Spectroscopic and Molecular-Level Characteristics of Dissolved Organic Matter in a Highly Polluted Urban River in South China

Quanhui Ye,‡ Zi-Ting Zhang,‡ Yu-Chen Liu,‡ Ying-Hui Wang,‡ Song Zhang,‡ Chen He,§ Quan Shi,§ Hai-Xiang Zeng,§ and Jun-Jian Wang‡

†Guangdong Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China
‡State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
§State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing 102249, China

ABSTRACT: China is undergoing unprecedented rapid urbanization, which is accompanied by the formation of more than 2000 black and odorous urban rivers. Dissolved organic matter (DOM) is an essential component in aquatic ecosystems and regulates the carbon and nutrient cycle, but its characteristics in highly polluted urban rivers are still not well understood. Here, water samples from the mainstream, tributaries, ponds, and reservoirs within the highly urbanized Maohou River watershed (Shenzhen, China) were analyzed using absorption and fluorescence spectroscopy, stable carbon isotope analysis, and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) to explore the impact of urbanization on DOM characteristics. The results showed elevated dissolved organic carbon concentrations in the disturbed waters (mainstream, tributary, and pond) compared to those in the reservoir water. Results of the spectroscopic, isotopic, and FT-ICR MS analyses suggest that disturbed water DOM had lower aromaticity, average molecular weight, and plant- and soil-derived terrestrial (e.g., lignin-like) contributions but greater contributions from microbial-derived protein-like and anthropogenic sulfur-containing compounds (e.g., synthetic surfactants) than those of reservoir DOM. These characteristics were strengthened from the upper to the lower mainstream, which implies greater anthropogenic impact on downstream waters. Point-source wastewater effluent input caused increases in DOM level in the mainstream water and abrupt changes in DOM characteristics with very strong anthropogenic characteristics. The distinct chemistry of urban DOM from natural organic matter highlights a potentially strong impact of rapid urbanization on carbon and nutrient cycle in aquatic ecosystems.

KEYWORDS: Urbanization, wastewater, Fourier transform ion cyclotron resonance mass spectrometry, molecular signatures, spectroscopic characters, dissolved organic matter

1. INTRODUCTION

Dissolved organic matter (DOM) is a highly complicated mixture and is actively involved in many critical biogeochemical processes. It serves as essential carbon and energy sources to fuel heterotrophic microorganisms and can be converted to greenhouse gases.1 DOM also serves as an important carrier for heavy metals2,3 and organic pollutants4 in aquatic environments and contributes to the toxic disinfection byproduct formation in water supply.5,6 Moreover, chromophoric DOM can absorb sunlight and reduce the depth of the euphotic zone, thereby altering the associated primary production and photobleaching processes.7 Consequently, DOM quantity and quality exert a remarkable impact on both water quality and aquatic ecosystems. As approximately 87% of land surface on the Earth are connected to oceans via rivers, riverine DOM is an essential mobile carbon pool that transports massive terrestrial materials to the estuaries.8

China is the largest developing country with an unprecedented rapid urbanization rate.9 The gross urban built-up area in China has increased from 35,500 km² in 2007 to 56,200 km² in 2017.10 The expanding population and intensified human activities in fast-growing urban areas have led to the formation of at least 2100 black and odorous rivers in China.11 Generally, riverine DOM in remote regions is composed of natural organic matter (NOM), primarily derived from terrestrial plants and soils as well as aquatic microbes. However, with...
additional DOM inputs from agricultural runoff, urban runoff, and industrial and domestic wastewater discharge, the DOM composition in urban rivers tends to be further complicated by rapid urbanization. For example, Meng et al. found that the levels of dissolved organic carbon (DOC) and tryptophan- and tyrosine-like fluorescent DOM components were significantly elevated in the Guangzhou section of the Pearl River in China. Similarly, Liu et al. showed that urbanized subcatchments had notably high abundance of protein-like components but low abundance of humic-like components relative to those of forested subcatchments in Dongjiang River of the Pearl River. These studies primarily focused on the properties and fluorescent components of DOM by means of ultraviolet−visible spectroscopy, fluorescent spectroscopy, and parallel factor analysis (PARAFAC). There is a paucity of information concerning the molecular-level characteristics of DOM in black and odorous rivers in highly urbanized areas. Understanding the molecular-level characteristics of DOM in highly urbanized areas is critical to provide novel insights into the source and fate of urban DOM and the impact of urbanization on regional carbon cycle and environmental quality.

In expanding urban areas, treated and untreated wastewater comprises an increasingly large fraction of surface waters in urbanized districts. Previous studies employed pyrolysis gas chromatography−mass spectrometry (Py-GC-MS) to examine the molecular-level composition of wastewater effluent organic matter (EfOM). For example, EfOM was more aliphatic and had higher concentrations of organic nitrogen and halogen compounds than NOM, and these organic signatures persisted in downstream river. In contrast, another Py-GC-MS study found that EfOM was dominated by aromatic fragments, such as benzene, phenol, and methyl-pyridine. However, Py-GC-MS is constrained to capture several hundreds of chemical fragments, and large parts of the chemical components of wastewater-derived DOM remain unresolved. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is one of the most powerful mass spectrometric techniques to unravel the diverse molecular compositions of the complex DOM. Gonsior et al. analyzed EfOM using FT-ICR MS and identified abundant sulfur-containing CHOS formulas (up to 90% of the total peak intensity), featured with many diverse synthetic surfactants (i.e., linear alkylbenzenesulfonates and sulfophenyl carboxylic acids). Recently, He et al. analyzed six mainstream DOM samples of the urbanized Daliao River in Northeast China based on FT-ICR MS and found that the peak intensity of CHOS (7.8−13.6% with an average of 11.5%) was significantly higher than those of the world’s ten largest rivers (average of 5.7%). However, little is known about how wastewater inputs change the molecular composition of DOM in highly polluted black and odorous urban rivers.

In this study, we collected 84 water samples (mainstream, tributary, pond, reservoir, and wastewater) from the highly urbanized Maozhou River watershed in Shenzhen, China. The...
Maozhou River is one of the most polluted black and odorous rivers in the Pearl River Watershed of China. The spectroscopic, carbon isotopic, and molecular-level characteristics of DOM were analyzed using absorption and fluorescence spectroscopy, stable carbon isotope analysis, and FT-ICR MS. The DOM characteristics and composition from different water types (mainstream, tributary, pond, and reservoir) and spatial locations (upper vs lower watersheds; wastewater-discharged locations vs nearby locations) were compared. This study will provide fundamental reference for future water management and ecological restoration of black and odorous rivers in China.

2. MATERIALS AND METHODS

2.1. Site Description and Water Sampling. Shenzhen is among the fastest-growing cities not only in China but also in the whole world during the past few decades. From being a small fishing village with tens of thousands of people 40 years ago, Shenzhen has developed into a large international metropolis with a population of tens of millions. Shenzhen had the highest gross domestic product in the Guangdong–Hong Kong–Macao Greater Bay Area in 2018. Shenzhen is often studied as a representative metropolis to understand the environmental consequences of rapid urbanization. The Maozhou River is the largest and longest river in Shenzhen, and 52.1% of its watershed was occupied by residential and industrial land, serving approximately 30% of Shenzhen’s population in 2017. The Maozhou River originates from the Yangtai Mountain, running through metropolitan Shenzhen (approximately 41.6 km) and draining approximately 388 km² before entering the Pearl River Estuary. The Maozhou River is composed of three main sections: the Dapig River in the upper watershed, the Yangchong River in the middle watershed, and the Dongbao River in the lower watershed. The sections of the Dapig River and the Yangchong River belong to the Shenzhen city, while the Dongbao River is the boundary between the cities of Shenzhen and Dongguan. The Maozhou River is one of the most polluted rivers in the Pearl River Watershed. More than 10 trillions of Chinese yuan (~1.5 trillions of US dollars) have been invested by the government to improve the water quality of the Maozhou River. However, there is no information about the DOM composition in the Maozhou River for river management and restoration in this watershed.

In January 2019, 84 surface water samples were collected using an acid-washed glass-bottle sampler (prerinsed with the water sample in triplicate) from various locations within the Maozhou River watershed. The sampling sites included 22 mainstream sites (from upper to lower: M1–M22), 43 tributary sites (from T1–1 to T21), 11 pond sites (P1–P11), six reservoir sites (R1–R6), and two wastewater samples (W1 and W2). Notably, W1 was discharged in a mainstream location between M18 and M19, while W2 was discharged in a tributary near T20–1. Sites M9 and M15 receive effluents from wastewater treatment plants, both of which use the anaerobic–anoxic–oxic treatment process and have wastewater loading of 150,000 m³/d. However, no wastewater effluent was collected there because no sewage pipe was found above the river water level at either site. Mainstream, tributary, and pond waters are highly disturbed by anthropogenic activities and categorized as disturbed waters, while the less impacted reservoirs are considered as reference sites. Photographs of several representative sampling sites are available in the Supporting Information (SI Figure S1). The whole sampling campaign was conducted within 5 days, and the collected water samples were immediately transported on ice to the laboratory for filtration. All samples were filtered through 0.7 μm glass fiber filters (precombusted for 4.5 h at 450 °C) and stored at 4 °C before further analyses.

2.2. Sample Analyses. The pH and electricity conductivity (EC) of the samples were measured using a HACH HQ440d benchtop meter equipped with Intelllicel PHC201 pH and CDC401 conductivity probes. DOC and total dissolved nitrogen (TDN) were measured by a total organic carbon (TOC)-L CSH/CSN analyzer with an ASI-L autosampler (Shimadzu, Japan). The stable carbon isotope composition of DOC (δ13CDOC) of the filtered water samples was determined using isotope ratio mass spectrometry (253 Plus, Thermo Scientific, Germany) at Southern University of Science and Technology. The sample preparation and instrument settings followed the procedures in Wang et al. δ13CDOC was corrected for blank and instrument fractionation and was finally calibrated against the Vienna Pee Dee Belemnite (VPDB) standard with a precision of <0.3‰.

The absorbance values at 254 nm (A254) and 365 nm (A365) of the filtered samples were measured using an ultraviolet–visible spectrometer (Horiba, Japan). The water samples (diluted to A254 < 0.1 at 25 °C) were further characterized by measuring emission–excitation matrices (EEMs) using an Aqualog spectrofluorometer (Horiba, Japan) [excitation (Ex): 240–550 nm with 3 nm interval; emission (Em): 246.58–828.17 nm with approximately 1 nm interval]. All EEMs were blank-subtracted, corrected for the instrument-related and inner-filter effects, and further normalized to the Raman unit. We used PARAFAC to decompose the EEMs into different components, and the model was validated based on the split-half analysis. All components were further compared to the previously reported models in the OpenFluor database. A four-component model was finally established, and the relative abundance of each component was determined as f_max/EF_max.

The spectroscopic indices were calculated to characterize the DOM. For example, the specific UV absorbance at 254 nm (SUVA254 in L/mg C/m), calculated by dividing A254 by the DOC concentration, was used to indicate DOM aromaticity. The E2/E3 ratio, having a negative correlation with the molecular weight (MW) of DOM, was calculated as the ratio of A254 to A365. The fluorescence index (FI), an indicator of the DOM source (approximately 1.8: microbial origin; approximately 1.2: terrestrial source), was calculated as the ratio of the fluorescence intensity at Em 470 nm to that at 520 nm when Ex was 370 nm. The humification index (HIX) was determined by dividing the regional area of 435 nm < Em < 480 nm by that of 300 nm < Em < 345 nm when Ex was 254 nm. The biological index (BIX), a metric for the contribution of newly produced autochthonous DOM content, was calculated as the ratio of Em intensity at 380 nm to the maximum Em between 420 and 435 nm when Ex was 310 nm.

To obtain molecular-level insights into the impact of urbanization on DOM chemistry, eight samples (M1, M6, M9, M13, M19, M22, W1, and R6) were further analyzed by FT-ICR MS. Values of the extracted (pH = 2) samples were prepared to achieve a target concentration of 45 µg C/mL. The DOM was solid-phase extracted (SPE) using PPL Bond Elut (Agilent) resins, following the method in...
Spencer et al. The SPE-DOM was eluted with 2 mL of methanol and dried under nitrogen gas. The extracts were reconstituted with methanol and analyzed in negative mode (ESI source) using a 9.4 T Apex-ultra FT-ICR MS (Bruker Daltonics) at the China University of Petroleum, Beijing. A total of 100 time-domain free induction decay signals were added for each sample before the Fourier transformation and a mass range of m/z 200–850 was recorded. The externally calibrated mass spectra were further calibrated with typical O5-class species peaks in Bruker Daltonics DataAnalysis 4.4. After calibration, a resolving power (m/Δm50%) at m/z 319 over 45,000, and a low absolute mass error (<1 ppm) were achieved when analyzing all samples. The raw data were converted by selecting all ions with a signal-to-noise threshold over 6 and an absolute intensity threshold of 100. The elemental formula is assigned to each monoisotopic mass using Formularity software with constraints of C0–100, H0–200, O0–50, N0–10, S0–2, H/C < 2.5, O/C < 1.2, N/C < 0.5, and S/C < 0.2. The formulas were assigned into four elemental groups (CHO, CHON, CHOS, and CHONS). Formulas with isotopomers (i.e., 13C, 18O, and 34S) were not discussed here. Double bond equivalents (DBE) and the modified aromaticity index (AImod) were calculated to evaluate the unsaturation degree and aromaticity of each compound, respectively. In addition, the molecular lability boundary for more labile contributions (MLB) and island of stability (IOS), two metrics for DOM lability, were calculated as respectively. Intensity-weighted methods were applied to calculate the average MW, DBE, AImod, and H/C and O/C ratios. Seven biochemical groups were assigned following the classification method of Wu et al. with small modifications, which include (1) lipid-like (0 ≤ O/C ≤ 0.2, 1.5 ≤ H/C ≤ 2.3, N/C ≤ 0.04, and S = 0), (2) protein-like (0.2 ≤ O/C ≤ 0.52, 1.5 ≤ H/C ≤ 2.2, and 0.178 ≤ N/C ≤ 0.44), (3) amino-sugar-like (0.52 ≤ O/C ≤ 0.7, 1.5 ≤ H/C ≤ 2.2, 0.07 ≤ N/C ≤ 0.182, and S = 0), (4) carbohydrate-like (0.7 ≤ O/C ≤ 1.1, 1.5 ≤ H/C ≤ 2.4, N = 0, and S = 0), (5) condensed-aromatics-like (0 ≤ O/C ≤ 0.25, 0.5 ≤ H/C ≤ 1.25, and S = 0), (6) lignin-like (0.25 ≤ O/C ≤ 0.67, 0.75 ≤ H/C ≤ 1.5, and S = 0), and (7) tannin-like (0.67 ≤ O/C ≤ 0.97, 0.53 ≤ H/C ≤ 1.5, and S = 0). The relative abundances of different biochemical groups were determined as the ratio of their signals to the total signal of all assigned formulas.

2.3. Statistical Analyses. Significant differences in chemical parameters between the disturbed waters and reservoir waters, and among the mainstream, tributary, pond, and reservoir waters, were performed using one-way analysis of variance (ANOVA) and Tukey’s post hoc test. Parameters that do not satisfy the assumptions (e.g., homogeneity of variance and normally distributed data) of one-way ANOVA were alternatively examined with the nonparametric Kruskal–Wallis test and Dunn’s post hoc multiple comparisons. The parameters of water quality, spectroscopic indices, and fluorescent components were standardized and evaluated with principal component analysis (PCA).

3. RESULTS AND DISCUSSION

3.1. Variations in General Water Quality. The pH values were 7.4 ± 0.2 (mean ± standard deviation) for mainstream, 7.7 ± 0.5 for tributary, 8.3 ± 0.3 for pond, and 7.3 ± 0.1 for reservoir waters (Figure 2A,B, SI Table S1). The EC values were higher in the mainstream (2.1 ± 3.4 mS/cm), tributary (3.1 ± 4.9 mS/cm), and pond waters (0.8 ± 0.5 mS/cm) than in the reservoir waters (0.1 ± 0.03 mS/cm) (Figure 2A,B, SI Table S1). These values are comparable to those commonly reported for Maozhou River and some other urban rivers (SI Table S2). Taking the mainstream, tributary, and pond waters as a whole (disturbed waters), the disturbed waters (pH: 7.7 ± 0.3; EC: 2.4 ± 4.3 mS/cm) had significantly higher pH and EC than those of the reservoir waters (pH: 7.3 ± 0.1; EC: 0.1 ± 0.03 mS/cm) (SI Table S1). The DOC and TDN levels of mainstream (DOC: 4.0 ± 0.8 mg/L; TDN: 15.4 ± 2.0 mg/L), tributary (DOC: 6.9 ± 6.4 mg/L; TDN: 12.4 ± 10.5 mg/L), and pond (DOC: 12.0 ± 6.0 mg/L; TDN: 14.6 ± 26.9 mg/L) waters were higher than those of reservoir waters (DOC: 3.2 ± 1.0; TDN: 0.9 ± 0.7 mg/L) (Figure 2C,D). The great EC values and TDN concentrations clearly indicate a high degree of water pollution in the Maozhou River. The DOC and TDN levels for tributary waters ranged from 1.5 to 29.0 mg/L and from 1.3 to 42.8 mg/L, respectively, with more than 10-fold differences. Similarly, the DOC concentration of pond waters ranged from 7.0 to 29.6 mg/L, and the TDN level varied markedly from 0.5 to 90.7 mg/L. The high variations in DOC and TDN for the tributary and pond waters indicate various degrees of anthropogenic carbon and nitrogen inputs.

The DOC and TDN concentrations displayed insignificant difference between the water types of mainstream and tributary waters, but great spatial variations from the upper to the lower watershed (Figure 2C,D). When the river went through the highly populous area (e.g., M11-M21 and T7-T21), both the DOC and TDN levels in the mainstream and tributary waters increased (SI Figure S2). Within mainstream waters, the DOC level increased from 2.8 to 5.9 mg/L from M1 to M21, and the TDN level from 13.6 to 18.5 mg/L. The decreased DOC (4.7 mg/L) and TDN (12.7 mg/L) concentrations in M22 may be...
Figure 3. Optical indices and percentages of four fluorescent components for the mainstream, tributary, pond, and reservoir waters. (A) SUVA$_{254}$: specific ultraviolet absorbance at 254 nm; (B) E2/E3; (C) Fl: fluorescence index; (D) HIX: humification index; (E) BIX: biological index; (F) C1: component 1; (G) C2: component 2; (H) C3: component 3; and (I) C4: component 4. The boxplot percentile distribution is the same as that in Figure 2. The lowercase letters indicate significantly different groups.

due to the mixing and dilution with seawater at the mouth of this river (as indicated by the high EC of 15.8 mS/cm of M22). Compared to the mainstream waters, the tributary waters had great variations in both DOC and TDN concentrations. Specifically, the DOC and TDN levels for the tributaries (T12–T17) of the lower watershed within Dongguan city could reach 29.0 and 42.8 mg/L, respectively, but no more than 5.9 mg/L and 8.1 mg/L for the tributaries in the upper watershed (SI Figure S2). Within a single tributary, increasing trends in both DOC and TDN concentrations were also observed from the upper to lower tributary (SI Figure S2). Taken together, our results suggest a continuous anthropogenic carbon and nitrogen inputs into the mainstream and tributary waters.

3.2. Variations in Spectroscopic Characteristics of DOM. To investigate the effect of the massive anthropogenic inputs on DOM chemistry, the spectroscopic indices and fluorescent components of the samples were determined. The SUVA$_{254}$ and HIX for the disturbed waters (SUVA$_{254}$: $1.87 \pm 0.46$ L/mg C/m; HIX: $2.61 \pm 0.66$) were significantly lower than those for reservoir waters (SUVA$_{254}$: $2.01 \pm 0.13$ L/mg C/m; HIX: $3.98 \pm 0.64$) (Figure 3A,D, SI Table S3). Low values of SUVA$_{254}$ and HIX commonly indicate low aromaticity and humification degree, respectively, which could be due to the large area of impervious surface in the watershed (52.1%) and low plant- and soil-derived terrestrial DOM inputs. In contrast, the E2/E3 and BIX values were significantly higher in the disturbed waters (E2/E3: $5.66 \pm 0.46$; BIX: $0.98 \pm 0.07$) than in the reservoir waters (E2/E3: $5.34 \pm 0.21$; BIX: $0.90 \pm 0.01$; Figure 3B,E, SI Table S3). High E2/E3 and BIX values are indicative of large DOM contributions from low-molecular-weight microbial sources. Thus, lower SUVA$_{254}$ and HIX and higher E2/E3 and BIX values collectively suggest more microbial-derived DOM in the disturbed waters. The Fl values for mainstream (2.08 ± 0.03), tributary (1.98 ± 0.14), and pond waters (1.85 ± 0.15) were also higher than those for reservoir waters (1.71 ± 0.06) (Figure 3C). Interestingly, Fl values for the disturbed waters were out of the typical range of 1.2–1.8 for natural water, implying that the former could be impacted by exogenous inputs (e.g., wastewater). Similar to the DOC and TDN values, which showed large variations, the spectroscopic indices also displayed no significant difference between mainstream and tributary waters but great spatial variations from upper to the lower watershed. SUVA$_{254}$ decreased from 1.95 to 1.55 L/mg C/m from M1 to M22, and HIX decreased from 3.89 to 2.13, whereas E2/E3 showed an opposite increasing trend from 5.45 to 7.11. Within tributary waters, similar decreasing trends were observed for SUVA$_{254}$ and HIX (SI Figure S2). These results suggest a continuously increasing microbial-derived DOM from upstream to downstream, which is consistent with the increasing DOC and TDN levels.

PARAFAC modeling is useful for resolving fluorescent DOM components. Four validated components (C1, C2, C3, and C4) were extracted (SI Figure S3, Table S4). C1 is assigned to the terrestrially derived humic-like component. C2 is a reprocessed humic-like component, which is often detected in wastewater. C3 is a tyrosine-like component, and C4 is regarded as a tryptophan-like component, and both are often detected in wastewater and autochthonous production (e.g., bacterial and algal materials). Significantly lower relative abundances for humic-like C1 and C2 were observed in the mainstream (C1: 21% ± 4%, C2: 21% ± 5%) than in the reservoir waters (C1: 29% ± 5%, C2: 26% ± 5%) (Figure 3F & G). In contrast, C3 is a tyrosine-like component, and C4 is regarded as a tryptophan-like component, and both are often detected in wastewater and autochthonous production. In the mainstream (C3: 26% ± 5%, C4: 32% ± 5%), tributary (C3: 28% ± 5%, C4: 34% ± 5%) and pond waters (C3: 24% ± 5%, C4: 21% ± 5%) compared to those in the reservoir waters (C3: 17% ± 5%, C4: 18% ± 5%) (Figure 3H). This result suggests that tyrosine-like and tryptophan-like DOM inputs possibly from wastewater, agricultural and urban runoff, and autochthonous
production have disturbed most of the mainstream, tributary, and pond samples. The total relative abundances of C3 and C4 gradually increased from the upper to the lower watershed, indicating that the downstream waters received increasing anthropogenic and microbial inputs in this highly urbanized watershed.

PCA was applied to summarize the majority of variables explaining DOM properties by reducing the dimensions of variables and clustering the sample data. The first two principal components explained 62.2% of the total variation in all parameters included (PC1: 42.7%, PC2: 19.5%). EC, TDN, FI, BIX, C3, and C4 had positive PC1 loadings, while DOC/TDN, HIX, and humic-like C1 and C2 exhibited negative PC1 loadings (Figure 4A). As the Maozhou River watershed was surrounded by the highly urbanized Shenzhen City and C3 and C4 displayed high PC1 loadings, PC1 can be associated with anthropogenic and autochthonous tyrosine-like and tryptophan-like DOM inputs. However, PC2 is associated with high DOC and TDN concentrations and low SUVA254 values. The reservoir water data clustered, which could be explained by similar HIX and percentages of the two humic-like components (C1 and C2). Compared to reservoir waters, the mainstream, tributary, and pond waters displayed high PC1 scores (Figure 4B), suggesting large tyrosine-like and tryptophan-like DOM inputs. Interestingly, within the mainstream waters, the PC1 score increased notably from negative to positive values from the upper to the lower mainstream (SI Figure S3), which is in good agreement with the increase in C3 and C4 components. In addition, the mainstream and reservoir samples had relatively stable PC2 scores, which reflects the relatively stable DOC, TDN, and SUVA254 levels. It suggests that the large water volumes of mainstream and reservoirs may have buffered possible disturbance to some extent. However, the large variations in PC2 scores of tributary and pond waters are associated with the great differences in DOC, TDN, and SUVA levels (Figures 2 and 4). This finding suggests that DOM quantity and quality in tributary and pond waters are highly susceptible to anthropogenic disturbance. It may be because, compared to the mainstream, the tributary and pond waters have much smaller water volumes that caused weaker dilution effect for anthropogenic inputs. In addition, the difference in water retention time, flow speed, and geomorphology might also affect the degradation and aggregation of DOM in different water bodies. Overall, our results suggest that tyrosine-like and tryptophan-like DOM inputs have largely impacted and contributed to the DOM pool in disturbed waters. The anthropogenic DOM, which features low aromaticity, small humification degree, low molecular size, and high microbial contributions, has a continuously increasing impact when river water flow through the highly populous area from the upper to the lower sites.

3.3. Impact of Point-Source Wastewater Effluents on Mainstream DOM. In addition to the aforementioned variations in water quality and DOM characteristics, the DOC and TDN levels for M9 and M15 (sites receiving effluents from wastewater treatment plants) sharply increased compared to M8 and M14, respectively (Figure 5A,B). This clearly indicates that wastewater effluents strongly elevate the DOM level in urban streams. Similar results have also been found in wastewater-discharged stream waters (e.g., Xihe River in Shenyang, China; DuPage River in DuPage County, USA). The wastewater effluents not only alter DOM quantity but also DOM chemistry, as indicated by the spectroscopic properties. Specifically, M9 and M15 had lower SUVA254 and HIX values but higher E2/E3 and FI values than

---

**Figure 4.** (A) Loadings of water properties on principal components 1 and 2 and the (B) scores of the mainstream, tributary, pond and reservoir water samples. EC: electrical conductivity; DOC: dissolved organic carbon; TDN: total dissolved nitrogen; SUVA254: specific ultraviolet absorbance at 254 nm; FI: fluorescence index; HIX: humification index; BIX: biological index; C1−C4: four fluorescent dissolved organic matter components.

**Figure 5.** Chemical properties of mainstream waters. (A) DOC: dissolved organic carbon; (B) TDN: total dissolved nitrogen; (C) SUVA254: specific ultraviolet absorbance at 254 nm; (D) E2/E3; (E) FI: fluorescence index; (F) HIX: humification index; (G) δ13C; (H) δ15N; C1−C4: four fluorescent dissolved organic matter components. The red arrows indicate the water receiving immediate discharge of treated wastewater.
those of M8 and M14, respectively (Figure 5C–F), i.e., having more similar SUVA$_{254}$, HIX, E2/E3, and FI values to the wastewater samples W1 and W2 (SI Table S3). Lower SUVA$_{254}$ and higher FI values have also been observed in some other riverine DOM near the wastewater-discharge region.$^{34,62}$ Interestingly, the total relative abundances of C3 and C4 for M9 and M15 increased sharply by 11% and 8% compared to those for M8 and M14, respectively. Previous studies showed that stream waters receiving wastewater effluents had obvious signatures related to proteins and amino sugars.$^{15,20}$ Thus, despite the possibility of autochthonous production, the sensitive increases in C3 and C4 fluorescent signals for M9 and M15 can be mainly attributed to wastewater effluent discharge. All these characteristics of wastewater effluents strongly support the impact of wastewater effluents on lowering the aromaticity, humification degree, and molecular size of DOM and increasing the microbial contributions at M9 and M15.

The $\delta^{13}$C values for M9 ($−26.6\%$) and M15 ($−26.4\%$) were higher than their respective upper mainstream sites M8 ($−26.7\%$) and M14 ($−27.0\%$) and those of reservoir waters (R1: $−28.5\%$; R2: $−28.4\%$; R6: $−28.3\%$) (Figure 5G). Higher $\delta^{13}$C values were usually found for microbial-derived DOC (approximately $−21\%$) than terrestrial DOC (approximately $−27\%$) in the natural environment.$^{35,64}$ Although the two wastewater samples we collected showed very high microbial contributions, they had relatively low $\delta^{13}$C values (W1: $−26.6\%$; W2: $−27.4\%$) similar to those for the plant- and soil-derived terrestrial DOM. Table 1. Molecular Characterization of the Solid-Phase-Extracted Dissolved Organic Matter As Determined by ESI FT-ICR MS.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Number of assigned molecules</th>
<th>Molecular weight (Da)</th>
<th>Average molecular formulas</th>
<th>Average H/C atom ratio</th>
<th>Average O/C atom ratio</th>
<th>Average DBE</th>
<th>Average AImod</th>
<th>% Molecular lability (MLB)</th>
<th>% Island of stability (IOS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R6</td>
<td>5015</td>
<td>387</td>
<td>C$<em>{18.3}$H$</em>{22.6}$O$<em>{8.4}$N$</em>{0.4}$S$_{0.1}$</td>
<td>1.24</td>
<td>0.46</td>
<td>8.2</td>
<td>0.26</td>
<td>19.06</td>
<td>14.17</td>
</tr>
<tr>
<td>W1</td>
<td>4499</td>
<td>358</td>
<td>C$<em>{16.7}$H$</em>{21.1}$O$<em>{7.3}$N$</em>{0.4}$S$_{0.4}$</td>
<td>1.27</td>
<td>0.44</td>
<td>7.36</td>
<td>0.24</td>
<td>12.89</td>
<td>10.50</td>
</tr>
<tr>
<td>M1</td>
<td>928</td>
<td>360</td>
<td>C$<em>{17.0}$H$</em>{21.5}$O$<em>{7.2}$N$</em>{0.4}$S$_{0.2}$</td>
<td>1.27</td>
<td>0.43</td>
<td>7.5</td>
<td>0.25</td>
<td>17.46</td>
<td>13.58</td>
</tr>
<tr>
<td>M6</td>
<td>5229</td>
<td>366</td>
<td>C$<em>{17.2}$H$</em>{21.8}$O$<em>{7.7}$N$</em>{0.3}$S$_{0.3}$</td>
<td>1.27</td>
<td>0.45</td>
<td>7.48</td>
<td>0.24</td>
<td>13.53</td>
<td>17.79</td>
</tr>
<tr>
<td>M9</td>
<td>5046</td>
<td>356</td>
<td>C$<em>{17.1}$H$</em>{21.7}$O$<em>{7.3}$N$</em>{0.4}$S$_{0.4}$</td>
<td>1.27</td>
<td>0.43</td>
<td>7.5</td>
<td>0.25</td>
<td>10.59</td>
<td>14.17</td>
</tr>
<tr>
<td>M13</td>
<td>5046</td>
<td>356</td>
<td>C$<em>{17.3}$H$</em>{21.9}$O$<em>{7.7}$N$</em>{0.3}$S$_{0.3}$</td>
<td>1.27</td>
<td>0.44</td>
<td>7.48</td>
<td>0.24</td>
<td>13.53</td>
<td>17.79</td>
</tr>
<tr>
<td>M19</td>
<td>5046</td>
<td>356</td>
<td>C$<em>{16.9}$H$</em>{22.0}$O$<em>{7.0}$N$</em>{0.4}$S$_{0.5}$</td>
<td>1.3</td>
<td>0.42</td>
<td>7.5</td>
<td>0.25</td>
<td>10.59</td>
<td>14.17</td>
</tr>
<tr>
<td>M22</td>
<td>4726</td>
<td>342</td>
<td>C$<em>{16.0}$H$</em>{21.2}$O$<em>{6.6}$N$</em>{0.3}$S$_{0.6}$</td>
<td>1.32</td>
<td>0.41</td>
<td>7.5</td>
<td>0.25</td>
<td>13.53</td>
<td>17.79</td>
</tr>
</tbody>
</table>

Collectively, our results support that wastewater effluent is an important point source of anthropogenic DOM in urban rivers and may greatly alter DOM aromaticity, humification degree, molecular size, and chemical composition.

3.4. Molecular-Level Characteristics for Anthropogenic DOM Inputs. FT-ICR MS was employed for six mainstream waters (M1, M6, M9, M13, M19, and M22), one wastewater effluent (W1), and one reservoir water sample (R6) to examine how wastewater inputs affect the molecular-level characteristics and compositions of DOM. The average DOM MW values of W1 (358 Da) and mainstream waters (342–367 Da) were notably lower than those of R6 (387 Da; Table 1). Additionally, the average DBE and AImod values in both W1 (DBE: 7.36; AImod: 0.24) and mainstream waters (DBE: 6.69–7.54; AImod: 0.21–0.25) were lower than those in R6 (DBE: 8.20; AImod: 0.26) (Table 1), indicating that the DOM of W1 and mainstream waters had relatively low unsaturation degree and aromaticity. Meanwhile, M9 (wastewater-discharge site) displayed lower average MW and aromaticity (indicated by the DBE and AImod values) relative to those of M6. All these results are consistent with the variations in E2/E3 and SUVA$_{254}$ (Table S2) and confirm that anthropogenic DOM had low MW and aromaticity. Also, W1 and mainstream waters had higher MLB$_{b}$ (W1: 13.93%; mainstream: 10.09–20.31%) and lower IOS (W1: 14.17%; mainstream: 13.58–17.79%).
mainstream: 10.50−17.46%) than those of R6 (MLBL: 8.90%; IOS: 19.06%; Table 1). As MLBL and IOS have a positive and negative relationship with DOM lability, respectively, this result indicates that the DOM lability of the wastewater and the disturbed mainstream water was much higher than that of reservoir water. Likewise, the point-source wastewater inputs increased DOM lability as indicated by the abrupt change in MLBL and IOS for M9 compared to that for M6. Such high lability is possibly due to the high protein-like composition in the wastewater compared to that of the natural plant- and soil-derived materials and may result in high greenhouse gas emission from the disturbed urban river waters.

The characteristics of assigned DOM formulas for W1 and mainstream waters were different from those for R6 (SI Figure S5), which shows that urbanization alters the DOM molecular composition. The DOM formulas were classified into seven biochemical groups. The lignin-like compounds always constituted the largest relative abundances (>39%), followed by tannin-like compounds (1.4−3.9%), and lipid-like and condensed aromatic-like compounds (Figure 6). The percentages of tannin-like, lipid-like, and protein-like compounds increased for M9 compared to those for M6. Lipid-like and protein-like compounds were considered as labile DOM pool, leading to high DOM lability for M9. Previous work by Sirivedhin and Gray also identified more protein precursors in water receiving wastewater effluent based on Py-GC-MS. All these results are consistent with the spectroscopic and isotopic findings and provide details of molecular-level alteration by urbanization disturbance and particularly point-source wastewater inputs.

In addition to the alteration in ubiquitous biochemical composition, surprisingly high amounts of anthropogenic components were present in the disturbed waters. Interestingly, the relative abundance of CHOS compounds in W1 (36.5%) and mainstream waters (25.8−50.9%) was 2- to 4-fold higher than that in R6 (12.3%) (Figure 7A). The relative abundance of CHOS compounds in W1 (50.9%) was 2- to 4-fold higher than that of R6 (12.3%) (Figure 7A). The relative abundances of diatoms and plant-like compounds belonging to the O3S and O5S classes were identified (Figure 7C,D), which include linear alkylbenzene sulfonates (LAS), dialkyl tetralin sulfonates (DATS), and dialkyl tetralin sulfonate intermediates (DATSI). The relative abundances of O3S + O5S, SAS + DATS, and O3S + SAS were all considerably higher for mainstream waters than for R6 (Figure 7C). Meanwhile, from M6 to M9, the proportions of LAS + DATS and SAS + DATSI increased by 73.7% and 48.0%, respectively (Figure 7C). This suggests that LAS, DATS, SAS, and DATSI are important molecular signatures for urbanization disturbance and point-source wastewater inputs. As the respective biodegradation metabolite of LAS and DATS, SPC and DATSI displayed 5.0−11.6-fold higher proportion than that for LAS and DATS. Clearly, LAS and DATS can be efficiently

Figure 6. Relative abundances of biochemical components of solid-phase extracted dissolved organic matter using Fourier transform ion cyclotron resonance mass spectrometric analyses based on the method in Wu et al. The red arrow indicates the wastewater-discharged site.

Figure 7. Elemental composition of the solid-phase extracted DOM and the relative abundances of different kinds of anthropogenic compounds. (A) Formulas; (B) O3S + O5S; (C) LAS + DATS; and (D) SPC + DATSI. LAS: linear alkylbenzene-sulfonates; DATS: dialkyl tetralin sulfonates; SPC: sulfolanyl carboxylic acids; DATSI: dialkyl tetralin sulfonate intermediates. The red arrow indicates the wastewater-discharged site.
removed in wastewater treatment process, but their biodegradation metabolites display slow further biodegradation and are thus more persistent in the highly polluted urban river. Previous study also found concentrations of SPC and DATSI were consistently higher in the effluents than in the influents.\textsuperscript{24} Thus, more attention should be given on the associated aquatic toxicity and ecosystem impacts of these metabolites. In addition to surfactant-related CHOS formulas, hundreds of additional CHOS compounds were identified and assigned, and these compounds occurred in M9 with over 50% higher abundance than in M6 (SI Table S6). It is possible that sulfur from wastewater react with CHO compounds and produce these sulfur-containing molecules.\textsuperscript{72} Interestingly, CHOS compounds from point-source wastewater tend to be labile, as indicated by the decreased relative abundance of CHOS compounds for M13 (27.9%) relative to that for M9 (32.0%). Meanwhile, the abundance of single CHOS molecule decreased by 25−50% from M9 to M13 (SI Table S6). One reason for this could be dilution effect due to the increased water volume from upper to middle mainstream (SI Figure S1). Another possible explanation is that CHOS compounds were susceptible to microbial or photochemical degradation\textsuperscript{73} during their transport.

3.5. Environmental Implications. Here, significantly elevated DOC levels were observed in the disturbed waters (mainstream, tributary, and pond) than in the reservoir waters. Massive anthropogenic inputs (i.e., point-source wastewater effluent) have greatly altered DOM chemistry, which featured with lower aromaticity, smaller humification degree, lower molecular size, and higher contributions from microbial-derived protein-like and anthropogenic sulfur-containing compounds (e.g., synthetic surfactants). This work provides molecular-level insight into DOM characteristics in addition to the DOM quantity in a highly polluted urban river, which will be instrumental for river restoration in fast-growing Chinese cities. DOM characteristics associated with its composition and lability helps to understand how urbanization impacts DOM pool and fate. CHOS compounds including synthetic surfactants significantly increased, suggesting their role as a significant contribution to DOM pool in the urban river.\textsuperscript{24,25} In spite of dilution effect, the decrease in the relative abundance of nitrogen- and sulfur-rich DOM from wastewater-discharged sites to lower sites implies the high lability of these compounds (as reflected by MLB\textsubscript{i} and IOS). On the one hand, DOM is susceptible to binding with inorganic contaminants (e.g., heavy metals)\textsuperscript{74,75} and organic pollutants,\textsuperscript{76} both of which displayed increased concentrations in urbanized rivers.\textsuperscript{27,78} This will impact the migration and environmental fate of these pollutants, and further investigations on the relationship between DOM and contaminants will help to better understand how urbanization affects urban streams as a whole. On the other hand, the nitrogen- and sulfur-rich DOM could be degraded into CO\textsubscript{2},\textsuperscript{79,80} via photodegradation or microbial degradation, which could greatly increase greenhouse emissions and impact the regional and possibly global carbon cycle as urbanization-impacted streams will continue to increase around the world.

4. CONCLUSIONS

Based on the spectroscopic, isotopic, and FT-ICR MS analyses, this study demonstrates that urbanization disturbance not only elevated the DOM quantity but also shifted the DOM chemistry. When (1) an urban river flows through a highly populous area from the upper to the lower watershed or (2) receives direct point-source wastewater inputs, the river DOM obtains strong signatures of low aromaticity, MW, and plant- and soil-derived terrestrial (e.g., lignin-like) contributions but great contributions from microbial protein-like and anthropogenic sulfur-containing compounds (e.g., synthetic surfactants). Notably, the Maozhou River contained 25.8−50.9% of CHOS of the identified formulas from FT-ICR MS, which is possibly the greatest number reported to date for river water. This result implies that in highly urbanized areas, river DOM should not be considered as “natural” and might have distinct eco-environmental functions (e.g., greenhouse gas emissions and pollutant binding and transport) compared to those of NOM. As China is expected to continuously expand its urban areas, anthropogenic signatures tend to increase in urban river DOM unless effective water management and river restoration are implemented.

**ASSOCIATED CONTENT**

**Supporting Information**
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsearthspacechem.9b00151.

Tables showing the general water quality of different types of water samples; previously reported pH, EC, and DOC concentrations in urban streams; spectroscopic indices and percentages of four fluorescent components of different types of water samples; identification of the four extracted fluorescent components; loadings in factor analysis; spatial variation of some representative CHOS molecules. Figures showing photos of sampling sites; tributary nutrient level and spectroscopic indices; loadings of the four components on the excitation and emission spectra; scores of PC1 for mainstream; van Krevelen diagrams (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**
*E-mail: wangjj@sustech.edu.cn.*

**ORCID**
Quan Shi: 0000-0002-1363-1237

**Author Contributions**
J.-J.W. designed this study. Q.H.Y. and J.-J.W. prepared the manuscript. All other authors contributed to the data interpretation and manuscript revision.

**Notes**
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We thank Dr. Kai Yu for his assistance in stable carbon isotope analysis. This work was funded by Guangdong University Featured Innovation Project (2018KTSCX199), National Science Foundation of China (41807360), National Key Research and Development Program of China (2018YFA0605800), Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control (No. 2017B030301012), and State Environmental Protection Key Laboratory of Integrated Surface-Water-Groundwater Pollution Control.
REFERENCES


