Persistence and removal of trace organic compounds in centralized and decentralized wastewater treatment systems

Natalie Mladenov[1*], Nathan Dodder[2,3], Lauren Steinberg[1], William Richardot[3], Jade Johnson[2], Bice Martincigh[4], Chris Buckley[5,§], Tolulope Lawrence[4], Eunha Hoh[2]

[1] Department of Civil, Construction, and Environmental Engineering, San Diego State University, San Diego, CA, USA 92182

[2] School of Public Health, San Diego State University, San Diego, CA, USA 92182

[3] San Diego State University Research Foundation, San Diego, CA, USA 92182

[4] School of Chemistry and Physics, University of KwaZulu-Natal, Westville Campus, Durban, 4000, South Africa

[5] Water, Sanitation & Hygiene Research & Development Centre, School of Engineering, University of KwaZulu-Natal, Durban, 4041, South Africa

* Corresponding author: nmladenov@sdsu.edu, +1(720)233-2007, 5500 Campanile Drive, San Diego State University, San Diego, CA, USA 92182

§ Deceased.

Highlights

- Non-targeted analysis (NTA) identified 111 common WWTP influent compounds
- 11 of the common influent compounds were persistent in the studied WWTP effluents
- Anaerobic treatment reduced chemical peak areas more than conventional treatment
- NTA revealed 5 new compounds not previously reported in WWTP effluent
- Physico-chemical properties did not differ among persistent and removed compounds

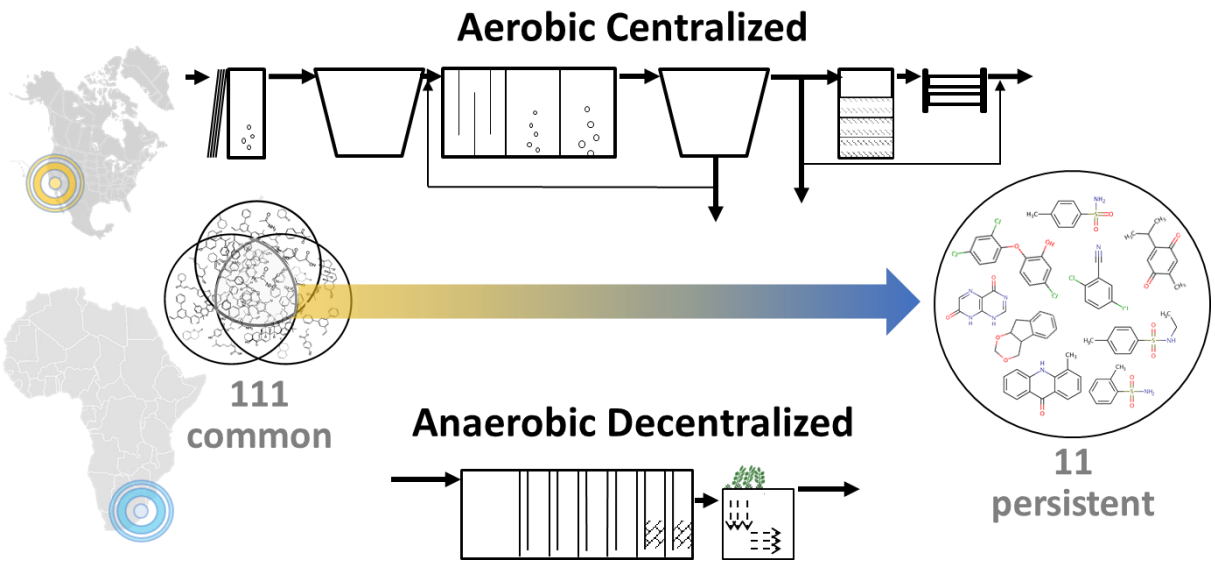
Abstract

The persistence of trace organic chemicals in treated effluent derived from both centralized wastewater treatment plants (WWTPs) and decentralized wastewater treatment systems (DEWATS) is of concern due to their potential impacts on human and ecosystem health. Here, we utilize non-targeted analysis (NTA) with comprehensive two-dimensional gas chromatography coupled with time of flight mass spectrometry (GC×GC/TOF-MS) to conduct an evaluation of the common persistent and removed compounds found in

two centralized WWTPs in the USA and South Africa and one DEWATS in South Africa. Overall, removal efficiencies of chemicals were similar between the treatment plants when they were compared for the number of chemical features detected in the influents and effluents of each treatment plant. However, the DEWATS treatment train, which has longer solids retention and hydraulic residence times than both of the centralized WWTPs and utilizes primarily anaerobic treatment processes, was able to remove 13 additional compounds and showed greater overall compound removal compared to the centralized WWTPs. Of the 111 common compounds tentatively identified in all three influents, 11 compounds were persistent in all replicates, including 5 compounds not previously reported in effluents of WWTPs or water reuse systems. There were no significant differences among the physico-chemical properties of persistent and removed compounds, but significant differences were observed among some of the molecular descriptors. These results have important implications for the treatment of trace organic chemicals in centralized and decentralized WWTPs and the monitoring of new compounds in WWTP effluent.

Keywords: non-targeted analysis; anaerobic baffled reactor; decentralized wastewater treatment systems (DEWATS); GC×GC/TOF-MS; pharmaceuticals and personal care products (PPCPs); molecular descriptors.

Graphical abstract



1. Introduction

The introduction of trace organic compounds to the environment through wastewater treatment and water reuse system effluent is a widespread problem because wastewater treatment technologies may be inefficient at removing them (Hamza et al., 2016) and many are considered chemicals of emerging concern (i.e., those compounds not included in current monitoring programs and may have adverse effects on humans and ecosystems). Studies seeking to evaluate the persistent and removed compounds in wastewater treatment plants (WWTPs) and water reuse systems have relied on targeted analysis of known compounds (see for example Anumol and Snyder, 2015; Taheran et al., 2016; Cecconet et al., 2017; Grandclément et al., 2017; Tran et al., 2018). As non-targeted analysis (NTA) approaches are increasingly being applied to this area of study, a wide diversity of compounds, many of which would be otherwise missed using compound specific techniques alone, is being uncovered (Blum et al., 2018). NTA is a class of full-scan mass spectrometry methods designed to acquire a molecular ion mass-to-charge (m/z) and/or fragmentation mass spectrum of every chromatographic peak, within the chromatographic, ionization, and sensitivity limitations of the instrument (Albergamo et al., 2019). The observed molecular features (distinct but unidentified compounds) may be used to categorize sample groups (Parry, 2016) or to prioritize chemical identifications (Köppe, 2020). Hug et al. (2013) developed NTA techniques for

screening compounds in wastewater effluent using liquid chromatography-high resolution mass spectrometry. Gago-Ferrero () developed a workflow to detect and identify suspect and unknown contaminants in Greek wastewater using liquid chromatography coupled to a quadrupole-time-of-flight mass spectrometer (LC-QTOF-MS). Blum et al. (2019) screened trace organic chemicals in wastewater using both gas chromatography and liquid chromatography mass spectrometry-based targeted and untargeted analysis, which were found to be complimentary methods for identifying different compounds in onsite sewage treatment facilities (OSSFs) in Sweden.

Despite WWTPs and water reclamation facilities not being designed specifically for the removal of trace organic compounds, there is some degree of removal of these chemicals in the treatment train (Ashfaq et al., 2017; Krzeminski et al., 2019). Although initially, much of the data on the removal of trace organic chemicals was compiled for systems treating synthetic wastewater or small bench-scale systems (Taheran et al., 2016; Cecconet et al., 2017; Grandclément et al., 2017; Tran et al., 2018), studies of the persistence and removal of chemicals of emerging concern in full-scale wastewater treatment systems are on the rise (Krzeminski et al., 2019; Soriano-Molina, 2019; Pérez et al., 2020). However, studies using non-targeted screening to evaluate trace organic compounds in WWTP effluent or through wastewater treatment processes have been conducted mainly in Europe (e.g., Hug et al., 2013; Singer et al., 2016; Blum et al., 2018; Tousova et al., 2018; Nurenburg et al., 2019; Lara Martin et al., 2020; Scholee et al., 2021; Tisler et al., 2021), the USA (e.g., Parry and Young, 2016; Kumar et al., 2021), and Asia (e.g., Choi et al., 2021; Qian et al., 2021), with studies in Africa and South America being less common.

In a review of emerging contaminants in South African waters by Gani et al. (2021) found only 41 studies focusing on emerging contaminants in water matrices, including surface waters, water treatment plants, and wastewater treatment or storage systems. Only two of those studies, K'oreje et al. (2012) and Gumbi et al. (2019), used nontargeted screening approaches to study emerging contaminants, but those were for rivers in Kenya and South Africa, and no studies using non-targeted screening to evaluate compound removal or discharge from WWTPs were identified in that review. One study has recently been published

using a non-targeted screening approach with ultrahigh performance liquid chromatography quadrupole time of flight tandem mass spectrometry (UHPLC-QTOF-MS) to evaluate emerging contaminants, focusing on azole antifungals, in a wastewater treatment plant in Pretoria, South Africa (Assress et al., 2019). In order to understand, at a global scale, the discharge to the environment of emerging contaminants that are not yet known or included in monitoring programs, more non-targeted screening studies are needed from data-scarce regions, like Africa.

Additionally, of the studies evaluating trace organic chemicals in real wastewater treatment or in non-potable water reuse facilities, the vast majority have focused on centralized systems, which primarily rely on activated sludge or other aerobic biological treatment processes (Soliman et al., 2007; Tran and Gin, 2017), and only a few have focused on mainstream anaerobic systems (Blum et al., 2019; Harb et al., 2019; Wang et al., 2020). In a recent review of municipal WWTPs around the world, Tran et al. (2018) found removal efficiencies of ECs to be wide ranging and called for more studies using alternatives to centralized treatment at the full-scale. Blum et al. (2017) compared the removal of chemicals in on-site systems, including septic systems, to removal in large-scale sewage treatment plants in Sweden using a combination of non-targeted screening and targeted analysis, and found greater removal of some of the more polar contaminants in the onsite soil beds compared to the conventional WWTPs. Another alternative to the conventional WWTP is decentralized wastewater treatment systems (DEWATS). DEWATS are non-sewered, waterborne treatment systems operating at scales larger than on-site systems but smaller and lower cost than centralized systems (Massoud et al., 2009) (Figure 1). DEWATS, which employ primarily anaerobic technologies such as anaerobic baffled reactors (ABRs), biogas chambers, and anaerobic filters, are increasingly being employed for the provision of sanitation especially in densely populated low-income areas around the world (Reynaud and Buckley, 2015). DEWATS are widely used in Africa, Asia, and Latin America for primarily residential wastewater treatment, but also for schools and universities, hospitals, and emergency sanitation (BORDA, 2017). Abafe et al. (2018) conducted a targeted analysis of the persistence and removal of 13 HIV antiretroviral drugs in effluents of both

centralized and decentralized wastewater treatment systems in South Africa using a surrogate standard based liquid chromatography with triple-quadrupole mass spectroscopy (LC-MS/MS) method. To our knowledge no studies have conducted a wide screening of trace organic compounds in DEWATS.

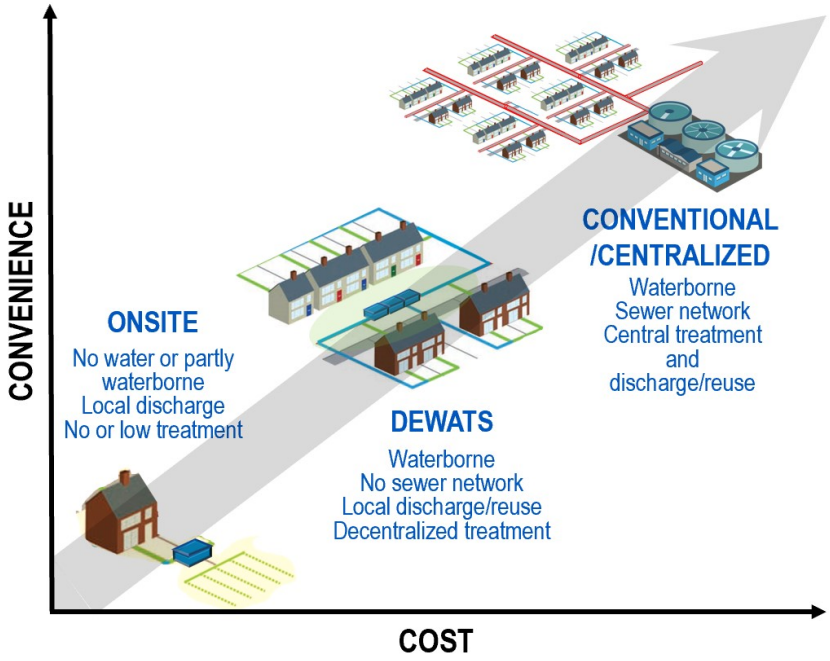


Figure 1. Position of DEWATS compared to onsite systems and centralized WWTPs in terms of convenience and cost. (Adapted from BORDA, 2020).

Therefore, the overarching goal of this study was to evaluate the persistence and removal of trace organic chemicals identified by NTA in both centralized and decentralized WWTPs, which use predominantly aerobic and anaerobic treatment technologies, respectively. Influent and effluent samples were collected from a centralized WWTP and a DEWATS in Durban, South Africa and a centralized WWTP in San Diego County, USA. We employed a NTA approach using comprehensive two-dimensional gas chromatography coupled with time of flight mass spectrometry (GC×GC/TOF-MS) to first focus on the identification of common compounds present in influent samples of all three WWTPs. From this set of common compounds, we then identified compounds that persisted in all effluents, and compounds that were completely removed through the aerobic centralized and anaerobic decentralized treatment trains.

131 Physico-chemical properties and molecular descriptors were further explored to evaluate significant
132 differences between persistent and removed compounds and among treatment types.

133 2. Methods

134 2.1 Sample collection

135 Treatment plants receiving primarily domestic wastewater were selected in the USA and South Africa,
136 and included two centralized WWTPs that employ conventional activated sludge treatment and one
137 decentralized facility that employs ABR-AF technology to achieve primary and secondary treatment and
138 constructed wetlands for polishing. The three treatment plants were selected to be geographically distinct
139 and to produce finished water used for non-potable water reuse.

140 The centralized treatment plant in San Diego, California, USA (US-C) is a medium sized conventional
141 aerobic activated sludge (CAS) treatment plant with advanced water purification processes, treating
142 > 19,000 m^3/d of primarily domestic wastewater (Table 1), and provides reclaimed water, meeting
143 Title 22 requirements for the State of California, USA (CCR, 2015) for irrigation and other non-potable
144 uses. Biological activated sludge treatment occurs in one of two aeration basins, fitted with baffle walls to
145 create selector zones (one anaerobic followed by two aerated zones). Following secondary clarification,
146 approximately 9,100 m^3/d of treated effluent feeds the recycled water facility, which includes four
147 continuously backwashing sand filters, and a chlorine contact basin. In order to meet reclaimed water
148 discharge limits (CCR, 2015), a side-stream (5,300 m^3/d) is diverted after the secondary treatment stage
149 for advanced purification processes, including microfiltration and reverse osmosis, and this stream is
150 combined with secondary treated effluent prior to chlorine disinfection. The discharge permit does not
151 require either of the two streams to be designed for nutrient removal. Final effluent is stored in reservoirs
152 until final distribution of the recycled water or final discharge to the ocean.

153 **Table 1.** Operating conditions of the three wastewater treatment facilities sampled.

Site	Treatment goal	Date sampled	Location	Average population served	Average flow rate (m^3/d)	HRT[1] (d)

US-C	Reclaimed water for landscape irrigation	5-6 February 2019	San Diego County, USA	40,000	19,873	0.14
SA-C	Nutrient removal for river discharge	19 – 21 August 2018	Durban, South Africa	181,695	23,006	0.63
SA-D	COD, TSS removal for agricultural water reuse	19 – 21 August 2018	Durban, South Africa	410	41.6	1.4

[1] Hydraulic residence time (HRT) for the activated sludge biological treatment processes for US-C and SA-C and for the ABR-AF biological treatment processes for SA-D.

The centralized WWTP in Durban in the KwaZulu-Natal province of South Africa (referred to as SA-C) treats > 23,000 m^3/d of wastewater (Table 1) from predominately residential and commercial land use areas, with light industrial usage. The SA-C facility implements aerobic activated sludge treatment and Bardenpho processes for nutrient removal. Biological processes are carried out in two zones, anoxic and anaerobic, which operate to reduce nitrogen and phosphorus simultaneously. Secondary treated effluent is diverted, passed through subsequent treatment processes including disinfection and maturation pond, and discharged into river outfall.

The DEWATS (referred to as SA-D), also located in Durban, South Africa, is a facility that treats 41.6 m^3/d from an urban neighborhood, which was approximately 410 residents at the time of this study. The domestic wastewater flows from the neighborhood by gravity to SA-D, which is situated at a lower elevation. Its treatment processes include a settling basin, three parallel ABR trains, each of which feeds two anaerobic coarse bed media filters, and two gravity-fed, vegetated constructed wetlands (vertical flow, followed by horizontal flow). Wetland effluent supplies water for experimental agricultural plots located at the DEWATS site, growing banana and other crops for scientific research.

Influent and treated effluent samples were collected in winter seasons (August for South Africa and February for the US; Table 1) as triplicates from the same 24-h composite samples (hourly sampling intervals) for US-C and SA-D and as a single 12-h composite sample (hourly sampling intervals) collected during daytime hours for SA-C (due to time constraints and site access limitations). Time and budget constraints allowed for sample collection on only one date at each of the WWTPs. The 24-h

composite influent and effluent samples for SA-D were collected 24 h apart to better track the same water parcel in SA-D, given its longer hydraulic residence time (HRT) (Table 1). All samples were collected in clean, pre-combusted, amber glass 1 L containers and either preserved with sodium azide and ascorbic acid and stored at 6 °C overnight, or extracted immediately. Field blanks consisted of an open container of LC/MS grade water left open during the sampling period near the effluent location. One field blank was collected at each site and processed in the same manner as the samples.

2.2 Physical-chemical analyses

Basic water quality measurements of pH and conductivity were measured on composite samples after collection. A Fisherbrand Accumet AP85 portable pH and conductivity meter (Lenexa, USA) was used for US-C. At SA-C, an Inolab pH/conductivity meter (Xylem Analytics, Weilheim, Germany) was used for conductivity measurements, and an 827 Lab Metrohm Swiss made pH meter (Herisau, Switzerland) was used for pH measurements. At SA-D, a 340i/Set WTW Wissenschaftlich pH and conductivity meter (Xylem Analytics, Weilheim, Germany) was used. Aliquots of influent and effluent samples were acidified with sulfuric acid to pH 2 and analyzed for total chemical oxygen demand (COD) using a HACH DR3900 spectrophotometer. Additional unfiltered and unacidified aliquots collected from US-C influent and effluent were analyzed for total suspended solids, performed according to Method 2540 (APHA, 2017) using a 1.6 μm pre-combusted glass fiber filter. Due to logistical challenges with sample transport, TSS concentrations were only measured on one of the three replicates for SA-D (using Method 2540, as described above) and were not measured for SA-C composite samples. Instead, average TSS concentrations for the week of sampling, provided by the SA-C staff, were used.

2.3 Sample preparation for non-targeted analysis

A solid-phase extraction (SPE) method was used to prepare samples for GC×GC/TOF-MS analyses. Prior to extraction, samples were filtered through 1.6 μm Whatman GF-A filters. The filtrate was run through Oasis HLB (Hydrophillic-Lipophillic Balance) cartridges (Waters, Milford, MA), which contain a universal polymeric reversed-phase sorbent, chosen for their ability to capture a wide range of analytes.

201 They have been successfully implemented for the environmental analysis of pharmaceuticals and personal
202 care products in wastewater, as well as pesticides in surface waters (Subedi et al., 2015; Luis Malvar et
203 al., 2018). Before loading the sample on to the cartridges, the cartridges were rinsed with 5 mL of
204 dichloromethane (DCM) and 5 mL of acetone and subsequently conditioned with 10 mL methanol and 15
205 mL LC/MS grade water. Filtered water samples (extraction volumes provided in in Supplemental Table
206 S1) were loaded to the cartridge at a flow rate of 3-5 mL/min (1-2 drops/s), and then the cartridges were
207 vacuum dried for an additional 10 minutes. The loaded cartridges were stored frozen and shipped on ice
208 to the U.S. where additional elution steps were completed.

209 Compounds were then eluted from the cartridges into pre-combusted test tubes using 5 mL of acetone and
210 5 mL DCM. The resulting solvent extract was dried by adding 5 g of sodium sulfate, pre-baked at 400˚C
211 for four hours, to each test tube. Then, 6 mL EnviroClean muffled sodium sulfate cartridges were loaded
212 onto the manifold with new pre-combusted test tubes for collection. The remaining solvent was pipetted
213 from the test tubes with sodium sulfate into the sodium sulfate cartridges for further drying, and flushed
214 with 2 mL of DCM. The extracts were then concentrated by evaporating with nitrogen gas, in a 40˚C
215 water bath, to 0.4 mL. The final extracts were stored at -20˚C before analysis.

216 **2.4 Non-targeted chemical analysis using GC×GC/TOF-MS**

217 Non-targeted analysis including chemical screening using a mass spectral library may be performed using
218 gas chromatography and liquid chromatography (LC)-based techniques, and the LC-based non-targeted
219 analysis is more common for water-soluble, semivolatile or nonvolatile organic pollutants in water
220 (Hollender et al., 2017). The utilization of GC×GC/TOF-MS for NTA may identify different chemicals
221 compared to alternate NTA methods (Blum et al., 2019), but studies have observed overlap among the
222 chemicals identified by GC and LC based methods. For example, the ENTACT study compared
223 identification of 1269 chemicals in mixtures by an LC/Q-TOF (+ESI and - ESI) and a GC×GC/TOF-MS
224 method (Ulrich et al., 2019). Successful identification of the chemicals overlapped by 40% among the two
225 methods. However, the GC×GC/TOF-MS method unexpectedly identified smaller molecular weight

compounds with higher water solubility, indicating the GC×GC/TOF-MS is suitable for WWTP sample analysis. Additionally, the > 200,000 compound NIST EI mass spectral library is compatible with the GC×GC/TOF-MS method, while libraries available for LC/electrospray based methods are considerably smaller (Blum et al., 2019; Schymanski et al., 2015).

Samples were analyzed by Pegasus 4D GC×GC/TOF-MS (LECO, St. Joseph, MI); detailed instrumental conditions are in Supplemental Table S2. LECO ChromaTOF software (version 4.50.8.0, optimized for the Pegasus) was used for data processing that, for each sample, generated a list of features with distinct chromatographic peaks and associated fragmentation mass spectra. Once processed, LECO's software add-in Statistical Compare was used to align the features across samples in a peak table based on retention time similarity and mass spectral similarity. Detailed data processing and Statistical Compare conditions can be found in Supplemental Table S3. The flowchart in Supplemental Figure S1 describes the data reduction strategy to select chromatographic features and then tentatively identified compounds to compare persistent and removed compounds among the treatment plants.

Chromatographic feature selection. Features from the initial peak table (raw instrumental data) were selected using the following criteria 1) signal to noise ratio (S/N) ≥ 50, 2) WWTP sample peak area ≥ 5 times blank peak area, 3) present in all influent sample replicates from at least one WWTP, and 4) present or absent in any combination of effluent samples. Compounds absent from at least one of the replicates are considered absent or removed; in other words, their presence in one or two of the triplicate samples is interpreted as false positive identification. Peak areas were normalized to the sample volumes extracted (given in Supplemental Table S1), and subsequent discussion refers to the normalized peak areas in all cases.

Tentatively identified compound selection and verification. The selected chromatographic feature set was further reduced using the confidence in tentative identifications resulting from matches against the 2014 National Institute of Standards and Technology (NIST) electron ionization mass spectral library. Compounds were considered tentatively identified if they met the following criteria: 1) Similarity score ≥

700 (out of a maximum of 999), 2) the top three most abundant ions in the matching NIST library spectrum were present in the corresponding experimental spectrum, and 3) the intensity of the most prominent fragment ions followed a rank order similar to that of the corresponding NIST library spectrum. The names of tentatively identified compounds are provided as the default names used by the NIST mass spectral library. In Tran et al. (2020), tentatively identified compounds were verified against synthetic standards using a similar GC×GC/TOF-MS method and search against the NIST mass spectral library, and the verification success rate was 94% (n = 30 of 32 compounds). In the current study, a set of 15 representative standards, covering a wide range of physicochemical properties, were verified. and the verification success rate (match of GC retention times and mass spectra) was 100% (Supplemental Table S4).

Assignment of removed and persistent compounds. Similar to Qian et al. (2021), changes in the normalized peak areas reflect the removal efficiencies or potential transformation and production of compounds within the WWTPs. The % change (or % removal) in normalized peak area was calculated as follows:

$$\%\ change = \frac{\bar{A}_{inf} - \bar{A}_{eff}}{\bar{A}_{inf}} \times 100 \qquad (1)$$

where $\bar{A}_{inf}$ is the average normalized peak area of the compound in influent replicate samples and $\bar{A}_{eff}$ is the average normalized peak area of the compound in effluent replicate samples.

Complete removal of compounds, in other words, no detection in the effluents, in this study technically indicates that the compounds were present under the limit of detection (LOD). However, considering that much larger volumes of the effluent samples compared to their corresponding influent samples were analyzed (4-10 times larger volume, see Table S1), no detection in the effluents suggests that a compound underwent a high degree of removal in its respective treatment system.

Compounds for which the normalized peak areas were undetectable in effluent or for which the % change was > 90% in all three treatment plants were considered common removed compounds (R). Although

most “R” compounds had 100% reduction of normalized peak areas from influent to effluent, it is important to note that some compounds had 10% or less of the peak areas still present in effluent, which we considered to be mostly removed. Those with detectable peak areas (i.e., with < 90% removal) in all three treatment plants were considered common persistent compounds (P). Tentatively identified compounds with > 90% removal in the decentralized treatment plant, which employs primarily anaerobic treatment, but present in all effluent samples of both of the centralized treatment plants, were referred to as “R-An.” Those compounds removed in both of the centralized treatment plants, which employ primarily aerobic treatment, but were present in effluent of the decentralized treatment plant, were referred to as “R-Ae.” For the remaining compounds, which persisted in one of the three WWTPs, the type of wastewater treatment in which they persisted is specified (eg., “P US-C” for those compounds that persisted in US-C but were removed in SA-D and SA-C). The level of persistence or removal of each compound is listed in Table 4.

In Tran et al. (2020), we estimated the LOD of the GC×GC/TOF-MS method by analyzing standard solutions at 0.1, 1, 10, 100, and 1000 ng/mL. Nine standard compounds representing halogenated and nonhalogenated contaminants were run. The lowest concentration at which the compounds generated a mass spectrum sufficient for identification was 10 ng/mL in a GC vial. For the WWTP sample analysis (concentrated by SPE, for example from 50-500 mL of wastewater to 0.4 mL), this corresponded to estimated LODs in the range of 8 – 80 pg/mL wastewater.

2.5 Physicochemical characteristics and molecular descriptors

Physicochemical properties of tentatively identified compounds were obtained from the Environmental Protection Agency (EPA) Comptox Chemicals Dashboard (Williams et al., 2017) and included atmospheric hydroxylation rate, bioconcentration factor, biodegradation half-life, boiling point, Henry’s Law coefficient, octanol air partitioning coefficient, soil adsorption coefficient, octanol water partitioning coefficient, melting point, vapor pressure, and water solubility all derived from the Opera model. Molecular descriptors related to shape (Geometric shape, Kier shape, Zagreb group, polar surface area)

and bonds (# hydrogen bond acceptors, # hydrogen bond donors, # atoms, # functional groups, # aromatic rings) were obtained using the cheminformatics libraries JOELib (Wegner, 2004) and ChemmineR (Backman et al., 2011). The property prediction modules of these two tools were accessed using the online service ChemMine Tools (Backman et al., 2011).

These tools required the input of SMILES strings for each compound. Out of the 111 compounds, SMILES strings were collected from the EPA Comptox Dashboard for all except 7 compounds. Five of those 7 compound's SMILES strings were found and retrieved from the PubChem search tool. The other 2 compounds (3-Oxo-androsta-1,4-dien-17β-spiro-2'-3'-oxo-oxetane and tricyclo[5.2.1.0(2,6)]dec-3-en-10-one) structures were manually drawn using PubChem Sketcher V2.4, which generated their SMILES strings. The physico-chemical properties and molecular descriptors of each of the tentatively identified compounds are presented in Supplemental Appendices A1 – A5 (Supplementary Information).

2.6 Data analyses

Hierarchical clustering analysis was conducted in R (version 3.6.1) using function hclust() and the centroid agglomeration method. The heatmap was created using package latticeExtra (version 0.6-28). Unpaired, unequal variance, two-sample t-tests (for comparisons of physico-chemical and structural characteristics of categories of compounds) or paired two-sample t-tests (for removal efficiency comparisons) were performed with the significance levels set to 0.05.

3. Results and discussion

3.1 Non-targeted chemical analysis

Chromatographic feature comparison. Due to the low sample size, one composite sample from influent and effluent of each WWTP, the selected chromatographic features from the GC×GC/TOF-MS analysis were compared as an aggregate sample comparison. In total, 2107 chromatographic features were detected in all influent replicates from at least one of the three WWTPs. Of the 1345 US-C influent

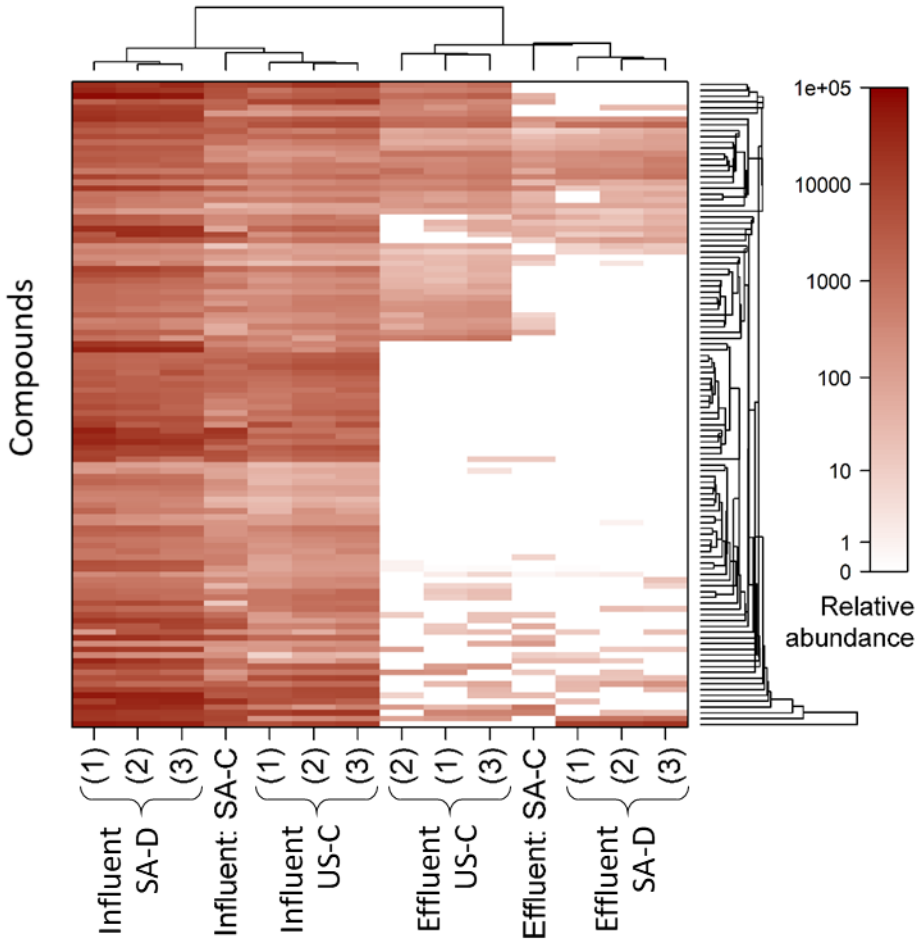
features, 1180 (88%) were found in all three replicates, and of the 1494 SA-D influent features, 1381 (92%) were found in all three replicates (Table 3), indicating similarity among the replicates. SA-C, with only one influent replicate and effluent replicate, had 444 features in the influent and 187 in the effluent. Table 3 shows the number of chromatographic features in the influent of the WWTPs, that persisted through treatment, were removed by the treatment, and were common among the WWTPs.

Table 2. Chromatographic features in samples from the three WWTPs. Influent compounds were present in all influent sample replicates at the WWTP, and any number of effluent replicates. Persistent compounds were present in all influent and effluent sample replicates at the WWTP. Removed compounds were present in all influent replicates and zero effluent sample replicates at the WWTP.

Type of feature	US-C	SA-D	SA-C[1]	US-C & SA-D & SA-C
Influent	1180	1381	444	317
Persistent	466	343	187	66
Removed	540	771	257	107

[1] SA-C had one replicate, and the other two WWTPs had three replicates each.

Tentatively identified compound comparison. To facilitate comparison among the WWTPs, a subset of contaminants was selected that met two criteria: 1) the contaminants were common to all influent replicates from all three WWTPs and 2) the contaminants were considered tentatively identified using a suspect screening approach (see criteria in the Methods section). These 111 “common influent” compounds and their treatment persistence or removal is shown in Figure 2, and the persistence and removal of these compounds is further described in Table 4. There were 65, 43, and 43 tentatively identified compounds present (with < 90% removal) in the effluent of US-C, SA-D, and SA-C, respectively. Using the definitions of persistent and removed described earlier, 11 compounds were persistent at all three WWTPs and 67 compounds were removed at all three WWTPs (Supplemental Table S4).



345

346 **Figure 2.** Heatmap showing tentatively identified chemical constituents common to all influent sample
347 replicates at all three WWTPs. Abundance (chromatographic peak area, *log* scaled) was normalized to the
348 sample volume. The axes were arranged by hierarchical clustering (Wilkinson et al., 2009), and the x-axis
349 parentheses indicate sample replicates.

Table 3. Numbers of tentatively identified compounds and average % removal (± standard deviation) of peak areas in samples from the three WWTPs.

Compound type/location	US-C	SA-D	SA-C[1]	All WWTPs
Influent	111	111	111	111
Persistent in all effluent replicates	37	14	26	11
Removed >90% in all effluent replicates	74	97	85	67
Compound peak areas, average % removal	81 ± 32	96 ± 16	89 ± 25	88 ± 25

[1] SA-C had one replicate, and the other two WWTPs had three replicates each.

Although the removal efficiency of total features was similar for the three WWTPs, the removal of tentatively identified compounds was greater in SA-D than in US-C (Table 4). Moreover, the decrease in normalized peak areas after treatment was significantly higher ($p < 0.01$ using a paired t-test) in SA-D (at 96% removal) compared to US-C (at 81% removal) and SA-C (at 88% removal). The greater removal in SA-D may be due to anaerobic conditions and additional sorption to wetland soils. Differences between the treatment systems that may influence compound removal, transformation, and recalcitrance are discussed in greater detail in subsequent sections. In addition, many of the SA-D influent compounds started with much higher peak areas than US-C and SA-C influent compounds. The close proximity of SA-D to the neighborhood that generates the wastewater for this treatment plant, and therefore shorter travel time for wastewater within the small sewershed, may mean that these organic compounds did not have much time to undergo degradation reactions during transport, and may explain the higher peak areas (and potentially concentrations) of these compounds in the SA-D influent than in the influent of the other two WWTPs (Figure 2).

3.2 Compounds found in influent of centralized and decentralized WWTPs

Tentatively identified influent compounds are presented in Supplemental Table S4 and categorized according to their source or use, which was inferred from a combination of "Product and Use Categories" and "Chemical Functional Use" data acquired from the Environmental Protection Agency (EPA)

Comptox Chemicals Dashboard and scientific articles with a description of a chemical's source or use. Their mass spectral data with their corresponding matching mass spectra from the NIST EI mass spectral library can be found in Supplemental Information (Appendix B). Most chemicals had sufficient information to categorize them as ingredients in pharmaceuticals, personal care products, or food, biocides, or industrial chemicals. We found the greatest number of tentatively identified compounds comprised personal care products (32%), whereas other studies (eg., Qian et al., 2021) using LC-based NTA report pharmaceuticals as the greatest fraction of WWTP influent compounds. Chemicals without source or use information in the Comptox database and which were listed in scientific articles as potential transformation products of a parent compound were categorized as "Unknown, may be transformation product." Chemicals without source or use information in the EPA Comptox Dashboard and which did not appear in searches of the peer-reviewed literature were considered "Unknown."

383 **Table 4.** Compounds found in all replicates of influent samples of the US-C, SA-D, and SA-C wastewater treatment plants and the respective CAS
384 numbers, % change in peak areas from influent to effluent, and persistence code[1], categorized by source or use[2].

Compound	Alternate name	CAS number	%change US-C	% change SA-D	% change SA-C	Persistence code[1]
A) Ingredient in pharmaceutical						
(S)-(-)-4-Isopropenyl-1-cyclohexene-1-carboxylic acid	(4S)-4-Prop-1-en-2-ylcyclohexene-1-carboxylic acid	23635-14-5	100%	100%	100%	R
1,3-Benzodioxole, 4-methoxy-6-(2-propenyl)-	Myristicin	607-91-0	100%	100%	100%	R
2-Methoxy-4-vinylphenol	2-Methoxy-4-vinylphenol	7786-61-0	66%	98%	81%	R-An
Acetaminophen	Acetaminophen	103-90-2	100%	100%	100%	R
Androst-2-en-17-one, (5α)-	Androst-2-en-17-one	963-75-7	100%	100%	100%	R-HMW
Benzenesulfonamide, 2-methyl-	o-Toluenesulfonamide	88-19-7	-9%	74%	76%	P
Benzenesulfonamide, N-ethyl-4-methyl-	N-Ethyl-4-methylbenzenesulfonamide	80-39-7	-32%	75%	-6%	P
Benzenesulfonanilide	Benzenesulfonanilide	1678-25-7	-5%	100%	-112%	R-An
Cannabinol	Cannabinol	521-35-7	100%	100%	100%	R-HMW
Cyclohexanemethanol, 4-hydroxy-α,α,4-trimethyl-	Terpin	80-53-5	100%	100%	100%	R
Diphenan	Diphenan	101-71-3	36%	100%	95%	R-An
Phenol, 3,4,5-trimethoxy-	3,4,5-Trimethoxyphenol	642-71-7	96%	100%	100%	R
Prasterone-3-sulfate	Prasterone sulfate	651-48-9	100%	100%	100%	R-HMW
γ-Tocopherol	gamma-Tocopherol	7616-22-0	100%	100%	100%	R-HMW
Cyclohexanone, 5-methyl-2-(1-methylethenyl)-	cyclohexanone, 5-methyl-2-(1-methylethenyl)-	529-00-0	100%	100%	100%	R
Benzyl chloride	Benzyl chloride	100-44-7	96%	100%	91%	R
2,2,2-Trichloro-1-phenylethanol	2,2,2-Trichloro-1-phenylethanol	2000-43-3	37%	100%	98%	R-An
B) Ingredient in personal care product						
1,7-Dimethylxanthine	1,7-Dimethylxanthine	611-59-6	100%	100%	100%	R
1H-Indole, 5-chloro-	5-Chloroindole	17422-32-1	100%	100%	100%	R
1-Penten-3-one, 1-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	1-(2,6,6-Trimethyl-1-cyclohexen-1-yl)pent-1-en-3-one	127-43-5	76%	100%	55%	R-An
2,8,9-Trioxa-5-aza-1-silabicyclo[3.3.3]undecane, 1-methyl-	Methylsilatrane	2288-13-3	100%	98%	93%	R
2-Butanone, 4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	Dihydro-beta-ionone	17283-81-7	99%	99%	99%	R
2-Butanone, 4-(4-hydroxyphenyl)-	4-(4-Hydroxyphenyl)butan-2-one	5471-51-2	99%	100%	100%	R
2H-Pyran-2-one, 6-heptyltetrahydro-	delta-Dodecalactone	713-95-1	100%	100%	100%	R
3-Pyridinol, 6-methyl-, acetate (ester)	2-Hydroxyacetanilide	614-80-2	100%	100%	100%	R
4-Acetyl-1-methylcyclohexene (6090-09-1)	4-Acetyl-1-methylcyclohexene	6090--09-1	88%	96%	100%	P US-C
Benzaldehyde, 4-methoxy-	-	123-11-5	78%	92%	88%	R-An
Benzene, 1,3-diethenyl-	Pentyl 2-hydroxybenzoate	2050-08-0	100%	100%	97%	R
Benzoic acid, 3-(1-methylethyl)-	Benzoic acid, 3-(1-methylethyl)-	5651-47-8	100%	100%	100%	R
Benzyl nitrile	Benzyl cyanide	140-29-4	94%	98%	96%	R
Benzylidenemalonaldehyde	2-Benzylidenemalonaldehyde	82700-43-4	40%	88%	79%	P
Bicyclo[2.2.1]heptane, 2,2-dimethyl-3-methylene-, (1S)- (5794-04-7)	(-)-Camphene	5794--04-5	97%	100%	98%	R
Butanenitrile, 4-(methylthio)-	Butanenitrile, 4-methylthio-	59121-24-3	100%	100%	100%	R
Cadala-1(10),3,8-triene	(1S)-4,7-dimethyl-1-propan-2-yl-1,2-dihydronaphthalene	21391-99-1	71%	100%	100%	P US-C
Cyclopenta[g]-2-benzopyran, 1,3,4,6,7,8-hexahydro-4,6,6,7,8,8-hexamethyl-	Cyclopenta[g]-2-benzopyran, 1,3,4,6,7,8-hexahydro-4,6,6,7,8,8-hexamethyl-	1222-05-5	41%	100%	82%	R-An
dl-α-Tocopherol	alpha-Tocopherol	10191-41-0	100%	100%	100%	R-HMW
Dodecanoic acid, 2-methyl-	Undecanoic acid, 2-methyl-	24323-25-9	100%	100%	100%	R

Compound	Alternate name	CAS number	%change US-C	% change SA-D	% change SA-C	Persistence code[1]
Ethylparaben	Ethylparaben	120-47-8	100%	100%	100%	R
Heptasiloxane, hexadecamethyl-	Hexadecamethylheptasiloxane	541-01-5	98%	99%	100%	R-HMW
Hydrocoumarin	3,4-Dihydrocoumarin	119-84-6	100%	100%	100%	R
Indano[2,1-d]1,3-dioxane,	indano[2,1-d]1,3-dioxane,	102688-70-0	8%	77%	42%	P
Methanone, (2,4-dihydroxyphenyl)phenyl-	2,4-Dihydroxybenzophenone	131-56-6	100%	100%	100%	R
Musk ketone	Musk ketone	81-14-1	-5%	100%	70%	R-An
ortho-Methoxyacetophenone	Ethanone, 1-(2-methoxyphenyl)-	579-74-8	94%	98%	86%	P SA-C
Oxiranecarboxylic acid, 3-methyl-3-phenyl-, ethyl ester, cis-	-	19464-95-0	57%	100%	100%	P US-C
Phenol, 2-ethyl-4,5-dimethyl-	Isopropyl-o-cresol	499-75-2	98%	100%	89.9%	P SA-C
Phenol, 2-methoxy-3-(2-propenyl)- (1941-12-4)	-	1941--12-4	100%	100%	99%	R
Phenol, 3,4-dimethoxy-	3,4-Dimethoxyphenol	2033-89-8	83%	100%	100%	P US-C
Phenol, 4-ethyl-2-methyl-	2-Propylphenol	644-35-9	100%	100%	100%	R
Pyrrolo[1,2-a]pyrazine-1,4-dione, hexahydro-3-(phenylmethyl)-	3-Benzyl-1-hydroxy-6,7,8,8a-tetrahydropyrrolo[1,2-a]pyrazin-4(3H)-one	14705-60-3	99%	100%	100%	R
Scopoletin	Scopoletin	92-61-5	100%	100%	100%	R
Triclosan	Triclosan	3380-34-5	89%	89%	93%	P
α-Tocopheryl acetate	D-alpha-Tocopheryl acetate	58-95-7	100%	98%	100%	R-HMW
C) Ingredient in food						
1H-Indol-4-ol, 3-methyl-	3-Methyl-1H-indol-4-ol	1125-31-1	99%	100%	98%	R
1H-Indole, 4-methoxy-3-cyanomethyl-	4-Methoxyindole-3-acetonitrile	4837-74-5	76%	100%	64%	R-An
2,4-Dithiapentane	Methane, bis(methylthio)-	1618-26-4	91%	100%	-17%	P SA-C
2-Butanone, 4-(4-hydroxy-3-methoxyphenyl)-	4-(4-Hydroxy-3-methoxyphenyl)-2-butanone	122-48-5	100%	100%	100%	R
3(2H)-Thiophenone, dihydro-2-methyl-	3(2H)-Thiophenone, dihydro-2-methyl-	13679-85-1	100%	98%	100%	R
3-Cyclohexene-1-acetaldehyde, α,4-dimethyl-	3-Cyclohexene-1-acetaldehyde, .alpha.,4-dimethyl-	29548-14-9	100%	100%	100%	R
3-tert-Butyl-4-hydroxyanisole	2-tert-Butyl-4-methoxyphenol	121-00-6	16%	100%	73%	R-An
Benzene, 1,2,3-trimethoxy-5-(2-propenyl)-	Elemicin	487-11-6	100%	100%	100%	R
Benzenemethanethiol	Benzenemethanethiol	100-53-8	100%	100%	100%	R
Cyclohexanol, 1-methyl-4-(1-methylethenyl)-	beta-Terpineol	138-87-4	99%	100%	-29%	P SA-C
Hexanoic acid, 2-ethyl-, anhydride	2-Ethylhexanoic anhydride	36765-89-6	-155%	100%	100%	P US-C
Neric acid	cis-Geranic acid	4613-38-1	100%	100%	100%	R
Pentanenitrile, 5-(methylthio)-	5-(Methylsulfanyl)pentanenitrile	59121-25-4	100%	100%	100%	R
Phenol, 2,6-dimethoxy-4-(2-propenyl)-	4-Allyl-2,6-dimethoxyphenol	6627-88-9	100%	100%	100%	R
Phenol, 2,6-dimethoxy-4-(2-propenyl)-	2,6-Dimethoxyphenol	91-10-1	46%	100%	100%	P US-C
Phenol, 2-methoxy-4-(1-propenyl)-	Isoeugenol	97-54-1	91%	100%	100%	R
Phenol, 4-ethyl-2-methoxy-	4-Ethyl-2-methoxyphenol	2785-89-9	100%	100%	100%	R
Piperine	Piperine	94-62-2	100%	100%	100%	R-HMW
Propanedioic acid, dihydroxy-, bis(1-methylethyl) ester	Benzoxazole, 2-methyl-	95-21-6	91%	99%	100%	R
Propylparaben	Propylparaben	94-13-3	100%	99%	100%	R
Theobromine	Theobromine	83-67-0	99%	97%	25%	P SA-C
Thymoquinone	2,5-Cyclohexadiene-1,4-dione, 2-methyl-5-(1-methylethyl)-	490-91-5	37%	77%	65%	P
D) Herbicide, pesticide						
1H-Indole-3-acetonitrile	1H-Indole-3-acetonitrile	771-51-7	77%	100%	100%	P US-C
8-Hydroxycarvotanacetone	8-Hydroxycarvotanacetone	7712-46-1	100%	95%	99%	R

Compound	Alternate name	CAS number	%change US-C	% change SA-D	% change SA-C	Persistence code[1]
Acetamide, N-phenyl-	Acetanilide	103-84-4	97%	85%	100%	R-Ae
Benzoic acid, 2,4-dihydroxy-3,6-dimethyl-, methyl ester	Methyl 3-methylorsellinate	4707-47-5	100%	100%	100%	R
Tepraloxydim	pteridine-4,7(3H,8H)-dione	33669-70-4	34%	74%	69%	P
E) Industrial chemical						
1H-Benzotriazole, 4-methyl-	4-Methyl-1,2,3-benzotriazole	29878-31-7	-67%	100%	100%	P US-C
1H-Indole-2,3-dione	Isatin	91-56-5	95%	100%	100%	R
2-Furanmethanol	Furfuryl alcohol	98-00-0	83%	100%	100%	P US-C
3H-1,2-Benzodithiol-3-one	Benzodithiolone	1677-27-6	100%	100%	96%	R
5,6-Dimethyl-1H-benzotriazole	2,3-Dimethylphenylisocyanate	1591-99-7	76%	100%	100%	P US-C
Benzeneacetic acid, α-oxo-, ethyl ester	Benzeneacetic acid, .alpha.-oxo-, ethyl ester	1603-79-8	97%	100%	100%	R
Benzenesulfonamide, 4-methyl-	4-Toluenesulfonamide	70-55-3	-45%	-36%	89.9%	P
Phenol, 4,4'-(1-methylethylidene)bis-	Bisphenol A	80-05-7	57%	87%	100%	P US-C & SA-D
Triphenyl phosphate	Triphenyl phosphate	115-86-6	42%	90%	100%	P US-C
F) Unknown, may be transformation product						
1,2-Cyclohexanedicarboxylic acid, diethyl ester	Diethyl cyclohexane-1,2-dicarboxylate	10138-59-7	100%	100%	100%	R
2,5-Dichlorobenzonitrile	2,5-Dichlorobenzonitrile	21663-61-6	-181%	73%	-326%	P
2,6-Dimethylphenyl isocyanate	Benzene, 2-isocyanato-1,3-dimethyl-	28556-81-2	100%	100%	99%	R
2-Cyclohexen-1-one, 3,5,5-trimethyl-4-(3-oxo-1-butenyl)-	3,5,5-Trimethyl-4-(3-oxobut-1-en-1-yl)cyclohex-2-en-1-one	20194-68-7	98%	100%	100%	R
3-Methyl-acridine (4740-12-9)	-	4740--12-9	80%	100%	30%	R-An
4-Methyl-acridone	4-Methyl-acridone	68506-36-5	-334%	76%	3%	P
Propanoic acid, 2-methyl-3-[4-t-butyl]phenyl-	-	66735-04-4	100%	100%	99%	R
Pyridine, 2-(1-methyl-2-pyrrolidinyl)-	2-(1-Methylpyrrolidin-2-yl)pyridine	23950-04-1	100%	100%	100%	R
G) Unknown						
1,2,3,4-Tetrahydroquinolin-8-ol, 2,2,4-trimethyl-	-	61855-47-8	96%	100%	100%	R
1H-Indole, 5-bromo-	1H-Indole, 5-bromo-	10075-50-0	100%	100%	100%	R
2,3-Dihydro-4-methoxyindole-2-one	4-Methoxyoxindole	7699-17-4	100%	100%	100%	R
3-Oxo-androsta-1,4-dien-17β-spiro-2'-3'-oxo-oxetane	-	0-00-0	100%	100%	100%	R-HMW
4-Acetyl-1,2,3,4-tetrahydro-2-oxoquinoline	-	92287-78-0	91%	97%	100%	R
Acetamide, N-acetyl-N,N'-1,2-ethanediylbis-	Acetamide, N-acetyl-N-[2-(acetylamino)ethyl]-	137706-80-0	48%	100%	100%	P US-C
Decanoic acid, 3-methyl-	3-Methyldecanoic acid	60308-82-9	100%	100%	100%	R
Formamide, (2-acetylphenyl)- (5257-06-7)	-	5257--06-7	99%	100%	99%	R
Oxepine, 2,7-dimethyl-	2,7-Dimethyloxepine	1487-99-6	56%	95%	60%	R-An
Phenol, p-(2-methylallyl)-	-	33641-78-0	77%	99%	83%	R-An
Preg-4-en-3-one, 17α-hydroxy-17β-cyano-	-	77881-13-1	100%	100%	100%	R-HMW
Tetrahydrofuran-2-one, 5-[1-hydroxyhexyl]-	-	87877-77-8	100%	42%	100%	R-Ae
trans-Dehydroandrosterone, trifluoroacetate	-	3798-17-2	100%	100%	100%	R-HMW
Tricyclo[5.2.1.0(2,6)]dec-3-en-10-one	-	0-00-0	73%	48%	89%	P

[1] Persistence code definitions: R = Removed from all WWTPs by both aerobic and anaerobic treatment; R-HMW = Compounds in the R dataset with highest molecular weights ($p << 0.001$) compared to other compounds in the dataset; R-An = Removed by anaerobic treatment and present in aerobic treated effluent; R-Ae = Removed by aerobic treatment and present in anaerobic treated effluent; P = Persistent in all WWTPs; P US-C = Persistent only in US-C; P US-C & SA-D = Persistent in both US-C and SA-D; P SA-C = Persistent only in SA-C.

[2] Some compounds may have multiple uses and may belong to more than one category; for the purposes of this study, compounds were listed according to their main source or use.

3.3 Removed compounds

Complete removal in all effluent replicates of the US-C, SA-C, and SA-D WWTPs occurred for 67 compounds (referred to as R), comprising a diverse group of chemicals, including ingredients in pharmaceuticals, personal care products, food, herbicides and pesticides, as well as unknown chemicals (Table S4). Some of the R compounds of this study have been identified in other studies of natural and treated waters. For example, α-tocopherol (vitamin E) was recently detected in a comprehensive screening of onsite wastewater treatment systems in Sweden and underwent up to 100% removal in char-fortified filter beds (Blum et al., 2019). 4-Ethyl-2-methoxyphenol was among a suite of residual organic pollutants detected in sediments contaminated with pulp and paper mill waste (Yadav and Chandra, 2018). Although the caffeine metabolite, 1,7-dimethylxanthine, was shown to persist in wastewater effluent of a WWTP employing activated sludge biotreatment (Gómez et al., 2007), this compound was completely removed in all three WWTPs of the current study.

Of the 67 R compounds, 64 were identified in the EPA Comptox Dashboard (Williams et al., 2017) and their physico-chemical properties were available. Although their biodegradability, soil adsorption, and water solubility spanned a range of values, two main categories of compounds emerge: those with low molecular weight (LMW) and generally shorter biodegradation half-lives, lower soil adsorption and higher water solubility and those with higher molecular weight (HMW; > 200 g/mol), longer half-lives (> 20 d), higher soil adsorption (log K_{oc} > 3.0 log L/kg), and low water solubility (log K_w below ~ -4 log mol/L; Figure 3). Removal of the HMW compounds (R-HMW) may have occurred through sorption to biosolids in the activated sludge and sedimentation treatment processes of the US-C and SA-C WWTPs and via sorption to the sludge blanket in the ABR or soil in the constructed wetlands of the SA-D treatment facility.

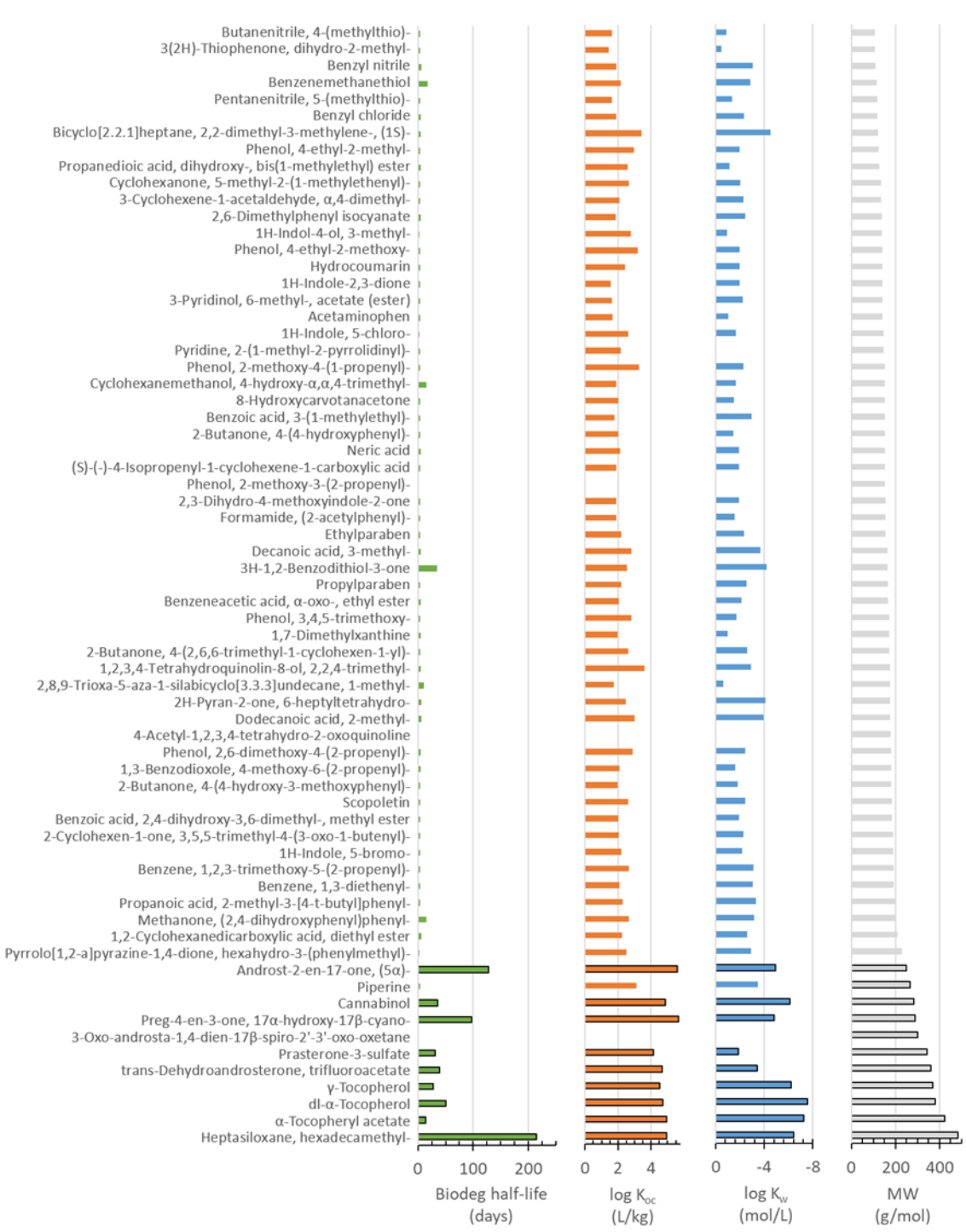


Figure 3. Biodegradation half-life, log soil adsorption (log K_{oc}), log water solubility (log K_w), and molecular weight (MW) values of 67 tentatively-identified, common removed compounds in centralized and decentralized WWTPs of this study. 11 high molecular weight compounds are outlined in black. Compounds without values were not identified in the EPA Comptox database.

3.4 Persistent compounds

Only 11 compounds persisted in the effluent of all three WWTPs (P compounds; Table 5). Of those, 10 were present in the EPA Comptox database, and their physico-chemical and structural characteristics could be evaluated. These persistent compounds were found in all categories of chemicals, including pharmaceuticals, personal care products, food, herbicides and pesticides, industrial chemicals, as well as unknown chemicals and transformation products. Although triclosan has been banned from consumer antiseptic wash products by the U.S. Food and Drug Administration (FDA, 2016), its use appears to still be ubiquitous since it was found in all three WWTP effluents. None of the persistent compounds appear in regulatory programs in the USA or South Africa, and only 6 compounds (4-methyl-benzenesulfonamide, 2-methyl-benzenesulfonamide, N-ethyl-4-methyl-benzenesulfonamide, tepraloxydim, benzylidenemalonaldehyde, and triclosan) have been previously reported in wastewater effluent, downstream of WWTPs, or in sludge, either directly or as derivatives of another compound (Kuster et al., 2009; Herrero et al., 2013; Herrero et al., 2014; Liu et al., 2015; Hrubik et al., 2016; Tohidi and Cai, 2017; Krichling, 2018; El Hayany et al., 2020).

Table 5. Persistent compounds found in effluent samples of the US-C, SA-D, and SA-C wastewater treatment plants and their respective CAS numbers, source or use category[1], and toxicity[2].

Persistent Compound	CAS RN	Source/use category[1]	Toxicity[2]
Triclosan	3380-34-5	Ingredient in personal care product	Irritant, environmental hazard
4-Methyl-acridone[§]	68506-36-5	Unknown, may be transformation product	No information
Tepraloxydim	33669-70-4	herbicide, pesticide	No information
Indano[2,1-d]1,3-dioxane[§]	102688-70-0	Ingredient in personal care product	No information
2,5-Dichlorobenzonitrile[§]	21663-61-6	Unknown, may be transformation product	Irritant
Benzylidenemalonaldehyde	82700-43-4	Ingredient in personal care product	No information
Thymoquinone[§]	490-91-5	Ingredient in food	Irritant
Benzenesulfonamide, 4-methyl-	70-55-3	Industrial chemical	Irritant, acutely toxic
Benzenesulfonamide, N-ethyl-4-methyl-	80-39-7	Ingredient in pharmaceutical	Irritant
Benzenesulfonamide, 2-methyl-	88-19-7	Ingredient in pharmaceutical	Irritant, human health hazard
Tricyclo[5.2.1.0(2,6)]dec-3-en-10-one[§]	0-00-0	unknown	No information

435
436 [1] Some compounds may have multiple uses and may belong to more than one category; for the purposes of this study, compounds
437 were listed according to their main source or use.
438 [2] Toxicity according to the Globally Harmonized System of Classification and Labelling of Chemicals (GHS).
439 [§] Not previously reported in WWTPs.
440

441 Two of the 11 P compounds, 4-methyl-acridone and 2,5-dichlorobenzonitrile, even increased substantially
442 from influent to effluent in the predominantly aerobic, centralized WWTPs (Supplemental Table S4).
443 From our review, these compounds were not previously reported in wastewater effluent. The compound
444 with the highest biodegradation half-life, 4-methyl-acridone, had a peak area increase of over 300% in the
445 US-C effluent and its removal was only 3% in the SA-C WWTP (Supplemental Table S4). In SA-D, the
446 removal of 4-methyl-acridone was much higher, at 76%, but, given the higher overall influent peak areas
447 in SA-D (described earlier), the effluent peak area was still higher than that of SA-C. Along with similar
448 compounds, acridone and 4-methylacridone have both been reported as metabolites of carbamazepine, a
449 well-known pharmaceutical found in wastewater (Leclercq et al., 2008; Zhou et al., 2016). However, 4-
450 methyl-acridone has not been reported as an emerging contaminant in wastewater, nor does the Comptox
451 database report any identifying source, use, or toxicity information for this compound. Therefore, this
452 compound may be a transformation product of compounds such as carbamazepine, and its formation
453 warrants further study.

454 The other compound that increased >180% from influent to effluent of US-C and SA-C, 2,5-
455 dicholorobenzonitrile (Supplemental Table S4), is also not well known in terms of its source. It can be
456 one of the isomers, 2,4-dicholorobenzonitrile (Kattan et al., 2014) or 3,5-dichlorobenzonitrile which is a
457 methyl orange dye byproduct (Ramjun et al., 2015). Or it can come from herbicides or pesticides since
458 other dihalogenated benzonitriles (e.g., dichlobenil) are active ingredients in a number of herbicides and
459 pesticides (McManus et al., 2014; Palofox et al. 2015).Many of the persistent compounds have been
460 reported to be acutely toxic, health hazards, irritants, corrosive, or environmental hazards (Table 5)
461 according to the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), a
462 United Nations system to identify hazardous chemicals. For example, tepraloxydim is a known

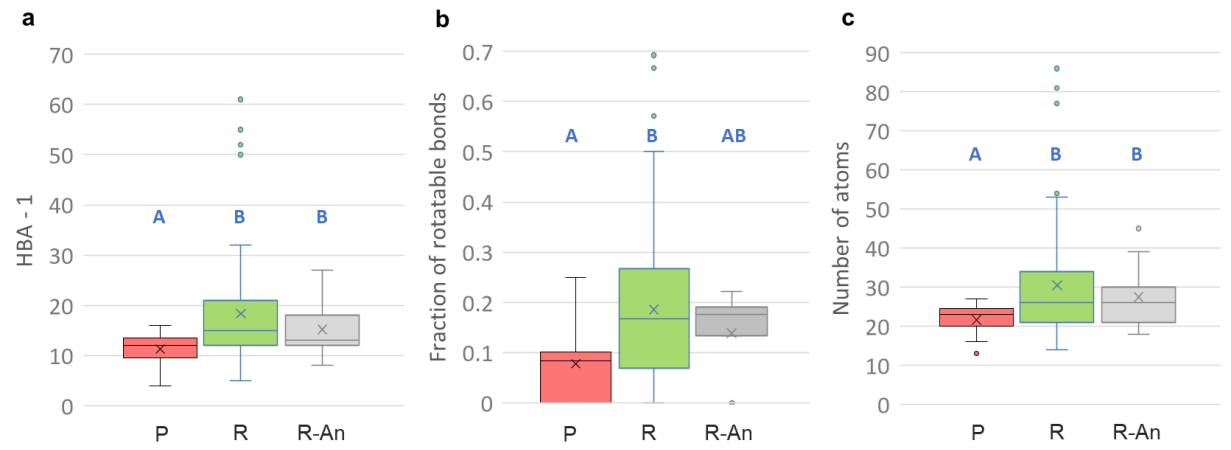
carcinogenic herbicide that has even been detected in drinking water (Sandín-España, et al., 2002). On the other hand, thymoquinone, a bioactive compound found in the spice Nigella sativa (black cumin), which has not been previously identified in WWTPs, has no toxicity information listed. Its strong anti-oxidant properties may even protect against toxicity (Nagi and Almakki, 2009). The lack of toxicological information available for these persistent effluent compounds highlights the need to further understand the characteristics and toxicity of new ubiquitous compounds revealed by NTA.

Due to the complete absence of some compounds and persistence or even increase of others in the treated effluents, we further explored whether there were any statistically significant differences among the 12 physico-chemical properties of the 67 completely removed (R and R-HMW) and 11 P compounds. Other than 4-methyl-acridone, most of the P compounds exhibited low biodegradation half-lives (from ~3 to 5 d) and their water solubility and other physico-chemical characteristics were not unlike those of the removed compounds (Figure 3). Furthermore, none of the physico-chemical properties, including those that might be expected to result in persistence or removal (e.g., biodegradation half-life, soil adsorption, octanol-water partitioning coefficient (log K_{ow}), or water solubility) were significantly different between R and P compounds. Due to the significantly higher molecular weight of the six R-HMW compounds, those compounds were excluded and a student t-test comparing only P and the remaining R compounds was performed. Even without the R-HMW compounds, there were no significant differences between the physico-chemical properties of P and R compounds. Other studies have shown that the removal of trace organic chemicals in wastewater treatment is a complex phenomenon and not solely a function of a compound's biodegradability or hydrophobicity (Berthod et al., 2017).

Prior investigations examined the use of molecular descriptors to model or predict different aspects of compound removal in wastewater and natural waters (e.g., Sathyamoorthy et al., 2013; Berthod et al. 2017; Nolte et al., 2018; Nolte et al., 2020). The body of knowledge on quantitative structural activity relationships (QSAR) and quantitative structural biodegradation relationship (QSBR) modeling is vast, with general consensus that molecular structure is one of the important terms in predicting the activity of

compounds but that characteristics of soil, sludge, or other environmental factors are also needed to enhance models, with dependence on contaminant structure alone being insufficient. Although developing such models is outside the scope of this study, the large number of ubiquitous compounds (111) shared among three different treatment plants as well as the similar removal of 67 of those compounds motivated a superficial assessment of whether there might be significant differences among molecular descriptors between persistent and removed compounds. Therefore, we investigated if molecular descriptors from the JOELib and Chemmine libraries, which were readily available to our study and include structural (shape and bond) characteristics and numbers of atoms and functional groups, have associations with contaminant persistence and removal in the WWTPs of our study. The molecular descriptors of each of the 111 compounds are given in Supplemental Appendices A2 – A5, and each descriptor is defined in the JOELib tutorial (Wegner, 2004).

The only significant differences ($p < 0.05$) between the R and P compounds were molecular descriptors related to bonds and atoms (Figure 4). These properties were still significantly different even when the six HMW compounds were excluded from the statistical analysis. Persistent compounds tended to have lower numbers of H bond acceptors-1, a lower fraction of rotatable bonds, and lower numbers of atoms (Figure 4). High HBA-1 values of removed compounds are corroborated to some extent by recent QSAR models. For example, hydrogen bonding has been identified as an important parameter that influenced the sorption of pharmaceutically active compounds in models of conventional biological treatment processes for wastewater (Sathyamoorthy and Ramsberg, 2013). In QSBR studies, an increase in the number of rotatable bonds (defined as any single bond bound to a non-terminal heavy (i.e., non-hydrogen) atom) was also found to increase the biodegradation of alkanes, alcohols, aldehydes and ketones, acids and phenols (He et al., 2015), but other descriptors were found to have a greater influence on biological degradation of the aromatic PAHs in the environment (Xu et al., 2015).



511

512 **Figure 4. Box-whisker plots showing the mean, median, minimum and maximum values, and**
513 **outliers of molecular descriptors: a)# of H bond acceptors-1, b) fraction of rotatable bondsand c)**
514 **number of atoms of 11 persistent (P), 67 completely removed (R) and 13 removed in anaerobic**
515 **DEWATS (R-An) compounds. Different letters show statistically significant differences (*p* <**
516 **0.05)RRP.**

517

518 Other studies have found the presence of strong electron donating groups (EDGs) to be associated with
519 greater removal in wastewater treatment processes, and the absence of EDGs with strong electron
520 withdrawing groups (EWGs) present to be associated with poor removal (Grandclément et al., 2017; Tran
521 et al., 2018). Review of the functional group information from Chemmine (Supplemental Appendix A5)
522 indicates that removed compounds in our study have, on average, a higher number of EDGs than P
523 compounds; however, removed compounds also had higher numbers of EWGs, and the differences were
524 not statistically significant in either case.

525 **3.5 Role of anaerobic decentralized treatment**

526 As noted earlier, although the number of distinct chromatographic features removed from influent to
527 effluent was similar among the three WWTPs (Table 3), the removal efficiency of the tentatively
528 identified compounds (Table 4) and their peak areas (Figure 1) was greater in the SA-D facility than in
529 the US-C and SA-C centralized WWTPs. An additional 13 tentatively identified chemicals (referred to as
530 R-An) had > 90% removal in all replicate samples of the SA-D facility, which employs primarily

anaerobic wastewater treatment, but were present in the effluents of US-C and SA-C. Given that composite effluent sample collection in SA-D was staggered by 24 h to account for its longer hydraulic residence time (compared to the other WWTPs), the greater reductions in normalized peak areas for SA-D should reflect real decreases due to the treatment process and characteristics rather than day-to-day variability. However, analysis of more samples collected over more dates would provide greater confidence in this result. Nevertheless, we explore potential reasons for the observed high removals of trace organic compounds in SA-D. Only 2 compounds were completely removed by US-C and SA-C but had < 90% removal in SA-D (R-Ae; Supplemental Table S4).

Although some compounds may be better suited to microbial degradation under anaerobic conditions, our review of physico-chemical properties and molecular descriptors did not reveal any properties that were significantly different between R-An and R compounds. Instead, the molecular descriptors, HBA-1 and number of atoms, that were higher in R compounds compared to P compounds were also significantly higher ($p < 0.05$) in the R-An compounds (Figure 4).

As far as the greater reduction of peak areas in SA-D samples, biodegradation rates are generally much slower under anaerobic conditions (Biel-Maeso et al., 2019), but longer residence times in anaerobic treatment processes are expected to influence compound removal and transformation. In a study of a conventional activated sludge WWTP, Völker et al. (2016) found that extended anaerobic conditions resulted in greater removal of endocrine disrupting chemicals. Harb et al. (2019) pointed out that although the longer SRT in anaerobic mainstream systems may be beneficial for biodegradation of trace organic chemicals, contact time with microorganisms is limited by system HRTs. In the SA-D facility, the HRT of the ABR biological treatment is twice as long as the Bardenpho treatment process of SA-C and more than 10 times longer than the conventional activated sludge process HRT (Table 1). Extending HRTs, particularly of anaerobic processes, should be further explored as a treatment modification to improve removal of trace organic chemicals.

555 One of the compounds that was removed in SA-D, but persisted in effluent of both the centralized plants
556 US-C and SA-C, 1,3,4,6,7,8-hexahydro-4,6,6,7,8,8-hexamethyl-cyclopenta-[γ]-2-benzopyran (HHCB), a
557 type of polycyclic musk that bioaccumulates in fish (Reiner and Kannan, 2011), has been shown to have
558 incomplete removal under aerobic activated sludge treatment (Simonich et al., 2002, Horii et al., 2007;
559 etc). The absence of HHCB in the effluent of the anaerobic SA-D plant may indicate that anaerobic
560 treatment or the use of constructed wetland treatment cells is effective for HHCB removal. Horii et al.
561 (2007) did demonstrate that HHCB had high sorption to sludge; therefore, sorption to the sludge blanket
562 in the ABR treatment step may contribute to its removal in the anaerobic decentralized plant.

563 ## 4. Conclusions

564 Despite the contrasting geographic locations (South Africa and USA) of the WWTPs examined in this
565 study, non-targeted analysis revealed 111 tentatively identified, ubiquitous compounds in influent
566 samples. Overall, in terms of the reduction in total numbers of chromatographic features identified using
567 GC×GC/TOF-MS, the performance of all three plants was similar. However, the removal efficiency of
568 the peak areas of the tentatively identified compounds was significantly greater in the SA-D facility than
569 in the US-C and SA-C centralized WWTPs, and may be due to the longer solids retention and hydraulic
570 residence times at SA-D.

571 It should be noted that although the non-targeted analysis of this study revealed the presence of new
572 chemicals and the removal of compounds that is consistent with other studies, there was limited access at
573 the South Africa WWTPs, and samples collected on multiple visits over longer periods would allow a
574 more robust comparison of ubiquitous compounds in wastewater. In addition, inclusion of a larger
575 number of centralized and decentralized WWTPs around the world is recommended to further investigate
576 geographic differences related to the use of PPCPs and other chemicals found in wastewater. Future
577 studies should utilize complementary NTA methods (both GC- and LC-based) to evaluate additional
578 compounds in the effluent of DEWATS and other mainstream anaerobic treatment systems, which are
579 thus far largely understudied.

Using our non-targeted screening approach, only 11 common compounds out of 111 tentatively identified common compounds were found to persist through the three WWTPs, which utilized diverse treatment processes, including both large scale activated sludge aerobic treatment with sedimentation and small scale ABR and AF treatment followed by constructed wetland cells. Although only 10% of the influent compounds persisted through the range of physical and biological treatment processes underway in the WWTPs, these 11 compounds included chemicals with human and environmental health hazards as well as compounds that had not been previously identified in WWTP effluent and that had unknown toxicities. Compared to the 67 compounds that were removed in all three WWTPs, persistent compounds were not significantly different in terms of their physico-chemical characteristics. However, structural features related to bonds and molecular shape characteristics did show significant differences that merit further study. Although limited to only to the three wastewater treatment plants of this study, these new results from NTA may inform future monitoring of trace organic chemicals WWTP effluent and begin to fill gaps in our knowledge of emerging chemicals found in DEWATS effluent.

Notes

The authors declare no competing financial interests.

Supplementary Material accompanies this paper.

Funding sources

This study was supported by the United States National Science Foundation (NSF CBET-1705901), the National Academy of Sciences (NAS), United States Agency for International Development (USAID) and the Department of Science and Technology (DST) of South Africa, under the Partnerships for Enhanced Engagement in Research (PEER) program (AID-OAA-A-11-00012), and the William E. Leonhard Jr. Endowment to N. Mladenov.

Acknowledgements

We are grateful for field and logistical support from H. Singh, C. Trees, S. Arredondo, A. Simonson, B. Pietruschka, K. L. Philp, B. Mthembu and lab assistance from L. Mollo, S. Chao, and C. Graves.

Dedication

We dedicate this work to Professor Chris Buckley, a leader in the WASH sector, who worked tirelessly to provide water, sanitation, and dignity to the underserved. Chris always connected like minds and created networks that would advance water and sanitation beyond existing paradigms. Chris was the spark that brought together this team of authors, and this research would not have been possible without him.

References

Abafe, O. A., Späth, J., Fick, J., Jansson, S., Buckley, C., Stark, A., Pietruschka, B., & Martincigh, B. S. (2018). LC-MS/MS determination of antiretroviral drugs in influents & effluents from wastewater treatment plants in KwaZulu-Natal, South Africa. Chemosphere, 200, 660-670.

Albergamo V, Schollée JE, Schymanski EL, Helmus R, Timmer H, Hollender J, de Voogt P. 2019. Nontarget Screening Reveals Time Trends of Polar Micropollutants in a Riverbank Filtration System. Environ Sci Technol. 53(13):7584–7594. doi:10.1021/acs.est.9b01750.

Anumol, T., & Snyder, S. A. (2015). Rapid analysis of trace organic compounds in water by automated online solid-phase extraction coupled to liquid chromatography–tandem mass spectrometry. *Talanta*, *132*, 77-86.

Ashfaq, M., Li, Y., Wang, Y., Chen, W., Wang, H., Chen, X., ... & Sun, Q. (2017). Occurrence, fate, and mass balance of different classes of pharmaceuticals and personal care products in an anaerobic-anoxic-oxic wastewater treatment plant in Xiamen, China. Water Research, 123, 655-667.

Assress, H. A., Nyoni, H., Mamba, B. B., & Msagati, T. A. (2019). Target quantification of azole antifungals and retrospective screening of other emerging pollutants in wastewater effluent using UHPLC–QTOF-MS. Environmental Pollution, 253, 655-666.

Backman, T. W., Cao, Y., & Girke, T. (2011). ChemMine tools: an online service for analyzing and clustering small molecules. Nucleic acids research, 39(Web Server issue), W486–W491. https://doi.org/10.1093/nar/gkr320

Berthod, L., Whitley, D. C., Roberts, G., Sharpe, A., Greenwood, R., & Mills, G. A. (2017). Quantitative structure-property relationships for predicting sorption of pharmaceuticals to sewage sludge during waste water treatment processes. Science of the Total Environment, 579, 1512-1520.

Biel-Maeso, M., González-González, C., Lara-Martín, P. A., & Corada-Fernández, C. (2019). Sorption and degradation of contaminants of emerging concern in soils under aerobic and anaerobic conditions. Science of The Total Environment, 666, 662-671.

Blum, K. M., Andersson, P. L., Ahrens, L., Wiberg, K., & Haglund, P. (2018). Persistence, mobility and bioavailability of emerging organic contaminants discharged from sewage treatment plants. Science of the Total Environment, 612, 1532-1542.

Blum, K. M., Gallampois, C., Andersson, P. L., Renman, G., Renman, A., & Haglund, P. (2019). Comprehensive assessment of organic contaminant removal from on-site sewage treatment facility effluent by char-fortified filter beds. Journal of hazardous materials, 361, 111-122.

Blum, K. M., Andersson, P. L., Renman, G., Ahrens, L., Gros, M., Wiberg, K., & Haglund, P. (2017). Non-target screening and prioritization of potentially persistent, bioaccumulating and toxic domestic wastewater contaminants and their removal in on-site and large-scale sewage treatment plants. Science of the total environment, 575, 265-275.

Bremen Overseas Research and Development Association (BORDA). (2017). Statistics on DEWATS implementation 1989 – 2017. BORDA, Bremen, Germany. 14 pp.

California Code of Regulations (CCR). (2015) Water Recycling Criteria, Title 22, Division 4, Chapter 3, California Code of Regulations: Sacramento, CA. https://govt.westlaw.com/calregs/Browse/Home/California/CaliforniaCodeofRegulations?guid=IE8ADB4F0D4B911DE8879F88E8B0DAAAE&originationContext=documenttoc&transitionType=Default&contextData=(sc.Default), Accessed 15 November 2020.

Cecconet, D., Molognoni, D., Callegari, A., & Capodaglio, A. G. (2017). Biological combination processes for efficient removal of pharmaceutically active compounds from wastewater: a review and future perspectives. *Journal of Environmental Chemical Engineering*, *5*(4), 3590-3603.

Choi, Y., Lee, J. H., Kim, K., Mun, H., Park, N., & Jeon, J. (2021). Identification, quantification, and prioritization of new emerging pollutants in domestic and industrial effluents, Korea: Application of LC-HRMS based suspect and non-target screening. Journal of Hazardous Materials, 402, 123706.

El Hayany, B., El Fels, L., Dignac, M. F., Quenea, K., Rumpel, C., & Hafidi, M. (2020). Pyrolysis-GCMS as a Tool for Maturity Evaluation of Compost from Sewage Sludge and Green Waste. Waste and Biomass Valorization, 1-14.

Gago-Ferrero, P., Schymanski, E. L., Bletsou, A. A., Aalizadeh, R., Hollender, J., & Thomaidis, N. S. (2015). Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS. Environmental science & technology, 49(20), 12333-12341.

Gani, K. M., Hlongwa, N., Abunama, T., Kumari, S., & Bux, F. (2020). Emerging contaminants in South African water environment-A critical review of their occurrence, sources and ecotoxicological risks. Chemosphere, 128737.

Gómez, M. J., Bueno, M. M., Lacorte, S., Fernández-Alba, A. R., & Agüera, A. (2007). Pilot survey monitoring pharmaceuticals and related compounds in a sewage treatment plant located on the Mediterranean coast. Chemosphere, 66(6), 993-1002.

Grandclément, C., Seyssiecq, I., Piram, A., Wong-Wah-Chung, P., Vanot, G., Tiliacos, N., ... & Doumenq, P. (2017). From the conventional biological wastewater treatment to hybrid processes, the evaluation of organic micropollutant removal: a review. Water Research, 111, 297-317.

Gumbi, B.P., Moodley, B., Birungi, G., Ndungu, P.G. (2019). Target, Suspect and Non-Target Screening of Silylated Derivatives of Polar Compounds Based on Single Ion Monitoring GC-MS. International Journal of Environmental Research and Public Health 16 (20), 4022.

Hamza, R. A., Iorhemen, O. T., & Tay, J. H. (2016). Occurrence, impacts and removal of emerging substances of concern from wastewater. Environmental Technology & Innovation, 5, 161-175.

Harb, M., Lou, E., Smith, A. L., & Stadler, L. B. (2019). Perspectives on the fate of micropollutants in mainstream anaerobic wastewater treatment. Current opinion in biotechnology, 57, 94-100.

He, J., Qin, W., Zhang, X., Wen, Y., Su, L., & Zhao, Y. (2013). Linear and nonlinear relationships between biodegradation potential and molecular descriptors/fragments for organic pollutants and a theoretical interpretation. Science of the total environment, 444, 392-400.

Herrero, P., Borrull, F., Pocurull, E., & Marcé, R. M. (2013). Efficient tandem solid-phase extraction and liquid chromatography-triple quadrupole mass spectrometry method to determine polar benzotriazole, benzothiazole and benzenesulfonamide contaminants in environmental water samples. Journal of Chromatography A, 1309, 22-32.

Herrero, P., Borrull, F., Pocurull, E., & Marcé, R. M. (2014). A quick, easy, cheap, effective, rugged and safe extraction method followed by liquid chromatography-(Orbitrap) high resolution mass spectrometry to determine benzotriazole, benzothiazole and benzenesulfonamide derivates in sewage sludge. Journal of Chromatography A, 1339, 34-41.

Hollender, J., Schymanski, E. L., Singer, H. P., & Ferguson, P. L. (2017). Nontarget screening with high resolution mass spectrometry in the environment: ready to go? Environmental Science & Technology, 51(20), 11505–11512.

698 Horii, Y., Reiner, J. L., Loganathan, B. G., Kumar, K. S., Sajwan, K., & Kannan, K. (2007). Occurrence
699 and fate of polycyclic musks in wastewater treatment plants in Kentucky and Georgia, USA.
700 Chemosphere, 68(11), 2011-2020.

701 Hrubik, J., Glisic, B., Tubic, A., Ivancev-Tumbas, I., Kovacevic, R., Samardzija, D., ... & Kaisarevic, S.
702 (2016). Toxicological and chemical investigation of untreated municipal wastewater: fraction-and
703 species-specific toxicity. Ecotoxicology and Environmental Safety, 127, 153-162.

704 Hug, C., Ulrich, N., Schulze, T., Brack, W., & Krauss, M. (2014). Identification of novel micropollutants
705 in wastewater by a combination of suspect and nontarget screening. Environmental pollution, 184, 25-32.

706 Kattan, D., Palafox, M. A., Kumar, S., Manimaran, D., Joe, H., & Rastogi, V. K. (2014). FT-IR, FT-
707 Raman spectra and other molecular properties of 2, 4-dichlorobenzonitrile: A interpretation by a DFT
708 study. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 123, 89-97.

709 K'oreje, K.O., Demeestere, K., De Wispelaere, P., Vergeynst, L., Dewulf, J., Van Langenhove, H. (2012).
710 From multi-residue screening to target analysis of pharmaceuticals in water: development of a new
711 approach based on magnetic sector mass spectrometry and application in the Nairobi River basin, Kenya.
712 Sci. Total Environ. 437, 153e164.

713 Krichling, J. (2018). Non-targeted analysis for discovery of chemicals of emerging concerns in treated
714 water for drinking and source investigation. MS Thesis. San Diego State University, 54 pp.

715 Krzeminski, P., Tomei, M. C., Karaolia, P., Langenhoff, A., Almeida, C. M. R., Felis, E., ... & Rizzo, L.
716 (2019). Performance of secondary wastewater treatment methods for the removal of contaminants of
717 emerging concern implicated in crop uptake and antibiotic resistance spread: A review. Science of the
718 Total Environment, 648, 1052-1081.

719 Kumar, N., Zhao, H. N., Awoyemi, O., Kolodziej, E. P., & Crago, J. (2021). Toxicity Testing of Effluent-
720 Dominated Stream Using Predictive Molecular-Level Toxicity Signatures Based on High-Resolution
721 Mass Spectrometry: A Case Study of the Lubbock Canyon Lake System. Environmental Science &
722 Technology, 55(5), 3070-3080.

723 Kuster, M., de Alda, M. L., & Barceló, D. (2009). Liquid chromatography–tandem mass spectrometric
724 analysis and regulatory issues of polar pesticides in natural and treated waters. Journal of
725 Chromatography A, 1216(3), 520-529.

726 Lara-Martín, P. A., Chiaia-Hernández, A. C., Biel-Maeso, M., Baena-Nogueras, R. M., & Hollender, J.
727 (2020). Tracing urban wastewater contaminants into the Atlantic Ocean by nontarget screening.
728 Environmental Science & Technology, 54(7), 3996-4005.

729 Leclercq, M., Mathieu, O., Gomez, E., Casellas, C., Fenet, H., & Hillaire-Buys, D. (2009). Presence and
730 fate of carbamazepine, oxcarbazepine, and seven of their metabolites at wastewater treatment plants.
731 *Archives of Environmental Contamination and Toxicology*, *56*(3), 408.

732 Liu, R., Ruan, T., Song, S., Lin, Y., & Jiang, G. (2015). Determination of synthetic phenolic antioxidants
733 and relative metabolites in sewage treatment plant and recipient river by high performance liquid
734 chromatography–electrospray tandem mass spectrometry. Journal of Chromatography A, 1381, 13-21.

735 Massoud, M. A., Tarhini, A., & Nasr, J. A. (2009). Decentralized approaches to wastewater treatment and
736 management: applicability in developing countries. Journal of environmental management, 90(1), 652-
737 659.

738 McManus, S. L., Moloney, M., Richards, K. G., Coxon, C. E., & Danaher, M. (2014). Determination and
739 occurrence of phenoxyacetic acid herbicides and their transformation products in groundwater using ultra

high performance liquid chromatography coupled to tandem mass spectrometry. *Molecules*, *19*(12), 20627-20649.

Nagi, M. N., & Almakki, H. A. (2009). Thymoquinone supplementation induces quinone reductase and glutathione transferase in mice liver: possible role in protection against chemical carcinogenesis and toxicity. Phytotherapy Research: An International Journal Devoted to Pharmacological and Toxicological Evaluation of Natural Product Derivatives, 23(9), 1295-1298.

Nolte, T. M., Pinto-Gil, K., Hendriks, A. J., Ragas, A. M., & Pastor, M. (2018). Quantitative structure–activity relationships for primary aerobic biodegradation of organic chemicals in pristine surface waters: starting points for predicting biodegradation under acclimatization. Environmental Science: Processes & Impacts, 20(1), 157-170.

Nürenberg, G., Kunkel, U., Wick, A., Falås, P., Joss, A., & Ternes, T. A. (2019). Nontarget analysis: A new tool for the evaluation of wastewater processes. Water research, 163, 114842.

Palafox, M. A., Bhat, D., Goyal, Y., Ahmad, S., Joe, I. H., & Rastogi, V. K. (2015). FT-IR and FT-Raman spectra, MEP and HOMO–LUMO of 2, 5-dichlorobenzonitrile: DFT study. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *136*, 464-472.Parry, E., & Young, T. M. (2016). Comparing targeted and non-targeted high-resolution mass spectrometric approaches for assessing advanced oxidation reactor performance. *Water research*, *104*, 72-81.

Parveen, Z., Brunhofer, G., Jabeen, I., Erker, T., Chiba, P., & Ecker, G. F. (2014). Synthesis, biological evaluation and 3D-QSAR studies of new chalcone derivatives as inhibitors of human P-glycoprotein. Bioorganic & medicinal chemistry, 22(7), 2311-2319.

Pérez, J. S., Arzate, S., Soriano-Molina, P., Sánchez, J. G., López, J. C., & Plaza-Bolaños, P. (2020). Neutral or acidic pH for the removal of contaminants of emerging concern in wastewater by solar photo-Fenton? A techno-economic assessment of continuous raceway pond reactors. Science of The Total Environment, 736, 139681.

Pitás, V., Somogyi, V., Kárpáti, Á., Thury, P., & Fráter, T. (2020). Reduction of chemical oxygen demand in a conventional activated sludge system treating coke oven wastewater. Journal of Cleaner Production, 273, 122482.

Qian, Y., Wang, X., Wu, G., Wang, L., Geng, J., Yu, N., & Wei, S. (2021). Screening priority indicator pollutants in full-scale wastewater treatment plants by non-target analysis. Journal of Hazardous Materials, 125490.

Ramjaun, S. N., Wang, Z., Yuan, R., & Liu, J. (2015). Can electrochemical oxidation techniques really decontaminate saline dyes wastewater?. *Journal of Environmental Chemical Engineering*, *3*(3), 1648-1653.

Reynaud, N., & Buckley, C. (2015). Field-data on parameters relevant for design, operation and monitoring of communal decentralized wastewater treatment systems (DEWATS). Water Practice and Technology, 10(4), 787-798.

Sandín-España, P., et al. (2002). Determination of herbicide tepraloxydim and main metabolites in drinking water by solid-phase extraction and liquid chromatography with UV detection. Chromatographia 55, 681-686.

Sathyamoorthy, S., & Ramsburg, C. A. (2013). Assessment of quantitative structural property relationships for prediction of pharmaceutical sorption during biological wastewater treatment. Chemosphere, 92(6), 639-646.

Schollée, J. E., Hollender, J., & McArdell, C. S. (2021). Characterization of advanced wastewater treatment with ozone and activated carbon using LC-HRMS based non-target screening with automated trend assignment. Water Research, 200, 117209.

Schymanski, E. L., Singer, H. P., Slobodnik, J., Ipolyi, I. M., Oswald, P., Krauss, M., ... & Hollender, J. (2015). Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. Analytical and bioanalytical chemistry, 407(21), 6237-6255.

Simonich, S. L., Federle, T. W., Eckhoff, W. S., Rottiers, A., Webb, S., Sabaliunas, D., et al. (2002). Removal of fragrance materials during U.S. and European wastewater treatment. Environmental Science & Technology, 36, 2839–2847.

Singer, H. P., Wössner, A. E., McArdell, C. S., & Fenner, K. (2016). Rapid screening for exposure to "non-target" pharmaceuticals from wastewater effluents by combining HRMS-based suspect screening and exposure modeling. Environmental science & technology, 50(13), 6698-6707.

Soliman, M. A., Pedersen, J. A., Park, H., Castaneda-Jimenez, A., Stenstrom, M. K., & Suffet, I. H. (2007). Human pharmaceuticals, antioxidants, and plasticizers in wastewater treatment plant and water reclamation plant effluents. Water Environment Research, 79(2), 156-167.

Soriano-Molina, P., Plaza-Bolaños, P., Lorenzo, A., Agüera, A., Sánchez, J. G., Malato, S., & Pérez, J. S. (2019). Assessment of solar raceway pond reactors for removal of contaminants of emerging concern by photo-Fenton at circumneutral pH from very different municipal wastewater effluents. Chemical Engineering Journal, 366, 141-149.

Taheran, M., Brar, S. K., Verma, M., Surampalli, R. Y., Zhang, T. C., & Valéro, J. R. (2016). Membrane processes for removal of pharmaceutically active compounds (PhACs) from water and wastewaters. *Science of the Total Environment*, *547*, 60-77.

Tisler, S., Liang, C., Carvalho, P. N., & Bester, K. (2021). Identification of more than 100 new compounds in the wastewater: Fate of polyethylene/polypropylene oxide copolymers and their metabolites in the aquatic environment. *Science of The Total Environment*, *761*, 143228.

Tohidi, F., & Cai, Z. (2017). Fate and mass balance of triclosan and its degradation products: Comparison of three different types of wastewater treatments and aerobic/anaerobic sludge digestion. Journal of Hazardous Materials, 323, 329-340.

Tousova, Z., Froment, J., Oswald, P., Slobodník, J., Hilscherova, K., Thomas, K. V., ... & Blaha, L. (2018). Identification of algal growth inhibitors in treated waste water using effect-directed analysis based on non-target screening techniques. Journal of hazardous materials, 358, 494-502.

Tran, C. D., Dodder, N. G., Quintana, P. J. E., Watanabe, K., Kim, J.H., Hovell, M.F., et al. (2020) Organic contaminants in human breast milk identified by non-targeted analysis. Chemosphere. 238: 124677. doi:10.1016/j.chemosphere.2019.124677

Tran, N. H., & Gin, K. Y. H. (2017). Occurrence and removal of pharmaceuticals, hormones, personal care products, and endocrine disrupters in a full-scale water reclamation plant. Science of the Total Environment, 599, 1503-1516.

Tran, N. H., Reinhard, M., & Gin, K. Y. H. (2018). Occurrence and fate of emerging contaminants in municipal wastewater treatment plants from different geographical regions-a review. Water Research, 133, 182-207.

Ulrich, E. M., Sobus, J. R., Grulke, C. M., Richard, A. M., Newton, S. R., & Strynar, M. J. (2019). EPA's non-targeted analysis collaborative trial (ENTACT): genesis, design, and initial findings. Anal. Bioanal. Chem. 411(853–866). doi:10.1007/s00216-018-1435-6.

U.S. Food and Drug Administration (FDA). (2016). Safety and Effectiveness of Consumer Antiseptics; Topical Antimicrobial Drug Products for Over-the-Counter Human Use. Final Rule 81 FR 61106, 61106-61130.

Völker, J., Castronovo, S., Wick, A., Ternes, T. A., Joss, A., Oehlmann, J., & Wagner, M. (2016). Advancing biological wastewater treatment: extended anaerobic conditions enhance the removal of endocrine and dioxin-like activities. Environmental Science & Technology, 50(19), 10606-10615.

Wang, X., Yu, N., Yang, J., Jin, L., Guo, H., Shi, W., ... & Wei, S. (2020). Suspect and non-target screening of pesticides and pharmaceuticals transformation products in wastewater using QTOF-MS. Environment International, 137, 105599.

Wegner (2004). JOELib Tutorial. A Java based cheminformatics/computational chemistry package. Dept. Cognitive Systems, University of Tübingen, Germany.

Wilkinson L, Friendly M. The History of the Cluster Heat Map. The American Statistician. (2009) 63: 179–184. doi:10.1198/tas.2009.0033.

Williams, A. J., Grulke, C. M., Edwards, J., McEachran, A. D., Mansouri, K., Baker, N. C., ... & Richard, A. M. (2017). The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. Journal of Cheminformatics, 9(1), 61.

Xu, P., Han, H., & Shi, Z. (2017). Development of QSBR models for anoxic biodegradability of polycyclic aromatic hydrocarbons by using SMLR and BP-ANN. Separation and Purification Technology, 178, 1-5.

Yadav, S., & Chandra, R. (2018). Detection and assessment of the phytotoxicity of residual organic pollutants in sediment contaminated with pulp and paper mill effluent. Environmental monitoring and assessment, 190(10), 581.