1 **Molecular characterization of low molecular weight dissolved**
2 **organic matter in water reclamation processes using Orbitrap**
3 **mass spectrometry**

4
5
6 Phanwatt Phungsai^{a,*}, Futoshi Kurisu^b, Ikuro Kasuga^a, and Hiroaki Furumai^b

7
8 ^a Department of Urban Engineering, School of Engineering, The University of Tokyo, 7-3-1
9 Hongo, Bunkyo, Tokyo 113-8656, Japan

10 ^b Research Center for Water Environment Technology, School of Engineering, The
11 University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan

12
13 * Corresponding author. Tel.: +81-0-8861-8520; Fax: +81-3-5841-6244.

14 E-mail address: phanwatt@env.t.u-tokyo.ac.jp

15
16
17 **Abstract**

18 Reclaimed water has recently become an important water source for urban use, but the
19 composition of dissolved organic matter (DOM) in reclaimed water has rarely been
20 characterized at the compound level because of its complexity. In this study, the

transformation and changes in composition of low molecular weight DOM in water
reclamation processes, where secondary effluent of the municipal wastewater treatment
plant was further treated by biofiltration, ozonation and chlorination, were investigated by
“unknown” screening analysis using Orbitrap mass spectrometry (Orbitrap MS). The
intense ions were detected over an m/z range from 100 to 450. In total, 2412 formulae with
various heteroatoms were assigned, and formulae with carbon (C), hydrogen (H) and
oxygen (O) only and C, H, O and sulfur (S) were the most abundant species. During
biofiltration, CHO-only compounds with relatively high hydrogen to carbon (H/C) ratio or
with saturated structure were preferentially removed, while CHOS compounds were mostly
removed. Ozonation induced the greatest changes in DOM composition. CHOS compounds
were mostly decreased after ozonation while ozone selectively removed CHO compounds
with relatively unsaturated structure and produced compounds that were more saturated and
with a higher degree of oxidation. After chlorination, 168 chlorine-containing formulae,
chlorinated disinfection by-products (DBPs), were additionally detected. Candidate DBP
precursors were determined by tracking chlorinated DBPs formed via electrophilic
substitution, half of which were generated during the ozonation.

Keywords

Unknown screening, Low molecular weight organic matter, Molecular characterization,
Orbitrap mass spectrometry, Disinfection by-products, Reclaimed water

41

42 **Abbreviations**

43 BOM, biodegradable organic matter; DATS, dialkyl tetralin sulfonate; DOM, dissolved
44 organic matter; DOC, dissolved organic carbon; DBE, double bond equivalent; DBE-O,
45 double bond equivalent minus oxygen; DBPs, disinfection by-products; EfOM, effluent
46 organic matter; ESI, electrospray ionization; FT-ICR-MS, Fourier-transform ion cyclotron
47 resonance mass spectrometry; H/C, hydrogen to carbon ratio; LAS, linear alkylbenzene
48 sulfonate; LMW, low molecular weight; MS, mass spectrometry; NOM, natural organic
49 matter; Orbitrap MS, Orbitrap mass spectrometry; O/C, oxygen to carbon ratio; SUVA,
50 Specific ultraviolet absorbance; SPC, sulfophenyl carboxylate; SPE, Solid phase extraction;
51 TOC, total organic carbon; TN, total nitrogen

52

53 **1. Introduction**

54 Remarkable increases in water consumption and water pollution have raised concerns about
55 the sustainability of water use in urban cities. Water stress has caused various problems for
56 natural water resources such as instability of water flow in rivers, spring water drying,
57 groundwater level decline and ecosystem deterioration (Furumai, 2008). Hence, sustainable
58 water use strategies are necessary to ensure long-term water supply with adequate water
59 quality as well as to protect against adverse ecological, economic and social effects.
60 Reclaimed water use is one of the countermeasures towards sustainable water use, and it has

61 been implemented for various applications such as toilet flushing water, river flow
62 maintenance, water front creation, plant irrigation and pavement cooling. The benefits of
63 reclaimed water are not only to secure water supplies while reducing demand for high
64 quality tap water but also to alleviate water pollution by reducing wastewater discharge to
65 natural water resources. Because wastewater increases proportionally with water
66 consumption, reclaimed water should be an abundant and stable water resource.

67 The quality of reclaimed water is the most important consideration for reuse applications,
68 and dissolved organic matter (DOM) is a central issue. Secondary or tertiary effluents, used
69 as influent streams in water reclamation plants, still contain DOM that is refractory to
70 biological treatment and are often referred to as effluent dissolved organic matter (EfOM).

71 EfOM is typically composed of a wide variety of compounds such as natural organic matter
72 (NOM) originating from drinking water, soluble microbial products produced during
73 biological treatment processes and synthetic organic compounds from domestic uses (Shon
74 et al., 2006). During and after the water reclamation process, DOM can cause various
75 problems such as adsorption and deposition of organic foulants in membrane treatment
76 processes (Xu et al., 2010), formation of disinfection by-products (DBPs) by providing
77 precursors, increasing demand for oxidants or disinfectants, and promoting microbial
78 regrowth in distribution systems by providing growth substrates (Thayanukul et al., 2013).

79

80 To tackle these problems, improved knowledge of DOM in reclaimed water is required.

81 Various characterization and quantification methods have extensively focused on bulk
82 chemical properties such as total and dissolved organic carbon (TOC and DOC), size
83 distribution, fluorescence excitation/emission, functional groups, and spectroscopic
84 properties. However, detailed descriptions at compound level are absent from those methods
85 (Sleighter and Hatcher, 2007; Zhang et al., 2012). An unknown screening (nontargeted
86 screening) by ultrahigh resolution and accurate mass spectrometry can reveal a broad
87 spectrum of a complex mixture and enable characterization of elemental compositions,
88 providing detailed information at the molecular scale. Elemental compositions of DOM can
89 provide useful information to understand reaction pathways, transformations and changes in
90 molecules during treatment processes. Molecular changes of DOM along with drinking
91 water treatment processes have been analyzed by electrospray ionization (ESI) coupled with
92 Fourier transform ion cyclotron resonance (FT-ICR-MS) using unknown screening
93 approach (Gonsior et al., 2014; Lavonen et al., 2013; Zhang et al., 2012). In addition,
94 compositions of EfOM were also evaluated by ESI-FT-ICR-MS (Gonsior et al., 2011; Tseng
95 et al., 2013). Little is currently known about changes in composition and transformation of
96 low molecular weight (LMW)-DOM in further treatment processes of water reclamation
97 plants.

98
99 In addition to FT-ICR-MS, a newly developed mass analyzer employing an ion trap in an
100 electrostatic field (Orbitrap) has been developed by Makarov (2000). Coupling of ESI and

Orbitrap Fourier-transform mass spectrometry (Orbitrap MS) is a promising technique for unknown screening of complex mixtures of DOM. With internal calibration, mass accuracy less than 2 ppm could be achieved (Krauss et al., 2010; Olsen et al., 2005) enable molecular formula identification for C, H, O and S containing molecules for < 500 Da molecules (Remucal et al., 2012).

In this study, unknown screening by ESI Orbitrap MS was applied to elucidate the molecular composition of LMW-DOM throughout water reclamation processes including secondary effluent, biofiltration, ozonation and chlorination. The transformation/composition changes of LMW-DOM in each treatment process were investigated by comparing compound intensities before and after treatments. Characteristics such as hydrogen to carbon (H/C) and oxygen to carbon (O/C) ratios and degree of unsaturation (double bond equivalent; DBE) of molecules were calculated to provide details of how LMW-DOM changed with each of the reclamation processes. In addition, efforts were made to identify DBPs, chlorine-containing formulae additionally formed after chlorination, and their potential precursors to determine the probable origins of those precursors.

2. Materials and methods

2.1 Reclaimed water samples

Water samples were collected from a water reclamation plant receiving secondary effluent from a municipal wastewater treatment plant in Japan in April 2014 and May 2015. The findings and discussion in this study were mainly based on the observations of 2014 samples, and reproducibility was tested on samples taken in 2015 when necessary. The treatment capacity of the plant was 960 m³/d. The treatment processes consisted of biofiltration for removal of remaining biodegradable organic matter (BOM), ozonation for removal of color and odor (4.5 g/m³ of ozone dosage) and chlorination for disinfection. Ceramic filter media was used as the microbial carrier in the biofiltration unit. The reclaimed water was stored in a 2000 m³ storage tank where sodium hypochlorite was added with dosage of 6 mg/L as free chlorine to achieve 1 mg/L of total chlorine concentration before distribution. Reclaimed water was storage for 24 h on average. We sampled the secondary effluent as well as process waters such as biofiltration, ozonation and chlorination waters in pre-combusted glass bottles (6 h, 550 °C). The water samples were immediately transferred to the laboratory within 1.5 h.

TOC and DOC concentrations were measured by TOC-L_{CSH} (Shimadzu, Japan) as a nonpurgeable organic carbon. Total nitrogen (TN) was determined with a TN unit equipped with TOC-L_{CSH}. Residual chlorine and pH were measured on-site using the colorimetric method with a Hach DR/890 colorimeter (Hach, USA) and model multi 3430 (WTW, Germany), respectively. Anion concentrations such as nitrate (NO₃⁻), and chloride (Cl⁻)

were evaluated using an ion chromatograph (IC 761, Metrohm, Switzerland) with SI-90 column (Shodex, Japan). Ammonia (NH_4^+) concentration was evaluated using a spectrophotometric method (Hach DR2800, Hach, USA). UV absorbance at 254 nm was measured using a 1 cm cell (UH3500 Hitachi, Japan) and specific UV-absorbance (SUVA) was calculated by the absorbance value at 254 nm divided by DOC concentration.

2.2 Solid phase extraction (SPE)

Five hundred milliliters of water samples were first filtered with 0.3 μm pre-combusted glass-fiber filters (GF-75, Advantec Toyo, Japan). The filtrates were acidified with high-purity grade hydrochloric acid to pH 2 and passed through an SPE cartridge (Bond Elut PPL, Agilent Technologies, USA) at a flow rate of 20 ml/min. The cartridge contained 1 g of a polystyrene-divinylbenzene modified with a nonpolar surface. It was previously conditioned with methanol (LC-MS grade) and equilibrated with acidified Milli-Q (0.01 M hydrochloric acid) before use. For the elution step, the cartridge was first rinsed with 20 ml of the acidified Milli-Q to eliminate retained salts, after which the retained organic matter was eluted with 10 ml of methanol. A blank sample for SPE (extraction blank) was also prepared with 500 ml Milli-Q water instead of the samples. The extracts were stored at $-20\text{ }^\circ\text{C}$ in the dark. To calculate DOC recovery in the SPE step, 1 ml of the extract was

evaporated to dryness by N₂ gas flow and redissolved in 20 ml of Milli-Q water for TOC measurement.

2.3 Unknown screening analysis

The unknown screening analyses were performed by Orbitrap MS (Exactive, Thermo Scientific, USA). First, the Orbitrap MS was externally calibrated by the calibration solution (Proteomass ESI, Sigma-Aldrich, USA) that covers m/z range of 50–2000. Additionally, a hexadecanoic acid peak ($[M-H]^- = 255.23295$) in the samples was used for lock mass as an internal calibrant (Sleighter et al., 2008). Average mass error of a naturally existed fatty acid was 0.40 ± 0.26 ppm. The operational conditions of ESI in negative ionization mode were as follows; spray voltage, 4.5 kV; sheath gas, 30 arbitrary units; sweep gas, 0 arbitrary units; capillary temperature, 250 °C; capillary voltage, 25 V; tube-lens voltage, 100 V; skimmer voltage, 30 V. The conditions of the Orbitrap were: resolution, 100000 at 200 m/z and 65000 at 500 m/z ; maximum injection time, 250 ms; automatic gain control target value, 500000 for ultimate mass accuracy of full scan analysis. The ESI is well-known to produce singly charged and mainly unfragmented molecules (Sleighter and Hatcher, 2007). The ionization condition was optimized to obtain as many ion peaks as possible by balancing desolvation and in-source fragmentation (Urai et al., 2014). In-source fragmentation of weakly bond molecules might be included in this study in

some extent. The m/z range was set at 100–1000. The samples were analyzed by flow injection mode at 200 $\mu\text{l}/\text{min}$ of a methanol mobile phase. In addition, the extraction blank was also measured to check background peaks that were subtracted in mass spectrum analysis. Six replicates of both samples and the extraction blank were analyzed for reproducibility and statistical testing for background subtraction.

Raw mass spectra of the samples and the extraction blanks were processed by Sieve software version 2.2 (Thermo Scientific) to screen background peaks, peak integration, isotope detection and component extraction. The m/z width was set at 2 ppm for peak integration. Peaks with intensities that were significantly higher ($p < 0.05$) than the extraction blank and with a signal to noise ratio ≥ 3 were retained. The charge state of each ion was automatically validated by the presence of isotopologues. Molecular formula were identified using compound identification algorithms (Kujawinski and Behn, 2006; Kujawinski et al., 2009) on Matlab software considering following elements: C_{0-39} , H_{0-72} , O_{0-20} , N_{0-2} , S_{0-2} , and Cl_{0-3} . Heuristic filtering of identification (7 golden rules) were also used in this algorithm to validate the identification (Kind and Fiehn, 2007). Mass error window for formula identification was set at 2 ppm and mass window allowed for identification by functional group relationship was 5 ppm. Peaks containing ^{35}Cl were additionally validated by the existence of ^{37}Cl isotopic peak. DBE minus oxygen (DBE-O), which was calculated by the following equation (1), of -10 to 10 was allowed (Peter

Herzprung and Wolf von Tümping, 2014). Assigned formulae with mass error > 5 ppm were excluded. Those peaks that did not meet the above criteria were excluded from further analysis.

$$\text{DBE-O} = (\text{C}+1) - 0.5 (\text{H}-\text{N}+\text{Cl}) - \text{O} \quad (1)$$

2.4 Calculation and statistical tests

Elemental composition of each molecule was converted to characteristics such as hydrogen to carbon (H/C) ratio, oxygen to carbon ratio (O/C), DBE and DBE-O. Ratios of H/C and O/C for each formula were used to plot van Krevelen diagrams to illustrate the possible type of compounds (Kim et al., 2003). DBE-O has been proposed to be a better indicator than DBE for indicating the true degree of unsaturation (Cortes-Francisco and Caixach, 2013; Gonsior et al., 2009). As DOM molecules contained carbonyl groups, DBE-O would be a better indicator for the degree of unsaturation of the carbon skeleton. The statistical differences of the mean characteristics between samples and groups of formulae were examined by analysis of variance (ANOVA) with post hoc tests using SPSS software (IBM). Hierarchical cluster analysis was applied using Ward's method to exam dissimilarity among water samples by considering the presence and absence of peaks.

3. Results and discussions

3.1 Basic water quality

The basic water qualities are given in Table 1 and Table S1 for samples taken in 2014 and 2015, respectively. TOC and DOC concentrations decreased along with the treatment processes, indicating the removal of organic matter. As SUVA is an indicator of hydrophobicity and hydrophilicity (Matilainen et al., 2011), reduction of SUVA after ozonation and chlorination implies the removal of hydrophobic compounds.

3.2 SPE

Carbon extraction efficiency of the PPL cartridge is one of the important factors influencing the number of peaks detected in unknown screening analysis. PPL cartridge was previously reported to achieved highest DOC recovery in natural water samples (Dittmar et al., 2008). The consistent results were obtained when comparing PPL, C18 and HLB cartridges for wastewater samples in this study by considering number of peaks and DOC recovery. The DOC recovery of PPL ratios were 62%, 63%, 47% and 46% for secondary effluent, biofiltration effluent, ozonation effluent and reclaimed water, respectively. Judging from the recovery ratio of SPE, about half of the organic matter (on a concentration basis) in samples was subjected to the analysis. The true recovery ratios might be higher if the evaporation of readily volatile organic molecules during the drying step is high. The recovery was lower in

the samples after ozonation. This is because ozone oxidizes some organic matter into very highly polar molecules that are poorly retained in the cartridge. The decrease in the SUVA value after ozonation also suggested this change. Similar DOC recoveries were obtained in water samples in 2015 (as shown in Supplementary Material (SM) Fig. S1).

3.3 Mass spectrum pattern

The mass spectra of each water sample obtained by Orbitrap MS under negative ionization mode were shown in Fig. 1. Analysis under positive ionization mode was not evaluated because the concentration of basic organic compounds was low in all water samples (SM Fig. S2) and the potential formation of sodium adducts increases complexity in molecular formula identification (Rostad and Leenheer, 2004). The mass spectra of secondary effluent and biofiltration were similar while clearly different patterns were observed in the ozonation effluent and reclaimed water. This implies that there were substantial effects of ozonation and chlorination on LMW-DOM. Mass spectra of all water samples revealed intense peaks spanning only the m/z range 100–450. Most of the ions (95.62%) were singly charged. This finding is consistent with previous reports utilizing negative ESI with FT-ICR-MS (Gonsior et al., 2014; Kim et al., 2003; Sleighter and Hatcher, 2007). The averaged neutral mass values (m/z plus weight of proton) were 281.8, 280.3, 283.3 and 264.0 Da for secondary effluent, biofiltration effluent, ozonation effluent and reclaimed water, respectively.

The samples taken in 2015 showed similar shifts in composition with treatment processes compared with those in 2014 (as shown in SM; Fig. S3). The pattern changed after ozonation, similarly to that observed for the samples in 2014. The similarity of mass spectra was evaluated by hierarchical cluster analysis of presence/absence of peaks in the samples (Fig. 2). The cluster analysis indicated that biofiltration induced small changes in LMW-DOM, while drastic change was induced by ozonation. Chlorination also changed LMW-DOM to some extent.

3.4 Elemental composition of LMW-DOM in water reclamation processes

Overall, 2412 molecular formulae out of 3825 total peaks (63.06%) were identified and over 60% of them had mass error <2 ppm (Fig. S4). Identified peaks with 2-5 ppm error were based on functional group relationship described elsewhere (Kujawinski and Behn, 2006). The unidentified peaks would likely be molecules that had element ratios beyond the identification conditions in this study or that contained other heteroatoms (e.g., P, Na and Br). The number of formulae in secondary effluent, biofiltration effluent, ozonation effluent and reclaimed water were 1215, 1122, 1047 and 985, respectively. Over 702 compounds were commonly found in all water samples.

Identified formulae were classified based on heteroatoms and the CHO-only class was the most abundant in number (Fig. 3). In terms of the peak intensity, abundant peaks were mainly attributed to CHOS and CHO-only classes (Fig. 4) possibly because the acidic properties of oxygen and sulfur enable them to readily lose protons (Kujawinski et al., 2009) and thus they could be easily ionized under negative ionization mode. In contrast, as N-containing compounds are difficult to deprotonate because of their high proton affinity low peak intensities were obtained for N-containing formulae. The abundance of CHO and CHOS formulae in secondary effluent were previously reported while CHON were highly suppressed under negative ionization mode (Gonsior et al., 2011). In addition, over two hundred of Cl-containing formulae were found in the reclaimed water (after chlorination). Those compounds in the reclaimed water will be discussed as DBPs in detail in Section 3.7. Sixty-nine chlorinated compounds were also observed in secondary effluent, even though the secondary effluent was not chlorinated. The existence of chloroorganic compounds in NOM samples has been observed previously (Gribble, 2003; Lavonen et al., 2013) and the presence of chloroorganics in effluent from activated sludge and membrane bioreactors before chlorination was previously reported (Fukushima et al., 2014). Those already present Cl-containing molecules could be attributed to (1) DBPs from tap water, (2) products from reactions of household chemicals with organic materials, (3) chlorine-containing health care products, (4) reactions with HCl in the acidification step used for SPE extraction, (5) chloroorganic compounds in NOM (Gribble, 2003; Lavonen et al., 2013), and (6) Cl-adduct

formation (Boutegrabet et al., 2012). In addition, a similar portion of LMW-DOM was detected in 2015, and over half of the identified LMW-DOM in 2014 was also present in 2015 samples (as shown in the SM; Fig. S5). The discussion of this study mainly focused on CHO, CHOS and Cl-containing formulae because of their abundance in reclaimed water under negative ionization mode.

The number of CHO and CHOS formulae in each treatment process as well as the averaged characteristics is summarized in Table 2. When comparing averaged characteristics between CHO and CHOS, CHO formulae had higher DBE_{avg} ($p < 0.0005$) than CHOS while the O/C_{avg} value was not significantly different. This implied that CHO formulae were more unsaturated than CHOS formulae. Similar findings were observed in 2015 (SM Table S2) in this study and were reported by Gonsior et al. (2011) in a study of secondary effluent from activated sludge processes. In comparison with Suwannee river NOM in the study by Cortes-Francisco and Caixach (2013) (Table 2), CHO formulae in secondary effluents detected by both Orbitrap MS and FT-ICR-MS had lower averaged DBE values, which implied an abundance of relatively saturated molecules in the effluents.

Van Krevelen diagrams have been widely used to visualize a large number of molecular formulae by plots between H/C and O/C ratios. The unique H/C and O/C ratios of each formula illustrate possible types of compounds such as lipid, lignin, protein, carbohydrate,

314 unsaturated compounds and condensed aromatic compounds (Hockaday et al., 2009). Van
 315 Krevelen diagrams of CHO and CHOS formulae detected in secondary effluent are shown
 316 in Fig. 5, and the other samples are provided in the SM (Figs. S6 and S7). For CHO class,
 317 the plots mainly fell within the areas of the unique H/C and O/C ratios of lignin ($H/C = 0.7\text{--}1.5$,
 318 $O/C = 0.1\text{--}0.67$), unsaturated hydrocarbons ($H/C = 0.7\text{--}1.5$, $O/C = 0\text{--}0.1$) and
 319 condensed aromatic or black carbon ($H/C = 0.2\text{--}0.7$, $O/C = 0\text{--}0.67$) molecules (Kim et al.,
 320 2004; Ohno et al., 2010). Plots of strong intensity were attributed to CHOS (Fig. 5b). The
 321 top 5 most abundant formulae were the same as the molecular formula of sulfophenyl
 322 carboxylates (SPCs) such as $C_9H_{10}O_5S$ (C₃-SPC), $C_{10}H_{12}O_5S$ (C₄-SPC), $C_{11}H_{14}O_5S$
 323 (C₅-SPC), $C_{12}H_{16}O_5S$ (C₆-SPC), and $C_{13}H_{18}O_5S$ (C₇-SPC) while formulae of longer chain
 324 SPCs such as C₈-SPC to C₁₃-SPC were also detected. SPCs are well known as
 325 biodegradation products of anionic surfactants namely linear alkylbenzene sulfonates (LASs)
 326 (Eichhorn et al., 2002). Molecular formulae of LASs such as $C_{17}H_{28}O_3S$ (C₁₁-LAS),
 327 $C_{18}H_{30}O_3S$ (C₁₂-LAS), and $C_{19}H_{32}O_3S$ (C₁₃-LAS) were also detected in secondary effluent.
 328 In addition, molecular formulae identical to dialkyl tetralin sulfonates (DATS, $C_{16}O_{24}O_3S$)
 329 and DATS intermediate (DATSI, $C_{15}O_{20}O_5S$), which are major co-products of LASs, were
 330 also observed. The formulae of SPCs, LASs and DATS were also detected in 2015 samples
 331 and SPCs-like formulae dominated in the 2015 samples (SM Fig. S8). The presence of
 332 LASs and their co-products in secondary effluent and reclaimed water were previously
 333 reported (Gonsior et al., 2011; Tseng et al., 2013; Urai et al., 2014).

3.5 Removal of LMW-DOM during biofiltration

Even if unknown screening by Orbitrap MS is not a quantitative analysis, it can be used in semi-quantitative way in comparing samples with similar matrices. As DOC concentration of samples were not very different (Table 1 and Table S1), the effect matrices to variation of intensity would be similar. To semiquantitatively compare the peak intensity of each compound before and after the treatment processes, the variations of intensity in six replicate analyses and reduction of intensity with concentration of the sample were examined. The histograms of variation in intensity showed a bell curve with low deviation and highly positive kurtosis values (SM Fig. S9) and absolute intensity of peaks were decreased with the decreasing of sample concentrations (SM Fig. S10). The majority (over 95%) of peaks had less than 30% variation. Thus, we set 30% change in the peak intensity after treatment processes as a criterion to semiquantitatively determine any decrease or increase in concentrations. With this criterion, approximately 95% confidence level of the change in LMW-DOM after treatment would be obtained.

In this treatment plant, biofiltration was applied to remove residual BOM in secondary effluent and to reduce ozone consumption in the following treatment. Organic carbon concentration decreased after biofiltration treatment probably due to the removal of remaining BOM (Table 1). Van Krevelen diagrams of molecular formulae (Fig. 6) showed changes of compounds during the biofiltration treatment. For CHO class, it was observed

that molecules with relatively high H/C values were removed by biofiltration (Fig. 6a). The averaged H/C of the compounds decreased by the treatment was significantly higher ($p < 0.0001$ in SM Tables S3 and S4) than those unchanged by the treatment. The same response was observed in 2015 samples. This implied that biofiltration preferentially removed saturated molecules and that unsaturated molecules were more resistant to biodegradation. DOM with low H/C ratios were previously reported to be resistant to biodegradation (Kim et al., 2006). In addition to the CHO-class molecules, most of the CHOS molecules were removed during biofiltration as they displayed similar saturated characteristics to CHO formulae. This result also indicated their relatively high biodegradability.

The intensities of C₁₂-, C₁₃- and C₁₄- LAS-like compounds were not decreased during biofiltration. As formulae of SPCs were observed, the LASs would have already been biodegraded into SPCs and other products at the municipal wastewater treatment plant. The removal of LAS and its co-products by biological treatment such as the activated sludge process could be up to 95% (Rapaport and Eckhoff, 1990). LAS-like compounds found in this study might be iso-LASs, which were reported to be resistant to biodegradation (Trehly et al., 1996). SPCs were reported to have less biodegradability compared with their parent LASs, but they could be further degraded into carbon dioxide and sulfate (Swisher, 1986; UNEP/DTIE, 2006). In this study, intensities of C₃- to C₇-SPCs-like compounds were decreased by over 60% to 80%, while intensities of C₈- to C₁₃-SPCs-like compounds were

decreased by 37% to 49%, respectively (SM Fig. S11a). This result indicated the high biodegradability of short-chain SPCs.

3.6 Effect of ozonation on LMW-DOM

Treatment reactions in ozonation would be caused by both ozone and the OH radical. Reactions of ozone with organic molecules generally occur at double bonds via addition reactions or at the aromatic ring via electrophilic attack; various ozone oxidation by-products such as carboxylic acids, aldehydes, ketones, keto aldehydes, alcohols, esters and hydroxy acids have been reported (von Gunten, 2003). Utilizing unknown screening by Orbitrap MS, over 200 formulae were found increased or newly formed after ozonation (Fig. 7). They should be considered as oxidation by-products. It is important to note that unchanged formulae might contain both increased and decreased compounds resulted in changed of intensity lower than a criterion of 30%. Over half of the CHOS formulae were decreased by over 30% in peak-intensity while only 26% of the CHO formulae were decreased by over 30% after ozonation treatment. This finding implies preferential oxidation of CHOS compounds over CHO ones. In addition, the plots of CHO by-products fall in the high O/C (high degree of oxidation) area while the CHOS by-products were broadly distributed. The DBE-O distribution of CHO (Fig. 8) was shifted to more negative values after ozonation treatment. This indicated that unsaturated CHO molecules (high DBE-O) were preferentially removed during ozonation and produced relatively more

saturated (low DBE-O) and oxidized compounds (high O/C ratio). The unsaturated molecules might contain multiple double bonds or an activated aromatic system available to be attacked by ozone.

The intensity of SPCs-like formulae was further decreased after ozonation by over 50% (SM Fig. S11b). About 50% removal of SPCs by ozonation was previously reported (Eichhorn et al., 2002).

3.7 Transformation and composition changes of LMW-DOM during chlorination

After chlorination, 81% and 72% of CHO and CHOS formulae, respectively, were decreased in intensity by over 30%. This implied a considerable effect of chlorine in reaction with LMW-DOM. In the reclaimed water, 168 chlorine-containing compounds, which were determined with high confidence level by matching isotopic patterns, were additionally detected or substantially increased in intensity after chlorination. These compounds should be classified as DBPs. The Cl-containing formulae contained one, two and three chlorine atoms with various heteroatoms such as CHOCl, CHONCl, CHOSCl and CHONSCl and the lists of chlorinated DBPs are provided in SM Table S6. Among these, formulae consistent with dichloroacetic and trichloroacetic acids, well known DBPs, were included. However, 26 chlorinated formulae were not found in ChemSpider, a chemical structure database (Pence

and Williams, 2010). The unidentified compounds detected in this study were classified as unknown DBPs. For the HOCl class, 10 formulae were the same as formulae of unknown chlorinated DBPs in previous studies on chlorination of NOM (Lavonen et al., 2013; Zhang et al., 2012). The number of chlorinated DBPs in reclaimed water in 2015 was 148 (SM Table S5). The difference could be because of the slightly higher chlorine dosage in 2015 and/or differences in the precursors in the ozonation effluent. In total, 97 DBP formulae were commonly detected in both 2014 and 2015 (SM Table S5).

In general, reactivity of hypochlorous acid ($^{\delta+}\text{Cl}-\text{OH}^{\delta-}$) is selective to limited sites of organic compounds such as reducing, nucleophilic and unsaturated sites because of the chemical structure and oxidizing power of Cl-O bond polarization. This induces three kinds of possible reactions of hypochlorous acid with organic compounds such as (1) oxidation reaction, (2) addition of HOCl to unsaturated bonds and (3) electrophilic substitution at nucleophilic sites (Deborde and von Gunten, 2008; Johnson and Jensen, 1986). In general, electrophilic attacks are fast enough to be significant in chlorination while oxidation and addition reactions are required longer retention time to be taking place (Deborde and von Gunten, 2008).

Electrophilic substitution induces a modification of the organic precursors, in which hydrogen atoms are displaced by chlorine atoms, resulting in chlorinated molecules with lower H/C while O/C and DBE-O remained unchanged. In addition, the addition reaction integrates HOCl into the precursors resulting in chlorinated molecules with increments of H/C and O/C

and reductions in the DBE-O value.

The averaged characteristics of CHO and CHOS formulae before chlorination and chlorinated DBPs were shown in Table 3. In comparison with the averaged values before chlorination, chlorinated DBPs had lower O/C_{avg} ($p < 0.005$ for CHO and CHOS), lower H/C_{avg} ($p < 0.005$ for CHO and $p < 0.05$ for CHOS) and higher $DBE-O_{avg}$ ($p < 0.005$ for CHO and $p < 0.05$ for CHOS). Low H/C_{avg} of DBPs implied that chlorinated DBPs were formed via electrophilic substitution of chlorine, or the reaction was selective to low H/C CHO molecules. In addition, as electrophilic substitution did not induce a change in the O/C and DBE-O value, the low O/C_{avg} and high $DBE-O_{avg}$ values implied the selective reaction of aromatic organic molecules with less oxygenated compounds probably due to the higher energy obtained from oxidation with more reduced compounds. This finding is in agreement with study of chlorination of drinking water treatment (Lavonen et al., 2013; Zhang et al., 2012).

It is also possible to extrapolate potential DBP precursors by assuming the possible chlorination reactions with organic matter. Lists of putative organic precursors of chlorinated DBPs (CHOCl class) formed by electrophilic substitution and addition reactions are given in Table 4. Only putative precursors of CHOCl were detected. Only 4 putative precursors were found for the chlorinated formula formed by an addition reaction while 18 precursors were found for those formed by electrophilic substitution. These findings supported the conclusion

that electrophilic substitution is a significant reaction in chlorination. All the precursors existed in the secondary effluent and detected in all the process water samples. About half of them were increased by the ozonation (Table 4). Other chlorinated DBPs, for which the precursor formulas were not found, would be formed by reaction with higher molecular weight precursors, as in the formation of organic halide and trihalomethanes from >1 kDa organic molecules (Hua and Reckhow, 2007; Sinsabaugh III et al., 1986). Hence, this unknown screening analysis provided a way to identify and study the fate of new DBPs and their LMW precursors by MS/MS analysis after chromatographic column separation.

3.8 Implications for Water Reclamation Plants

The findings in this study showed that Orbitrap MS is a promising tool for the unknown screening approach to qualitatively and semiquantitatively analyze a wide range of complex LMW-DOM in water reclamation processes. This also provides the possibility of detecting and tracking a large number of pollutants in a single analysis by a combination of chromatographic separation and MS/MS analysis when databases are available for those pollutants containing isotopic matches, fragment searches and the MS/MS library. The combination of MS/MS analysis with chromatographic column separation would make analysis more quantitative by reducing ion suppression by separation of the ion suppression molecules from susceptible molecules (Annesley, 2003). It is a great challenge to select the most appropriate column and pretreatment and elution conditions for unknown screening

analysis to cover most of the compounds and to avoid irreversible adsorption of the compounds to the sorbent. In addition, direct injection or incorporating a pre-concentration column (Olaf et al., 2014; Sleighter et al., 2009) would make routine online monitoring for a large number of pollutants in one analysis at either water reclamation or water purification plants a more realistic option. To achieve that outcome, it is necessary to establish databases of those commonly found organic compounds in wastewater streams so as to track their abundance as well as to identify newly emerging compounds. Automated software would be necessary to achieve such monitoring.

4. Conclusions

- Over 2000 elemental compositions and behaviors of LMW-DOM in complex organic mixtures from water reclamation processes were investigated by the unknown screening analysis using ESI Orbitrap MS. Most peaks were detected in the m/z range 100–450.
- LMW-DOM in water reclamation processes consisted of formulae with various heteroatoms such as CHO, CHON, CHOS, CHONS and Cl-containing formulae. CHO and CHOS formulae were most abundant in term of intensity.
- Van Krevelen diagrams revealed that CHOS formulae were relatively saturated while CHO formulae fell within the areas of lignin-like, unsaturated carbon and

black carbon-like molecules.

- Biofiltration preferentially removed molecules that were relatively reduced or highly saturated (high H/C). Ozonation induced the largest changes to LMW-DOM by removing wide ranges of compounds and producing more oxidized molecules (high O/C).
- After chlorination, over 100 chlorine-containing molecular formulae were additionally detected with high confidence level, and these were considered as chlorinated DBPs. Formulae of potential organic precursors were tracked, and those precursors originated not only from secondary effluent but also from treatment processes such as biofiltration and ozonation.

Acknowledgments

This study was conducted as a part of a Core Research for Evolutional Science and Technology (CREST) project supported by the Japan Science and Technology Agency, entitled ‘Development of well-balanced urban water use system adapted to climate change’. The authors also extend their acknowledgement to the Sumitomo Foundation for their financial support.

References

Annesley, Thomas M. 2003. Ion suppression in mass spectrometry. *Clinical chemistry*, 49

511 (7), 1041-1044.

512 Boutegrabet, L., Kanawati, B., Gebefugi, I., Peyron, D., Cayot, P., Gougeon, R. D., et al.
513 2012. Attachment of chloride anion to sugars: mechanistic investigation and
514 discovery of a new dopant for efficient sugar ionization/detection in mass
515 spectrometers. *Chemistry*, 18 (41), 13059-13067. doi: 10.1002/chem.201103788

516 Cortes-Francisco, N., and Caixach, J. 2013. Molecular characterization of dissolved organic
517 matter through a desalination process by high resolution mass spectrometry. *Environ*
518 *Sci Technol*, 47 (17), 9619-9627. doi: 10.1021/es4000388

519 Deborde, M., and von Gunten, U. 2008. Reactions of chlorine with inorganic and organic
520 compounds during water treatment-Kinetics and mechanisms: a critical review.
521 *Water Res*, 42 (1-2), 13-51. doi: 10.1016/j.watres.2007.07.025

522 Dittmar, Thorsten, Koch, Boris, Hertkorn, Norbert, and Kattner, Gerhard. 2008. A simple
523 and efficient method for the solid-phase extraction of dissolved organic matter
524 (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods*, 6, 230-235.

525 Eichhorn, Peter, Knepper, Thomas P, Ventura, Francesc, and Diaz, Alfredo. 2002. The
526 behavior of polar aromatic sulfonates during drinking water production: a case study
527 on sulfophenyl carboxylates in two European waterworks. *Water Res*, 36 (9),
528 2179-2186.

529 Fukushima, T., Hara-Yamamura, H., Urai, M., Kasuga, I., Kurisu, F., Miyoshi, T., et al.
530 2014. Toxicity assessment of chlorinated wastewater effluents by using
531 transcriptome-based bioassays and Fourier transform mass spectrometry (FT-MS)
532 analysis. *Water Res*, 52, 73-82. doi: 10.1016/j.watres.2014.01.006

533 Furumai, Hiroaki. 2008. Rainwater and reclaimed wastewater for sustainable urban water
534 use. *Physics and Chemistry of the Earth, Parts A/B/C*, 33 (5), 340-346. doi:
535 10.1016/j.pce.2008.02.029

536 Gonsior, M., Schmitt-Kopplin, P., Stavklint, H., Richardson, S. D., Hertkorn, N., and
537 Bastviken, D. 2014. Changes in dissolved organic matter during the treatment

538 processes of a drinking water plant in sweden and formation of previously unknown
539 disinfection byproducts. *Environ Sci Technol*, 48 (21), 12714-12722. doi:
540 10.1021/es504349p

541 Gonsior, M., Zwartjes, M., Cooper, W. J., Song, W., Ishida, K. P., Tseng, L. Y., et al. 2011.
542 Molecular characterization of effluent organic matter identified by ultrahigh
543 resolution mass spectrometry. *Water Res*, 45 (9), 2943-2953. doi:
544 10.1016/j.watres.2011.03.016

545 Gonsior, Michael, Peake, Barrie M, Cooper, William T, Podgorski, David, D'Andrilli,
546 Juliana, and Cooper, William J. 2009. Photochemically induced changes in dissolved
547 organic matter identified by ultrahigh resolution Fourier transform ion cyclotron
548 resonance mass spectrometry. *Environ Sci Technol*, 43 (3), 698-703.

549 Gribble, Gordon W. 2003. The diversity of naturally produced organohalogenes.
550 *Chemosphere*, 52 (2), 289-297. doi: 10.1016/s0045-6535(03)00207-8

551 Hockaday, William C, Purcell, Jeremiah M, Marshall, Alan G, Baldock, Jeffery A, and
552 Hatcher, Patrick G. 2009. Electrospray and photoionization mass spectrometry for
553 the characterization of organic matter in natural waters: a qualitative assessment.
554 *Limnology and Oceanography: Methods*, 7, 81-95.

555 Hua, Guanghui, and Reckhow, David A. 2007. Characterization of disinfection byproduct
556 precursors based on hydrophobicity and molecular size. *Environ Sci Technol*, 41 (9),
557 3309-3315.

558 Johnson, J Donald, and Jensen, James N. 1986. THM and TOX formation: routes, rates, and
559 precursors. *Journal (American Water Works Association)*, 156-162.

560 Kim, Sunghwan, Kaplan, Louis A, and Hatcher, Patrick G. 2006. Biodegradable dissolved
561 organic matter in a temperate and a tropical stream determined from ultra-high
562 resolution mass spectrometry. *Limnology and Oceanography*, 51 (2), 1054-1063.

563 Kim, Sunghwan, Kaplan, Louis A., Benner, Ronald, and Hatcher, Patrick G. 2004.
564 Hydrogen-deficient molecules in natural riverine water samples—evidence for the

565 existence of black carbon in DOM. *Marine Chemistry*, 92 (1-4), 225-234. doi:
566 10.1016/j.marchem.2004.06.042

567 Kim, Sunghwan, Kramer, Robert W, and Hatcher, Patrick G. 2003. Graphical method for
568 analysis of ultrahigh-resolution broadband mass spectra of natural organic matter,
569 the van Krevelen diagram. *Analytical chemistry*, 75 (20), 5336-5344.

570 Kind, T., and Fiehn, O. 2007. Seven Golden Rules for heuristic filtering of molecular
571 formulas obtained by accurate mass spectrometry. *BMC Bioinformatics*, 8, 105. doi:
572 10.1186/1471-2105-8-105

573 Krauss, M., Singer, H., and Hollender, J. 2010. LC-high resolution MS in environmental
574 analysis: from target screening to the identification of unknowns. *Anal Bioanal*
575 *Chem*, 397 (3), 943-951. doi: 10.1007/s00216-010-3608-9

576 Kujawinski, Elizabeth B, and Behn, Mark D. 2006. Automated analysis of electrospray
577 ionization Fourier transform ion cyclotron resonance mass spectra of natural organic
578 matter. *Analytical chemistry*, 78 (13), 4363-4373.

579 Kujawinski, Elizabeth B., Longnecker, Krista, Blough, Neil V., Vecchio, Rossana Del,
580 Finlay, Liam, Kitner, Joshua B., et al. 2009. Identification of possible source
581 markers in marine dissolved organic matter using ultrahigh resolution mass
582 spectrometry. *Geochimica et Cosmochimica Acta*, 73 (15), 4384-4399. doi:
583 10.1016/j.gca.2009.04.033

584 Lavonen, E. E., Gonsior, M., Tranvik, L. J., Schmitt-Kopplin, P., and Kohler, S. J. 2013.
585 Selective chlorination of natural organic matter: identification of previously
586 unknown disinfection byproducts. *Environ Sci Technol*, 47 (5), 2264-2271. doi:
587 10.1021/es304669p

588 Makarov, Alexander. 2000. Electrostatic axially harmonic orbital trapping: a
589 high-performance technique of mass analysis. *Analytical chemistry*, 72 (6),
590 1156-1162.

591 Matilainen, A., Gjessing, E. T., Lahtinen, T., Hed, L., Bhatnagar, A., et al. 2011. An

592 overview of the methods used in the characterisation of natural organic matter
 593 (NOM) in relation to drinking water treatment. *Chemosphere*, 83 (11), 1431-1442.
 594 doi: 10.1016/j.chemosphere.2011.01.018

595 Ohno, Tsutomu, He, Zhongqi, Sleighter, Rachel L, Honeycutt, C Wayne, and Hatcher,
 596 Patrick G. 2010. Ultrahigh resolution mass spectrometry and indicator species
 597 analysis to identify marker components of soil-and plant biomass-derived organic
 598 matter fractions. *Environ Sci Technol*, 44 (22), 8594-8600.

599 Olaf, Scheibner., Atrizia, van Baar., Florian, Wode., Uwe, Dünnebier., Kristi, Akervik.,
 600 Jamie, Humphrie., et al. 2014. A Strategy for an Unknown Screening Approach on
 601 Environmental Samples Using HRAM Mass Spectrometry. United States: Thermo
 602 Fisher Scientific.

603 Olsen, Jesper V, de Godoy, Lyris MF, Li, Guoqing, Macek, Boris, Mortensen, Peter, Pesch,
 604 Reinhold, et al. 2005. Parts per million mass accuracy on an Orbitrap mass
 605 spectrometer via lock mass injection into a C-trap. *Molecular & Cellular Proteomics*,
 606 4 (12), 2010-2021.

607 Pence, Harry E, and Williams, Antony. 2010. ChemSpider: an online chemical information
 608 resource. *Journal of Chemical Education*, 87 (11), 1123-1124.

609 Peter Herzsprung, Norbert Hertkorn, and Wolf von Tümpling, Mourad Harir, Kurt Friesen,
 610 Philippe Schmitt-Kopplin. 2014. Understanding molecular formula assignment of
 611 Fourier transform ion cyclotron resonance mass spectrometry data of natural organic
 612 matter from a chemical point of view. *Anal Bioanal Chem*, 406, 7977-7987. doi:
 613 10.1007/s00216-014-8249-y

614 Rapaport, RA, and Eckhoff, WS. 1990. Monitoring linear alkyl benzene sulfonate in the
 615 environment: 1973–1986. *Environmental Toxicology and Chemistry*, 9 (10),
 616 1245-1257.

617 Remucal, C. K., Cory, R. M., Sander, M., and McNeill, K. 2012. Low molecular weight
 618 components in an aquatic humic substance as characterized by membrane dialysis

619 and orbitrap mass spectrometry. *Environ Sci Technol*, 46 (17), 9350-9359. doi:
620 10.1021/es302468q

621 Rostad, Colleen E., and Leenheer, Jerry A. 2004. Factors that affect molecular weight
622 distribution of Suwannee river fulvic acid as determined by electrospray
623 ionization/mass spectrometry. *Analytica Chimica Acta*, 523 (2), 269-278. doi:
624 10.1016/j.aca.2004.06.065

625 Shon, H. K., Vigneswaran, S., and Snyder, S. A. 2006. Effluent Organic Matter (EfOM) in
626 Wastewater: Constituents, Effects, and Treatment. *Critical Reviews in*
627 *Environmental Science and Technology*, 36 (4), 327-374. doi:
628 10.1080/10643380600580011

629 Sinsabaugh III, Robert L, Hoehn, Robert C, Knocke, William R, and Linkins, Arthur E.
630 1986. Precursor size and organic halide formation rates in raw and coagulated
631 surface waters. *Journal of environmental engineering*, 112 (1), 139-153.

632 Sleighter, R. L., and Hatcher, P. G. 2007. The application of electrospray ionization coupled
633 to ultrahigh resolution mass spectrometry for the molecular characterization of
634 natural organic matter. *J Mass Spectrom*, 42 (5), 559-574. doi: 10.1002/jms.1221

635 Sleighter, Rachel L, McKee, Georgina A, Liu, Zhanfei, and Hatcher, Patrick G. 2008.
636 Naturally present fatty acids as internal calibrants for Fourier transform mass spectra
637 of dissolved organic matter. *Limnology and Oceanography: Methods*, 6 (6),
638 246-253.

639 Sleighter, Rachel L., McKee, Georgina A., and Hatcher, Patrick G. 2009. Direct Fourier
640 transform mass spectral analysis of natural waters with low dissolved organic matter.
641 *Organic Geochemistry*, 40 (1), 119-125. doi: 10.1016/j.orggeochem.2008.09.012

642 Swisher, Robert Donald. 1986. *Surfactant biodegradation* (Vol. 18): CRC Press.

643 Thayanukul, P., Kurisu, F., Kasuga, I., and Furumai, H. 2013. Evaluation of microbial
644 regrowth potential by assimilable organic carbon in various reclaimed water and
645 distribution systems. *Water Res*, 47 (1), 225-232. doi: 10.1016/j.watres.2012.09.051

- Trehy, Michael L, Gledhill, William E, Mieux, James P, Adamove, John E, Nielsen, Allen M, Perkins, Herbert O, et al. 1996. Environmental monitoring for linear alkylbenzene sulfonates, dialkyltetralin sulfonates and their biodegradation intermediates. *Environmental Toxicology and Chemistry*, 15 (3), 233-240.
- Tseng, Linda Y, Gonsior, Michael, Schmitt-Kopplin, Philippe, Cooper, William J, Pitt, Paul, and Rosso, Diego. 2013. Molecular Characteristics and Differences of Effluent Organic Matter from Parallel Activated Sludge and Integrated Fixed-Film Activated Sludge (IFAS) Processes. *Environ Sci Technol*, 47 (18), 10277-10284.
- UNEP/DTIE. 2006. SIDS INITIAL ASSESSMENT REPORT. OECD Initial Assessment Reports For High Production Volume Chemicals including Screening Information Datasets.
- Urai, M, Kasuga, I, Kurisu, F, and Furumai, H. 2014. Molecular characterization of dissolved organic matter in various urban water by using Orbitrap Fourier transform mass spectrometry.
- von Gunten, Urs. 2003. Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Res*, 37 (7), 1469-1487. doi: 10.1016/s0043-1354(02)00458-x
- Xu, Pei, Bellona, Christopher, and Drewes, Jörg E. 2010. Fouling of nanofiltration and reverse osmosis membranes during municipal wastewater reclamation: Membrane autopsy results from pilot-scale investigations. *Journal of Membrane Science*, 353 (1-2), 111-121. doi: 10.1016/j.memsci.2010.02.037
- Zhang, H., Zhang, Y., Shi, Q., Ren, S., Yu, J., Ji, F., et al. 2012. Characterization of low molecular weight dissolved natural organic matter along the treatment trait of a waterworks using Fourier transform ion cyclotron resonance mass spectrometry. *Water Res*, 46 (16), 5197-5204. doi: 10.1016/j.watres.2012.07.004

Figure legends

Fig. 1. Negative-ion ESI Orbitrap mass spectra (m/z 100–600) of (a) secondary effluent, (b) biofiltration effluent, (c) ozonation effluent and (d) reclaimed water

Fig. 2. Hierarchical cluster analysis of the water samples based on the presence and absence of the peaks in each sample

Fig. 3. Element compositions of the identified DOM in the processed water samples

Fig. 4. Plots of m/z of identified DOM and log-scaled intensities

Fig. 5. Van Krevelen diagrams, plots based on H/C and O/C ratios of the molecular formulae detected in secondary effluent, of (a) CHO and (b) CHOS formulae. Bubble sizes of the dots indicated the abundance relative to the highest intensity among the peaks. Note that dots might be overlapping at the same position.

Fig. 6. Van Krevelen diagrams of (a) CHO and (b) CHOS formulae in secondary effluent and/or biofiltration effluent. The plots were categorized by the change in peak intensity by the treatment. dot: newly formed; triangle: increased by more than 30%; filled cross: decreased by more than 30%; diamond: unchanged (less than 30 % change)

692

693 **Fig. 7.** Van Krevelen diagram of (a) CHO and (b) CHOS formulae in biofiltration effluent
694 and/or ozonation effluent. The plots were categorized by the change in peak intensity by the treatment.
695 dot: newly formed; triangle: increased by more than 30%; filled cross: decreased by more than 30%; diamond:
696 unchanged (less than 30 % change)

697

698 **Fig. 8.** Histograms of DBE-O values of the formulae detected before and after ozonation:
699 (a) CHO and (b) CHOS compounds.

700

701 **Table legends**

702 **Table 1.** Basic water quality characteristics of the samples in 2014

703

704 **Table 2.** Averaged elemental compositions of DOM in secondary effluent and in water
705 reclamation processes identified by negative ESI Orbitrap MS and FT-ICR-MS.

706

707 **Table 3.** Averaged values for molecular weight, H/C ratio, O/C ratio, and DBE-O of CHO and
708 CHOS compounds before chlorination and chlorinated compounds after chlorination.

709

710 **Table 4.** Chlorinated DBPs and putative precursors by electrophilic substitution and addition
711 reaction

1 **Table 1.** Basic water quality characteristics of the samples in 2014

Water samples	TOC (mg/L)	DOC (mg/L)	NO ₃ ⁻ (mg/L)	NH ₄ ⁺ (mg/L)	pH	SUVA (L/mg.m)	Free Cl ₂ (mg/L)	Cl ⁻ (mg/L)
Secondary effluent	4.56	3.72	35.9	0.72	6.1	2.80	-	49.0
Biofiltration effluent	3.79	3.42	37.9	0.06	7.0	2.81	-	48.7
Ozonation effluent	3.18	3.01	39.2	0.10	7.3	1.46	-	48.6
Reclaimed Water	3.50	3.27	39.0	0.02	7.4	1.25	0.86	50.8

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

Table 2. Averaged elemental compositions of DOM in secondary effluent and in water reclamation processes identified by negative ESI Orbitrap MS and FT-ICR-MS.

Class	Sample	No. of formulae	Averaged value for the compounds in the samples				Reference
			MW	H/C	O/C	DBE	
CHO	Secondary effluent	582	299.93	1.21	0.55	6.41	This study
	Biofiltration effluent	576	297.86	1.21	0.55	6.42	
	Ozonation effluent	569	298.69	1.21	0.58	6.29	
	Reclaimed water	457	277.36	1.22	0.58	5.87	
	Secondary effluent	979	356.4	1.23	0.47	7	Gonsior et al. (2011)**
	SRNOM	844	-	1.04	0.50	9.5	Cortes-Francisco and Caixach (2013)*
CHOS	Secondary effluent	165	264.52	1.29	0.56	4.82	This study
	Biofiltration effluent	130	266.60	1.27	0.55	4.87	
	Ozonation effluent	118	268.15	1.29	0.59	4.67	
	Reclaimed water	91	262.14	1.33	0.61	4.37	
	Secondary effluent	1197	314.4	1.48	0.4	4.7	Gonsior et al. (2011)**

*Asterisks denote formulae identified by *Orbitrap MS and **FT-ICR-MS

Table 3. Averaged values for molecular weight, H/C ratio, O/C ratio, and DBE-O of CHO and CHOS compounds before chlorination and chlorinated compounds after chlorination.

Sample	Formula Class	No. of Formulae	Averaged value for the formulae in classes		
			H/C _{avg}	O/C _{avg}	DBE-O _{avg}
Oz Effluent	CHO	569	1.21	0.58	-1.36
DBPs in Rm	CHOC1	98	1.11**	0.37**	1.60**
Oz Effluent	CHOS	118	1.29	0.59	-1.28
DBPs in Rm	CHOSCl	15	1.20**	0.43*	0.80*

*Asterisk denotes significant difference in a characteristic of formulae, as a group, between ozone effluent and chlorinated DBPs * $p < 0.05$ and ** $p < 0.005$.

Table 4. Chlorinated DBPs and putative precursors by electrophilic substitution and addition reaction

Putative precursor	Intensity relative to secondary effluent (%)				Chlorinated DBPs	
	Secondary effluent	Biofiltration effluent	Ozonation effluent	Reclaimed water	m/z (error in ppm)	Chemical formula
Electrophilic substitution						
C ₁₀ H ₁₀ O ₄	100	77	37	20	227.0115 (0.61)	C ₁₀ H ₉ O ₄ Cl
C ₁₁ H ₁₈ O ₄	100	67	53	34	247.0731 (4.78)	C ₁₁ H ₁₇ O ₄ Cl
C ₁₀ H ₁₀ O ₃	100	87	58	35	211.0166 (0.84)	C ₁₀ H ₉ O ₃ Cl
C ₁₆ H ₂₄ O ₇	100	92	67	36	361.1058 (0.41)	C ₁₆ H ₂₃ O ₇ Cl
C ₁₂ H ₁₆ O ₂	100	85	72	37	259.0291 (2.65)	C ₁₂ H ₁₄ O ₂ Cl ₂
C ₉ H ₆ O ₃	100	73	82	64	194.9857 (1.31)	C ₉ H ₅ O ₃ Cl
C ₁₀ H ₁₄ O ₄	100	73	86	39	231.0430 (0.04)	C ₁₀ H ₁₃ O ₄ Cl
C ₁₀ H ₁₂ O ₄	100	84	88	43	229.0271 (0.92)	C ₁₀ H ₁₁ O ₄ Cl
C ₁₁ H ₁₄ O ₆	100	81	97	43	308.9945 (2.29)	C ₁₁ H ₁₂ O ₆ Cl ₂
C ₁₀ H ₁₂ O ₃	100	91	99	46	213.0320 (1.85)	C ₁₀ H ₁₁ O ₃ Cl
C ₉ H ₁₂ O ₄	100	88	108	54	217.0270 (1.24)	C ₉ H ₁₁ O ₄ Cl
C ₉ H ₁₂ O ₂	100	91	124	57	185.0369 (3.13)	C ₉ H ₁₁ O ₂ Cl
C ₁₀ H ₁₄ O ₆	100	80	143	78	296.9938 (0.05)	C ₁₀ H ₁₂ O ₆ Cl ₂
C ₉ H ₁₀ O ₆	100	88	166	108	280.9626 (0.34)	C ₉ H ₈ O ₆ Cl ₂
C ₁₀ H ₁₂ O ₇	100	77	179	107	310.9734 (1.27)	C ₁₀ H ₁₀ O ₇ Cl ₂
C ₇ H ₆ O ₅	100	115	207	141	236.9362 (0.49)	C ₇ H ₄ O ₅ Cl ₂
C ₁₀ H ₁₄ O ₇	100	86	220	122	312.9892 (1.49)	C ₁₀ H ₁₂ O ₇ Cl ₂
C ₈ H ₁₀ O ₅	100	77	244	163	252.9677 (0.21)	C ₈ H ₈ O ₅ Cl ₂
Addition reaction						
C ₁₁ H ₈ O ₇	100	72	90	54	355.9702 (1.99)	C ₁₁ H ₁₀ O ₉ Cl ₂
C ₁₁ H ₁₆ O ₆	100	73	93	64	400.0095 (3.12)	C ₁₁ H ₁₉ O ₉ Cl ₃
C ₈ H ₈ O ₅	100	71	143	90	287.9804 (0.63)	C ₈ H ₁₀ O ₇ Cl ₂
C ₁₁ H ₁₆ O ₃	100	90	146	81	248.0815 (4.79)	C ₁₁ H ₁₇ O ₄ Cl ₁

Fig. 1. Negative-ion ESI Orbitrap mass spectra (m/z 100–600) of (a) secondary effluent, (b) biofiltration effluent, (c) ozonation effluent and (d) reclaimed water

Fig. 2. Hierarchical cluster analysis of the water samples based on the presence and absence of the peaks in each sample

Fig. 3. Element compositions of the identified DOM in the processed water samples

Fig. 4. Plots of m/z of identified DOM and log-scaled intensities

Fig. 5. Van Krevelen diagrams, plots based on H/C and O/C ratios of the molecular formulae detected in secondary effluent, of (a) CHO and (b) CHOS formulae. Bubble sizes of the dots indicated the abundance relative to the highest intensity among the peaks. Note that dots might be overlapping at the same position.

Fig. 6. Van Krevelen diagrams of (a) CHO and (b) CHOS formulae in secondary effluent and/or biofiltration effluent. The plots were categorized by the change in peak intensity by the treatment. dot: newly formed; triangle: increased by more than 30%; filled cross: decreased by more than 30%; diamond: unchanged (less than 30 % change)

Fig. 7. Van Krevelen diagram of (a) CHO and (b) CHOS formulae in biofiltration effluent and/or ozonation effluent. The plots were categorized by the change in peak intensity by the treatment. dot: newly formed; triangle: increased by more than 30%; filled cross: decreased by more than 30%; diamond: unchanged (less than 30 % change)

Fig. 8. Histograms of DBE-O values of the formulae detected before and after ozonation: (a) CHO and (b) CHOS compounds.

Graphical abstract