WAX: mixed-mode weak anion exchange cartridge. Length of error bars indicate ± 1 standard deviation of three replicates. Vertical bars indicate the arbitrary value for "acceptable" recovery (>75%).

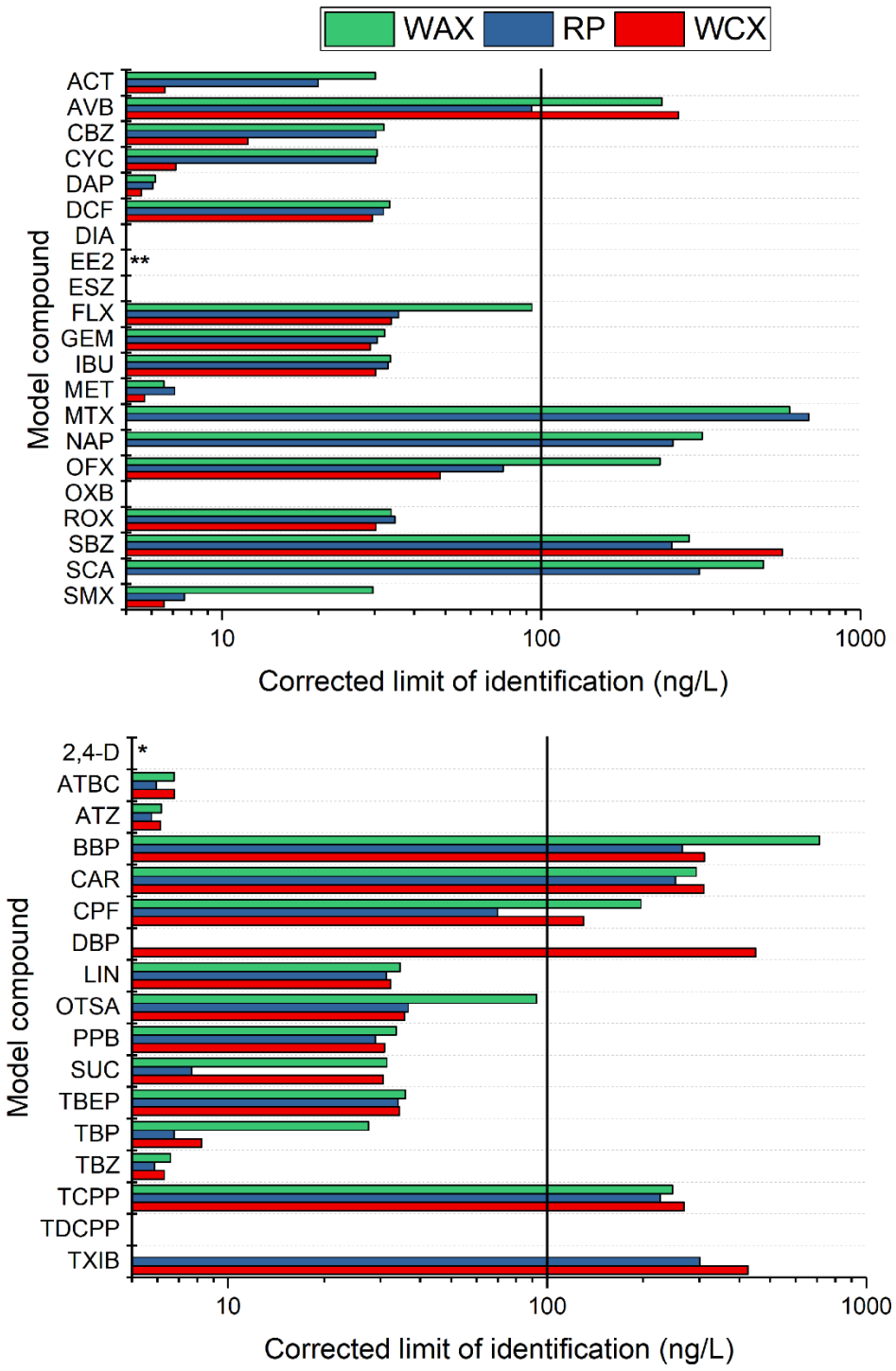

Figure 6. Limits of identification (LOI) for PPCPs (top) and pesticides and consumer product additives (bottom) in spiked surface water samples. Values were corrected for recovery. The single asterisk (*) indicates that 2,4-D could not be recovered by any cartridge and double asterisk (**) indicates that EE2 was not detected by the instrument. The other compound with nil values (DIA, ESZ, OXB and TDCPP) were detected by the instrument, but not identified by the MF algorithm. Vertical bars indicate the target value for acceptable LOI of TOCs in surface waters (<100 ng L^{-1}).

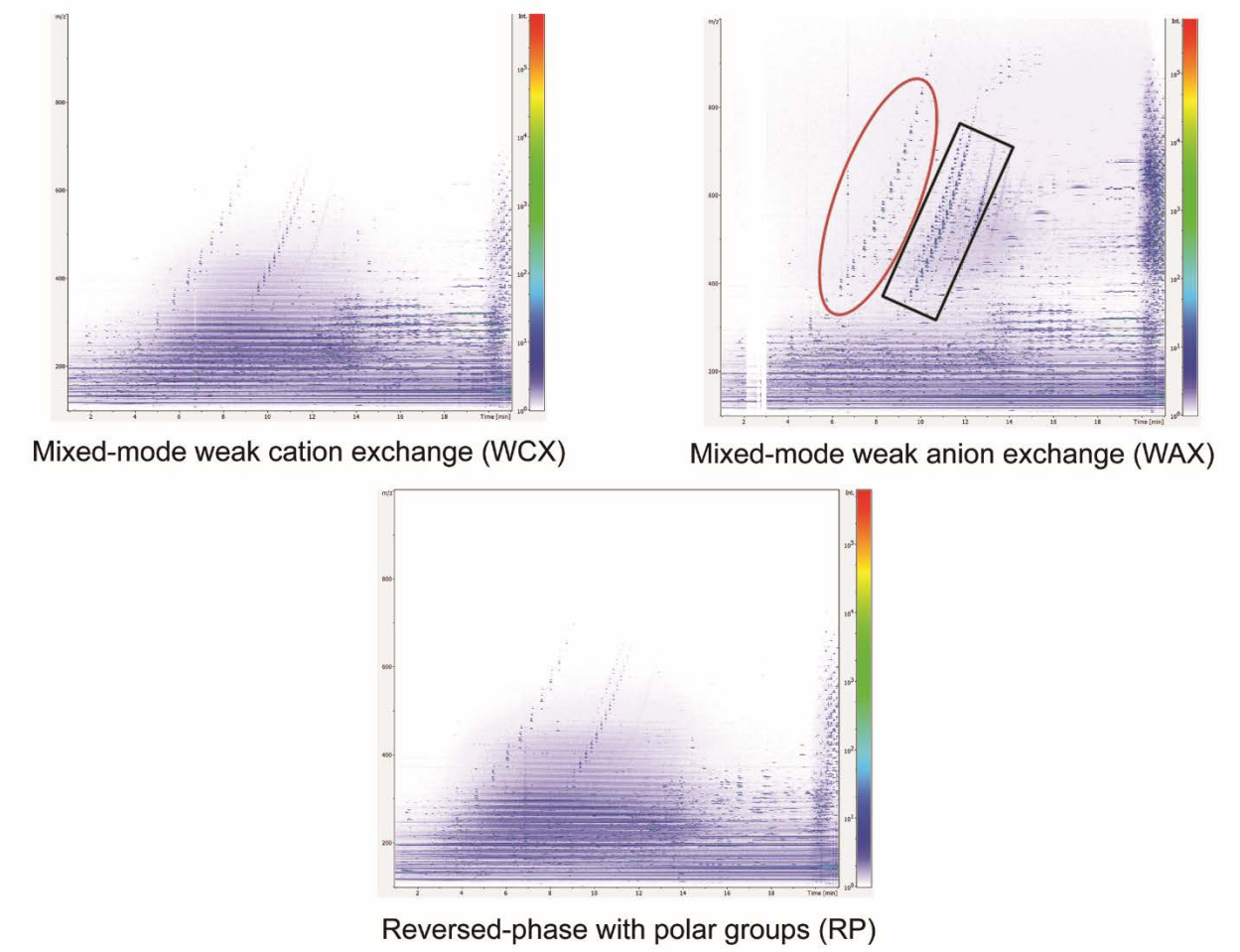

Figure 7. Survey view chromatograms of non-spiked river water samples extracted with three different SPE cartridges and analysed by LC-QqTOFMS. For chromatogram, ion *m/z* is presented in the ordinate (*m/z* 100 to 1000) and retention time in the abscissa (1 to 21 min). Intensity is given by the color scale at the right (10^0 to 10^6 counts). For the WAX survey view chromatogram, compounds inside the red ellipse correspond to ethylene glycol oligomers (n=5 to 17). Compounds inside the black rectangle are separated by 44 and 15 Da and may be related to ethylene glycol oligomers.

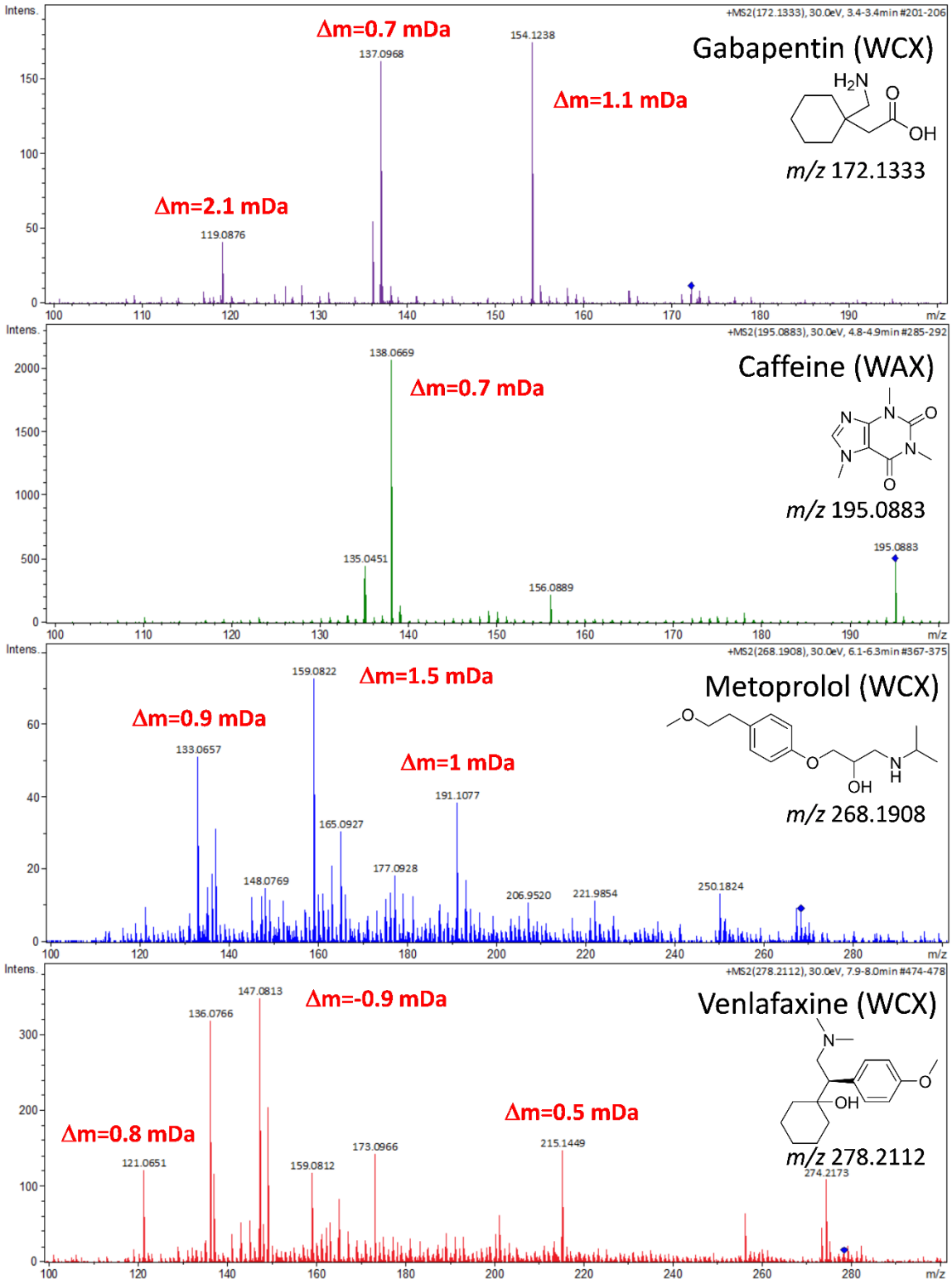

Figure 8. Tandem mass spectra of four tentative candidates. Δm indicates the mass difference in mDa between the experimental and library accurate masses.