Available online at www.sciencedirect.com

ScienceDirect

www.journals.elsevier.com/jes

JES JOURNAL OF ENVIRONMENTAL SCIENCES

www.jesc.ac.cn

Nontargeted identification of peptides and disinfection byproducts in water

Yanan Tang[1], *Ying Xu*[2], *Feng Li*[1], *Lindsay Jmaiff*[1], *Steve E. Hrudey*[1], *Xing-Fang Li*[1,*]

1. *Division of Analytical and Environmental Toxicology, Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, AB T6G 2G3, Canada. E-mail: yanan3@ualberta.ca*
2. *Department of Computer Science, University of Alberta, Edmonton, AB T6G 2E8, Canada*

ARTICLE INFO

Article history:
Received 9 July 2015
Revised 25 August 2015
Accepted 27 August 2015
Available online 29 October 2015

Keywords:
Nontargeted detection
Water analysis
Mass spectrometry
Peptide
Disinfection byproducts

ABSTRACT

A broad range of organic compounds are known to exist in drinking water sources and serve as precursors of disinfection byproducts (DBPs). Epidemiological findings of an association of increased risk of bladder cancer with the consumption of chlorinated water has resulted in health concerns about DBPs. Peptides are thought to be an important category of DBP precursors in water. However, little is known about the actual presence of peptides and their DBPs in drinking water because of their high sample complexity and low concentrations. To address this challenge and identify peptides and non-chlorinated/chlorinated peptide DBPs from large sets of organic compounds in water, we developed a novel high throughput analysis strategy, which integrated multiple solid phase extraction (SPE), high performance liquid chromatography (HPLC) separation, and non-target identification using precursor ion exclusion (PIE) high resolution mass spectrometry (MS). After MS analysis, structures of candidate compounds, particularly peptides, were obtained by searching against the Human Metabolome Database (HMDB). Using this strategy, we successfully detected 625 peptides (out of 17,205 putative compounds) and 617 peptides (out of 13,297) respectively in source and finished water samples. The source and finished water samples had 501 peptides and amino acids in common. The remaining 116 peptides and amino acids were unique to the finished water. From a subset of 30 putative compounds for which standards were available, 25 were confirmed using HPLC-MS analysis. By analyzing the peptides identified in source and finished water, we successfully confirmed three disinfection reaction pathways that convert peptides into toxic DBPs.

Introduction

Drinking water safety is an ongoing and ever challenging issue for the water industry and government agencies. Disinfection plays a key role to prevent waterborne diseases and ensure water safety. During disinfection processes, thousands of disinfection byproducts (DBPs) are produced through reactions between various organic compounds in source water and disinfectants (*e.g.*, chlorine and monochloramine). Disinfection poses a health risk tradeoff because many organics in water and their DBPs have toxic properties, and some are animal carcinogens (*e.g.*, nitrosamines) (Komaki et al., 2014; Li et al., 2015; Richardson et al., 2007). Epidemiological studies have shown an association between consumption of chlorinated water and

* *Corresponding author*. E-mail: xingfang.li@ualberta.ca (Xing-Fang Li).

http://dx.doi.org/10.1016/j.jes.2015.08.007

increased risk of bladder cancer (Villanueva et al., 2004). Peptides represent a large category of organics in water; they are commonly present in water as degradation products of biomaterials (Hayes et al., 2008; Spoof et al., 2003). A few are known to be toxic and can cause serious disease in humans and other species. For instance, microcystins are cyclic heptapeptides that are hepatotoxic and may be tumor promoters (Humpage et al., 2000; Ito et al., 1997). Besides their bioactivities and potential impacts on human health, peptides are important organic precursors of DBPs (Brosillon et al., 2009), as they can be modified during water disinfection to produce non-chlorinated/chlorinated peptide DBPs. Peptide DBPs could be biologically significant because it is known that some chloro-amino acids are mutagenic in bacterial assays and induce micronuclei in mammalian cells (Garland et al., 2009; Grant, 2009; Tavtigian et al., 2008). Chloro-amino acids can be produced by the reaction of aqueous chlorine with amino acids in laboratory experiments. Despite their ubiquitous presence and the specific attention on a number of toxic peptides, there has been no systematic study on the identification and characterization of this large group of potential DBP precursors. This is largely due to the lack of sensitive and high throughput analytical techniques (Hernandez et al., 2012; Lauby-Secretan et al., 2015).

To develop a novel analytical technique to systematically analyze peptides and non-chlorinated/chlorinated peptide DBPs in water, it is critical to address two specific challenges. First, peptides in water usually occur at extremely low concentrations, often from low parts-per-trillion (ppt, ng/L) to parts-per-billion (ppb, μg/L) (Ueno et al., 1996). Second, because of the diversity of peptides, the potential formation of their DBPs is complex. To address these challenges, we developed a high throughput liquid chromatography tandem mass spectrometry (LC-MS/MS) strategy (Fig. 1), which applied a non-target analysis technique to collectively identify peptides and their resulting DBPs in water samples. Specifically, the strategy consists of multiple solid phase extraction (SPE) of complementary retention mechanisms for sample enrichment and high performance liquid chromatography (HPLC) separation with non-target identification provided by high resolution mass spectrometry (MS). A database search against the Human Metabolome Database (HMDB) and bioinformatics analysis was used to identify possible peptides, amino acids, and non-chlorinated/chlorinated peptide DBPs in drinking water.

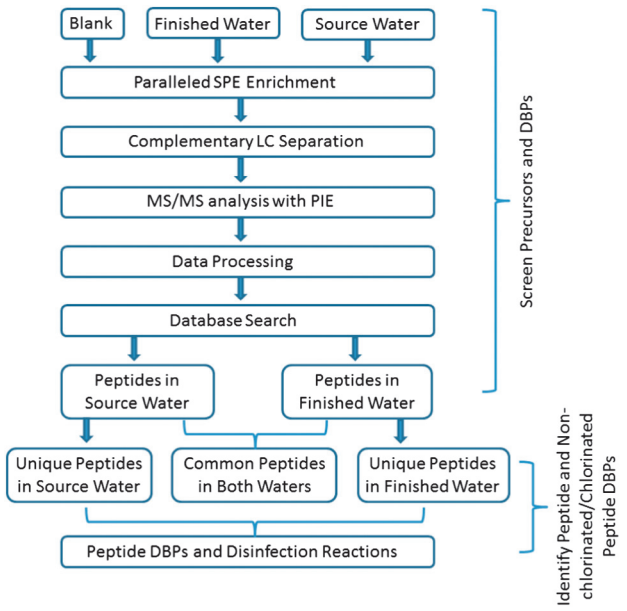

Fig. 1 – Experimental workflow of water sample analysis with the liquid chromatography tandem mass spectrometry (LC-MS/MS) strategy.

1. Materials and methods

1.1. Chemicals and reagents

All chemicals and reagents were purchased from Sigma-Aldrich (Oakville, ON, CA) unless stated otherwise. Methanol (MeOH), acetonitrile (ACN), and LC-MS grade water were purchased from Fisher Scientific Canada (Edmonton, AB). S-(1,2,dichlorovinyl)-glutathione was purchased from CacheSyn Inc (Mississauga, ON, CA); Asp-Pro, His-Pro, N-nitrosoproline, and aflatoxin B1 were purchased from Cedarlane (Burlington, ON, CA). 1,1′-Ethylidenebistryptophan was ordered from LGC Standards (Teddington, UK); bradykinin, leukotriene F4, and pentosidine were purchased from Cayman Chemicals (Ann Arbor, MI, USA).

1.2. Biosafety

The researchers involved in this study were trained in chemical safety and biosafety. All reagents were handled in the fume hood. Experiments were performed in a Biosafety Level 2 (BSL-2) laboratory. The waste materials were disposed of according to the biosafety and chemical safety procedures.

1.3. Samples and controls

During water treatment, the source water (plant influent, prior to any water treatment process) goes through the following steps: alum coagulation, sedimentation, sand and activated carbon filtration, and chlorine/chloramine/UV disinfection, before being distributed as finished water (plant effluent, after drinking water treatment). Samples of source water and finished water were collected from a water treatment plant. Throughout this study, LC-MS grade water was used as the negative control, and LC-MS grade water spiked with mixed peptide standards was used as the positive control. In each set of experiments, the samples and the controls were analyzed in parallel.

1.4. SPE

An Oasis HLB (Waters, Mississauga, ON, CA), a Bond Elut ENV and a Bond Elut C18 (Agilent, Mississauga, ON, CA) cartridge were used to extract organic compounds from water samples. The SPE cartridges were conditioned before use with MeOH and H_2O. The extraction process of water samples with SPE cartridges was performed according to the manufacturers' instructions. The details of SPE extraction and sample reconstruction are described in the Appendix A: Supporting Information.

1.5. Liquid chromatography separation

The compounds enriched using SPE extractions were separated on an Agilent 1290 Infinity Binary LC System (Agilent, Mississauga, ON, CA) before MS analysis. Two columns, AdvanceBio Peptide Mapping C18 column (150 × 2.1 mm, 2.7 μm; Agilent, Mississauga, ON, CA) and TSKgel Amide-80 hydrophilic interaction liquid chromatography (HILIC) column (250 × 2 mm, 5 μm; Tosoh Bioscience, King of Prussia, PA, USA), were used for separation. The flow rate of LC separation was 100 μL/min. LC gradients were optimized respectively for each column. The details of the optimized LC gradient for each column are described in the Appendix A: Supporting Information.

1.6. Mass spectrometry analysis

The organic compounds present in water samples, separated on C18 and HILIC LC systems, were analyzed using a TripleTOF® 5600 system (AB Sciex, Concord, ON, CA). All compounds were analyzed in positive ionization mode. The MS analysis conditions are illustrated in the Appendix A: Supporting Information. The precursor ion exclusion (PIE) strategy was applied as reported to eliminate redundant identifications (Wang and Li, 2008).

1.7. Data processing

A Java program was coded to help process the generated LC-MS/MS data. The *m*/*z* value of each precursor ion was extracted. A mass tolerance of ±0.005 Da was applied when grouping the precursor ions by their charge states and retention times with a window of ±30 sec. The extracted accurate masses were searched in the HMDB, which contains chemical structures, clinical characteristics, and molecular biology information for 41,993 organic compounds, including proteins and peptides, toxins, environmental pollutants, drugs and drug metabolites (Wishart et al., 2013). The search results were exported to an Excel file, providing information that included the HMDB ID, compound name, theoretical mass, adduct, experimental mass, and derivative for each query.

1.8. Identification validation and bioinformatic analysis

The measured accurate masses of the precursor ions were searched against HMDB to identify putative compounds. When the measured tandem mass spectrometry (MS/MS) spectra of the precursor ions matched those of the theoretical fragments of the putative compounds, these compounds were considered to be detected. The detected putative compounds were further confirmed with available standards using the same LC-MS analysis conditions as the water sample analysis.

Five peptide standards, Gly–Tyr, Val–Tyr–Val, Leu Enkephalin, Met Enkephalin, and Angiotensin II, were repeatedly analyzed using the LC-MS method and their retention times were evaluated for inter-day variations. The data are included in the Appendix A: Supporting Information. In these experiments, we noticed that the hydrophilic peptides, such as Val–Tyr–Val, had retention time shifts up to 100 sec, which is due to partial retention of hydrophilic peptides on the C18 reverse phase HPLC column. Similarly, the hydrophobic peptides also experienced larger retention time shifts on the HILIC column because of poor retention of hydrophobic peptides on the HILIC column. This observation was previously reported (Schmidt et al., 2008). Therefore, the retention time window for validation of detection of putative compounds was set to be 60 sec (Gika et al., 2007; Schmidt et al., 2008), and the *m*/*z* tolerance window was 0.005 Da. A peptide was considered as having been detected only when all three criteria were satisfied: its retention time was within 60 sec, its accurate mass had less than 0.005 Da difference from that of the standard at the same charge state as the standard. Identified peptides and compounds were searched in the literature and databases to evaluate their bioactivities for toxicological relevance.

2. Results and discussion

Fig. 1 illustrates the overall workflow of our strategy. It consisted of comparative analysis of blank, source water, and finished water using multiple SPEs for sample preconcentration, HPLC separation, comprehensive advanced mass spectrometry analysis, and a relevant database search. The combination of these techniques confronts the challenges of trace analysis of complex peptides and their DBPs. Manual interpretation of MS/MS spectra and validation with chemical standards were performed to confirm the existence of putative identifications in water.

2.1. Water sample preparation

SPE enrichment is a critical step to concentrate analytes present in water because of sample complexity and trace concentrations (often at ng/L). We applied three commonly used SPE cartridges (Oasis HLB, Bond Elut ENV, and Bond Elut C18) of complementary retention mechanisms in parallel, to achieve efficient enrichment of a variety of organics in water. The Oasis HLB cartridge is composed of a hydrophilic–lipophilic balanced polymer and can retain polar and nonpolar organic compounds. The Bond Elut C18 cartridge is packed with a C18-coated silica reverse stationary phase and can retain mainly hydrophobic compounds, while the Bond Elut ENV cartridge is packed with modified styrene-divinylbenzene polymers and can retain hydrophilic and polar organic compounds. Parallel SPEs provide fractionation of different organics that can enhance subsequent MS detection.

Fig. 2 shows representative chromatograms of LC-MS grade water (blank), finished water, and source water analyzed using SPE (HLB)-HPLC (C18)-MS/MS analysis. The peaks detected in the blank are excluded from the list of candidate compounds detected in the finished water and source water. Similarly, chromatograms containing a large number of peaks were detected using the other two SPEs with HPLC-MS/MS and HILIC-MS/MS analyses. However, each chromatogram resulting from the different SPE and LC-MS/MS methods used contains some identical peaks and some unique peaks (Appendix A Figs. S1–S5).

Fig. 3a shows that a total of 8519 putative organic compounds were detected in a source water sample following enrichment

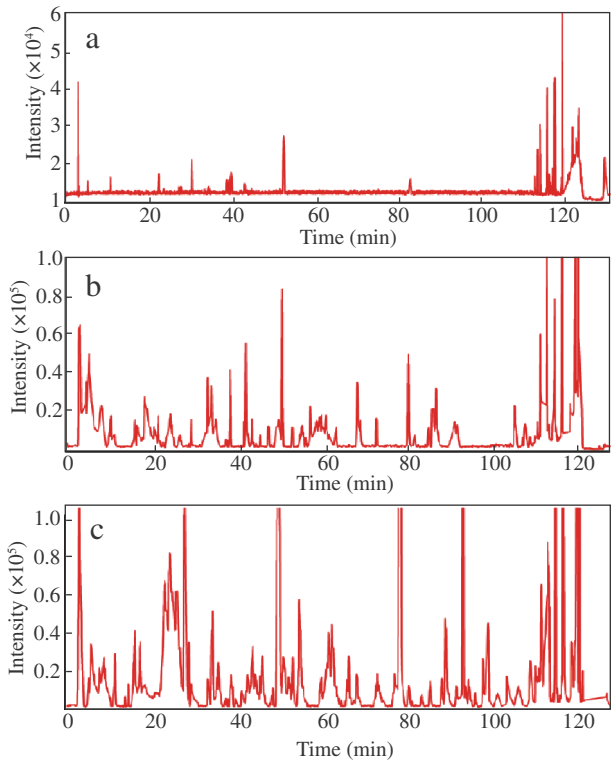

Fig. 2 – Base peak intensity (BPI) chromatograms of water samples extracted with Oasis HLB SPE cartridge: (a) BPI of LC-MS grade water (negative control); (b) BPI of finished water; (c) BPI of source water. SPE: solid phase extraction; LC-MS: liquid chromatography–mass spectrometry.

with the three SPE cartridges from a source water sample when analyzed by C18 LC-MS/MS. The use of individual Oasis HLB, Bond Elut C18, or Bond Elut ENV cartridges led to the identification of 5099, 3112, and 4461 putative organic compounds, respectively. Approximately 40% of the total unique putative compounds were identified from each SPE. Additional putative compounds in the same water sample were identified using HILIC-MS/MS, with ~45% unique identifications resulting from each SPE (Fig. 3b). Fig. 3 clearly demonstrates the advantage of using parallel SPE with complementary retention mechanisms to maximize the identification of unknown compounds in water samples.

Once DBPs and precursors were enriched by the three SPEs, each extract was separated using two complementary HPLC techniques: reversed phase HPLC and HILIC. The complementary HPLC techniques improved the retention and separation of various compounds of different hydrophobicity. The improved separation greatly enhanced the subsequent MS detection using high resolution tandem MS (MS/MS). In the present study, we identified 8519 and 10,952 putative compounds in the source water sample using C18 LC-MS/MS and HILIC-MS/MS analyses. While 2266 putative compounds were detected by both methods, each method detected over 70% unique identifications in the same water sample (Fig. 4a). Through the analysis of the finished water sample with both the C18 LC-MS/MS and the HILIC-MS/MS method, we observed over 50% unique identifications by each method (Fig. 4b). This indicated that the two LC separation techniques retained different organic compounds of diverse hydrophobicity, and the combination of these two LC separations greatly improved the detection of organic compounds in water. Combining the application of multiple SPEs and LC columns of diverse binding mechanisms before MS/MS analysis was effective.

2.2. Mass spectrometry analysis with PIE strategy

To improve the MS detection of organics with trace level concentrations in water, it is vital to optimize the spectral acquisition efficiency of mass spectrometry. The major factors governing the spectral acquisition efficiency are the acquisition speed, the spectral quality, and the frequency of spectral redundancy (Wang and Li, 2008). The dynamic exclusion created by reducing repeated scans on previously acquired ions can effectively increase the acquisition

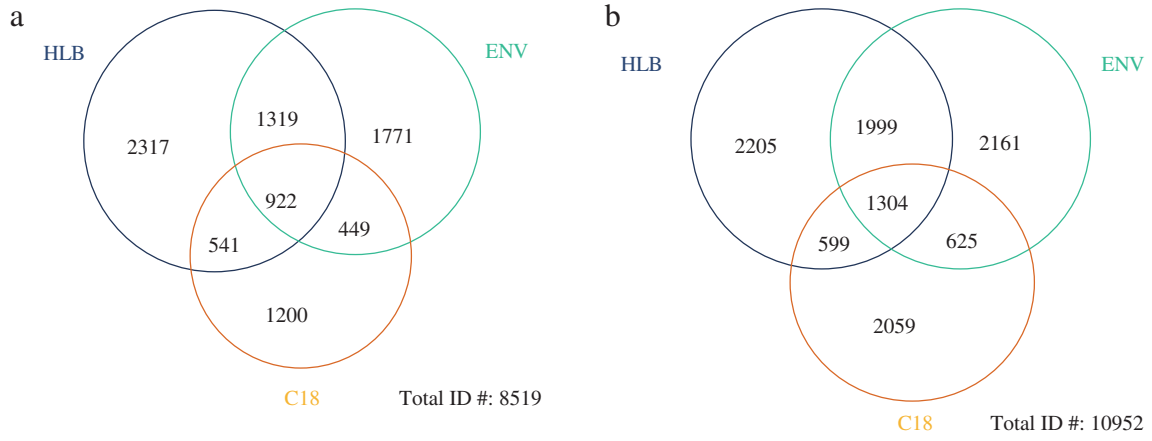

Fig. 3 – Putative identification numbers detected in source water after three SPE enrichments: Oasis HLB, Bond Elut ENV and Bond Elut C18. (a) Putative identification numbers in source water after three SPE enrichments detected from C18 LC-MS/MS experiment. (b) Putative identification numbers in source water after three SPE enrichments detected from HILIC LC-MS/MS experiment.

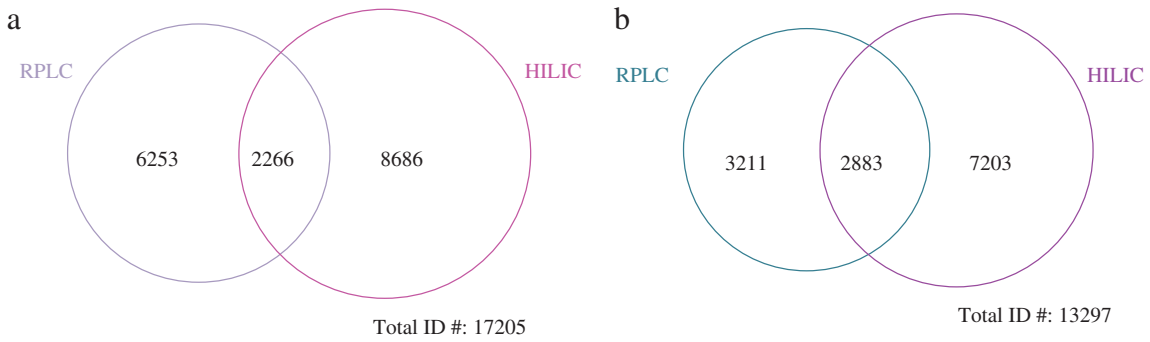

Fig. 4 – Putative identification numbers detected in source water and finished water samples. (a) Putative identification numbers in source water analyzed by C18 LC-MS/MS and HILIC LC-MS/MS. (b) Putative identification numbers in finished water analyzed by C18 LC-MS/MS and HILIC LC-MS/MS.

efficiency in a LC-MS/MS run. However, dynamic exclusion alone is not sufficient for the acquisition of spectral data of all compounds in a complex sample due to the co-elution of multiple compounds, the formation of multiple adduct ions from one compound, and the limited analysis time and number of MS/MS experiments in one LC-MS/MS run.

To address this challenge, we integrated the PIE strategy in the HPLC-MS/MS analysis. From each HPLC run, the high resolution tandem mass spectrometer collects a huge number of MS/MS spectra. Setting a lower threshold for triggering MS/MS would allow the collection of MS/MS spectra from more precursor ions, thus potentially detecting more compounds. However, more precursor ions mean less time spent on collecting each MS/MS spectrum, resulting in lower intensity and poorer quality of all MS/MS spectra. To overcome this problem, we applied the PIE strategy. The PIE strategy excludes selected preset ions so that more time is devoted to the detection of other ions. By excluding the highly abundant precursor ions that were analyzed in the first LC-MS/MS run, PIE allowed the MS instrument to spend more time collecting MS/MS spectra for ions of lower abundance.

To perform PIE, we first collected MS/MS spectra of ions with high abundance in the first LC-MS/MS acquisition. Then, an exclusion list of the *m/z* values of these precursor ions along with their retention time was generated and entered into the ion exclusion program of MS acquisition. These ions were subsequently excluded from MS/MS analysis in the duplicate LC-MS/MS run. Thus, more analysis time could be spent analyzing ions with lower abundance, allowing enhanced identification of low abundance compounds in the second LC-MS/MS run.

Analysis of source water and finished water by LC-MS/MS with PIE showed an increase of 30%–40% in the number of compounds detected in duplicate analyses compared to a single analysis without using the PIE approach. Table 1 shows the increase of putative identifications in the source water from C18 LC-MS/MS analysis with the implementation of PIE. In the source water sample enriched by the Oasis HLB cartridge, 1256 unique identifications were found from the second LC-MS/MS run with PIE, corresponding to a 32.7% increase compared to a single LC-MS/MS analysis. Similarly, 41.8% and 27.8% increases in the number of identified putative compounds were obtained with PIE in the source water extracts from Bond Elut ENV and Bond Elut C18 cartridges, respectively.

Additionally, the combination of complementary C18-MS/MS and HILIC-MS/MS analyses of a sample resulted in the further increase of the number of compounds detected. The preliminary results demonstrated the promise of our new technique, which integrates complementary HPLC separations with high resolution MS and MS/MS analysis of a sample to enhance the detection of a variety of compounds.

Using the optimized method, we identified a total of 17,205 and 13,297 putative organic compounds in the source water sample and the finished water sample respectively by searching against the HMDB. For different analytical emphasis, the LC-MS/MS data generated above can be searched in different databases. To identify peptides and their DBPs, we searched the LC-MS/MS data against the HMDB, as it contains information about chemical structures, clinical characteristics, and bioactivities of more than 41,000 biomolecules (Wishart et al., 2013).

After identifying a large number of putative organics in water, we manually examined their fragmentation patterns in MS/MS, and compared their accurate masses, as well as their retention times, with chemical standards to validate their identifications. As a proof of principle, we selected 30 putative compounds of interesting bioactivities, including peptides, toxins, and DBPs, to demonstrate the feasibility of our strategy. The experimental MS/MS spectra of all selected compounds were compared with their theoretical fragments. If over 50% of the intensive peaks in the MS/MS spectrum of a precursor ion matched with its theoretical fragments, this MS/MS spectrum was considered a good match, and the identification of this precursor ion was considered to be potentially valid. Otherwise the identification was rejected and not considered further. This

Table 1 – Comparison of identifications in replicate C18 LC-MS/MS analyses of source water with the implementation of PIE.

	Oasis HLB	Bond Elut ENV	Bond Elut C18
No. of IDs in the 1st run	3843	3145	2436
No. of IDs in the 2nd run with PIE	3388	3026	1724
Unique IDs in the 2nd run	1256	1316	676
Increases from the 1st run	32.7%	41.8%	27.8%
Total IDs	5099	4461	3112

PIE: precursor ion exclusion.

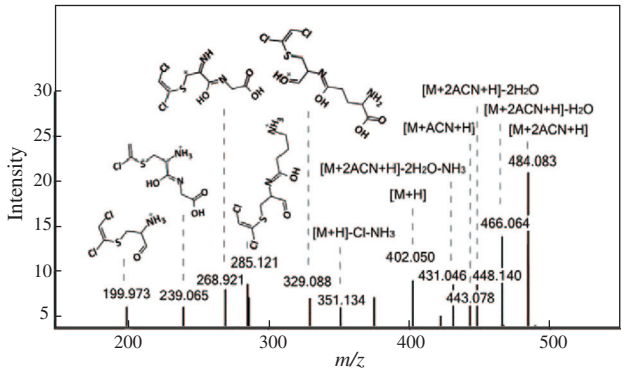

Fig. 5 – Manual interpretation of a MS/MS spectrum, precursor ion = 484.0828 Da, compared with the theoretical fragments and adduct ions of S-(1,2-dichlorovinyl)-glutathione. MS/MS: tandem mass spectrometry.

judgmental threshold was adopted for this scanning process to avoid false positive identifications.

As illustrated in Fig. 5, fragmentation peaks of S-(1,2-dichlorovinyl)-glutathione (precursor ion *m/z* = 484.0828 Da) in the MS/MS spectrum were manually matched with its theoretical fragments. With a mass tolerance window of 0.2 Da for fragment ions, 12 out of 15 peaks in the MS/MS spectrum, including the most intensive peaks, were matched with the adduct ions and fragments of this compound. Therefore, the identification of S-(1,2-dichlorovinyl)-glutathione was considered to be potentially valid. If one precursor ion had several putative identifications, its MS/MS spectrum would be manually interpreted and compared with the theoretical fragments of all putative compounds. The first ranked identification that had the highest number of matched peaks would be considered to be potentially valid (Table 2). After confirmation from the MS/MS spectra interpretation, the 30 compounds were further validated with chemical standards. To ensure the accuracy of validation, chemical standards of the 30 compounds were separated and analyzed under the same LC-MS conditions as the water sample analysis. Accurate *m/z* values, charge states, and retention times were used to compare the selected compounds with their chemical standards. The criteria for positive identification of a compound in water were: the *m/z* difference between the standard and the expected compound in water was <0.005 Da, the retention time tolerance window was 60 sec (Gika et al., 2007; Schmidt et al., 2008), and the charge state was the same. The retention time tolerance window was selected based on the replicated LC-MS experiments of five peptide standards (Appendix A Table S1). By this validation step, 25 compounds (out of 30) were confirmed in the source and finished water samples. Specifically, 19 organic compounds, including 12 peptides, were confirmed in the source water and 14 organic compounds, including 8 peptides, were confirmed in the finished water (Appendix A Table S2).

2.3. Comparative analysis of source water and finished water for identification of new DBPs

To identify putative peptide precursors and DBPs, we compared the candidate compounds detected in finished water with those in source water. Rather than identifying the compounds in both samples, we focused on the unique identifications in the finished water, as DBPs could belong to this group of compounds. For example, a comparison of the 17,205 putative organic compounds in source water with the 13,297 in finished water showed that over 70% (9701) of the putative identifications in finished water were identical to

Table 2 – Examples of putative identifications' prioritization using LC-MS/MS data. Putative identifications are ranked with the numbers of matched peaks in MS/MS spectra.

Query *m/z* (Da)	Name	Adduct	Exact mass (Da)	Number of matched peaks
435.2007	1,1′-Ethylidenebistryptophan	$M + H^+$	435.2027	12
	Melledonol	$M + H^+$	435.2013	6
	Moschamine	$M + 2ACN + H^+$	435.2027	5
	6′-Apiosyllotaustralin	$M + ACN + H^+$	435.1973	4
	1-Isothiocyanato-8-(methylthio)octane	$2M + H^+$	435.1991	2
	Myristicanol A	$M + H^+$	435.2013	2
	Epoxyfumitremorgin C	$M + ACN + H^+$	435.2027	2
	Erlotinib	$M + ACN + H^+$	435.2026	2
295.1247	Aspartame	$M + H^+$	295.1216	11
	Glutamyl-methionine	$M + NH4^+$	295.1197	7
	Methionyl-glutamate	$M + NH4^+$	295.1197	7
	Diosbulbinoside D	$M + 2ACN + 2H^+$	295.1232	6
395.1189	Aflatoxin B1	$M + 2ACN + H^+$	395.1237	11
	(E)-Antibiotic BE 23372M	$M + 2ACN + H^+$	395.1237	8
	3′,4,4′-Trihydroxypulvinone	$M + 2ACN + H^+$	395.1237	8
	3-Hydroxy-8,9-dimethoxycoumestan	$M + 2ACN + H^+$	395.1237	8
	8-Hydroxy-3-methoxy-1-methylanthra quinone-2-carboxylic acid	$M + 2ACN + H^+$	395.1237	7
	Aloe emodin w-acetate	$M + 2ACN + H^+$	395.1237	7
	De-O-methyldihydrosterigmatocystin	$M + 2ACN + H^+$	395.1237	7
	Wairol	$M + 2ACN + H^+$	395.1237	7

Table 3 – Chlorinated peptides detected in source water and finished water.

Source water	HMDB ID	Name	Molecular mass
	HMDB30457	Cyclochlorotine	571.1601
	HMDB60504	S-(1,2-dichlorovinyl)glutathione	401.0215
Finished water	HMDB ID	Name	Molecular mass
	HMDB01309	m-Chlorohippuric acid	213.0193
	HMDB01885	3-Chlorotyrosine	215.0349
	HMDB30399	N-(carbethoxyacetyl)-4-chloro-L-tryptophan	352.0826
	HMDB60358	2,3-Dihydro-2-S-glutathionyl-3-hydroxy bromobenzene	479.0362
	HMDB60504	S-(1,2-dichlorovinyl)glutathione	401.0215
	HMDB60506	S-(2,2-Dichloro-1-hydroxy)ethyl glutathione	419.0321

HMDB: Human Metabolome Database.

those detected in the source water. However, nearly 30% of the putative identifications were unique in finished water. Among all putative compounds, we found 625 peptides in source water (Appendix A Table S3) and 617 peptides in finished water (Appendix A Table S4). We identified 501 peptides in both the source water and the finished water samples, while the remaining 116 peptides were unique in finished water. As well, 124 unique peptides were detected exclusively in source water.

Among the 625 peptides detected in source water, we found two chlorinated peptides. These two chlorinated peptides were also found in finished water. Another four chlorinated peptides were detected among the 116 unique peptides found in finished water (Table 3). These findings most likely indicate that chlorination of peptides in water occurred during the chlorine disinfection process. Besides chlorination, we observed two nitrosation reactions on peptides after water disinfection, as shown in Fig. 6, by examining the 116 and 124 unique peptides detected in finished water and source water. One was S-nitrosation, wherein glutathione is transformed to S-nitrosoglutathione in the disinfection process (Choi and Valentine, 2003; Wang et al., 2002). The second N-nitrosation reaction pertained to dipeptides containing proline in source water. As reported, these dipeptides containing proline were transformed to N-nitrosoproline and N-nitrosopyrrolidine (Nebelin et al., 1980). This approach allows for the identification of non-chlorinated and chlorinated DBPs which have not been reported previously. It is novel and effective for advancing DBP research. The methodology can help increase knowledge of the potential precursors and resulting DBPs of biological molecules, which is important for the design of water treatment processes.

Acknowledgments

This project was supported by grants from the Natural Sciences and Engineering Research Council of Canada, Alberta Health, and Alberta Innovates—Energy and Environment Solutions.

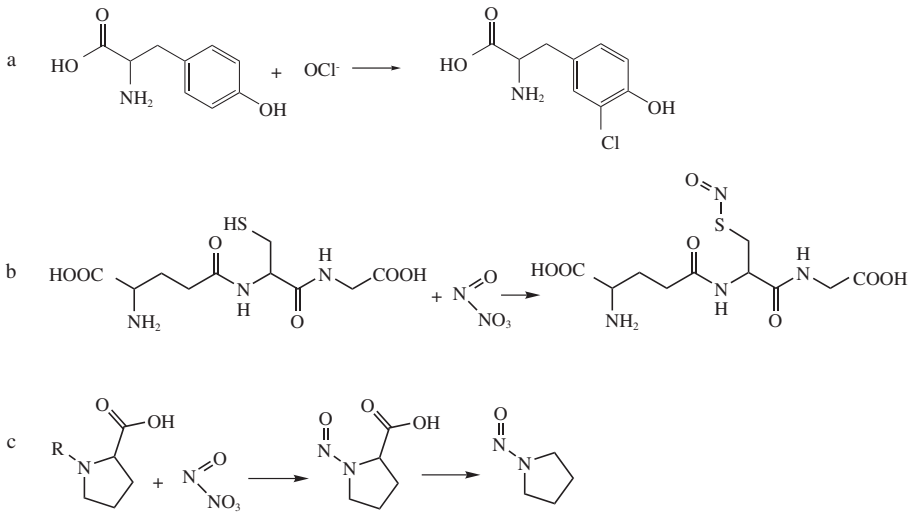

Fig. 6 – Three disinfection reactions on peptides in source water and finished water samples. (a) Chlorination of tyrosine. (b) S-nitrosation of glutathione during chlorine disinfection of water to produce S-nitrosoglutathione. (c) N-nitrosation of proline containing dipeptides during chlorine disinfection to produce N-nitrosoproline and N-nitrosopyrrolidine.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jes.2015.08.007.

REFERENCES

Brosillon, S., Lemasle, M., Renault, E., Tozza, D., Heim, V., Laplanche, A., 2009. Analysis and occurrence of odorous disinfection by-products from chlorination of amino acids in three different drinking water treatment plants and corresponding distribution networks. Chemosphere 77 (8), 1035–1042.

Choi, J., Valentine, R.L., 2003. N-nitrosodimethylamine formation by free-chlorine-enhanced nitrosation of dimethylamine. Environ. Sci. Technol. 37 (21), 4871–4876.

Garland, C.F., Grant, W.B., Boucher, B.J., Cross, H.S., Garland, F.C., Gillie, O., Gorham, E.D., Heaney, R.P., Holick, M.F., Hollis, B.W., Moan, J.E., Peterlik, M., Reichrath, J., Zittermann, A., 2009. Open letter to IARC Director Christopher P. Wild Re: IARC Working Group Report 5: vitamin D and cancer. Dermato-endocrinology 1 (2), 119–120.

Gika, H.G., Theodoridis, G.A., Wingate, J.E., Wilson, I.D., 2007. Within-day reproducibility of an HPLC-MS-based method for metabonomic analysis: application to human urine. J. Proteome Res. 6 (8), 3291–3303.

Grant, W.B., 2009. A critical review of vitamin D and cancer: a report of the IARC Working Group. Dermato-endocrinology 1 (1), 25–33.

Hayes, T.M., Hayes, M.H.B., Skjemstad, J.O., Swift, R.S., 2008. Compositional relationships between organic matter in a grassland soil and its drainage waters. Eur. J. Soil Sci. 59 (4), 603–616.

Hernandez, F., Sancho, J.V., Ibanez, M., Abad, E., Portoles, T., Mattioli, L., 2012. Current use of high-resolution mass spectrometry in the environmental sciences. Anal. Bioanal. Chem. 403 (5), 1251–1264.

Humpage, A.R., Hardy, S.J., Moore, E.J., Froscio, S.M., Falconer, I.R., 2000. Microcystins (cyanobacterial toxins) in drinking water enhance the growth of aberrant crypt foci in the mouse colon. J. Toxicol. Environ. Health A 61 (3), 155–165.

Ito, E., Kondo, F., Terao, K., Harada, K., 1997. Neoplastic nodular formation in mouse liver induced by repeated intraperitoneal injections of microcystin-LR. Toxicon 35 (9), 1453–1457.

Komaki, Y., Marinas, B.J., Plewa, M.J., 2014. Toxicity of drinking water disinfection byproducts: cell cycle alterations induced by the monohaloacetonitriles. Environ. Sci. Technol. 48 (19), 11662–11669.

Lauby-Secretan, B., Scoccianti, C., Loomis, D., Benbrahim-Tallaa, L., Bouvard, V., Bianchini, F., Straif, K., 2015. Breast-cancer screening—viewpoint of the IARC Working Group. N. Engl. J. Med. 372 (24), 2353–2358.

Li, J., Wang, W., Moe, B., Wang, H., Li, X.-F., 2015. Chemical and toxicological characterization of halobenzoquinones, an emerging class of disinfection byproducts. Chem. Res. Toxicol. 28 (3), 306–318.

Nebelin, E., Pillai, S., Lund, E., Thomsen, J., 1980. On the formation of N-nitrosopyrrolidine from potential precursors and nitrite. IARC Sci. Publ. (31), 183–193.

Richardson, S.D., Plewa, M.J., Wagner, E.D., Schoeny, R., Demarini, D.M., 2007. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. Mutat. Res. 636 (1-3), 178–242.

Schmidt, A., Gehlenborg, N., Bodenmiller, B., Mueller, L.N., Campbell, D., Mueller, M., Aebersold, R., Domon, B., 2008. An integrated, directed mass spectrometric approach for in-depth characterization of complex peptide mixtures. Mol. Cell. Proteomics 7 (11), 2138–2150.

Spoof, L., Vesterkvist, P., Lindholm, T., Meriluoto, J., 2003. Screening for cyanobacterial hepatotoxins, microcystins and nodularin in environmental water samples by reversed-phase liquid chromatography-electrospray ionisation mass spectrometry. J. Chromatogr. A 1020 (1), 105–119.

Tavtigian, S.V., Greenblatt, M.S., Goldgar, D.E., Boffetta, P., 2008. Assessing pathogenicity: overview of results from the IARC Unclassified Genetic Variants Working Group. Hum. Mutat. 29 (11), 1261–1264.

Ueno, Y., Nagata, S., Tsutsumi, T., Hasegawa, A., Watanabe, M.F., Park, H.D., Chen, G.C., Chen, G., Yu, S.Z., 1996. Detection of microcystins, a blue-green algal hepatotoxin, in drinking water sampled in Haimen and Fusui, endemic areas of primary liver cancer in China, by highly sensitive immunoassay. Carcinogenesis 17 (6), 1317–1321.

Villanueva, C.M., Cantor, K.P., Cordier, S., Jaakkola, J.J., King, W.D., Lynch, C.F., Porru, S., Kogevinas, M., 2004. Disinfection byproducts and bladder cancer: a pooled analysis. Epidemiology 15 (3), 357–367.

Wang, N., Li, L., 2008. Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. Anal. Chem. 80 (12), 4696–4710.

Wang, P.G., Xian, M., Tang, X., Wu, X., Wen, Z., Cai, T., Janczuk, A.J., 2002. Nitric oxide donors: chemical activities and biological applications. Chem. Rev. 102 (4), 1091–1134.

Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., Bouatra, S., Sinelnikov, I., Arndt, D., Xia, J., Liu, P., Yallou, F., Bjorndahl, T., Perez-Pineiro, R., Eisner, R., Allen, F., Neveu, V., Greiner, R., Scalbert, A., 2013. HMDB 3.0—the Human Metabolome Database in 2013. Nucleic Acids Res. 41 (Database issue), D801–D807.