# CHARACTERIZATION OF DISSOLVED ORGANIC MATTER IN WASTEWATER USING LIQUID CHROMATOGRAPHY-HIGH RESOLUTION MASS SPECTROMETRY

**Yaroslav Verkh**


Universitat de Girona

DOCTORAL THESIS

# CHARACTERIZATION OF DISSOLVED ORGANIC MATTER IN WASTEWATER USING LIQUID CHROMATOGRAPHY-HIGH RESOLUTION MASS SPECTROMETRY

Yaroslav Verkh

January 2019

Universitat de Girona

DOCTORAL THESIS

# CHARACTERIZATION OF DISSOLVED ORGANIC MATTER IN WASTEWATER USING LIQUID CHROMATOGRAPHY-HIGH RESOLUTION MASS SPECTROMETRY

The thesis contains four appendices

Yaroslav Verkh

January 2019

Water Science and Technology Doctoral Programme

Supervised by: Prof. Mira Petrovic

Tutor: Prof. Claudia Fontas Rigau

Presented to obtain the degree of PhD at the University of Girona

*To my family*

*It is not important to be better than someone else, but to be better than yesterday.*

*Kano Jigoro*

# DECLARATION

This dissertation is the result of my own work and does not include results which are the outcome of work done in collaboration, except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university or institution for any degree, diploma, or other qualification.

Signed:

Date: 11.07.18

Yaroslav Verkh, M.Sc.

Girona

# THESIS SUPERVISION CERTIFICATE

Dr Mira Petrovic, ICREA Research Professor at Institut Català de Recerca de l'Aigua (ICRA),

I DECLARE:

That the thesis titled *characterization of dissolved organic matter in wastewater using liquid chromatography-high resolution mass spectrometry*, presented by Yaroslav Verkh to obtain a doctoral degree, has been completed under my supervision and meets the requirements to opt for an International Doctorate.

For all intents and purposes, I hereby sign this document.

Girona, 29.10.2018

# ABSTRACT

Individual hazardous chemicals and substance mixtures with synergistic toxicity effects occur in the dissolved organic matter (DOM) of wastewater and negatively impact human health. Yet a large number of chemicals and their treatment by-products in wastewater makes the tracking of individual compounds nearly impossible and demands new analytical strategies.

The thesis describes the development and evaluation of non-targeted and suspect analysis methods aimed at the transformation of DOM and micro-contaminants of interest during wastewater treatment using liquid chromatography-high-resolution mass spectrometry (LC-HRMS) data.

On one hand, a non-targeted method to track transformations of DOM in a multiphase wastewater treatment using LC-HRMS data was developed. LC-MS signals were extracted, aligned, and had their isotopologues clustered and elemental composition predicted using open license software MZmine 2 in a way that conceptually prioritized the detection of anthropogenic compounds. After, unreliable signals were removed using own methodology in R computing environment. This data analysis revealed that the secondary treatment removed 67% detected DOM signals, while 24% new signals appeared relatively to the influent DOM. The number of large molecules (> 450 Da) decreased and the number of unsaturated molecular features of the effluent organic matter (OM) increased. Van Krevelen plots revealed the distribution of unsaturation and heteroatoms.

The non-targeted methodology was applied to quadrupole-time of flight (QTOF) and LTQ-Orbitrap HRMS data to test the influence of instrumental setup on analytical conclusions. The Orbitrap setup showed a higher amount of detected molecular features and a higher rate of molecular formula assignment for the detected molecular features. The QTOF setup uncovered a subset of high molecular weight features due to the stronger stability of resolving power in the QTOF setup at higher *m/z*.

On the other hand, this thesis approached the transformation of chemicals in wastewater DOM by analyzing micro-contaminants and their confirmed, as well as tentative, transformation products (TPs) to supplement the developed non-targeted analysis. The suspect screening of micro-contaminants, using Thermo Exact Finder software and an own post-processing pipeline in R, yielded 11 pharmaceuticals and their TPs in a multistage WWT using own detection and data processing method. Hereby, secondary treatment was efficient in removing chemicals, while the removal in tertiary treatment was less consistent. Additionally, the screening method was applied to a novel granular bio-activated carbon-ultrafiltration (BAC-UF) system. The analysis of 11 pharmaceuticals and their TPs in four time series showed that BAC had led to a high removal of chemicals in the first months of the reactor runtime, while the efficiency dropped in the latter months due to BAC filter saturation.

# RESUMEN

El destino de los microcontaminantes en el tratamiento de aguas residuales ha ido ganando atención debido al impacto ecológico y toxicológico de estas sustancias en la salud humana. Adicionalmente a los productos químicos peligrosos individuales, las mezclas de sustancias como las que ocurren en la materia orgánica disuelta (MOD) en las aguas residuales, pueden producir efectos de toxicidad sinérgicos. Sin embargo, el gran número de compuestos naturales y antropogénicos, así como sus productos de transformación de tratamiento (PTs) provocan que el seguimiento de los compuestos individuales sea casi imposible, exigiendo nuevas estrategias para realizar un seguimiento eficaz de transformaciones MOD.

Esta tesis describe un método desarrollado y probado para rastrear transformaciones de MOD en aguas residuales, utilizando un análisis no dirigido de cromatografía de líquidos y espectrometría de masas de alta resolución (LC-MS). Potencialmente, el análisis prioriza la detección de compuestos antropogénicos. Se extrajeron y alinearon las señales de LC-MS, se agruparon sus isotopólogos y se predijo la composición elemental utilizando MZmine 2. Posteriormente, las señales no fiables se eliminaron usando una metodología propia en el entorno informático R. Este análisis de datos reveló que el tratamiento secundario eliminó el 67% de las señales de MOD detectada, mientras que un 24% de nuevas señales aparecieron.

Se observó una reducción en el número de moléculas grandes (> 450 Da) y un aumento en las moléculas insaturadas de la materia orgánica del efluente. Los gráficos de Van Krevelen revelaron la distribución de la insaturación y los heteroátomos. Además, el método se aplicó a los datos de cuadrupolo-tiempo de vuelo (QTOF) y de Orbitrap HRMS para probar la influencia de la configuración instrumental en las conclusiones analíticas. Orbitrap destacó en términos de precisión y mayor poder de resolución, por ejemplo, mostrando una mayor cantidad de moléculas detectadas y la tasa de asignación de fórmula molecular. QTOF mostró ventajas al incluir un grupo de moléculas de alto peso debido a la mayor estabilidad del poder de resolución de QTOF a mayor *m/z*.

Esta tesis reveló la transformación de sustancias químicas en aguas residuales buscando microcontaminantes específicos y sus PTs confirmados, y su potencial uso para complementar el desarrollado análisis del tipo “non-target”.

La transformación de microcontaminantes en MOD se exploró mediante exámenes de 11 productos farmacéuticos y sus PTs, utilizando métodos propios de detección y procesamiento de datos. El tratamiento secundario fue eficaz para eliminar los productos químicos, mientras que la eliminación en el tratamiento terciario varió en la filtración de arena, la radiación UV y la cloración. Adicionalmente, el método de selección se aplicó a un nuevo sistema de carbón bioactivado granular con una ultrafiltración posterior (CBA-UF). La transformación de 11 productos farmacéuticos y sus PTs se exploró en cuatro series temporales. CBA condujo a una alta eliminación de productos químicos en los primeros meses de ejecución del reactor, mientras que la eficiencia disminuyó en los últimos meses debido a la saturación del filtro CBA.

# RESUM

El destí dels micro-contaminants en tractament d'aigua ha passat a un primer pla degut al impacte eco-toxicològic d'aquestes substancies en la salut humana. A més de químics perillosos individuals, mescles de substancies que ocorren en la matèria orgànica dissolta (DOM en angles) en aigües residuals poden produir efectes tòxics sinergètics. No obstant, el gran nombre de compostos naturals i antropogènics, així com els productes del seu tractament (TPs) comporta que el seguiment de components individuals sigui pràcticament impossible, el que demana noves estratègies de seguiment per les transformacions de DOM.

Aquesta tesi descriu un mètode desenvolupat i testejat per seguir transformacions de DOM de aigua residual mitjançant anàlisis no-específic de cromatografia liquida – dades d'espectrometria de massa de alta resolució (LC-HRMS en angles) que prioritza conceptualment la de compostos antropogènics. Senyals LC-MC han sigut extretes i alineades; els seus isotopòlogs han sigut agrupats i la seva composició elemental ha sigut predita usant MZmine 2. Després, senyals que no eren fiables han sigut eliminades utilitzant una metodologia pròpia desenvolupada en el llenguatge computacional R. Aquest anàlisi ha revelat que el tractament secundari elimina 67% de les senyals DOM detectades, mentre que 24% de senyals noves apareixen. Una reducció en el nombre de molècules grans (> 450 Da) i un increment en característiques moleculars no-saturades del efluent de materia orgànica (OM) ha sigut observat. Gràfics "Van Krevelen" han revelat la distribució de insaturació i heteroàtoms. Addicionalment, el mètode va ser aplicat a temps-quadrupole de vol (QTOF en angles) i dades Orbitrap HRMS per testejar la influencia de la preparació instrumental en les conclusions analítiques. Orbitrap va ser superior en termes de precisió i poder de resolució, per eixample mostrant un nombre major de característiques moleculars detectades i un major ritme de assignació de formules moleculars. QTOF va demostrar avantatges en incloure un subset de característiques de alt pes molecular degut a la estabilitat més forta de resolució de a majors *m/z*.

Aquesta tesi ha revelat la transformació de químics en aigua residual mitjançant la cerca de micro-contaminants particulars i els seus TPs confirmats i temptatius per suplementar el anàlisi no-específic desenvolupat.

La transformació de micro-contaminants individuals a la DOM ha sigut explorat mitjançant l'escanejat de 11 farmacèutics i els seus TPs utilitzant el procés de detecció i processament de dades propi. El tractament secundari va ser eficient eliminant químics, mentre que la eliminació del tractament terciari va variar en filtres d'arena, irradiació UV i cloració. Addicionalment, el mètode de escanejat va ser aplicat a un nou sistema de carbó-ultrafiltració bio-activat granular (BAC-UF). La transformació de 11 farmacèutics i els seus TPs va ser explorada en 4 series temporals. BAC va comportar una gran eliminació de químics durant els primers mesos de funcionament del reactor però la eficiència va caure durant els últims mesos degut a la saturació del filtre.

# FINANCIAL ACKNOWLEDGMENTS

This work has been supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642904 - TreatRec ITN-EID project.

TreatRec

EUROPEAN INDUSTRIAL DOCTORATE

# ACKNOWLEDGMENTS

In life, we do not meet many people that truly shape who we are and guide us to become a better person. I am happy to say that Mira Petrovic (ICRA, Spain) was one of them. She was the greatest supervisor, supporting and leading me on the often thorny path of a doctorate. Mira was always helping with academic and personal matters, no matter the distance between us. Her integrity, honesty, and kindness will inspire me for the years to come.

Thanks to Marko Rozman (Ruđer Bošković Institute, Croatia) for populating my brain with ideas that became a part of this thesis, mutual work, and all the trips we shared.

The samplings by Lluis Corominas (ICRA, Spain) and Gianluigi Buttiglieri (ICRA, Spain) is greatly acknowledged. Many thanks to Mira Celic (ICRA, Spain) for the working up and measurement of samples used in this study and many hours of help in the office. Many thanks to Lucia Gusmaroli (ICRA, Spain) for the working up of samples that, unfortunately, barely escaped the scope of this thesis. A big thank you to Núria Cáceres (ICRA, Spain) for showing me the ropes and the help in the lab.

I thank Antoni Ginebreda (IDAEA, Spain) for the advice on the application of statistical tools. And thanks to Wolfgang Gernjak (ICRA, Spain) for the expertise on tertiary treatment. Their motivation was crucial in keeping me on the topic that makes up a large part of this thesis.

Big thanks to Alessio Fenu (Aquafin, Belgium) for his expert opinion on biological activated carbon. Moreover, thanks for the talks when the lunchtime was especially lonely, the Belgian beer, and morra tournaments.

Thanks to Bozo Zonja (IDAEA, Spain) for providing a database of chemicals used in the thesis and his patience in explaining micro-contaminants, transformation products, and software packages to a novice.

Huge thank you to TreatRec fellows Luca Sbardella (Aquafin, Belgium), Pau Gimeno (ICRA, Spain), Pau Juan Garcia (Atkins, UK), and Sara Johannson (Lequia, Spain). Thank you for your support and kindness throughout these years. Special thanks to Luca for providing the samples used in this thesis and to Pau Juan Garcia for the abstract translation.

Many thanks to Barbara Kasprzyk-Hordern and her group of researchers at the University of Bath, UK for providing the equipment and materials used to produce a large chunk of this thesis. Special thanks to Shaun Rijksteen (University of Bath, UK) who spend endless hours calibrating and testing the equipment to squeeze out the last bit of data from my samples.

Many thanks to Vera Jones (Atkins, UK) for the nonstop help in Bristol and navigating me through the complex world of consultancies. Thanks to Arthur Thornton, Carlos Constantino, and Peter Daldorph (Atkins, UK) for training and valuable input on micro-contaminant analysis.

Endless thanks to Bego Morejón for the understanding, patience, and making me smile everytime. Without you, this Ph.D. and my life would have been very lonely.

Thank you to Kayla Friedman and Malcolm Morgan (University of Cambridge, UK) for the Microsoft Word thesis template used to produce this document.

# CONTENTS

# LIST OF PUBLICATIONS DERIVED FROM THE THESIS

Verkh Y, Rozman M, Petrovic M. A non-targeted high-resolution mass spectrometry data analysis of dissolved organic matter in wastewater treatment. Chemosphere. 2018 Jun 1;200:397–404.

Yaroslav Verkh, Marko Rozman, Mira Petrovic. Extraction and cleansing of data for a non-targeted analysis of high-resolution mass spectrometry data of wastewater. MethodsX. 2018 Apr 17.

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS AND ACRONYMS

| Abbreviation | Full expression |
|---|---|
| %H | percentage of hydrogen |
| %wt | weight percent |
| CV | coefficient of variation |
| DBP | disinfection by-product |
| DOM | dissolved organic matter |
| ESI | electro-spray ionization |
| HILIC | hydrophilic interaction liquid chromatography |
| HLB | hidrophilic-lipophilic balanced |
| HRMS | high-resolution mass spectrometry |
| L | litre |
| LC | liquid chromatography |
| log P | logarithmic partition coefficient |
| M | molar concentration |
| MAX | Mixed-mode, strong Anion-eXchange |
| MCX | Mixed-mode, strong Cation-eXchange |
| MeOH | methanol |
| MS2 | tandem mass spectrometry |
| $Na_2EDTA$ | Ethylenediaminetetraacetic acid disodium salt |
| NI | negative ionization |
| NOM | natural organic matter |
| OM | organic matter |
| PCA | principal component analysis |
| PI | positive ionization |
| RMD | relative mass defect |
| SPE | solid phase extraction |
| TP | transformation product |
| UV | ultra violett |
| Δ | difference |

# LIST OF APPENDICES

# 1 INTRODUCTION

## 1.1 Motivation

Wastewater **dissolved organic matter** (DOM) represents a complex, heterogenic mixture of polysaccharides, proteins, lipids, nucleic acids, soluble microbial products, and anthropogenic organic chemicals. Among these chemicals, the complement of anthropogenic substances used by humans and effectually disposed of in wastewater encompasses thousands of compounds. It includes, among others, surfactants, personal care products, pharmaceuticals, biocides, pesticides, and industrial chemicals [1]. Additionally, there is a wide range of biologically active **transformation products** (TPs), intermediates, metabolites [2], and tertiary treatment **disinfection by-products** (DBPs) [3]. Some of these compounds can be hazardous even at a low concentration and the confirmed discharge of micro-contaminants into groundwater and water bodies for indirect potable reuse puts the end-consumer at risk [4–7].

Meanwhile, the detection of thousands potentially dangerous DOM constituents and their transformation in wastewater treatment remains a challenge, despite many methodologies to detect environmental contaminants in wastewater. This gap in knowledge impedes the innovation of treatment technologies and public health measures in respect to micro-contaminants [8,9]. For this reason, analysis of DOM and of particular micro-contaminant transformation was explored in this thesis.

## 1.2 The scope of the thesis

The standard suite of wastewater treatment consists of a grid to remove large objects and a primary treatment to settle suspended solids. It is followed up by a secondary biological treatment aiming to reduce the amount of nutrients phosphorus and nitrogen, but also **organic matter** (OM) containing the fraction of anthropogenic trace contaminants. In the last decades, tertiary treatment was added aiming to improve the quality of water by removing pathogens as in chlorination or UV-treatment or particles/DOM as in various filtration techniques. Tertiary treatment technologies were also being explored concerning the removal of micro-contaminants since the secondary biological treatment has shown to underperform in this respect.

Despite the wide applicability and robustness of various secondary and tertiary treatment setups, the prediction of micro-contaminant fate in WWT remains a scientific and engineering endeavor which has to account for the quality of influent water, economic levers, and climatic conditions [10]. The composition of effluent wastewater depends upon the type of wastewater (municipal, industrial, hospital's effluent, runoff from fields, etc.) and the treatment processes in wastewater treatment plants (WWTP) [11,12]. Accordingly, the behavior of DOM and the fraction of micro-contaminants in it varies widely. A second notable challenge is a uniform fate prediction for micro-contaminants in wastewater due to a wide range of their physiochemical properties.

The preferable analytical method to detect and characterize micro-contaminants in wastewater is mass spectrometry (MS), usually with a prior chromatographic separation of WW constituents [13]. Such separation of analytes prior to MS with ultra-high performance liquid chromatography (UHPLC) improves the resolution of signals [19]. The high precision of various novel HRMS analyzers as Fourier transform ion cyclotron resonance (FT ICR), Orbitrap or QTOF allowed reaching new levels of substance identification in environmental samples [14–16]. Such confident detection of small molecules is achieved using the chemical's exact mass, isotopic pattern, and the distribution of isotopic intensities [17]. The identification is even more confident when confirmation by tandem MS or internal standards (IS) is included in the analysis [18].

Different MS methods were developed for the analysis of WWT constituents. Initially, those methods focused on the detection of a number of contaminants and occasionally on their TPs [20]. These highly creative and functional approaches uncovered the presence of precarious trace substances, as illicit and legal drugs, but more importantly their fate and decomposition into potentially hazardous TPs in wastewater [21,22]. However, such screening methods omit thousands of DOM constituents that are present in the influent or are created during the treatment process. An HRMS analysis yields $10^3 – 10^5$ signals due to wastewater complexity, which makes a manual structural identification of so many unique substances nearly impossible. Therefore such HRMS methodologies generally uncover only a small fraction of compounds and omit the unknown majority of wastewater DOM [23]. Yet, even without a tentative structural identification of particular substances, a large number of signals with assigned elemental compositions can be used to uncover physiochemical changes in wastewater treatment.

While wastewater is a complex mixture and such analysis of unknowns is not yet widely represented in literature, the author's hope is that it will become useful in pinpointing incidents, flaws, and benefits in WWTP as it has already shown to do in other fields [26]. Yet, it should be kept in mind that HRMS analyzers themselves have the potential to affect the outcome of analyses due to the differences in resolution of masses or the ability to detect compounds based on their physiochemical properties. These differences are less pronounced in screening for particular compounds [27] but have to be kept in mind and explored further in the analysis of unknown DOM constituents [28].

The researcher community produced numerous reviews of excellence on the topics of transformation and fate of micro-contaminants and DOM in wastewater treatment [29–37]. In the current introduction, the author streamlined the literature review according to the scope of the thesis, personal interests of the author, and his proficiencies. The following sections of the Introduction chapter cover the transformation of DOM in wastewater treatment, a brief description of advances in the treatment of micro-contaminants, HRMS instrumentation, and HRMS workflows for detection of micro-contaminants such as the targeted, suspected or non-targeted search for trace substances,

## 1.3 The context of the thesis

### 1.3.1 Micro-contaminants, their TPs, and metabolites

The research interest in micro-contaminants has been increasing in the last decades due to the environmental importance of these compounds [38]. Many of them are not eliminated fully in WWTP and thus are discharged directly into the environment where they can accumulate [39–41]. Many organic substances have adverse effects on ecosystems and the human population once released into the environment either by themselves or in complex mixtures [42–45]. This complexity of impacts leads to various toxicity endpoints [46–48] which underlines the importance of micro-contaminant monitoring already in WWTP [30].

Substance classes, as legal pharmaceuticals, illicit drugs or personal care products, which show difficult elimination patterns in WWTP, take a special role in the realm of environmental micro-contaminants [49,50]. This is because various **anthropogenic substances** were designed to hinder or change bio-elimination mechanisms in living organisms, while similar metabolic processes are occurring in the biological treatment of wastewater [51].

Moreover, besides these high-level treatment targets, the influent in WWTP already contains metabolites of chemicals derived from human activity [52]. These include **Phase I metabolites**, representing products of biochemical transformation reactions. The metabolisms of the parent chemical aims at a better water solubility and excretion from the organism [53]. **Phase II metabolites** are conjugates of the parent compound with a conjugating moiety such as glucuronic acid. The conjugation reduces the toxicity and speeds up the elimination of the compound from the body [54,55]. Phase II metabolites can deconjugate and release the active parent compound in the sewer or WWTP [56,57]. Further, chemicals and their metabolites undergo biotic and abiotic reactions in the sewer network before entering WWTP [58,59].

The treatment at WWTP itself produces TPs by partial biodegradation of chemicals in secondary biological treatment [60,61]. Many TPs and metabolites retain the bioactive properties of their parents when the bioactive part of the molecule remains intact, which makes them a potential danger down the pipe [62]. As a rule of thumb, TPs show a decrease of bioactivity compared to their parents.

However, the lack of monitoring in wastewater and the subsequent unsupervised release into the environment where they can accumulate enhances the importance of their research [63].

There are multiple pathways of removal in biological treatment as the conventional activated sludge treatment. Organic chemicals, especially hydrophobic compounds with a high adsorption affinity, might be removed by adsorption onto sludge. Meanwhile, the elimination of trace contaminants in the water phase includes decomposition of compounds with metabolic and co-metabolic reactions initiated in and by bacteria. The biological treatment also produces conjugated TPs, which while reducing biochemical activity, serve as reservoirs of contaminants once they are released into the environment.

Additionally, to the secondary treatment, tertiary treatment technologies were introduced to improve the quality of effluent water by removing pathogens, suspended solids, and DOM. They were also applied to reduce the impact of micro-contaminants. However, tertiary treatment inadvertently introduced TPs in processes as UV-irradiation [64], Fenton-like processes [65–67], ozonation [68–72], and chlorination [73–85]. Tertiary treatments as sand or activated carbon filters [86–90] and membranes [8,9,91] are physical treatment methods, but can also introduce TPs by chemical or biological processes within, such as DOM transformation by a biofilm adsorbed onto the granules or membranes. Such TPs or DBPs are observed in tertiary treatment and in some cases even show a higher toxicity than the parent compounds [92]. The physiochemical mechanisms of DBP formation differ between various tertiary treatment technologies and are crucially different compared to the metabolic logic of TP formation in biological treatment [3]. Such variability adds complexity to the task of DBP detection and elimination [93].

Tertiary treatment is an indispensable tool to increase the quality of the effluent WW despite the challenges of DPB formation and removal. For example, filtration techniques are a viable option for tertiary treatment since they show a good removal of particulate OM, DOM, and micro-contaminants combined with a moderate cost of operation [86,87,89,90,94]. Among these, the application of granular activated carbon (GAC) treatment has shown to remove micro-contaminants of interest, for example, pharmaceuticals [95–97].

Here, the high adsorption capacity of GAC was responsible for the removal of micro-contaminants [98]. Additionally, contaminants were being removed by biological activity of the biomass on the surface of carbon granules fed by OM retained in the filter bed [99].

A GAC filter bed with such bioactivity is described as **biological activated carbon** (BAC) and shows a higher micro-contaminant removal compared to filtering technologies as sand filtration [88]. The combination of adsorption and biological activity in a BAC filter removes a broader range of micro-contaminants and drives the viability of big scale application of the technology [100,101].

However, the treatment and disposal of the carbon granules and the concentrate collected during the filtration remain a challenge. In addition, microorganisms attached to carbon granules can be washed out from the BAC filter causing a health risk down the pipe. The installation of an ultrafiltration (UF) treatment as an additional filtration step after the BAC has shown to prevent the washing out of bacteria [102].

Despite the challenges, BAC has proven to be an efficient and sustainable removal technology for micro-contaminants of interest reaching high removal efficiencies [103,104]. Hereby the treatment parameters as the adsorbent composition, pH, temperature regulate the removal efficiency [105,106]. For example, the time of operation correlates negatively with adsorption on GAC and the removal of trace contaminants, which drops at longer reactor operation times [107]. In addition, the quality of influent water, the individual adsorption affinity of a micro-contaminant, its charge, and hydrophobicity control the removal [108].

The tracking of organic trace contaminants and their TPs is a complex task constricted by the treatment process, the properties of the contaminant, and the environmental variables. The physicochemical properties of chemicals as hydrophobicity, solubility or volatilization are crucial to understand the elimination patterns of each contaminant [109]. WWTP properties as the sludge or hydrological retention time, size of particles, and kinetics of biodegradation make the elimination site-dependent [36,94]. Additionally, environmental variables as temperature, pH, and redox conditions influence the removal patterns on a daily or seasonal scale [110].

### 1.3.2 DOM as a reservoir for micro-contaminants

There is a large number of unknown factors in the removal of anthropogenic OM, despite the ingenious treatment technologies described in the previous chapter. Overlooking potentially hazardous constituents in WWTP effluent limits the understanding of the impact of effluent OM on the environment. Therefore, monitoring of the entire molecular complement or a sub-complement of wastewater DOM offers a possibility for a more comprehensive evaluation of the organic content in wastewater and a deeper understanding of the treatment processes and DOM transformation [33]. This new understanding will allow to study the shortcomings of the treatment processes themselves and propose evidence-based improvement strategies similar to other fields of OM research [111].

Currently, the efficiency of DOM removal at a WWTP is evaluated through measurements of the chemical oxygen demand (COD), biological oxygen demand (BOD) and the total organic carbon (TOC). Additional specialized technologies for the prioritized fractions of DOM include the assessment of aromaticity using UV absorbance ($SUVA_{254}$), size exclusion chromatography (SEC) to identify mass/size distributions of C-or N-containing constituents or excitation-emission-matrix fluorescence used to identify substance classes in natural OM (NOM) [112–114]. These techniques reveal the chemical characteristics to a certain extent and the abundance of organics in WWTP influent and effluent, Yet, they do not provide information on the presence of unique organic substances and need an effort to be combined into one data stream. Thus there is a need for strategies to assess the quality of wastewater treatment, especially in respect to anthropologic chemicals [115].

## 1.3.3 DOM detection. From target analysis and suspect screening to non-target analysis

### 1.3.3.1 DOM sample preparation

Wastewater DOM often requires a pre-concentration prior to HRMS acquisition, since the minute concentrations of trace contaminants might be below the detection limit of the instrument. Additionally, the extraction of solely a part of wastewater DOM has the benefit of avoiding damage or contamination of the analytical equipment down the line.

The solid-liquid phase extraction, more commonly called **solid phase extraction** (SPE) is prevalent in the field of wastewater analysis, due to the facility and the swiftness of the procedure compared to other extraction methods. It has a pre-concentration ability of 20 -1000 times.

The choice of SPE for pre-concentration allows for a high level of specificity when extracting a certain compound group and applying a compound-specific, targeted LC-MS method [116]. Hereby, the choice of the sorbent (solid phase) defines which complement of DOM will be selectively retained. The retention depends on the sorbent's physical and structural properties. The widely spread choices for the SPE cartridge sorbents are Oasis HLB (Hydrophilic-lipophilic balance), Oasis MCX (Mixed Cation Exchange), Oasis WCX (Weak Cation Exchange), Oasis MAX (Mixed Anion Exchange), Oasis WAX (Weak Anion Exchange) fabricated by Waters™, and Strata-X, -X-A, X-AW, -X-C, -X-CW by Strata™, as well as GCB (Graphitized Carbon Black) by Sepra™.

The balanced Oasis HLB/Strata-X cartridges are used for medium polar compounds at various pH. Meanwhile, Oasis MCX/WCX/Strata-X-C/Strata-X-W and Oasis MAX/WAX/Strata-X-A/Strata-X-AW are more suitable for basic and acidic compounds respectively. A graphene-based sorbent as GCB is an alternative to previously mentioned polymer-based sorbents and is capable to retain non-polar compounds as pesticides. However, alkyl-bonded silica phases (such as $C_{18}$) and balanced sorbents, based in the case of Oasis HLB on a copolymer of lipophilic divinylbenzenevinylpyrrolidone and hydrophilic N-vinylpyrrolidone, are prevalent for targeted LC-MS methods [117].

With an increased interest in non-targeted LC-MS analyses, came the need for a pre-concentration which did not target a specific subset of DOM, but aimed at the holistic extraction of DOM constituents while still removing equipment damaging contaminants as inorganic salts. Previously alkyl-bonded silica phases and solid phases as Bond Elut™ PPL and ENV based on styrene divinylbenzene polymers have shown an efficient extraction of NOM [118].

In a further step, a solution for the challenge of insufficient extraction was proposed when multiple solid phase materials and extraction protocols were combined into one extraction procedure. This has led to the improvement of extraction yields and the scope of extractable DOM making SPE alluring for a non-targeted analysis of DOM [86,119]. Hereby, the samples with the neutral pH had been loaded onto four connected in-line SPE cartridges with respectively balanced, acidic, basic, and non-polar properties. Subsequently, the retained DOM was extracted from each cartridge in a respective pH/solvent-specific procedure and the combined DOM has been brought back to neutral pH.

It has to be kept in mind that a measurement without pre-concentration, although it prevents a loss of DOM that is inevitable in any SPE pre-concentration, is not always the best option, due to a strong matrix effect [120]. A pre-concentration step prior to HRMS acquisition might reduce matrix-dependent signal suppression of trace contaminant mixtures. The signal suppression or sometimes enhancement is caused by the interaction of analyte ions with unwanted DOM constituents (matrix) within the sample and their competition for the HRMS analyzer [121]. However, a pre-concentration reduces the matrix effect only when the relevant fraction of wastewater DOM is being concentrated. Otherwise, the matrix is being pre-concentrated as well.

Another way to reduce matrix-dependent signal suppression is a chromatographic separation of mixtures prior to HRMS. Yet, often DOM analyses do not include a chromatographic separation that might enhance the resolution of the spectral data. A direct infusion, without a separation on a chromatographic column, can simplify the procedure. However, retention time serves as an additional variable to distinguish molecules and an advantage of LC over an injection without a separation [122].

The separation on a chromatographic column is beneficial for the reduction of the matrix effect compared to a direct infusion, since it simplifies the mixture thus increasing the chance of detecting low-intensity signals [123,124]. This is a common practice for analyses of micro-contaminants and their select TPs. Non-polar, volatile trace contaminants and TPs are separated by gas chromatography (GC) while water-soluble, medium polar compounds are separated by LC. In more recent cases hydrophilic interaction liquid chromatography (HILIC) was applied to separate mixtures of highly polar compounds [63,125]. Enantioselective chromatography is applied in cases where a mixture of stereoisomers has to be separated to reveal the active chemical [126].

### 1.3.3.2 LC-MS analytical workflows

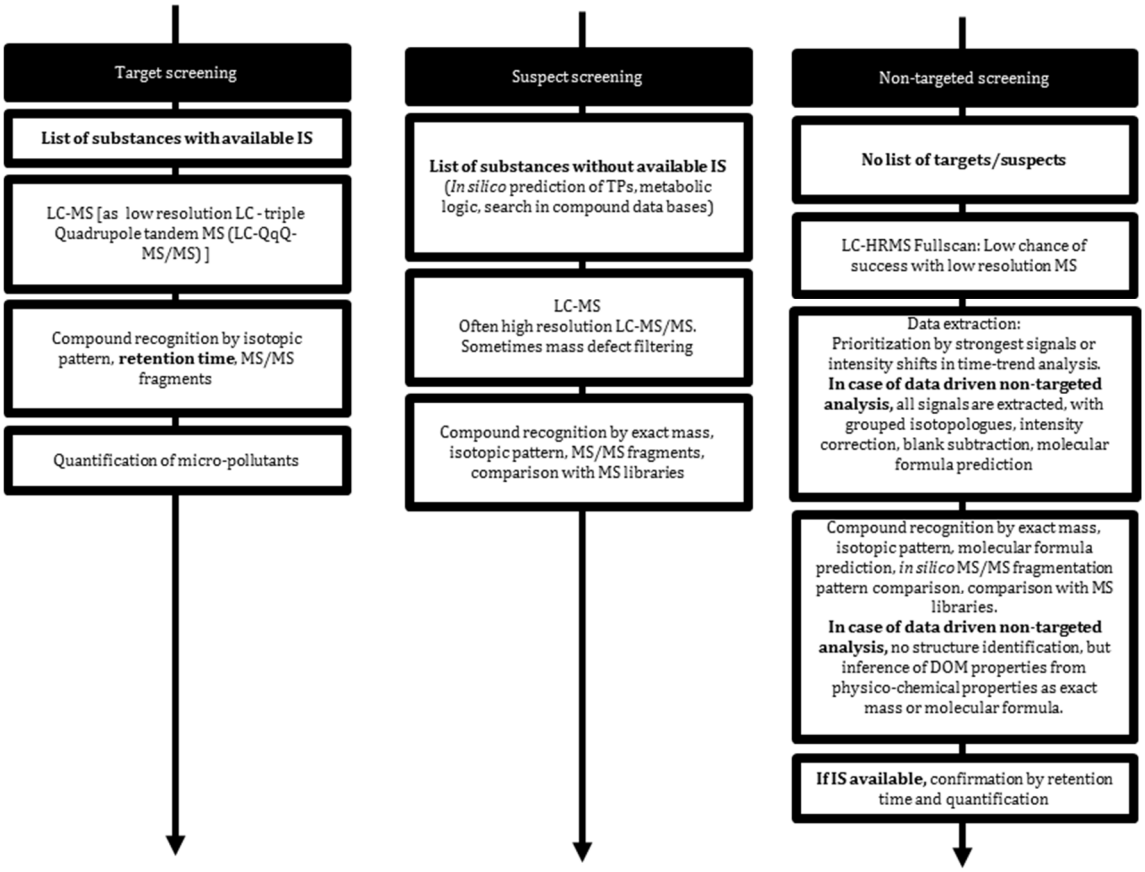



**Figure 1.1 Simplified schemes of the target, suspect, and non-targeted screenings for organic compounds using LC-MS.**

The ability of high-resolution mass spectrometry (HRMS) to identify small amounts of organic chemicals from increasingly complex mixtures provides a large amount of information on wastewater DOM and micro-contaminants of interest within. The workflows broadly separate into three major groups, namely targeted, suspect, and non-targeted screenings [127].

While the first two approaches involve a prioritization previously to the analysis by building lists of targets and tentative suspects, the non-targeted approach is rather based on observation of shifting signal intensities in time or space series analysis (Figure 1.1). HRMS approaches for the detection of trace contaminants and their TPs vary depending on the level of available data and required outcomes, with substances gaining more identification certainty moving from a non-targeted to a targeted screening protocol (Figure 1.2) [128].

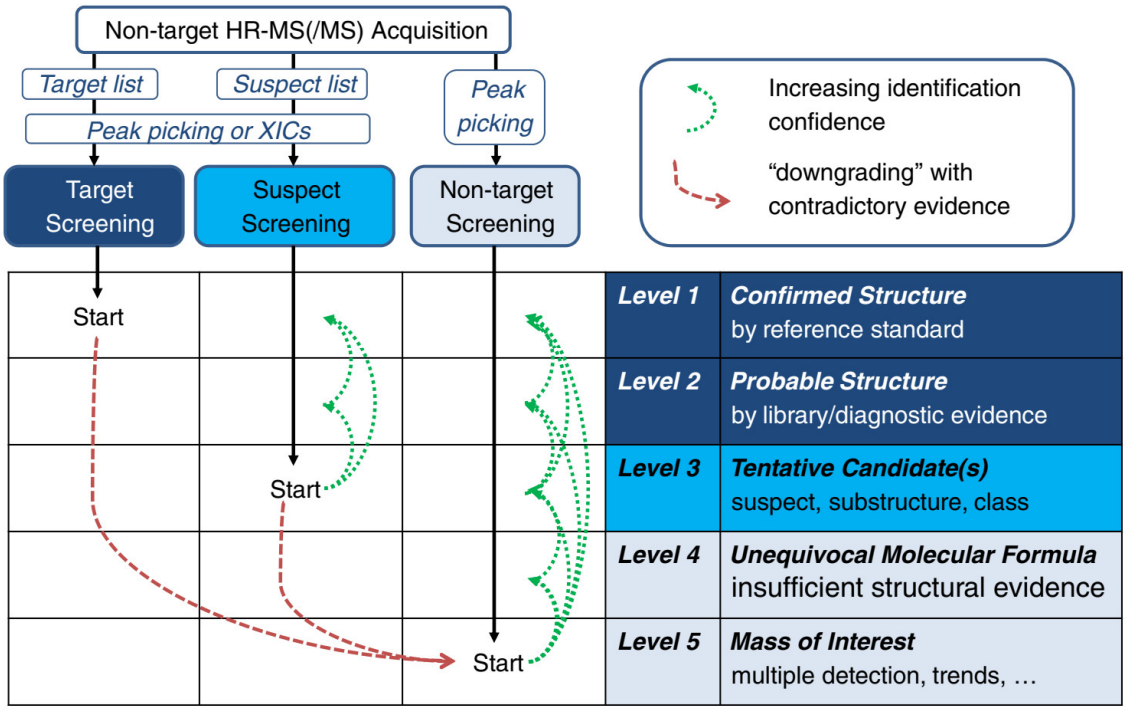



**Figure 1.2 "*Matrix of identification approach versus identification confidence*". A schematic representation of requirements and possibilities in the targeted, suspect or non-targeted screening for micro-contaminants. Reprinted from "*Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis*" by Schymanski et al., 15 May 2015 [35].**

The target analysis was used to encounter and monitor the concentration of micro-contaminants of interest, such as antibiotics and pharmaceutical compounds [129,130]. The suspect screening was applied in recent years to encounter TPs of pharmaceuticals, flame retardants, benzotriazoles using both the strategy of building a TP list using *in silico* fragmentation tools and confirming compounds derived from metabolic logic by MS/MS fragment comparison [131]. Non-targeted analysis of time series and a close combination with the suspect screening approach elucidated novel environmental micro-contaminants, such as pharmaceuticals and TPs, by prioritizing temporal shifts and prioritization by the strongest signal [132,133].

The targeted screening allows obtaining quantitative results on many chemicals in one run, given that the analytical standards corresponding to these substances match the target signals in the same experiment [14,130,134]. Usually, the protocol of targeted micro-contaminant analysis relies on LC-MS setups allowing low limits of detection and good fragmentation patterns while the precision of signals is of secondary importance. LC-triple quadrupole low-resolution tandem MS is well suitable for this task and therefore a common choice for a target analysis of micro-contaminants and TPs for which analytical standards are available.

In a suspect screening, exact monoisotopic masses of suspects are compared to candidates from manually compiled or internet databases [16,18,135,136]. IS are often not available, especially for non-commercial compounds as TPs. Therefore the routinely applied strategies are confirmation by isotopic pattern and structural confirmation by fragments in tandem MS [137–140]. In particular, HRMS suspect screening identifies the structure of dominant signals using databases of chemicals and *in silico* prediction by linking the analyte's exact mass, isotopic pattern, and fragment information from tandem MS to the predictions [18].

The post-targeted analysis is a special case of suspect screening in which an already acquired HRMS scan is explored for suspects after the measurement using suspect databases which can contain > 1000 entries [23,27]. Herein, the acquisition of tandem MS or comparing to the analytical standard of found suspects in a repeated data acquisition increases the level of confirmation.

Meanwhile, in some cases, the line between the suspect and non-targeted workflows is blurry as the research moves on both levels of identification [141]. However, a non-targeted screening for TPs involves the search for unknown TPs by comparing the isotopic pattern, structural information, or retention time of parent compounds to the signals found in MS data [19,119,132,142]. Here, the workflows can go beyond the identification of specific compounds and encompass whole groups of TPs as in unknown DBPs of bromination or chlorination [143–146]. Such non-targeted HRMS signals are prioritized due to the occurrence of intensity patterns in time series or sampling sites [133,147,148].

### 1.3.4 Non-targeted analysis of DOM

Recent attempts in wastewater analytics took advantage of the large complement of signals detectable by HRMS using rather a data-driven than a selective approach to the analysis of wastewater DOM [24]. Such non-targeted workflow aims to elucidate insights on wastewater DOM transformation based on a large number of detected signals. The currently almost impossible task of structural identification for thousands of compounds is omitted and the signals are analyzed on the level of molecular features with an exact mass, an average retention time, isotopic pattern and possibly an assigned molecular formula [25].

Such data exploration reduces the challenge of manual data treatment present in a suspect screening. This HRMS data treatment emerged from the fields of petroleomics [149], biological research [150], and characterization of natural organic matter (NOM) [28,151–154]. It was then applied for the organic matter in processed water [155–157] and wastewater [158–161]. Hereby the strategies range from identifying changes for a wastewater treatment sequence within one corresponding hydrological retention time to encompassing multiple time measurements over a range of sampling sites [162].

The protocol of a data-driven non-targeted analysis involves the extraction of ion chromatograms, a grouping of isotopologues/adducts and the assignment of elemental composition to the representative signal. The signals then are being tracked throughout all samples by using alignment algorithms based on the non-linear comparison of their *m/z* and retention time. Such data extraction yields a list of **molecular features** which correspond to compounds and have a mass, a retention time (when chromatography used in data acquisition) and if assigned, an elemental composition. A large number of obtained molecular features together with variables derived from their physio-chemical properties allows then to infer transformation of DOM using a set of analytical tools described further below in this chapter.

The mass of molecular features as an intrinsic property of chemicals offers the possibility to track transformations of DOM in subsets of signals. Also mass defects are widely applied in HRMS analysis. For example, the **Relative Mass Defect (RMD)** which is defined as the difference of an atom's exact mass from the integer-rounded mass (nominal mass) and divided by the exact mass.

A large positive RMD in organic molecules often means a large number of H atoms in a molecule. This observation allows for interesting applications in non-targeted analysis of DOM. For example, a correlation between the **percentage of hydrogen** in a formula (%H) and RMD of naturally occurring compounds predicted %H for molecular features without an assigned elemental composition [163]. This tool offers a quality test of the elemental composition assignment by comparing whether the RMD of the raw data *m/z* and the %H from the estimated molecular formula correlate well.

Also, the identification of the homolog series using a non-targeted mass defect analysis was used to map surfactants in wastewater [164]. Homologs occur in NOM [152] but more importantly in classes of anthropogenic substances like surfactants, polyfluorinated compounds or chlorine substitute series [15]. The **Kendrick Mass Defect (KMD),** which sets the exact mass of a chosen molecular fragment to a nominal value (for example 14.015 Da to 14.000 Da for -$CH_2$-) is able to identify homologous series for various structural moieties as -$C_2H_4O$- or -$H/+Cl$. The pattern recognition in Kendrick plots previously elucidated reactions, heteroatom distributions, and series of novel compounds [15,152,165,166]. The plot helped to assign elemental formulae to high masses by extrapolation from chemical homologs with a lower mass, thus avoiding the statistical error of formula prediction for masses above 400 Da.

Additionally to the mass, the extraction of chemical properties of DOM from molecular formulae has a big focus in the non-targeted analysis. As such, the **van Krevelen diagram** is an important part of DOM exploration where the atomic ratio X/C, with X being an element of interest, is plotted against H/C. In petroleomics and NOM chemistry the correlation between the areas in this plot and functional classes of compounds led to the elucidation of the chemical composition of organic matter [144,167–169].

A comparison of multiple samples using the van Krevelen plot revealed the transformation of orgainic matter (for example oxidation of DOM) [156,170,171]. The van Krevelen plot applied to a wastewater treatment revealed a possible transformation of detected DOM (Figure 1.3) [25]. On the other hand, heterogeneous DOM (found in wastewater or eutrophic river) leads to a less structured distribution of points obscuring the graphic nature of the van Krevelen plot. Such difficulty of exploring the graphical nature of the van Krevelen plot of the heterogeneous OM was experienced outside the field of water chemistry [172].

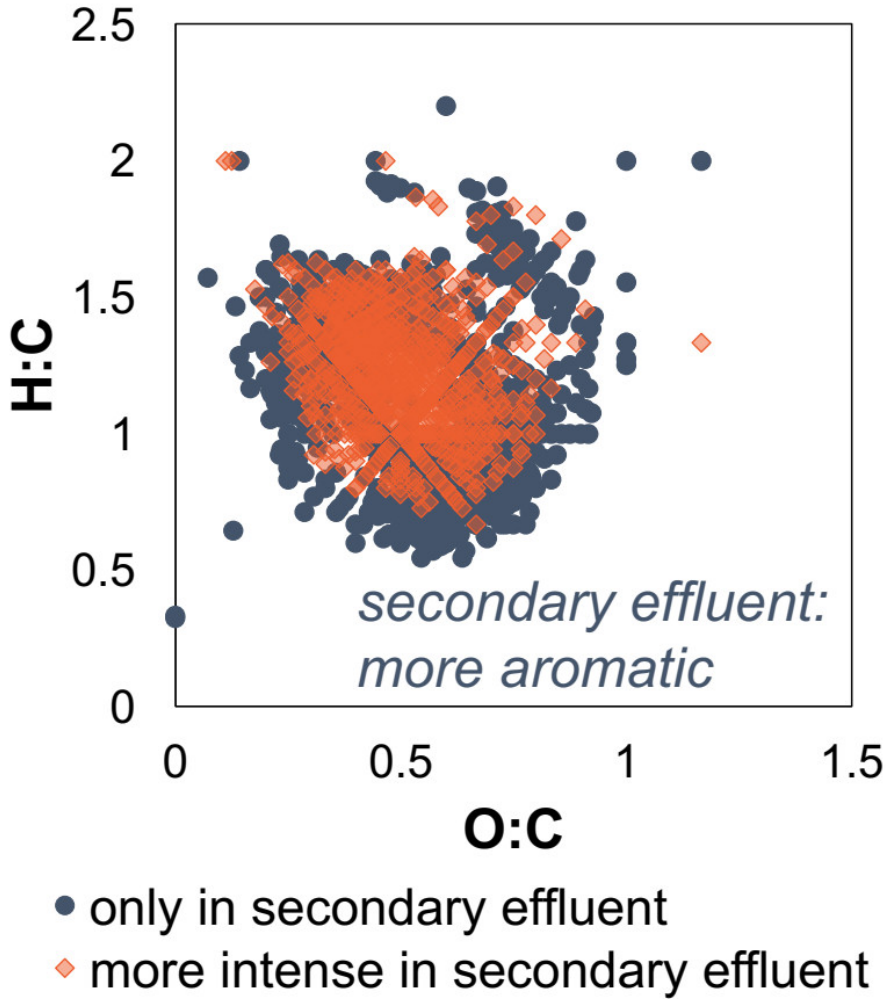



**Figure 1.3 Van Krevelen plot of DOM molecular features in secondary effluent. Reprinted from "*The Effect of Advanced Secondary Municipal Wastewater Treatment on the Molecular Composition of Dissolved Organic Matter*" by Maizel et al., 1 October 2017 [25].**

Also, the monitoring of **double bond equivalents (DBE)** in wastewater treatment can estimate the quality of the process, to recognize the hydrophobicity-altering reactions as hydrolysis or oxidation, and conceptually to even support the process control [173]. DBE reflect the level of unsaturation by double bonds in an organic molecule using only the counts of H, C, O, N, and halogen atoms in a molecule [174,175].

However, DBE do not always apply to predict aromaticity since they can include double bonds with heteroatoms. Other models to predict unsaturation of a molecule as DBE divided by the number of C atoms, **DBE minus oxygen atoms (DBE-O)**, and the aromaticity index were proposed [176]. DBE-O correlate especially well with the saturation of oxygen-rich organic compounds, but they represent a more abstract measure of unsaturation than DBE.

Finally, various **multivariate statistical tools** sieve through large amounts of data and infer DOM transformations processing all available data from a non-targeted data extraction. As such, principal component analysis (PCA) is a multivariate statistical tool which previously reduced many variables observed in HRMS analysis to a few principal components and revealed correlations between DOM quality parameters [177–179]. In addition, partial least squares discriminant analysis (PLS-DA) was previously applied in the non-targeted analysis to evaluate properties such as LC-MS setup performance [180,181].

### 1.3.5 Importance of MS instruments in non-targeted analysis

Despite the novelty of HRMS instruments, the availability of brands and instrumental setups on the market is increasingly diverse driving the importance of choosing the right instrument [182]. A crucial distinction between machines is the mass analyzer, as Quadrupole, FT ICR, Time-of-Flight (TOF) or Ion Trap. **Hybrid analyzers** such as Quadrupole (Q) – TOF or C Trap – Orbitrap expand the capabilities of acquisition by combining the properties of analyzers.

The analyzer selects the mass range of detected chemicals, the speed of processing, the sensitivity, and the cost of operation. For example, Orbitrap analyzers offer a higher resolving power, defined as the ratio of mass to full width at half maximum (FWHM), than QTOF, but at the same time, QTOF setups show a smaller decrease of resolving power with increasing *m/z* of the chemical as observed in Orbitrap or FTICR [183–185].

The second important parameter of HRMS instruments is the ionization method. In the case of LC, it is liquid-to-gas phase switching, which for organic analytes narrows down to Electron Impact Ionization, Electrospray Ionization (ESI), Atmospheric Pressure Chemical Ionization, and Laser Desorption. The ionization method narrows down the polarity, size, and desired fragmentation of detected compounds. Wastewater LC-MS makes an extensive use of ESI since it fits well the profile of medium polar trace contaminants encountered in water that can be separated by LC. Yet, parameters as resolution and precision should not be abused, independent of the choice of the analyzer and the detector. Peer-reviewed literature happens to report anecdotally low ppm error intervals of *m/z* or precise isotopic pattern distributions which, while valid for select signals, can lead to bad extraction yields in non-targeted data [186–190].

HRMS instruments show a comparably high rate of targeted identification between the instruments [27]. Yet, the similarity of data produced in non-targeted data mining is to the author's best knowledge not extensively explored in the complex matrix of wastewater. Hereby, the challenge of standardizing the MS response between multiple LC-MS setups is faced by correcting the intensity of detected signals with an internal and/or external mass calibration by standards [178,191]. The Retention Time Index homogenizes GC methods among multiple machines by injecting a set of specific carbohydrate compounds [192]. However, a methodology for LC is still in development [33].

The insight that machines do not perform similarly given the same parameters has implications for the application of the non-targeted methods in wastewater research. Comparative studies of the influence of the HRMS instrumental setup on the DOM analysis were performed in adjacent scientific disciplines. Non-targeted analyses of natural DOM and effluent of oil sands treatment were compared between FTICR-MS and LC-QTOF-MS or LTQ-Orbitrap respectively [193,194]. The studies could establish a comparable result for various system parameters. However, they also found a low ***m/z* detection bias** in FTICR-MS compared to Orbitrap MS, and the lower resolution of QTOF-MS led to ambiguous results in positive ionization mode.

The results of data-driven HRMS analyses also depend on the software used for data extraction and cleaning up. MS vendors provide **commercial software packages** optimized for the output of their spectrometers together with corresponding help and training, e.g. Waters MassLynx®, Thermo Scientific Compound Discoverer 2® or Agilent MassHunter® [195]. **Data processing software with a non-commercial license** as MZmine [196–198], XCMS, OpenMS [199], and enviMass [200] often allow a larger suite of accepted data input and output formats while having the advantage of being accessible without a commercial license. Despite a large number of options, all software packages require a careful selection of data extraction settings. Additionally, many were conceived as tools for non-targeted analysis in biological analysis and have to be adjusted to the conditions in wastewater samples [201].

## 1.4 Filled gaps

The previous sections outlined the current state of knowledge on the fate of DOM and micro-contaminants in the wastewater treatment. A comprehensive analysis of wastewater DOM as a reservoir for micro-contaminants is still in the first stages of exploration, despite the multitude of developed approaches to analyze select micro-contaminants and sophisticated methods to detect suspected micro-contaminants and TPs. The challenges of a non-targeted data analysis remain the inclusion of noise signals into the datasets and the exclusion of true DOM signals. These arise from the necessarily complex HRMS data acquisition, extraction, and cleansing methodology.

The use of LC-ESI-MS additionally reduces the discovered DOM, since it mostly detects medium-polar compounds [202]. Therefore, it is crucial to maximize the output of analysis by expanding the options of non-targeted data extraction, given current limitations. This thesis concentrates on the **development of a methodology** for a non-targeted analysis that prioritizes the detection of anthropogenic compounds by combining previously researched approaches from metabolomics, petroleomics, and micro-contaminant research.

Previously, the output of non-targeted analysis was shown to vary depending on the instrumentation and conditions during HRMS data acquisition. The reliability and flexibility of analysis, together with its price will drive the applicability of non-targeted analysis

of wastewater, especially in the business sector. It is important to map out whether different HRMS setups perform non-targeted analysis at a sufficiently high level of comparability since the specifications and prices of spectrometers vary largely. For example, FT ICR-MS remains difficult to access due to the price of the instruments, despite its superior mass resolution. At the same time, QTOF and LTQ-Orbitrap instrumental setups gained a lot of trust in environmental analytics due to a good combination of the price, the sensitivity, and the resolving power [203]. The comparison of non-targeted analyses based on these instruments was explored in adjacent fields as petroleomics and metabolomics [181,204,205]. To explore the LC-QTOF and LC-LTQ-Orbitrap influence on wastewater non-targeted analysis, the own non-targeted methodology was applied using both instrumental setups in this thesis. This allowed to additionally test the robustness of the methodology's.

At the same time, the LC-HRMS suspect screening for TPs remains a reliable way to map out transformation paths of chemicals in wastewater [206]. A suspect screening has the ability to zoom in on known micro-contaminants and extract quantitative or semi-quantitative information on their fate in wastewater treatment [125].

A suspect screening for known and unknown TPs was applied in this thesis to the same wastewater samples as in non-targeted approach. The analysis mapped out the transformation of known micro-contaminants in a multistage wastewater treatment. It also allowed comparing the performance of both analyses in regard to the transformation of anthropogenic DOM constituents.

Additionally, this work explored the application of a novel long-term BAC-UF pilot WWTP as a stand-alone tertiary treatment system for the removal of select micro-contaminants and their TPs. The author's hope is that the exploration of BAC in regard to trace contaminant will help the BAC-UF operators to prepare the reactor system for a full-scale challenge of micro-contaminant removal.

## 1.5 Thesis roadmap

The aim of the thesis and the four sub-points are described in Chapter 2.

Chapter 3 describes the materials, the sample preparation, the data acquisition and the cleaning up of data. In particular, Chapter 3.3.1 describes the experimental procedure and testing of the developed non-targeted methodology in a multistage WWT and corresponds to the results in Chapter 4.1. The experimental procedure of spectrometer comparison in respect to non-targeted analysis in Chapter 3.3.2 corresponds to the results in Chapter 4.2. Chapter 3.4 describes the experimental procedure of suspect screenings for micro-pollutants performed on two sets of WWT samples and corresponds to the results in Chapter 4.3.

Chapter 4 “Results and Discussion” presents the results and their discussion of the four sub-topics described in Chapter 2. Chapter 4.1 involves the non-targeted analysis of wastewater, while Chapter 4.2 discusses the robustness of non-targeted analysis by comparing the output of two LC-HRMS setups. Chapter 4.3 describes the results of suspect screenings. It contains in Chapter 4.3.1 a suspect screening for micro-contaminant TPs with wastewater samples used for non-targeted analysis in Chapter 4.1. The Chapter 4.3.2 presents results in which the removal profiles of micro-contaminants and their TPs are explored for a novel BAC-UF pilot plant. The Results and Discussion part is concluded by Chapter 4.4 which offers an insight into the comparison of experimental procedures used in the non-targeted and the suspect screenings and also explores the pre-concentration challenges encountered and tackled throughout the thesis.

Finally, Chapter 5 offers a general summary of the previously presented results and offers an outlook into the future development of the field.

# 2 OBJECTIVES

The aim of this thesis was to develop and evaluate a methodology with a deep insight into the transformation of DOM and micro-contaminants of interest during wastewater treatment. The work was subdivided into four specific objectives according to the current challenges in the field:

Development and validation of a methodology for **non-targeted analysis** of wastewater DOM focusing on the fraction of anthropogenic chemicals.

Evaluation of **MS instrument influence** on the output of non-targeted methodology. Data acquired with two LC-HRMS setups (QTOF and Orbitrap analyzers) were used for the task.

Zooming in on the transformation of micro-contaminants by using an own methodology for a **suspect screening** for pharmaceuticals and their TPs. Drawing parallels to non-targeted analysis.

Elucidation of abatement profiles for pharmaceuticals and their TPs in a **novel BAC-UF pilot WWTP** at multiple measurements and time points as a performance test for the explored suspect screening methodology.

# 3 METHODOLOGY

## 3.1 Reagents and standards

HPLC grade solvents methanol, water, ethyl acetate, and acetonitrile were purchased from Fisher (Germany) and the buffer solution was prepared using HPLC grade 98-100% formic acid (Merck, Germany). Membrane and glass fiber filters were from Merck Millipore Ltd (Germany). Milli-Q water was purified in-house. Ethylenediamine tetra-acetic acid disodium salt solution ($Na_2EDTA$) was from Panreac (Spain). $N_2$ of 99.9990% purity for drying was from Abelló Linde S.A. SPE cartridges used in this work were Oasis hydrophilic-lipophilic balanced (HLB), mixed-mode, strong anion-exchange (MAX), and mixed-mode, strong cation-exchange (MCX) from Waters Corporation (USA).

All pharmaceutical standards were of high purity grade (>90%). Isotopically labeled compounds, used as internal standards acetaminophen-$d_4$, antipyrine-$d_3$, atenolol-$d_7$, azaperone-$d_4$, bezafibrate-$d_4$, carbamazepine-$d_{10}$, cimetidine-$d_3$, citalopram-$d_4$, cprofloxacin-$d_8$, diclofenac-$d_4$, dexamethasone-$d_4$, diazepam-$d_5$, diltiazem-$d_3$, fluoxetine-$d_5$, glyburide-$d_3$, indomethacin-$d_4$, ketoprofen-$d_3$, lincomycin-$d_3$, meloxicam-$d_3$, ofloxacin-$d_3$, ronidazole-$d_3$, sertraline-$d_3$, simvastatin-$d_6$, sulfamethoxazole-$d_4$, sulfapyridine-n-acetyl-$d_4$, trimethoprim-$d_3$, valsartan-$d_8$, venlafaxine-$d_6$, verapamil-$d_6$, warfarin-$d_5$, xylazine-$d_6$, bezafibrate-$d_4$, chloramphenicol-$d_5$, citalopram-$d_4$, glyburide-$d_3$, meloxicam-$d_3$, valsartan-$d_8$, and warfarin-$d_5$ were purchased from Sigma-Aldrich, CDN Isotopes, and Toronto Research Chemicals. The surrogate standards sulfadimethoxine-$d_6$ and sulfadoxine-$d_3$ were purchased from Sigma–Aldrich.

## 3.2 Software

LC-Orbitrap-MS system was controlled via Aria software, version 1.6, under Xcalibur 2.2 (Thermo Fisher Scientific, USA). Custom built MZmine 2 [207] was used for data extraction in which the atomic ratios in formula prediction were set to H/C < 3.2, O/C < 1.2, N/C < 1.3, and S/C < 0.8.

The formula prediction yielded output in form of "Form*X* Iso*Y* Mass*Z*", where *X* is a neutral formula, *Y* is isotopic pattern score between 0 and 1, and *Z* is a neutral monoisotopic mass of the predicted formula in Da. Details for modifying MZmine 2 are described in Verkh et al. [208].

R ≥ 3.2.3 system for statistical computation [209] was used with own package MZmineR ≥ 0.0.2.1 [210]. R packages used in the thesis were ChemmineR ≥2.28.0 and ChemmineOB ≥ 1.14.0 from Open Babel [211–213], tidyverse ≥ 1.1.1 [214], and Rdisop ≥ 1.30.0 [215–218]. Images were created with R package ggplot2 ≥ 2.2.1 [219] and processed in Inkscape 0.92.1 [220]. Molecular structures as SMILES were compiled using MolView [221] and plotted in R with package depict 0.1.1 [222]. Exact Finder 2.5.4 (Thermo Fisher Scientific) was used for compound detection in suspect screening. Mass Frontier™ 7.0.3 (High Chem™) was used for *in silico* prediction of $MS^2$ fragments and their recognition in spectra.

## 3.3 Non-targeted analysis

The non-targeted screening follows a series of steps shown in Figure 3.1.

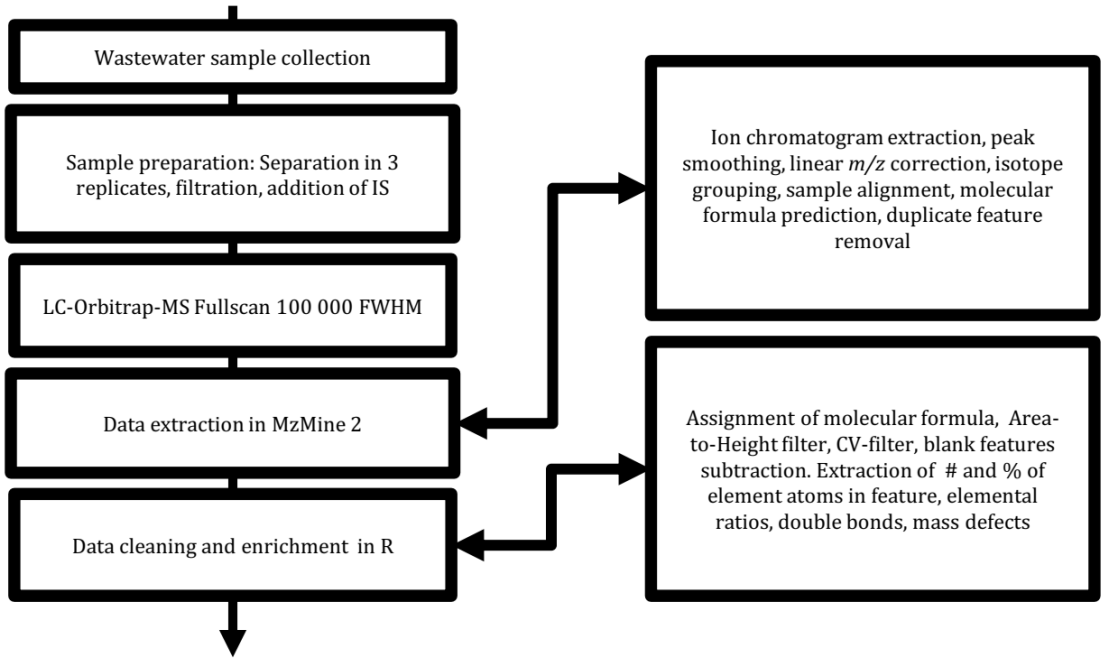



**Figure 3.1 Flow scheme of applied non-targeted analysis.**

### 3.3.1 Methodology development and analysis of a multi-stage WWTP

#### 3.3.1.1 Sampling at Castell d'Aro

The samples were collected from the municipal WWTP at Castell d'Aro, Spain with 175000 people equivalents and a flow of 35000 $m^3 \cdot d^{-1}$. This facility consists of conventional primary and secondary biological treatment, and tertiary treatment with three steps: sand filtration, UV-treatment, and chlorination. 24 h composite secondary influent and effluent samples with a corresponding hydrological retention time correction were taken on 07.06.2017.

Additionally, six grab samples of two different time series of tertiary treatment with 24 h between them were taken on 07.06.2017 and 08.06.2017. A grab sample of secondary effluent was taken which corresponded in time to a time series of tertiary treatment samples. The motivation behind the grab sampling is the short residence time of the tertiary treatment which renders composite sampling difficult.

Wastewater samples were collected in pre-rinsed 1 L PET bottles and stored at 4 °C during transportation before being frozen. The samples were filtered under vacuum using 1.0 µm and 0.45 µm Hydrophilic Polyvinylidene Fluoride Durapore® membrane filters. A mixture of 32 detected internal standards (IS, isotopically labeled pharmaceuticals and antibiotics) was used to evaluate the ion suppression caused by the matrix, and to estimate variations in the instrumental response from injection to injection (Annex Table 7.1). The concentration of IS was 10 ppb with the exception of 5 ppb for Lincomycin. A correction of the matrix induced ion suppression using an intensity normalization of spectra by averaged IS intensity was not attempted. A normalization leads to worse results in replicates compared to the unaltered spectra.

#### 3.3.1.2 Validity of tertiary grab samples

The samples were extracted as described in Section 3.3. Two grab samples of each tertiary treatment step were used to check the consistency of the sampling, before comparing the results of secondary and tertiary treatment. The sample pairs were compared using four variables. Mass, retention time (RT), logarithmic intensity (Int) served as raw LC-MS data outputs and the count of N atoms as a measure of elemental properties.

The sample pairs were considered significantly similar upon exceeding a p-value of 0.05 in hypothesis testing. The grab samples of the sand filtration ($p_{mass} = 0.71$, $p_{RT} = 0.54$, $p_{Int} = 0.30$, $p_N = 0.75$) and the chlorination ($p_{mass} = 0.87$, $p_{RT} = 0.62$, $p_{Int} = 0.95$, $p_N = 0.54$) were significantly similar. The UV-treatment samples had an intensity difference of 1.2 times, but otherwise were significantly similar ($p_{mass} = 0.86$, $p_{RT} = 0.71$, $p_{Int} = 0.03$, $p_N = 0.44$). Grab samples were largely stable over the 24 h period due to the hypothesis testing.

Composite secondary effluent and grab secondary effluent differed in mass and retention time distribution ($p_{mass} = 0.03$, $p_{RT} < 0.001$, $p_{Int} = 0.06$, $p_N = 0.71$). The difference, given the same data extraction, shows that care has to be exercised when comparing molecular features between composite samples and grab samples in the study.

Differences in the treatment series of grab secondary effluent and tertiary treatment corresponding to the same time series were explored using one-way ANOVA tests. The stages did not differ significantly in their properties by exceeding the p-value of 0.05, except considering the retention time. ($anova_{mass}$: $F(3, 4763) = 1.09$, $p = 0.35$; $anova_{RT}$:

$F(3, 4763) = 2.83$, $p = 0.04$; $anova_{Int}$: $F(3, 4763) = 0.80$, $p = 0.49$; $anova_{N}$: $F(3, 2099) = 0.31$, $p = 0.82$; homogeneity of variance was confirmed.)

This shows that the detected complement of features was largely stable during the tertiary treatment. DOM and small molecules have been shown to react in advanced treatment technologies [3,223]. However, the tertiary treatments of sand filtration, UV, and chlorination are designed to remove the suspended particles of biomass from the secondary treatment and to remove biological constituents and can underperform considering small molecules [88,143].

#### 3.3.1.3 LC-HRMS analysis

LC-HRMS analysis was performed on an LTQ-Orbitrap Velos™ coupled with the Aria TLX-1 HPLC system (Thermo Fisher Scientific, USA). This system comprised a PAL autosampler and two mixing quaternary pumps (eluting pump and loading pump). The chromatographic separation was achieved on Acquity UPLC® BEH $C_{18}$ (2.1 mm × 50 mm, 1.7 µm particle size, Waters UK) chromatographic column in both the positive ionization (PI) and the negative ionization (NI) modes. 20 µL of the filtered sample was injected into the LC-MS.

A solvent gradient with 0.1% formic acid in acetonitrile and an aqueous solution of formic acid 0.1% was used in both PI and NI modes. The chromatographic resolution was performed in the following way: solvent (A) solution of formic acid 0.1% in acetonitrile and solvent (B) aqueous solution of formic acid 0.1% at a flow rate of 0.4 $mL \cdot min^{-1}$ at 25 °C. Initial conditions 5% A; 0.02 – 10.02 min, 5 – 95% A; 10.02 – 15.02 min, 95% A; 15.02 – 15.10 min, 95 – 5% A; 15.10 – 18.00 min, 5% A.

The hybrid linear ion trap-Fourier Transform Mass Spectrometry analyzer (LTQ-Orbitrap Velos™, Thermo Fisher Scientific) was equipped with a diverter valve and an ESI source. Mass calibration and mass accuracy checks were performed with LTQ ESI Positive Ion Calibration Solution (Thermo Fisher Scientific) and mass accuracy was always within the error of ±5 ppm. The ionization voltage was set at 3.0 kV with the sheath gas flow at 45, auxiliary gas flow at 20, S-Lens RF level at 60% and the capillary temperature and the source heater temperature at 350 °C.

The full scan spectra were acquired within a range of 100 to 1000 *m/z* at a set resolving power of 100,000 full width at half maximum (FWHM).

#### 3.3.1.4 Extraction and cleaning up of non-targeted data

##### *3.3.1.4.1 MZmine 2.26 parameters in PI mode*

Mass Detection: "Exact mass" detector with an intensity noise level of $1 \cdot 10^3$ a.u.

FTMS Shoulder peaks filter: Lorenzian extended peak model function with a mass resolution of $1 \cdot 10^5$ FWHM.

Chromatogram builder: A minimum time span of 0.03 min (2 sec); a minimum signal height of $1 \cdot 10^3$ a.u.; a mass tolerance of extraction 5 ppm.

Chromatogram smoothing: Filter width of 7 points.

Chromatogram deconvolution: Noise amplitude algorithm was applied with a signal duration of 0.05 – 8.00 min, a minimal signal height of $3 \cdot 10^3$ a.u. and an amplitude of noise $1 \cdot 10^3$ a.u.

Custom database search for IS: mass tolerance < 5 ppm, retention time tolerance < 0.3 min. A visual check of the IS signals confirmed their validity.

A personal script in R corrected the *m/z* of signals in exported XML peak lists using linear models based on IS. The linear models were calculated for two *m/z* ranges over and under 400 Da. The script removed IS outliers > 4 * interquartile range and omitted the correction where less than 3 IS signals defined the model.

Isotopic peaks grouper: 60 ppm mass tolerance, 0.03 min of time tolerance and a maximum charge of 2, representative peak: lowest *m/z* and monotonic shape.

Filters: peaks in an isotope pattern ≥ 2, peak per row ≥ 1, points per peak 7-500.

Adduct search: retention time tolerance < 0.03 min, adducts in PI LC-ESI-MS: $[M+NH_4]^+$, $[M+Na]^+$, $[M+2Na]^{2+}$, $[M+K]^+$, $[M+2K]^{2+}$, $[M+CH_3OH]^+$ , mass tolerance 5 ppm and maximum adduct signal height of 50%. Subsequently these adducts were removed from the feature list.

RANSAC aligner: Mass tolerance 5 ppm, retention time tolerance 2.00 min, retention time tolerance after correction 1.00 min. Number of iterations was set by the software, 60% of points in the model were considered for validation, fit threshold of 1 min. A linear model was not assumed. Same charge was required.

*m/z* and RT gap filler: mass tolerance of 10 ppm.

List of aligned features was separated by charge to contain either charge 1 or 2.
Duplicate peak filter: 3 ppm mass tolerance, 0.03 min retention time tolerance.
Formula prediction: The ionization type was $[M+H]^+$ and the mass tolerance of prediction 5 ppm. Atomic ranges used for prediction: $C_{1-80}$, $H_{1-100}$, $O_{0-20}$, $N_{0-15}$, $S_{0-4}$, $Cl_{0-4}$, $Br_{0-4}$. Heuristic rules and RDBE restrictions of 0–40 RDBE were applied. A custom-built MZmine version allowed the atomic ratios H/C < 3.2, O/C < 1.2, N/C < 1.3, S/C < 0.8. Isotopic pattern comparison enhanced the prediction with a mass tolerance of isotopes set to 5 ppm, a minimum absolute intensity of $6 \cdot 10^2$ a.u. and a minimum match score of 60%. Depending on the data, the charge varied between 1 and 2. Features with charge 2 and *m/z* > 500 Da did not receive an elemental formula. The custom-built formula prediction yielded output in form of "Form*X* Iso*Y* Mass*Z*", where *X* is neutral formula, *Y* is isotopic pattern score between 0 and 1, and *Z* is neutral monoisotopic mass of the predicted formula in Da. The best formula candidate was picked in R using the output according to the algorithm described below.
Export of csv with *m/z*, retention time, and all IDs of features. Heights, areas, and charges of features were exported.

#### *3.3.1.4.2 MZmine 2.26 parameters in NI mode*

Mass Detection: "Exact mass" detector with an intensity noise level of $2.5 \cdot 10^3$ a.u.
FTMS Shoulder peaks filter: Lorenzian extended peak model function with a mass resolution of $1 \cdot 10^5$ FWHM.
Chromatogram builder: A minimum time span of 0.03 min (2 sec); a minimum signal height of $2.5 \cdot 10^3$ a.u.; a mass tolerance of extraction 5 ppm.
Chromatogram smoothing: Filter width of 7 points.
Chromatogram deconvolution: Noise amplitude algorithm was applied with a signal duration of 0.05 – 8.00 min, a minimal signal height of $7.5 \cdot 10^3$ a.u. and an amplitude of noise $2.5 \cdot 10^3$ a.u.
Custom database search for IS: mass tolerance < 5 ppm, retention time tolerance < 0.3 min. A visual check of the IS signals confirmed their validity.

A personal script in R corrected the *m/z* of signals in exported XML peak lists using linear models based on IS. The linear models were calculated for two *m/z* ranges > 400 Da and < 400 Da. The script removed IS outliers > 4 * interquartile range and omitted the correction where less than 3 IS signals defined the model.
Isotopic peaks grouper: 60 ppm mass tolerance, 0.03 min of time tolerance and a maximum charge of 2, representative peak: lowest *m/z* and monotonic shape.
Filters: peaks in an isotope pattern ≥ 2, peak per row ≥ 1, points per peak 7-500.
Adduct search: retention time tolerance 0.03 min, adducts in NI LC-ESI-MS: $[M-H_2O-H]^-$, $[M-2(H_2O)-H]^-$, $[M-2H+K]^-$, $[M-2H+Na]^-$, $[M+Cl]^-$, $[M-H+HAc]^-$, $[M+Br]^-$, $[M-H+FA]^-$, mass tolerance 5 ppm and maximum adduct signal height of 50%. Subsequently, the adducts were removed from the feature list.
RANSAC aligner: Mass tolerance 5 ppm, retention time tolerance 2.00 min, retention time tolerance after correction 1.00 min. Number of iterations was set by the software, 80% of points in the model were considered for validation, fit threshold of 1 min. A linear model was not assumed. Same charge was required.
*m/z* and RT gap filler: mass tolerance of 10 ppm.
List of aligned features was separated by charge to contain either charge 1 or 2.
Duplicate peak filter: 3 ppm mass tolerance, 0.03 min retention time tolerance.
Formula prediction: The ionization type was $[M+H]^-$ and the mass tolerance of prediction 5 ppm. Atomic ranges used for prediction: $C_{1-80}$, $H_{1-100}$, $O_{0-20}$, $N_{0-15}$, $S_{0-4}$, $Cl_{0-4}$, $Br_{0-4}$. Heuristic rules and RDBE restrictions of 0–40 RDBE were applied. A custom-built MZmine version allowed the atomic ratios H/C < 3.2, O/C < 1.2, N/C < 1.3, S/C < 0.8. Isotopic pattern comparison enhanced the prediction with a mass tolerance of isotopes set to 5 ppm, a minimum absolute intensity of $1 \cdot 10^3$ a.u. and a minimum match score of 60%. Depending on the data, the charge varied between 1 and 2. Features with charge 2 and *m/z* > 500 Da did not receive an elemental formula. The custom-built formula prediction yielded output in form of "Form*X* Iso*Y* Mass*Z*", where *X* is neutral formula, *Y* is isotopic pattern score between 0 and 1, and *Z* is neutral monoisotopic mass of the predicted formula in Da. The best formula candidate was picked in R using the output according to the algorithm described in section 3.3.1.4.3.

Export of csv with *m/z*, retention time, and all IDs of features. Heights, areas, and charges of features were exported. Picking of the best candidate for the molecular formula of the molecular feature, cleaning up of data and calculation of elemental properties were performed in R.

#### *3.3.1.4.3 Data cleaning up in R*

IS were removed from datasets.

Picking the best candidate for the molecular formula: The candidate with the highest isotopic pattern score was selected unless the candidate with the next highest isotopic score had a deviation in isotopic score < 10% and absolute ppm error < 1 ppm compared to the first candidate. In that case, the algorithm picked the second candidate formula.

Avoiding "square" peak extraction artifacts in MZmine: Features with an area-to-height ratio > 30 were removed. This value was chosen for this dataset through observation.

Filtering by CV: Features with a CV > 30% were removed.

Baseline correction: Features with a sample-to-blank area ratio < 3 were removed. Features with a sample-to-blank area ratio > 3 were corrected by the area of the blank samples if the CV of the blank samples was < 30%.

Molecular features in datasets of PI and NI modes were combined and the neutral mass was calculated for features from *m/z*, considering the charge of features, the ionization mode, and loss/gain of hydrogen.

Duplicate removal: PI and NI features were united and the duplicates removed for the mass deviation < 5 ppm and retention time deviation < 0.5 min under the assumption that the retention time is similar in both PI and NI modes, because both were recorded under the same chromatographic method.

The elemental properties as the number of atoms in a molecule, atomic ratios of type X/C (where X is H, N, O, S), double bond equivalents (DBE), and DBE-O were calculated using the neutral molecular formula of a molecular feature, if a molecular formula was assigned.

DBE was calculated using the equation $DBE = C + 1 + (N - X - H)/2$ where X is the sum of halogen atoms in a molecule and DBE-O was calculated by subtracting the number of O atoms in a molecule from DBE.

#### 3.3.1.5 Non-targeted data extraction evaluation

The mass tolerance for the processing of LC-MS spectra was set to 5 ppm in MZmine. *m/z* mass error < 5 ppm for data extraction was experimentally confirmed with IS (Appendix 1: Table 7.1). The mass error is < 2 ppm for DOM proxies in these data. But, the precision falls increasingly with the loss of intensity, so the established threshold of 5 ppm is more realistic considering the literature suggestions about the negative impact of too stringent tolerance margins [187,188,224].

Signals with signal-to-baseline ratio < 3 were ignored in MZmine to exclude instrumental noise. Orbitrap–MS is a Fourier transform based technique and signals show satellites that arise during the conversion of current to spectral data. These satellite signals were ignored under the assumption of $1 \cdot 10^5$ mass resolution in the measurement to ensure exclusion of noise and faster computation. Margins of retention time < 0.5 min and *m/z* < 5 ppm were used to exclude duplicate molecular features detected in both modes, thanks to the same chromatographic method in PI and NI modes.

The duration boundaries of the extracted ion chromatograms were 0.05 – 8.00 min. The peak duration was focused on extracting all signals in the spectra while controlling for the column bleed in the applied chromatographic method. The retention time in the MS signal extraction helped to distinguish molecular features with a similar *m/z*. This is shown on the example of compounds from secondary influent, which without retention time would require a mass resolution > $5 \cdot 10^{-5}$ Da to be distinguished (Figure 3.2).

A baseline correction can improve the extraction of ion chromatograms with elevated baselines. A mathematical baseline correction was not applied, to avoid an uncontrolled loss of information. The correction was applied only to relevant chromatographic regions by subtracting signal intensity of solvent blank from the matching signal in a sample.

Only monoisotopic signals with at least one consecutive isotopologue were analyzed. Grouping of isotopes into a molecular feature represented by one monoisotopic mass reduces the data stream and adds isotopic information. Ungrouped signals were removed to makes sure that only molecular features with an isotopic pattern are used for the formula prediction.

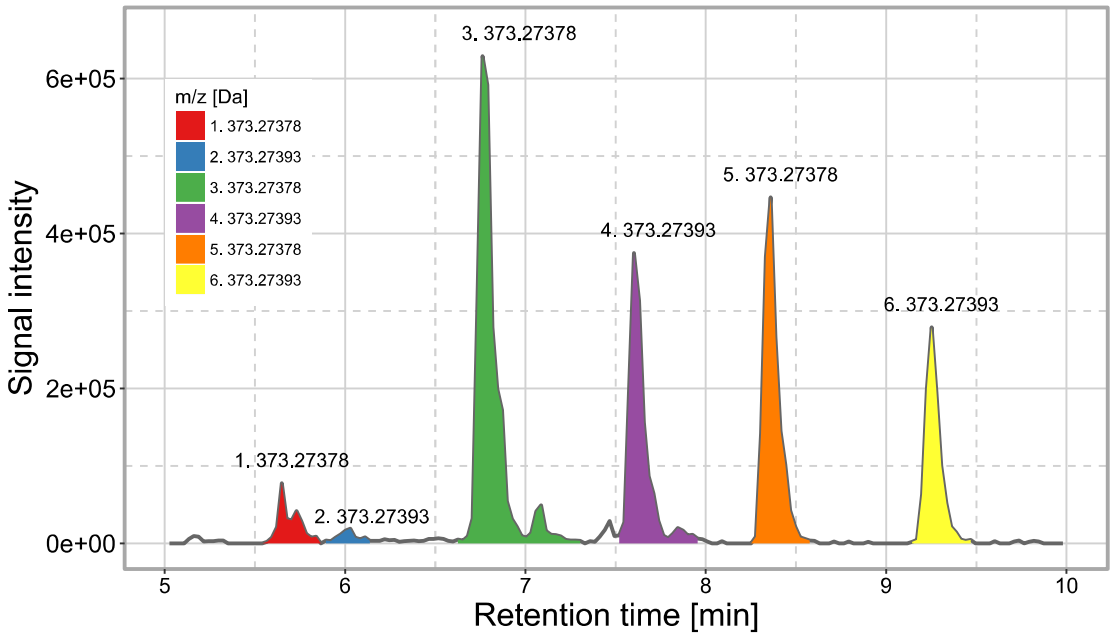



**Figure 3.2 Extracted ion chromatograms of signals with *m/z* 373.2738 ± 0.0001 Da in secondary influent as detected in MZmine and distinguished by retention time.**

Only protonated or deprotonated ions with a charge ±1 or ±2 were analyzed. ESI is a soft ionization method and wastewater DOM largely consists of small molecules, which cannot accumulate many charges. Molecular features that corresponded to their adducts $[M+NH_4]^+$, $[M+Na]^+$, $[M+2Na]^{2+}$, $[M+K]^+$, $[M+2K]^{2+}$, $[M+CH_3OH]^+$ in PI mode and $[M-H_2O-H]^-$, $[M-2(H_2O)-H]^-$, $[M-2H+K]^-$, $[M-2H+Na]^-$, $[M+Cl]^-$, $[M-H+HAc]^-$, $[M+Br]^-$, $[M-H+FA]^-$ in NI mode were removed. Potential LC-MS adduct targets were derived from Keller et al. [225]. The intensity of these adduct features was not added to the intensity of $[M-H]^-$/$[M+H]^+$molecular features because this does not offer a significant improvement of the data [24]. The IS showed prevalent ionization $[M+H]^+$ in PI and $[M-H]^-$ in NI modes.

Matrix induced ion suppression becomes more problematic in complex LC-MS samples, which leads to non-linear intensity deviations among samples. The suppression was calculated by comparing intensities of IS in samples to the respective intensity in the solvent blank (Figure 3.3, Figure 3.4 and Figure 3.5).

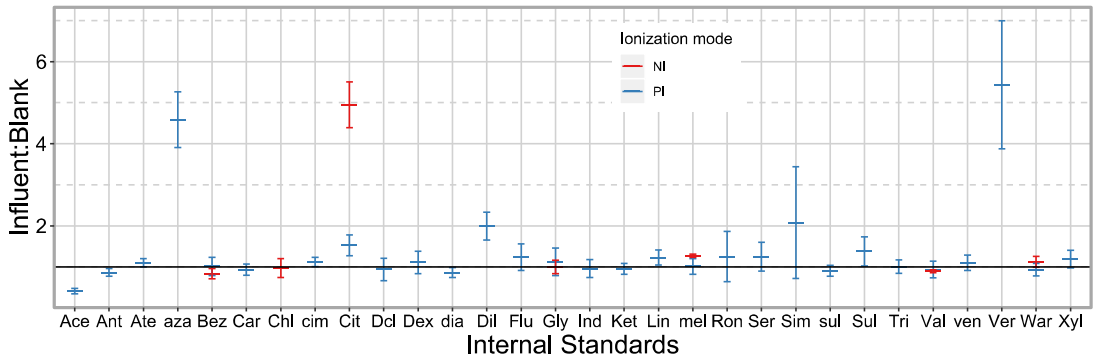



**Figure 3.3 Ratio secondary influent-to-blank for area intensities of IS with a corresponding error of standard deviation assuming no covariance. For clarity, Ciprofloxacin-$d_8$ (44.8 ± 9.4) and Ofloxacin-$d_3$ (24.7 ± 4.6) were omitted.**

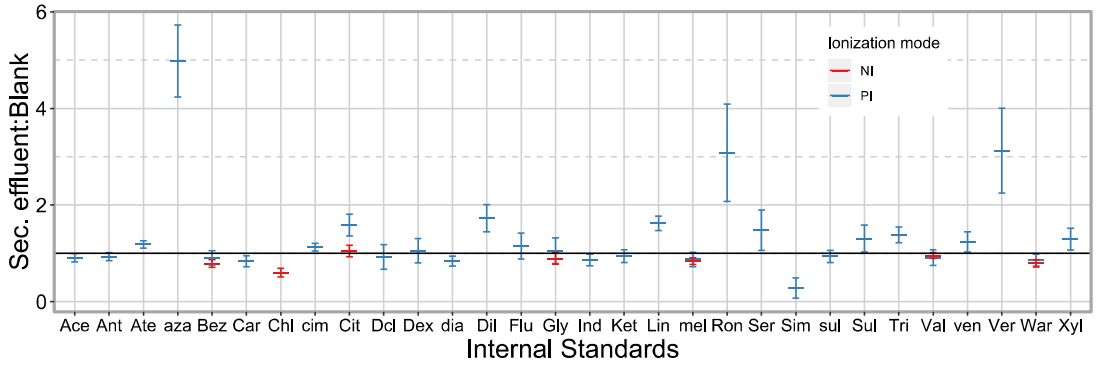



**Figure 3.4 Ratio secondary effluent-to-blank for area intensities of IS with a corresponding error of standard deviation assuming no covariance. For clarity, Ciprofloxacin-$d_8$ (56. 5 ± 11.0) and Ofloxacin-$d_3$ (29.4 ± 5.0) were omitted.**

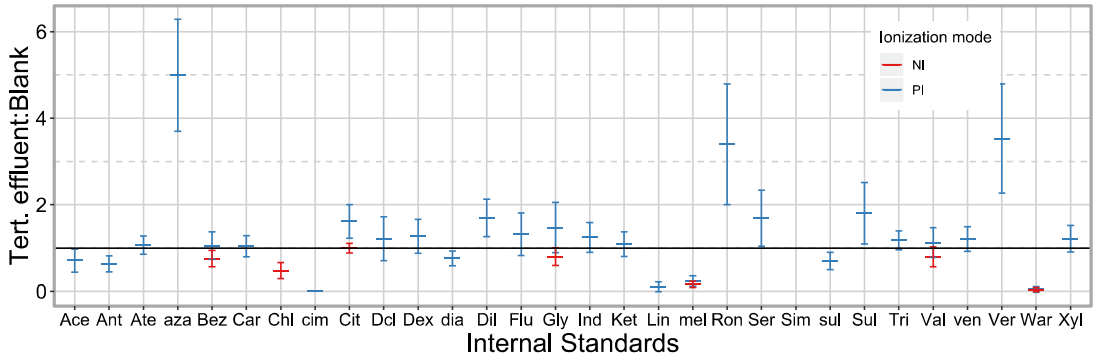



**Figure 3.5 Ratio tertiary effluent-to-blank for area intensities of IS with a corresponding error of standard deviation assuming no covariance. For clarity, Ciprofloxacin-$d_8$ (54.1 ± 12.0) and Ofloxacin-$d_3$ (25.3 ± 5.3) were omitted.**

The influence of the matrix effect can be reduced by standardizing signal intensity [226]. A correction of the matrix effect using an intensity normalization of spectra was not attempted. A normalization leads to worse results in replicates compared to the unaltered spectra. The automatic standardization requires a set of IS throughout the entire spectrum. The amount of IS in the study did not cover the range and not enough signals had the required high intensity for normalization. CV of features among the replicates of a sample has to decrease after a successful standardization while here the CV increased.

Molecular formulae were predicted for molecular features and, where discovered, assigned. The formulae were predicted using the atomic ranges: $C_{1-80}$, $H_{1-100}$, $O_{0-20}$, $N_{0-15}$, $S_{0-4}$, $Cl_{0-4}$, $Br_{0-4}$, isotopic pattern score > 60%, and mass error < 5 ppm.

The prediction was only allowed for features with neutral mass < 1000 Da to prevent a high rate of false assignments for heavy molecules. The atomic ratios of generated formulae were H/C < 3.2, O/C < 1.2, N/C <1.3, and S/C < 0.8 which corresponds to 99.7% of registered small molecules [227]. The presented workflow prioritized the isotopic pattern score to select the best candidate for the molecular formula. In R a molecular formula candidate was selected with the smallest absolute mass deviation from 10% of highest isotopic pattern scores to represent the elemental composition of a molecular feature.

The ranges for the DOM elements CHONPS, which are common for organic matter, and halogens FClBrI that can occur in synthetic organics, were derived from a database of 14631 micro-contaminants [228]. These helped to establish suitable parameters for prediction of molecular formula of DOM and in particular for organic micro-contaminants of anthropogenic origin (Figure 3.6).

Common elements C, H, O, and N have narrow distributions and were included in the prediction ranges at values approaching their maxima. S, P, Cl, Br, F and I have a median close to zero and many outliers. S, Br, and Cl are important in the wastewater treatment and they were included in the formula prediction at sensibly low values.

A balanced atomic range reduced appearance of wrong formulae, but contained enough atoms to predict most formulae, when F and P were excluded. The lack of characteristic isotopic information of F and P in MS are responsible for the challenging prediction. This leads to a false assignment of F and P to substances that contain none.

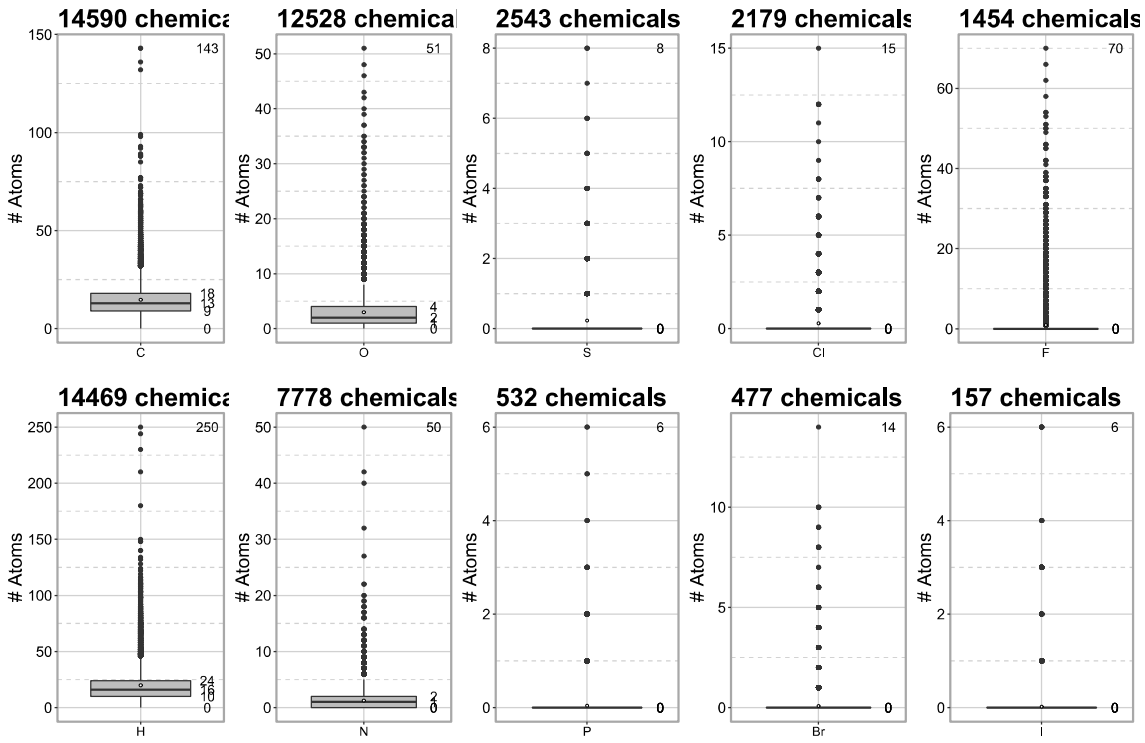



**Figure 3.6 Modified boxplot distributions of the most common small molecule atoms extracted from a database of 14631 contaminants [228].**

79 pharmaceuticals, which covered major composition-relevant atomic properties of micro-contaminants, were measured in PI and NI modes using the same spectrometric conditions as the samples in this study to check the applicability of the formula prediction (Appendix 1: Table 7.2). The pharmaceuticals' samples went through the same signal detection, *m/z* correction, isotope grouping, and formula prediction.

A subsequent targeted screening for $[M+H]^+$ in PI and $[M-H]^-$ in NI modes with *m/z* error < 5 ppm was used to identify the molecular features of the pharmaceuticals within the non-targeted data. 69 out of 79 (87%) molecular formulae were predicted correctly (Appendix1: Figure 7.1). The observation shows that a correct prediction is challenging for molecules with high masses and a low signal intensity. Pharmaceuticals that contained F could not receive a correct formula under the applied settings.

The sensibility of the predicted molecular formulae in the wastewater data was checked with a linear model of predicted percentage of hydrogen (%H) vs. RMD derived from raw data, for 1000 > RMD > -100 (Figure 3.7).

"Outliers" with RMD ≤ -100 shown in orange and violet have median$_{H/C}$ = 2.1, median$_{Mass}$ = 849.6 Da. The remaining molecular features corresponded well to the fit ($R^2$ > 90%) and the absence of random outliers implied a good prediction of the hydrogen content for molecular features.

Most of the features in the "outlier" group with RMD ≤ -100 were saturated, large compounds (median$_{H/C}$ = 2.1, median$_{Mass}$ = 849.6 Da). The deviation of heavy unsaturated molecular features from the linearity stems from a mathematical problem of using the integer or rounded nominal mass to calculate the RMD. The same set of outliers fits the linearity well depending on the calculation used.

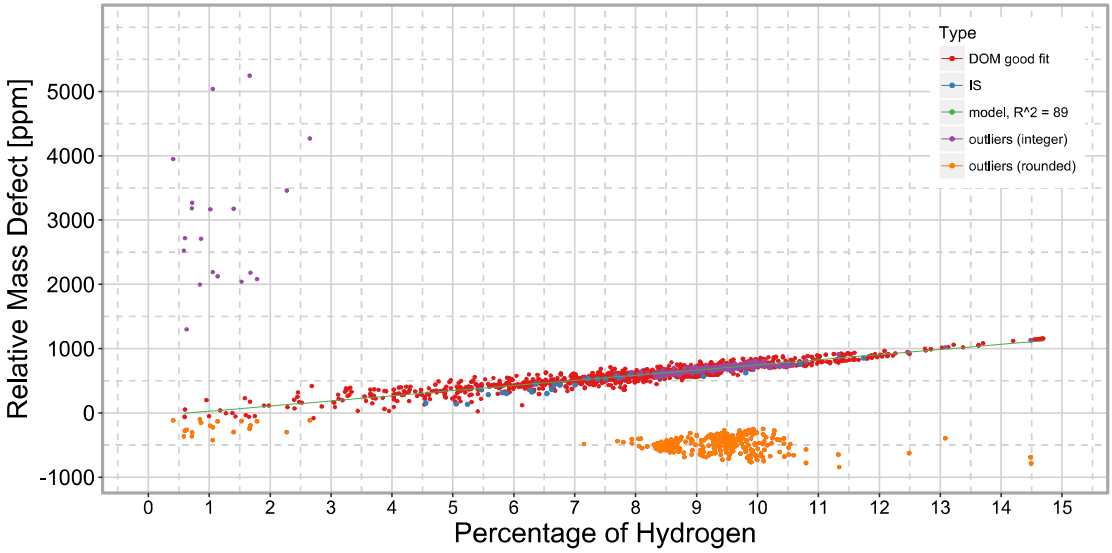



**Figure 3.7 Percentage of hydrogen in feature' molecular formula plotted against feature's RMD for influent molecular features. The same 317 outliers are shown in violet where integer nominal mass was used to calculate RMD and in orange where rounded nominal mass was used. Features shown in red were used to build the linear model. IS are marked in blue.**

A gap-filling algorithm within MZmine looked for signals in the feature list of combined samples which were potentially missed during peak extraction. The algorithm reconstructed omitted chromatographic peaks for a molecular feature in "empty" samples by scanning the *m/z* and retention time region of LC-MS spectra corresponding to the detected peaks in other samples. The methodology later avoids noise recognized as peaks

to enter the final list of molecular features by removing samples with CV > 30 % intensity deviation.

Some features showed duplicates appearing during peak recognition and alignment of samples. To avoid the removal of false duplicates from the dataset the *m/z* and retention time tolerances were set smaller than in previous MZmine modules to 3 ppm and 0.03 min respectively.

The baseline was corrected by subtracting the intensity of a signal in the solvent blank from the matching signal in the sample, but only for signals with an intensity $Ratio_{Sample:Blank} > 3$, while signals with a lower ratio were ignored. This approach accounts for intensity deviations caused by the matrix suppression [158].

The coefficient of variation (CV) of feature intensity among triplicates of a sample filtered out random noise with a conservative value of 30%. The CV of IS in the blank, influent and effluent of the secondary treatment were mostly below the 30% threshold independent of the sample matrix (Figure 3.8). The 2 outliers out of 32 IS were ronidazole-$d_3$ and simvastatin-$d_6$. The presented results are for the combined data of PI and NI modes unless stated otherwise.

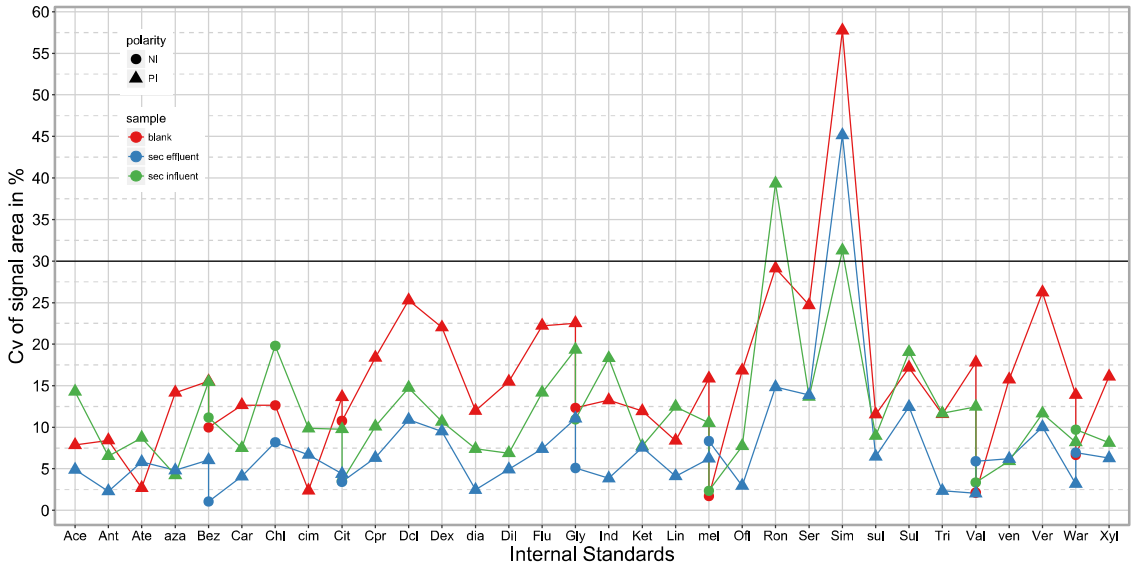



**Figure 3.8 Coefficient of variation (CV) of IS signal intensity in PI samples of secondary treatment. The black line marks the CV-cutoff of 30% applied in the study.**

### 3.3.1.6 Analytical tools in non-targeted analysis

The estimation of the statistically significant differences between the properties of molecular features of any two samples was performed under the assumption of independent data unless stated otherwise.

The statistical significance was tested with a two-sided Welch two-sample t-test for variables with a distribution resembling a normal one or skewed distributions with sample sizes > 50 for a confidence interval of 95% and α of 0.05. Skewed distributions were log-transformed where necessary using $\log_2$ to obtain a normal distribution. Differences in location shifts of non-normally distributed variables were tested with a Wilcoxon rank-sum test with continuity correction.

Herein presented KMD series recognition was programmed in R and tested with the data of surfactants (-$CH_2$- and -$C_2H_4O$- moieties) and halogenated substances (-$CF_2$- and -*H/+Cl* moieties). Both positive and negative cases determined the precision of the own KMD homolog detection tool. 400 $[M+H]^+$-ions in 46 series of surfactants and polyfluorinated substances from the NORMAN database [229,230] and a polychlorobiphenyl (PCB) substitution series [231] tested the positive case using a KMD error of 5 ppm, retention time filter of 2 min, and a rounding of the KMD to 3 digits.

26 fake homologs arranged in 6 series and 100 random signals in the *m/z* range of (100; 1001) Da tested the negative case. Fake homologs were artificial signals with the same KMD as in real homolog series. Yet the mass difference between the homologs of the true series and the fake series did not correspond to n times the mass of a KMD moiety for $n \geq 1$.

The test assigned all homologs in all homologous series for a KMD precision of ±0.0005 Da and at least three homologs in a series (Figure 3.9). Fake homologs built separate series and did not infiltrate the true homologous series. The random noise did not infiltrate the true series or built random series. The recognition worked not only for addition/elimination structural moieties, as -$CH_2$-, but also for substitution reactions.

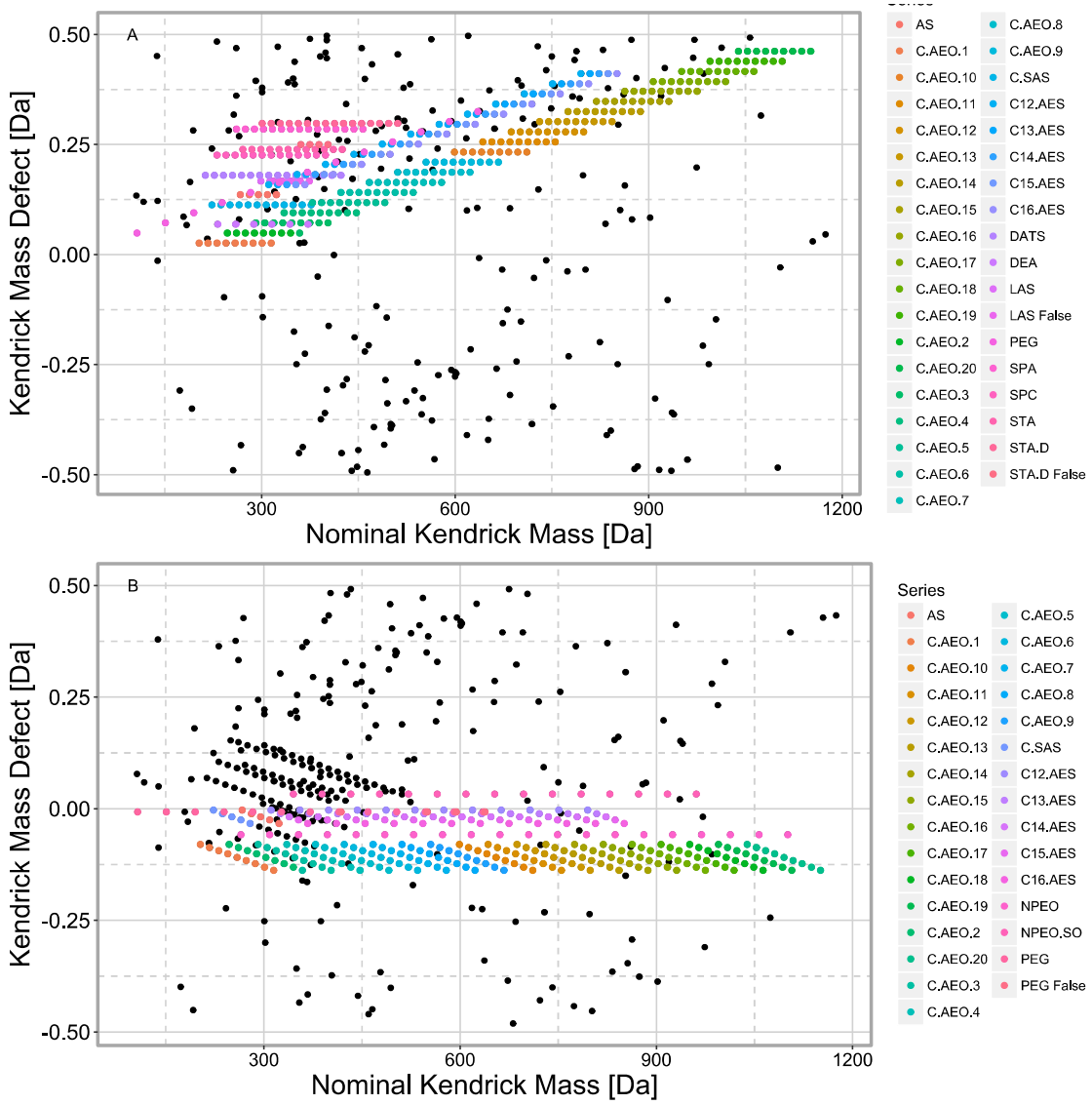



**Figure 3.9 Evaluation of KMD series recognition with real surfactant homologs [229], fake artificial homologs and random noise for structural units (A) -$CH_2$- and (B) -$C_2H_4O$-**

PCB substances showing a substitution of H for Cl were identified (Figure 3.10). Very strong defects that appear in molecules with heavy atoms built series correctly. Yet, they might have a deviating absolute KMD value from the theorized KMD as was shown here with the fake PFAA homologs deviating strongly from the real PFAA homologs (Figure 3.11).

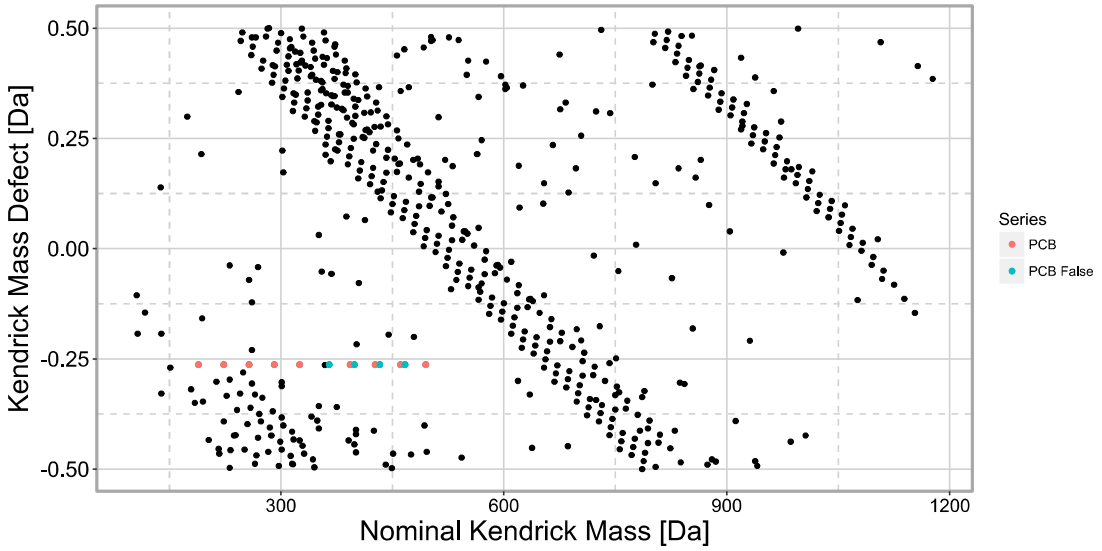



**Figure 3.10 Evaluation of KMD series recognition with real polychlorinated and polyfluorinated substances [231], fake artificial homologs and random noise for substitution structural unit *-H/+Cl*.**

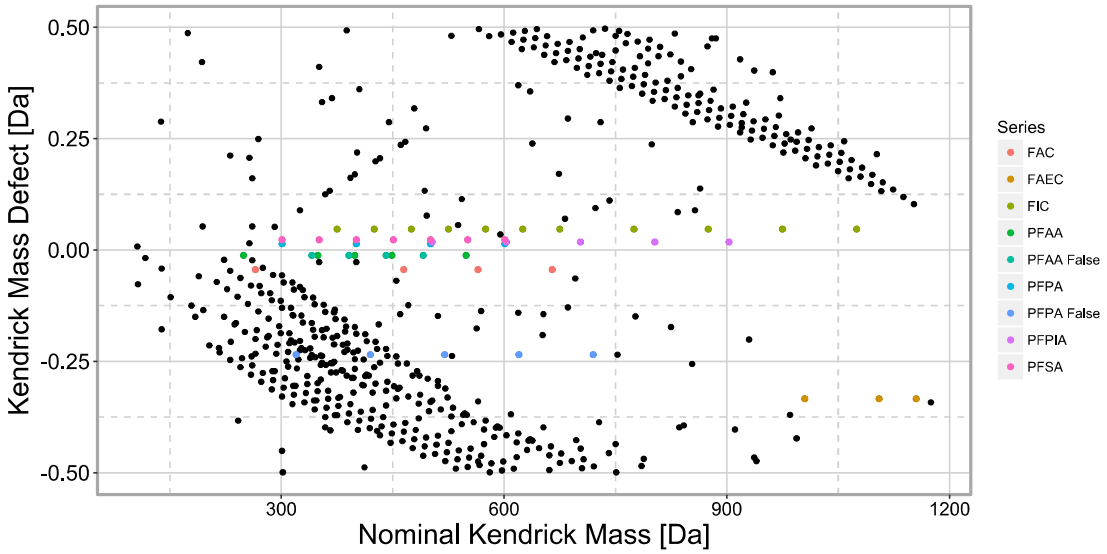



**Figure 3.11 Evaluation of KMD series recognition with real polychlorinated and polyfluorinated substances [230], fake artificial homologs and random noise for the structural unit *-$CF_2$-*.**

In presented wastewater data, *-$CH_2$-* and *-$C_2H_4O$-* series were constricted to at least seven homologs per series in the influent and at least three homologs per series in the effluent of the secondary treatment.

Additionally, the effluent only contained series with KMD detected in the influent, to prioritize the comparison of the KMD homologs among the samples.

An own script in R helped to extract gains or losses of simple structural addition/elimination and de/conjugation moieties between parent compounds and bio-TPs. The moieties were estimated using metabolic logic [179,232]. The algorithm looked for the mass difference corresponding to the exact monoisotopic mass of the structural moiety between the features in the influent sample and the subset of features appearing in secondary and tertiary effluents. The search was constrained to a chromatographic window of 2 min between the parent and TP, and a mass error of 0.005 Da.

## 3.3.2 Comparison of non-targeted analysis on LC-MS setups

### 3.3.2.1 Sampling and pre-treatment of DOM for the setup comparison

Influent, primary treatment effluent and secondary biological treatment effluent were taken from a municipal WWTP in Celrà, Spain on 5$^{th}$ of October 2016. After sampling, wastewater was immediately filtered through 0.7 μm glass fiber filters and stored in acid-washed polyethylene bottles at 4 °C in the dark.

Subsequently, the samples were pre-concentrated following a modified procedure from Dittmar et al. [118]. SPE was carried out for wastewater samples in triplicate with cartridges connected manually in series containing 200 mg Oasis HLB, 200 mg MAX, and 200 mg MCX sorbents. Milli-Q water was extracted to prepare a blank sample.

The cartridges were conditioned with 5 mL MeOH and 5 mL water. Then 100 mL sample was loaded onto the cartridges, washed with 5 mL water and dried under nitrogen stream. The samples were eluted with 6 mL ethyl acetate: MeOH (1:1) with ammonia and 3 mL ethyl acetate:MeOH (1:1) with formic acid. The fractions were combined, pH was set to 7, and the solvent was evaporated under a nitrogen stream. The residue was dissolved in 1 mL water:MeOH (9:1) and IS added with 100 ppb final concentration.

#### 3.3.2.2 Measurement with both LC-MS setups

The chromatographic separation was the same for both LC-MS setups, including the same set of samples, column, the same solvent quality and proportions, and the same elution profile, unless deliberately stated otherwise in the description below.

The chromatographic separation was achieved on Acquity UPLC® HSS $T_3$ (2.1 mm × 50 mm, 1.8 µm particle size, Waters UK) in PI mode. The chromatographic resolution with LC-Orbitrap-MS was derived from Gros et al. [130] to preferentially separate anthropogenic compounds, and performed in the following way in PI mode: solvent (A) methanol and solvent (B) 10 mM aqueous solution of formic acid/ammonium formate (pH = 3.2) at a flow rate of 0.2 mL·min$^{-1}$. The NI mode spectra were measured, however not used in further evaluation to concentrate on the rather anthropogenic DOM fraction.

The gradient elution was performed in the following way: initial conditions 5 % A; 0.0 – 10.0 min, 5 – 95 % A; 10.0 –10.2 min, 100 % A; 10.2 – 13.0 min, 100 % A; 13.0 – 13.2 min, 5 % A; 13.2 – 15.4 min, 5 % A (equilibration of the column). HPLC conditions in NI mode: solvent (A) acetonitrile and solvent (B) 5 mM aqueous solution of ammonium acetate/ ammonia (pH = 8.0) at a flow rate of 0.2 mL min$^{-1}$. A gradient elution was performed in the following way: initial conditions 1 % A; 0.0 – 6.7 min, 60 % A; 6.8 – 10.4 min, 100 % A; 10.4 – 13.7 min, 100 % A; 13.7 – 15.0 min, 1 % A; 15.0 – 17.0 min, 1 % A (equilibration of the column). In LC-QTOF-MS 1 min was added at the beginning of the chromatographic run to account for the 1 min long calibration prior to each acquisition.

Settings of LTQ-Orbitrap Velos™ HRMS are described in Section 3.3.1.3.

Given the same samples, chromatographic columns, and method as in the measurement with LC- LTQ-Orbitrap Velos™ HRMS described in this section, QTOF-UHPLC mass spectra were recorded using a MaXis HD quadrupole electrospray time-of-flight (ESI-QTOF) mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) coupled to an Ultimate 3000 UHPLC (Thermo Fisher Scientific, California, USA) with an injection volume of 20 µL. Bruker MaXis QTOF spectrometer was calibrated using a range of sodium formate clusters introduced by 10 µL loop-injection prior to the chromatographic run. The mass calibrant solution consisted of 3 parts 1 M NaOH to 97 parts of 50:50

water:isopropanol with 0.2% formic acid. The observed mass and isotopic pattern matched the corresponding theoretical values. Calibration was performed using Data Analysis 4.3 (Bruker Daltonik GmbH, Bremen, Germany). The Bruker Maxis QTOF did not have a specific desired resolution option in the setup but was operated in high resolution mode. The Bruker Maxis QTOF scan range was set to 100-1000 *m/z*, and ramp range to 200.0-700.0 V. The capillary voltage was set to 4500 V, end plate offset to 500 V, nebulizing gas at 4 bar, drying gas at 12 $L \cdot min^{-1}$ at 220 °C and sample time 1.00 s. Transfer Funnel 1 radio frequency was 150.0 Vpp, quadrupole ion energy 8.0 eV, and low mass 100.00 *m/z*. collision cell energy was 4.0 eV, radio frequency 300 Vpp, transfer time 75.0 µs, and pre-pulse storage 5.0 µs. Native Bruker file format was converted to mzXML format for further data processing in MZmine.

#### 3.3.2.3 Non-targeted data extraction for MS setup comparison

Data extraction of Orbitrap PI mode spectra in MZmine followed the protocol in Section 3.3.1.4.1. The extraction of MaXis QTOF PI mode data follows the same protocol, while the deviations from it are noted further below.

> Mass Detection: "Exact mass" detector with an intensity noise level of $5 \cdot 10^2$ a.u.
> FTMS Shoulder peaks filter: omitted.
> Chromatogram builder: A minimum time span of 0.05 min (2 sec); a minimum signal height of $5 \cdot 10^2$ a.u.; a mass tolerance of extraction 10 ppm.
> Chromatogram deconvolution: Noise amplitude algorithm was applied with a signal duration of 0.05 – 8.00 min, a minimal signal height of $1.5 \cdot 10^2$ a.u. and amplitude of noise $5 \cdot 10^2$ a.u.

The FTMS peak shoulder filter was omitted since QTOF is not a Fourier transform analyzer. The ppm error in chromatogram building and deconvolution was raised to 10 ppm to improve the peak recognition and shape and to avoid "cleaved" peaks.

## 3.4 Suspect screening

The suspect screening follows a series of steps shown in Figure 3.12.

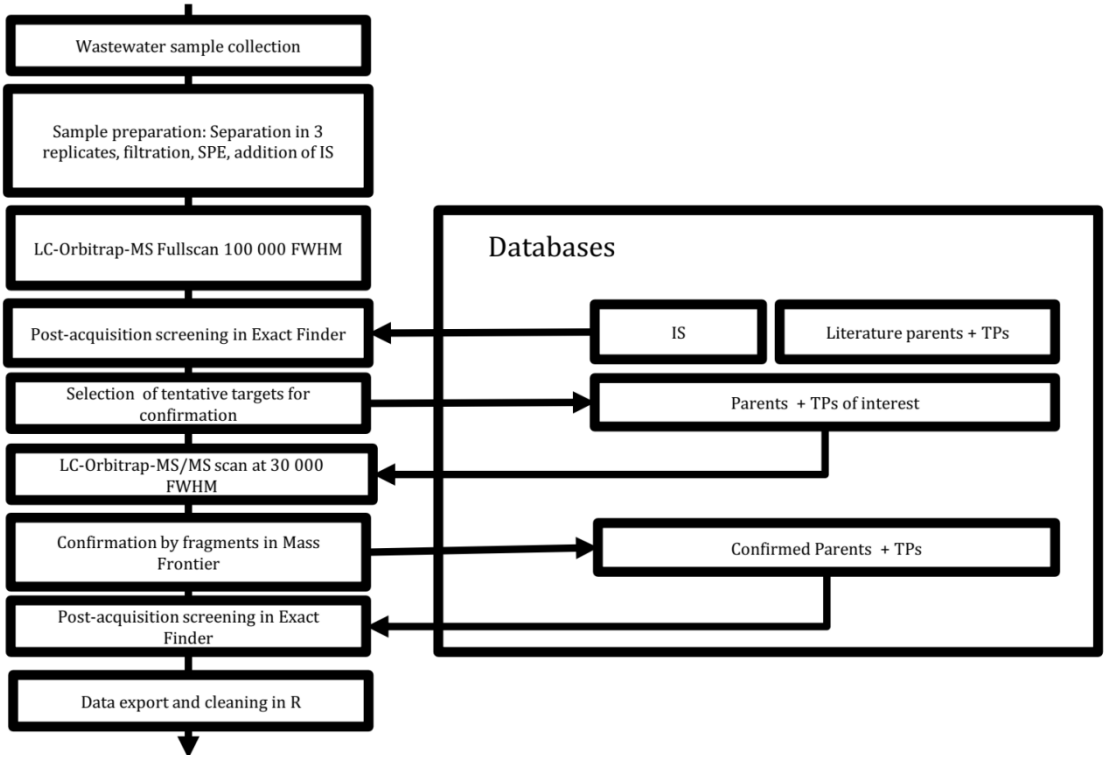



**Figure 3.12 Flow scheme of applied suspect screening.**

### 3.4.1 Suspect libraries

A library of biodegradation products was acquired from Bozo Zonja, CSIC, Barcelona, Spain [233]. A home-made library of DBPs was compiled using literature sources on the degradation of various micro-contaminants in tertiary treatments of tertiary wastewater treatment [3]. The library of biodegradation products included 213 compounds: 42 contaminants and their 171 transformation products and metabolites. The library of DBP series included 189 compounds: 26 contaminants and their 163 transformation products and metabolites.

## 3.4.2 Suspect screening in multistage WWTP described in section 3.3.1

### 3.4.2.1 Multistage WWTP sample preparation

The same set of samples was used as described in Section 3.3.1.1. The samples were separated into three replicates each and subsequently filtered through 2.7 and 1.0 µm glass fiber filters.

3 mL of a 0.1 mol·$L^{-1}$ $Na_2EDTA$ solution per 100 ml of sample was added to obtain a final concentration of 0.1 wt%. Surrogate standard consisting of Sulfadimethoxine-$d_6$ and Sulfadoxine-$d_3$ was added prior to SPE with a final concentration of 200 ng·$L^{-1}$ for influent and 100 ng·$L^{-1}$ for effluent wastewaters. SPE cartridges were conditioned with 6 mL MeOH followed by 6 mL of water at a flow rate of 1 mL·$min^{-1}$.

100 mL of influent and 200 mL of effluent wastewater were loaded onto cartridge at a flow rate of 1 mL·$min^{-1}$. After, the cartridges were rinsed with 6 mL HPLC grade water, at a flow rate of 2 mL·$min^{-1}$, and were air-dried for 30 min. The content of cartridges was eluted with 6 mL MeOH at a flow rate 1 mL·$min^{-1}$. Extracts were evaporated to dryness under a nitrogen stream and reconstituted with 1 mL of 1:9 MeOH:water solution. Finally, 10 µL of a 1 ppm IS mixture was added to achieve a final concentration of 10 ppb (Appendix 2: Table 7.3).

### 3.4.2.2 Suspects' data acquisition and processing

A full-scan LC-HRM spectrum with a set resolution of 100 000 FWHM within an *m/z* range 100-800 *m/z* was acquired in PI ESI mode using spectrometric and chromatographic conditions described in chapter 3.3.1.3. The obtained spectra were screened in Exact Finder for 42 and 26 parents with their respective 171 biodegradation and 163 disinfection TPs. A full-scan LC-HRM spectrum was acquired in NI ESI, however the data was not processed further due to the lack of data bases of NI mode transformation pathways and suspect fragments in Thermo Mass Frontier™.

The Exact Finder screening settings were: no RT limits, mass tolerance < 5 ppm, 5 smoothing points, chromatogram view widths 0.75 min, threshold override > 300 a.u., S/N ratio threshold > 3, RT used for confirmation with an override window of 30 s but is

ignored where not present. The isotopic pattern score was calculated for isotopic pattern > 70 % with allowed mass deviation < 5 ppm and allowed intensity deviation < 30 % of base peak height.

The tentative suspects were filtered in R after the screening in Exact Finder, to only include "Confirmed" substances with peak score > 60%. The signal with the highest count of detected isotopologues was being selected as the representative in case of multiple suspects for a substance. Potential 58 parents with 202 TPs were discovered.

After, 21 parents Acetaminophen, Atenolol, Azithromycin, Bezafibrate, Carbamazepine, Diazepam, Ketoprofen, Ofloxacin, Sulfamethoxazole, Verapamil, Venlafaxine, Valsartan, Acebutolol. Atenolol, Endoxifen, Erlotinib, Estradiol, Iopamidol, Methadone, Prometryne, Propyphenazone, Terbutryne and their 84 TPs were selected for $MS^2$ confirmation screening.

Data-dependent LC-HRM spectra were acquired with a set resolution of 30 000 FWHM in PI ESI mode with the suspects mentioned above as targets for $MS^2$ fragmentation and using spectrometric and chromatographic conditions described in chapter 3.3.1.3. In the data-dependent acquisition, the ions were isolated in the ion trap with a width of 2.0 *m/z,* default charge state 1, a collision-induced dissociation activation type (Q = 0.250 and an activation time of 30 ms), and normalized collision energy of 35 eV.

The acquired $MS^2$ spectra were screened for suspect fragments in Mass Frontier™. The fragments were searched in wastewater samples in which the substance was previously detected in the Exact Finder screening described above. A suspect was considered confirmed when apart from the exact mass in $MS^1$ spectrum at least one characteristic fragment was detected in time-corresponding $MS^2$ spectra (Appendix Table 7.4). The fragments were predicted *in silico* using the Mass Frontier™ built-in prediction from the structure (FiSh filter) under the option "general fragmentation and rearrangement rules" which include the typical fragmentation reactions in organic chemistry. *In silico* fragmentation of $[M+H]^+$, $[M+NH_4]^+$ or $[M+Na]^+$ was chosen depending on the species detected in $MS^1$. The quality of the $MS^2$ spectra, fragment prediction, and their assignment in the spectra were controlled by fragmentation of IS corresponding to the parent compounds of interest. The IS were detected as, for example, Bezafibrate-d4 (Fragments 280.10,

320.13, 348.13 Da) corresponding to Bezafibrate (Fragments at 276.08, 316.11, and 344.10 Da).

The confirmed compounds were then rescreened in Exact Finder using 100 000 FWHM LC-HRMS data with following settings: no RT limits, mass tolerance < 5 ppm, 5 smoothing points, chromatogram view widths 0.75 min, threshold override > 300 a.u., S/N ratio threshold > 3, RT used for confirmation with an override window of 50 s. The isotopic pattern used for identification > 30 % with allowed mass deviation < 10 ppm and allowed intensity deviation < 50 % of base peak height. Where Exact Finder was omitting true signals because of stringent rules, those were extracted manually.

A search for tentative TPs was carried out for metabolic reactions noted in Table 3.1for the confirmed parent compounds. Hereby the reactive moiety was subtracted/added from/to the parent formula to construct suspected TPs creating in total 101 TP suspects.

**Table 3.1 Biotransformation reactions used to construct TP suspects.**

| Transformation moiety name | Added or subtracted molecular moiety | Moiety mass [Da] |
|---:|---:|---:|
| hydroxylation | +O | +15.9949 |
| demethylation | -$CH_2$ | −14.0157 |
| deethylation | -C2$H_4$ | −28.0313 |
| dehydrogenation | +$H_2$ | −2.0157 |
| hydrogenation | -$H_2$ | +2.0157 |
| dehydration | -$H_2O$ | −18.0106 |
| chlorine_reduction | -Cl/+H | −33.9611 |
| acetylation | +$C_2H_2O$ | +42.0106 |
| deacetylation | -$C_2H_2O$ | −42.0106 |
| glucuronidation | +$C_6H_8O_6$ | +176.0320 |
| deglucuronidation | -$C_8H_8O_6$ | −176.0320 |
| sulfonation | +$SO_3$ | +79.9568 |
| desulfonation | -$SO_3$ | −79.9568 |

Subsequently, the 100 000 FWHM full scan was explored for suspects in Exact Finder. Constrictive options were used to exclude false hits. The suspects were searched in PI mode as adducts $[M+H]^+$, $[M+NH_4]^+$ and $[M+Na]^+$. Retention time was constricted to ± 3 min from the corresponding parent. The intensity threshold was set to $10^5$ after observing the median confirmed TP intensity in secondary effluent being $10^6$.

Isotopic pattern score was > 60 % with ≥ 3 isotopes. Suspects observed in the blank were removed from further processing. The results were checked manually in Exact Finder to include not recognized signals and make sure that the signals were integrated well. 9 tentative TPs were found and received special naming is differentiate them from confirmed literature TPs.

For compounds with 3 replicates in a sample with a CV of intensity > 30 % the replicate with the highest deviation among the samples was being omitted in the calculation of average intensity in R. Subsequently, the matrix suppression of chemicals was normalized using correction factors derived from IS of the parent. IS with the closest RT to the parent was used for correction when the parent lacked own IS. Signal intensity of IS was checked manually to avoid misrepresentation in correction matrix. The IS signal intensities in samples were compared to signal intensity in secondary effluent. These ratios served as matrix correction factors by multiplying the detected intensity of substances by the respective correction factor. The matrix correction factors also accounted for the lower sample volume of influent samples prior to extraction.

## 3.4.3 Suspect screening in BAC-UF treatment

### 3.4.3.1 BAC-UF reactor setup and sampling

Two reactors comprised the industrial-scale BAC-UF pilot plant (Figure 3.13). Secondary effluent from a municipal WWTP in Aartselaar (Belgium) designed for biological removal of nutrients was the influent of the pilot. Details of the reactor setup can be found in Appendix 3.

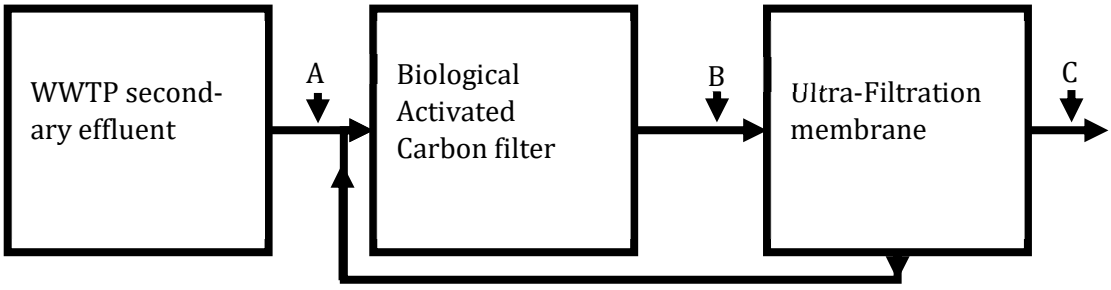



Figure 3.13 Scheme of the pilot BAC-UF treatment. A corresponds to the sampling point of influent, B to BAC effluent and C to UF effluent. Centrate of UF was recirculated into BAC after sampling point A.

24 h composite samples were collected considering the empty bed contact time and the hydraulic retention time. The samples correspond to the influent of BAC before the recirculation of centrate (A), effluent of BAC tank (B), and effluent of UF tank denoted as permeate (C).

The samples were collected in amber glass bottles, pre-filtered with 0.45 µm filters (Whatman, U.K.), and stored at -25 °C. The samples were collected on days 02.05.16, 03.05.16, 16.12.16, and 25.01.17 and named here for convenience “day 0”, “day 1”, “day 229”, and “day 269”, respectively.

#### 3.4.3.2 BAC-UF sample preparation

Samples were filtered through 1 µm glass fiber and 0.45 µm membrane filters. 0.1 M $Na_2EDTA$ was added until a final concentration of 0.1wt%. The surrogate standard was added until a final concentration of 100 $ng \cdot L^{-1}$. Oasis HLB cartridges (60 mg, 3 mL) were conditioned with 5 mL MeOH and 5 mL water at a flow rate of 2 $mL \cdot min^{-1}$. 100 mL sample was loaded onto the cartridges at a flow rate of 2 $mL \cdot min^{-1}$. The loaded cartridges were washed with 6 mL water and air dried for 30 min. The content of cartridges was eluted with 6 mL MeOH at a flow rate of 1 $mL \cdot min^{-1}$, evaporated to dryness using $N_2$, and reconstituted in 1 mL MeOH: water (10:90). 10 ppb IS were added (Appendix 3: Table 7.3).

### 3.4.3.3 Suspects' data acquisition and processing

LC-HRMS acquisition was performed with equipment and software described in Section 3.4.2.2. $MS^1$ and $MS^2$ settings were the same as in Section 3.4.2.2, however, the chromatographic separation was achieved on Acquity UPLC® HSS $T_3$ (2.1 mm × 50 mm, 1.8 µm particle size, Waters UK) in PI mode. 20 µL of sample was injected into LC-MS using the solvents and gradients as described by Gros et al. [130].

The obtained spectra were screened for 42 parents and their 171 biodegradation TPs. 11 parents and their 50 TPs were selected for $MS^2$ confirmation screening: Amoxicillin, Atenolol, Azithromycin, Bezafibrate, Carbamazepine, Codeine, Iopromide, Penciclovir, Ranitidine, Sulfamethazine, and Valsartan.

A data-dependent LC-HRM spectrum was acquired with a set resolution of 30 000 FWHM in PI ESI mode. A full-scan LC-HRM spectrum was acquired in NI ESI, however the data was not processed further due to the lack of data bases of NI mode transformation pathways and suspect fragments in Thermo Mass Frontier™. The acquired $MS^2$ spectra were screened for fragments of suspects using Mass Frontier™ 7.0.3 (High Chem™) (Appendix Table 7.6). The confirmed compounds were then rescreened in Exact Finder using 100 000 FWHM LC-HRMS data. The exact $MS^2$ conditions and applied Exact Finder method can be found in section 3.4.2.2.

In R the replicates were summarized to yield an average intensity for a sample. For compounds with 3 replicates in a sample with CV > 30 % the replicate with the highest deviation among the samples was being omitted in the calculation of summary. Subsequently, the matrix suppression of chemicals was normalized using correction factors derived from an IS corresponding to the parent.

Where isotopically labeled IS corresponding to the parent was absent, an IS with closest retention time was picked for normalization of compound family. The correction factors used to correct the corresponding parent and TPs were derived from the factor of IS intensity in samples to the signal intensity of IS in permeate sample of "269 days" series.

# 4 RESULTS AND DISCUSSION

## 4.1 Non-targeted methodology development and analysis

### 4.1.1 Distribution of substances

**Table 4.1 The number of the total detected molecular features, the percentage of all features in a fraction compared to a number of influent features, and only the features with an assigned molecular formula in fractions of the samples. The fractions of both secondary and tertiary effluent are relating to the influent.**

| Fraction | Total features | % of influent | Features with assigned formula |
|---|---|---|---|
| $Influent_{Total}$ | 2409 | 100 | 1495 |
| $Influent_{Disappeared}$ | 1617 | 67 | 1069 |
| $Secondary\ effluent_{Total}$ | 1047 | 44 | 515 |
| $Secondary\ effluent_{Appeared}$ | 255 | 11 | 89 |
| $Secondary\ effluent_{Recalcictrant}$ | 792 | 33 | 426 |
| $Tertiary\ effluent_{Total}$ | 1106 | 46 | 541 |
| $Tertiary\ effluent_{Appeared}$ | 379 | 16 | 161 |
| $Tertiary\ effluent_{Recalcictrant}$ | 727 | 30 | 380 |

In the secondary treatment, 1617 of 2409 (67%) detected molecular features were removed and 255 of 1047 (24%) new features appeared in the secondary treatment (Table 4.1). Biological treatment involves both metabolic and co-metabolic reactions that form TPs.

This process explains a large number of removed and appeared molecular features. The tertiary treatment removed a minute portion of detected molecular features but added new 379 of 1106 features (34% of tertiary effluent).

Compared to the conventional COD, BOD or TOC measurements, which give an average estimate of the DOM content, the presented method tracks changes at the level of subsets of DOM constituents. This approach can benefit specialized treatments for which the appearance or removal of DOM fractions is more important than the bulk removal. For example, the tracking of fractions of soluble microbial products might improve the understanding of membrane fouling in reverse osmosis and nanofiltration [234]

### 4.1.2 Mass changes

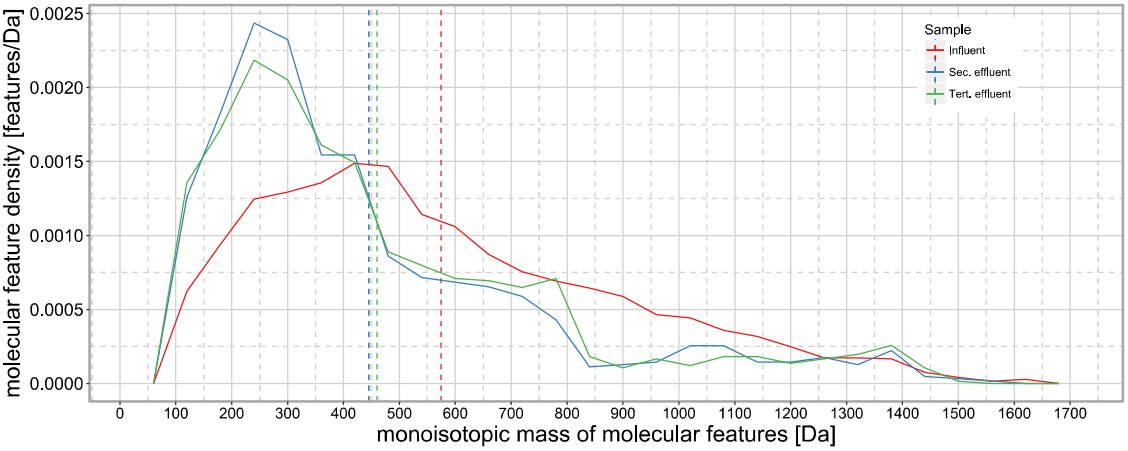



**Figure 4.1 Density distribution of feature masses in secondary influent, secondary, and tertiary effluents. The dashed lines show means of distributions.**

Mass changes within the subsets were explored after recognizing interesting subsets of features, since mass is the most intrinsic property of molecular features in HRMS. The secondary effluent showed a mean mass drop of 129 Da ($p < 0.001$) compared to the influent (Figure 4.1). There is no significant mass drop between the secondary and tertiary treatments ($p = 0.27$).

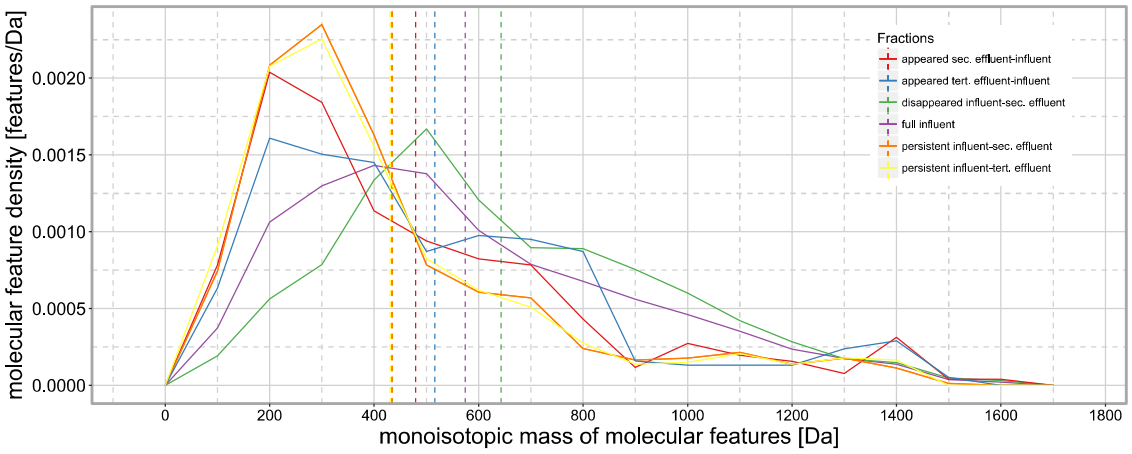



**Figure 4.2 Density distribution of molecular feature masses in the subsets of features in secondary influent and effluent, and tertiary effluent compared to the full complement of the influent. The dashed lines show means of distributions.**

However, the mean mass of appeared features dropped by 163 Da compared to the disappeared ones ($p < 0.001$) (Figure 4.2). The drop was considerably larger for this transformed DOM compared to the bulk DOM in secondary treatment, showing that major changes can be masked by bulk properties. Meanwhile, features detected both in influent and effluent of tertiary treatment do not show a mean mass change which hints at their recalcitrant nature. Often, SEC combined with a carbon detector is used to estimate the weight distributions of DOM and the bulk rejection estimation with experimental ease [234]. However, the SEC spectra lack the high resolution of LC-HRMS, the possibility to extract subsets of signals of interest and the possibility to correlate masses of individual substances in these fractions to atomic properties such as unsaturation or atomic ratios of DOM.

### 4.1.3 Intensity changes

The possibility to connect feature subsets to the spectrometric intensity of individual molecular features is also a powerful incentive for non-targeted analysis. The mean absolute intensity decreased by 3.38 times in the secondary effluent ($p < 0.001$). Biological treatment is known to achieve high TOC removal efficiencies for the entire DOM [32], yet many medium polar substances of environmental concern are either recalcitrant or are transformed without mineralization [115].

A high number of molecular features can be detected in HRMS of effluent wastewater even while controlling DOC and BOD [171]. Therefore, it is interesting that the mean of the absolute intensity differences of 792 features, detected before and after secondary treatment, dropped by 2.5 times ($p < 0.001$ for paired data) in effluent compared to the influent. This value is lower than for the overall intensity decrease hinting at the recalcitrant nature of these constituents.

Together with a high removal of molecular features in the secondary treatment, the data agree with previous work (Park et al., 2009). A fraction of recalcitrant features showed an intensity increase, which can be connected to a lower matrix effect in the effluent. There is no significant difference in feature intensity between the secondary and tertiary effluents of the measured samples ($p = 0.66$), again hinting at their recalcitrant nature.

### 4.1.4 Chemical profile

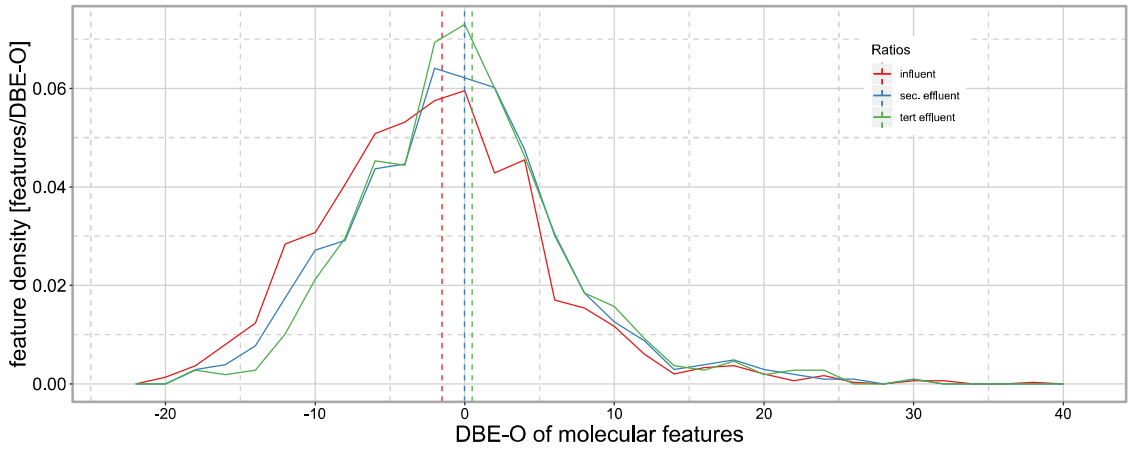



**Figure 4.3 Density distribution of DBE-O of molecular features with assigned molecular formulae in influent and effluent of secondary treatment, and effluent of tertiary treatment. The dashed lines show means of distributions.**

The feature subset analysis was supplemented with DBE-O exploration. The DBE-O increased by 1.5 ($p < 0.001$) in secondary effluent (Figure 4.3). Moreover, the appearing molecular features show a significant increase of DBE-O of 4 compared to the disappearing ($p < 0.001$). These fractions are interesting for the estimation of metabolic activity and the observations are supported by the documented accumulation of aromaticity in biological treatment [235]. The established estimation of aromaticity in wastewater DOM using $SUVA_{254}$ allows recognizing the presence of aromatic compounds.

The presented method, however, estimates the unsaturation of DOM and expands the analysis to not UV-active molecules in subsets of molecular features.

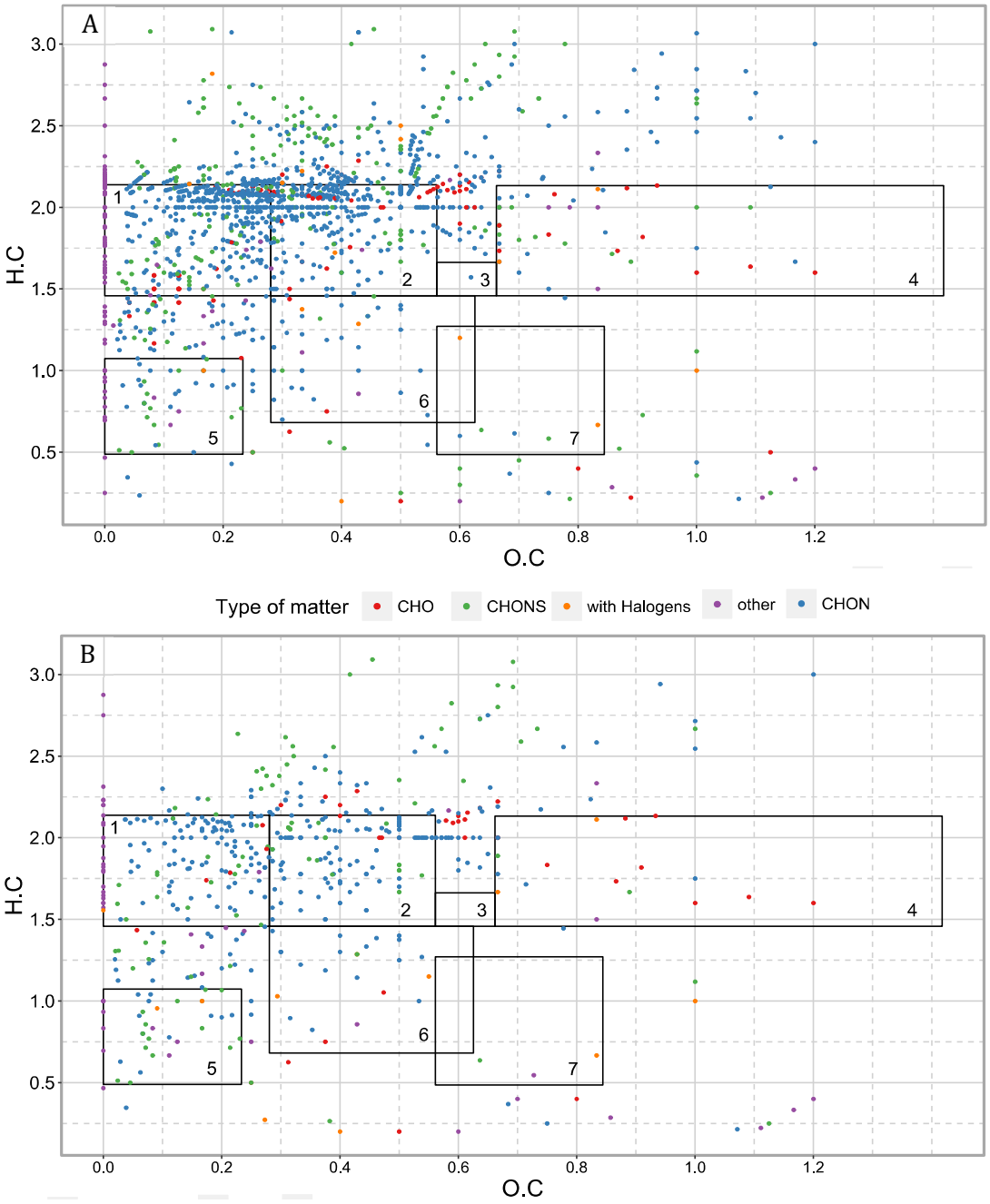



**Figure 4.4 Van Krevelen plot of molecular features in secondary influent and effluent. Elemental composition shown as color, where "other" signifies the not explicitly mentioned combinations of elements C, H, O, N, S. The rectangles, reproduced from Maizel et al. [25], depict NOM regions of 1 - lipids, 2 - proteins, 3 - amino sugars, 4 - carbohydrates, 5 - condensed hydrocarbons, 6 - lignin, and 7 - tannin.**

Additionally to the explored unsaturation by DBE-O, the chemical properties of DOM were investigated using the van Krevelen plot. It assessed types of wastewater DOM in secondary treatment depending on the characteristic regions in the diagram (Figure 4.4). No significant difference was observed between secondary and tertiary effluents (Figure 4.5).

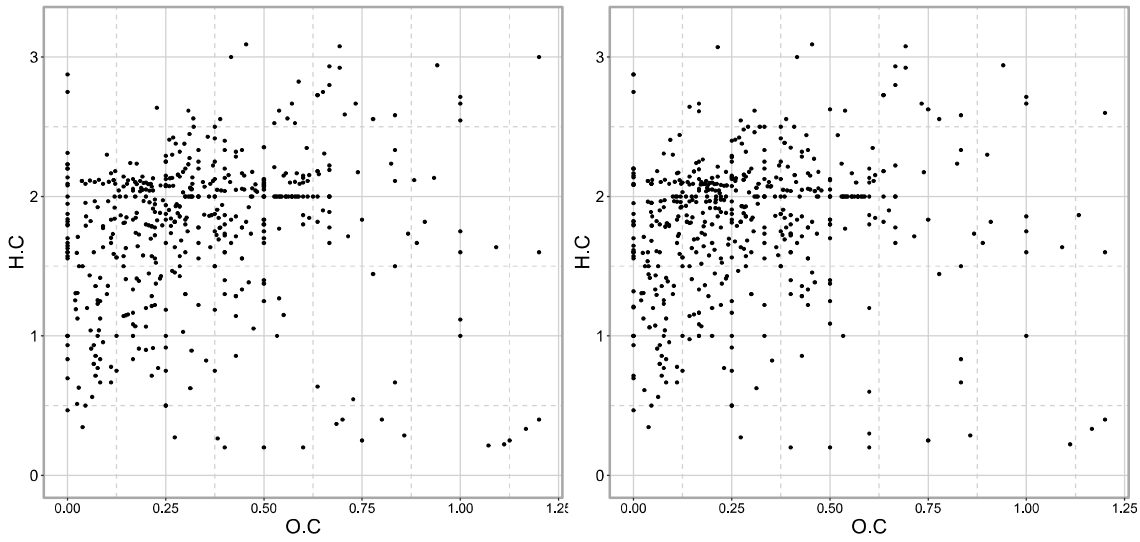



**Figure 4.5 Van Krevelen plots of the secondary (left) and tertiary (right) treatment for the features with assigned elemental composition.**

The diagram deviates from a distribution in a previous study of a biological treatment [25], where the majority of molecular features fell within the seven regions. Here, the van Krevelen plot showed many features with H/C > 2, while the literature's features did not exceed H/C ≤ 2, which is caused by different HRMS acquisition and approach to the prediction of elemental composition. Different to the literature, the plot here showed features with H/C < 0.5 and O/C up to 1.2, which would correspond to highly oxygenated, unsaturated matter.

Here, the different MS acquisition and data treatment might have led to highlighting different subsets of DOM. In the presented analysis, DOM is largely aliphatic and the oxygen content dropped with increasing unsaturation. The overall feature density decreased in the secondary treatment. However, none of the seven regions displayed a disproportional change of feature density in course of treatment compared to other regions.

This indicates that DOM composition stayed largely constant throughout the treatment. N/C in the van Krevelen plot correlates with H/C, providing a link between the unsaturation and the content of nitrogen (Figure 4.6). The N/C of DOM did not display regional changes in the secondary treatment, showing that the detected molecular features were not subject to a transformation detectable in the patterns of the plot.

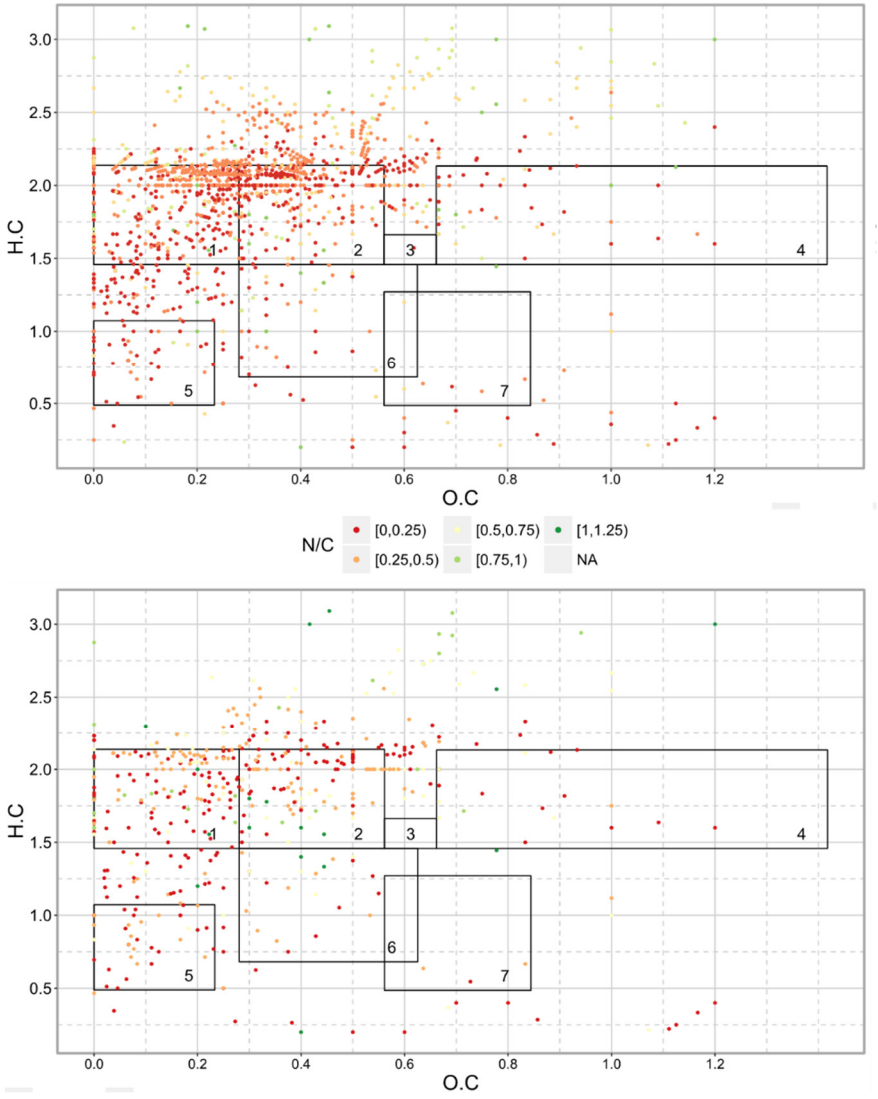



**Figure 4.6 Van Krevelen plot of molecular features of the secondary influent and effluent with N/C ranges as color. The rectangles reproduced from Maizel et al [25], depict NOM regions of 1 - lipids, 2 - proteins, 3 - amino sugars, 4 - carbohydrates, 5 - condensed hydrocarbons, 6 - lignin, and 7 - tannin.**

In Figure 4.4 the CHONS group exhibited a spread to either higher aliphatic or lower aromatic H/C values compared to CHO and CHON which suggests that there is a level of clustering by the elemental composition which has to be explored in the future. The applied LC-MS acquisition, as well as data treatment, prioritized compounds of anthropogenic origin.

Many molecular features in the analysis did not belong to the regions of DOM applied by Maizel et al., and they did not follow the literature's observations for NOM and effluent DOM. The author suggests that due to the shifted focus of the analysis a complement of features was revealed that might be of anthropogenic origin and therefore did not look like NOM in the van Krevelen diagram. This focus on the anthropogenic fraction in the presented van Krevelen analysis adds to the extensively researched field of NOM and effluent DOM.

Additionally to the van Krevelen plot, the gains or losses of simple structural moieties were used to improve the knowledge of transformations in the treatment. Reaction pairs between parent compounds and biotransformation products were explored in the subsets of influent features and features appearing in secondary and tertiary effluents since these are the most relevant fractions to monitor TPs (Table 4.2).

This estimation recorded chemical reactivity with 212 and 290 potential transformation pairs for the monitored moieties between the influent and the appeared features in the secondary and tertiary effluent, respectively. Losses of -$CH_2$-, -$C_2H_4$- and -$C_2H_4O$- are connected with the degradation of oligomeric surfactants, as shown previously for -$CH_2$- [179].

**Table 4.2 Number of pairs between parents in influent and biotransformation products in the subset of appeared molecular features in secondary and tertiary effluents, respectively, corresponding to gains/losses of molecular moieties in transformation reactions.**

| Reaction | Moiety | Secondary effluent | Tertiary effluent |
|---|---|---|---|
| Demethylation | - $CH_2$ | 22 | 46 |
| Deethylation | - $C_2H_4$ | 17 | 22 |
| Deethoxylation | - $C_2H_4O$ | 26 | 37 |
| Hydroxylation | + O | 19 | 23 |
| Alcohol Oxidation | + O/ - 2H | 39 | 39 |
| Hydrogenation | + $H_2$ | 6 | 11 |
| Dehydrogenation | - $H_2$ | 26 | 33 |
| Dehydration | - $H_2O$ | 23 | 22 |
| Acetylation | + $C_2H_2O$ | 20 | 29 |
| Deacetylation | - $C_2H_2O$ | 5 | 10 |
| Glucuronidation | + $C_6H_8O_6$ | 1 | 1 |
| Deglucuronidation | - $C_6H_8O_6$ | 0 | 3 |
| Chlorine substitution | + Cl/ - H | 2 | 4 |
| Sulfonylation | + $SO_3$ | 4 | 9 |
| Desulfonylation | - $SO_3$ | 2 | 1 |

The author compared the percentages of the total features ratios for 6 structural moieties discovered both here and by Schollée et al. in a secondary treatment under the consideration of the retention time filter. Our analysis yielded a lower overall amount of detected potential transformations which can be attributed to different data processing procedures. The differences Δ between the subsets of transformation pairs of gain/loss moieties with a retention time filter in my data and by Schollée et al. in literature [179] were compared (Table 4.3) using the equation

*(1)* $\Delta(\text{reaction}) = \left(\frac{\text{pairs(reaction)}}{\text{pairs(sum)}}\right)_{here} - \left(\frac{\text{pairs(reaction)}}{\text{pairs(sum)}}\right)_{\text{Schollée et al}}$

The comparison yielded differences of < 7% ($\Delta(+ H_2)$= -6.2; $\Delta(- H_2)$= 5.3; $\Delta(- H_2O)$= 7.1; $\Delta(- CH_2)$= 0.4; $\Delta(- C_2H_4)$= -4.4; $\Delta(+ O)$= -2.3).

**Table 4.3 The comparison of presented algorithm of extraction of transformation pairs to the previously published results by Schollée et al. The table shows the type of proposed transformation reaction, the added/eliminated structural moiety, the amount of discovered transformation pairs in both approaches and their percentage in respect to the total number of discovered pairs, and the difference of these percentages between both approaches.**

| Reaction | Moiety | # of Pairs$_{here}$ | # of Pairs$_{Schollée\ et\ al.}$ | Pairs$_{here}$% | Pairs$_{Schollée\ et\ al.}$% | Δ% |
|---|---|---|---|---|---|---|
| Hydrogenation | + $H_2$ | 6 | 465 | 0.05 | 0.12 | -6.19 |
| Dehydrogenation | - $H_2$ | 26 | 715 | 0.23 | 0.18 | 5.32 |
| Dehydration | - $H_2O$ | 23 | 535 | 0.20 | 0.13 | 7.12 |
| Demethylation | - $CH_2$ | 22 | 770 | 0.19 | 0.19 | 0.42 |
| Deethylation | - $C_2H_4$ | 17 | 785 | 0.15 | 0.19 | -4.37 |
| Hydroxylation | + O | 19 | 773 | 0.17 | 0.19 | -2.30 |
| Sum of all pairs | | 113 | 4043 | | | |

The deviation < 7%, suggests that despite a different total amount of discovered pairs the results in both analyses can be compared. The presented analysis explored structural moieties not described by Schollée et al. as + O/- 2H and + Cl/- H. Transformation pairs for the alcohol oxidation were found as expected due to the metabolic nature of secondary treatment. Chlorination traces were not found using Cl substitution, despite a chlorination step. The complicated sample matrix caused by a direct injection of the samples into LC-MS can lead to the expectedly weak signals going undetected. Another possibility is that the mechanisms of chlorination are more complicated than the substitution of a single hydrogen. Such building of reaction pairs is a first step to estimate the chemical activity of a treatment system. For example, the elimination/addition on a double bond can assess the mechanisms in the shifts of unsaturation.

### 4.1.5 Kendrick mass defect

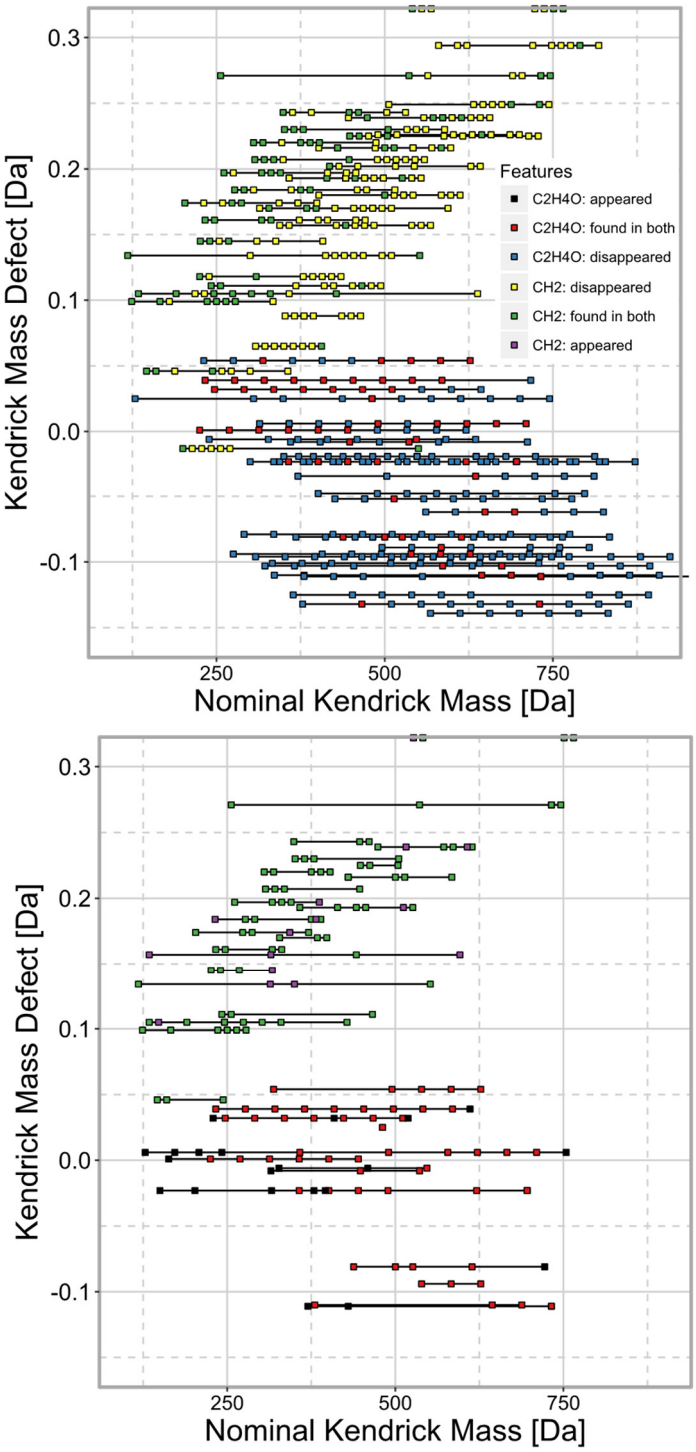



**Figure 4.7 31 -$CH_2$- and 38 -$C_2H_4O$- homolog series in secondary influent (left) with more than six molecular features in a series and their development in secondary effluent (right) for PI mode data.**

The exploration of transformation pairs above aided in selecting interesting moieties for the KMD analysis. The KMD analysis uncovered 31 homologous -$CH_2$- series in the secondary treatment (Figure 4.7), which in a municipal wastewater treatment often correspond to surfactants. The plot showed a removal of heavy constituents. This removal of heavy homologs, however, was not observed for 38 series of -$C_2H_4O$- homologs. Most -$C_2H_4O$- mass series were completely removed, while a few were almost completely recalcitrant in the secondary effluent.

There are more -$CH_2$- and -$C_2H_4O$- homologous series in raw data than shown in the plots, but the algorithm prioritized pronounced series by involving a strict set of rules. The exclusion of fake homologs was ranked over losing true homologs. Additionally, the minimum length of a homolog series for -$CH_2$- and -$C_2H_4O$- was < 7 in secondary influent being a high number of homologs in one series. In case of abundant surfactant series, this prioritization improves the interpretation using the graphic nature of the Kendrick plot. This becomes evident in the graphically heavy plot of 111 -*O*- homolog series with minimum 3 homologs in secondary influent (Figure 4.8).

The author takes into account that the mechanism of TP formation for DOM is more complicated than a sequential cleavage of -$CH_2$- [236,237]. Also, the detection of homolog series is not straightforward and depends both on the quality of acquired data and the restrictions of the picking algorithm [164]. In this study, Kendrick plots helped to uncover prominent homologous series of prescreened structural moieties of interest which will be used for further prioritized research.

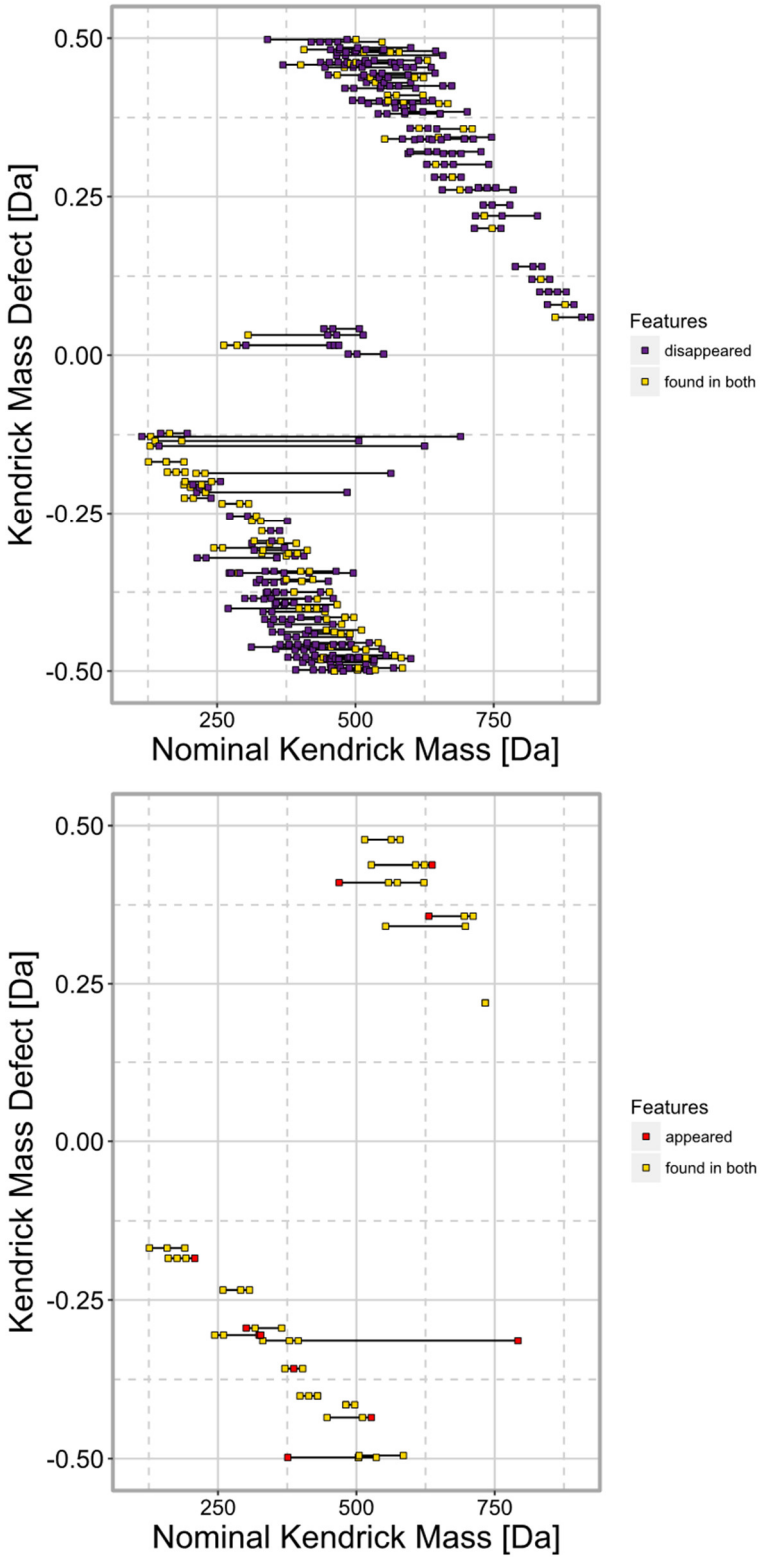



**Figure 4.8 111 -*O*- KMD homolog series in the secondary influent with more than two molecular features in a series and their development in the secondary effluent for PI mode data.**

### 4.1.6 Benefits and challenges

The focus of the study was not DOM enrichment but data analysis. While being aware of the possible shortcomings of the applied pre-treatment and LC-HRMS approach, the analysis detects significant changes of various DOM attributes during WWT. The fractions of high interest, such as anthropogenic organics, are retained by the procedure, as signified by the occurrence of IS throughout the chromatographic range. The focus of data treatment is to extract many features while reducing inclusion of noise. Although noise signals cannot be avoided in big datasets , properly selected settings lower their amount. The application of retention time to distinguish molecular features facilitates the task.

Similarly, the prediction of elemental composition for DOM is challenging, especially for high-molecular-weight compounds. Predictions were improved by using heuristic rules and isotopic information. The prediction of P- and F-containing formulas remains difficult even having a low mass error [238].

## 4.2 Comparison of Orbitrap and QTOF LC-MS setups using non-targeted DOM analysis

### 4.2.1 The scope of the study

The presented research compared the quality of non-targeted analysis with respect to the spectrometer setup. The used grab wastewater samples of primary influent and effluent as well as secondary effluent were pre-concentrated with multiphase SPE and high-resolution data was acquired using Orbitrap-MS and QTOF-MS as described in Section 3.3.2.2. To simplify the comparison of the setups, the Thermo LTQ-Orbitrap Velos$^{TM}$ HR mass spectrometer setup will be colloquially named “Orbitrap” and the Bruker Maxis QTOF HR mass spectrometer setup will be named “QTOF” in this thesis chapter.

The decision to do multiphase SPE was motivated by widening the exploration of DOM. Hereby the non-targeted analysis of directly injected samples was compared to pre-concentrated samples.

Grab sampling applied here does not ensure a representative analysis of wastewater DOM for the system. However, the scope of the study was to evaluate the mass spectrometer performance concerning the non-targeted analysis of wastewater. In this respect, the application of grab samples fits the purpose. The comparison of Orbitrap- and QTOF-MS is explored by analyzing PI mode data which has the larger impact on the methodology compared to NI mode data. Although NOM is more detectable in NI mode, this study aims to focus on anthropogenic compounds preferentially detected in PI mode.

### 4.2.2 Detected complement

The complement of detected features is consistently lower for the QTOF setup under given experimental conditions (Table 4.4). This confirms the expectation, given the higher resolving power of Orbitrap-MS that leads to a higher number of separate signals. The fraction of identified formulae-to-all detected features was higher in Orbitrap samples due to the higher resolution of the instrument.

Orbitrap showed a higher precision of isotopologue signals needed to produce a valid molecular formula candidate by generating a valid isotopic pattern score in MZmine.

More importantly, the set accuracy of formula candidates < 5 ppm lowered the rate of molecular formula assignment for an instrument with a lower resolution.

**Table 4.4 Number of all detected molecular features, features with assigned molecular formula and the fraction of assigned formulae in all detected signals for entire samples and their subsets.**

| Samples and Subsets | Detected signals | Assigned formulae | Percentage of assigned formulae |
|---|---|---|---|
| Influent (Orbitrap) | 7798 | 3724 | 48 |
| Influent (QTOF) | 6736 | 2287 | 34 |
| Primary effluent (Orbitrap) | 6731 | 3273 | 49 |
| Primary effluent (QTOF) | 4658 | 1718 | 37 |
| Secondary effluent (Orbitrap) | 3626 | 1881 | 52 |
| Secondary effluent (QTOF) | 3046 | 1141 | 37 |
| Influent $(Orbitrap)_{disappeared}$ | 5518 | 2383 | 43 |
| Influent $(QTOF)_{disappeared}$ | 4591 | 1379 | 30 |
| Secondary effluent $(Orbitrap)_{appeared}$ | 1346 | 540 | 40 |
| Secondary effluent $(QTOF)_{appeared}$ | 901 | 233 | 26 |
| Secondary effluent $(Orbitrap)_{recalcitrant}$ | 2280 | 1341 | 59 |
| Secondary effluent $(QTOF)_{recalcitrant}$ | 2145 | 908 | 42 |

### 4.2.3 Distribution of masses

The mass ranges had right-skewed distributions in both Orbitrap- and QTOF MS data. However, Orbitrap consistently detected less high molecular weight features compared to QTOF. This property of Orbitrap analyzer was previously noted in a comparison of non-targeted analysis with Orbitrap and FTICR-MS [28].Both instrumental setups identified mass shifts for detected DOM in course of wastewater treatment.

There is no significant difference between the average masses of detected molecular features in influent and primary effluent in Orbitrap ($p = 0.84$) while in QTOF the average masses show a drop of 55.99 Da ($p < 0.001$). The feature density distributions of influent and primary treatment samples are similar for Orbitrap data (Figure 4.9).

However, equivalent QTOF samples show that the percentage of molecular features with mass > 500 Da dropped after primary treatment. Primary treatment removes suspended solids and large particles. It removes to a lesser extent DOM and micro-contaminants within by adsorption to suspended solids [239,240]. A higher sensitivity of QTOF for the subset of large molecules detected a change in the fraction of molecules which was more susceptible to adsorption in primary treatment.

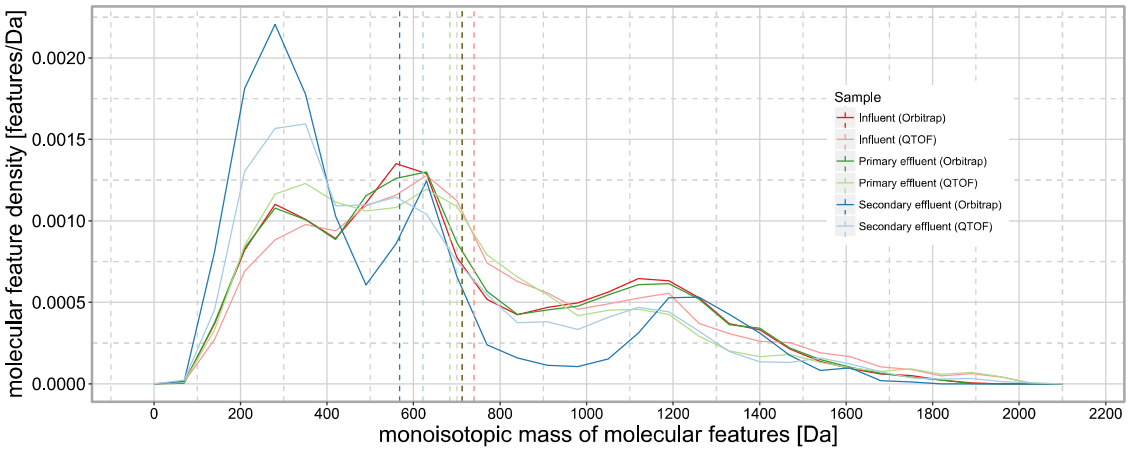



**Figure 4.9 Density distribution of molecular feature masses of Orbitrap-MS and QTOF-MS data. The dashed lines indicate the average mass of distributions.**

The significant decrease of average mass between influent and secondary effluent was 145.19 Da ($p < 0.001$) for Orbitrap and 118.07 Da ($p < 0.001$) for QTOF. Orbitrap data of secondary effluent shows an increase of feature density in the region of small masses and a pronounced loss of density in the region of 700 -1200 Da. QTOF also shows an increase of small molecules < 500 Da, however here molecular feature density decreases in the entire region of larger masses after treatment.

#### 4.2.3.1 Mass Fractions

The average masses of molecular feature subsets as removed from the influent and as appeared after secondary treatment show a drop of 172.22 Da ($p < 0.001$) for Orbitrap and 141.04 Da ($p < 0.001$) for QTOF samples, implying a decomposition of DOM according to metabolic and co-metabolic reactions in secondary treatment. Both QTOF and Orbitrap show a prevalent disappearance of molecules with mass ≈ 600 Da and appearing signals in the range 200 - 400 Da. Additionally, Orbitrap shows a fraction of appearing features at 1200 – 1400 Da observed in the non-targeted analysis described in Section 4.1 (Figure 4.2) which indicate a fraction of appearing large treatment by-products.

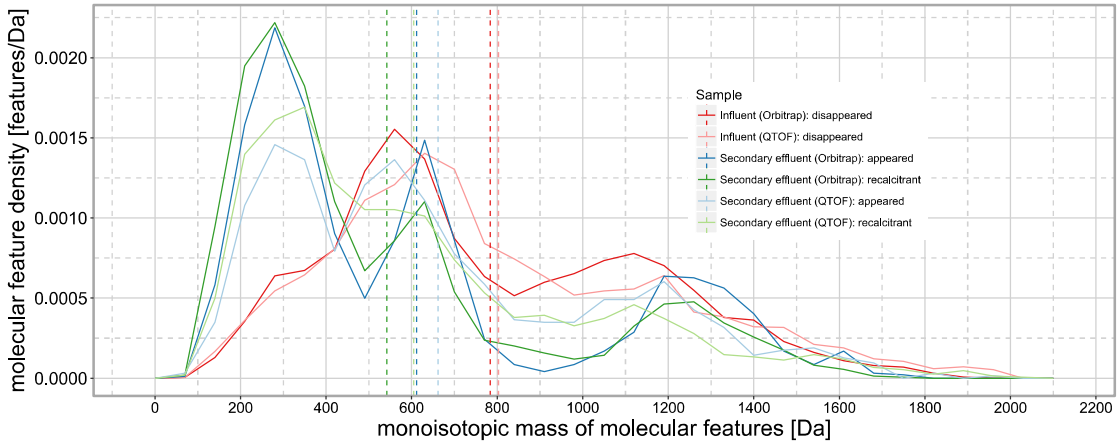



**Figure 4.10 Density distribution of molecular feature masses in subsets of features removed from influent, appeared in secondary effluent or encountered in both samples of Orbitrap-MS and QTOF-MS data. The dashed lines indicate the average mass of distributions.**

### 4.2.4 Intensity changes

The recalcitrant fraction after secondary treatment has a similar amount of molecular features and their mass distribution in both QTOF and Orbitrap data, accounting for the differences that were already described above. The ratio of secondary effluent-to-influent absolute intensity displays a median rise of 0.3 and 0.4 times for Orbitrap and QTOF recalcitrant molecular features, respectively. Upon $\log_2$-transformation, the ratio becomes a bimodal distribution with local maxima at 2 and 0.03 times.

This demonstrates a strong removal for a subset of molecular features, however with an intensity increase for the remaining ones. The intensity rise can be caused by a reduced matrix in the secondary effluent sample. Such intensity rise would increase the influent-to-effluent ratio, since the intensity was not corrected in respect to matrix-assisted ion suppression. The possibility of accumulation of compounds already present in the influent, e.g. products of metabolism in humans or sewers, is viable as well, however difficult to test in the presented setup.

### 4.2.5 Chemical profile

The unsaturation of detected DOM was explored by means of DBE-O. DBE-O distributions in influent and primary effluent did not reveal a significant difference in both Orbitrap ($p = 0.925$) and QTOF data ($p = 0.054$) (Figure 4.11). However, there was a significant DBE-O increase of 1.7 ($p < 0.001$) in Orbitrap and 0.9 ($p < 0.05$) in QTOF data when comparing influent and secondary effluent. QTOF-MS was showing higher DBE-O throughout the entire mass range with exception of features with mass 100-200 Da where DBE-O was lower in all QTOF samples.

DBE-O distributions of QTOF-MS molecular features had a higher average DBE-O compared to the distributions of Orbitrap-MS data. The reason for such difference in behavior will be explored in the future. Interestingly, DBE-O increased in secondary treatment as was also observed in the non-targeted analysis above (Section 4.1.4), showing that this might be an intrinsic property of secondary biological treatment which can be used for future non-targeted fingerprinting of WWTP.

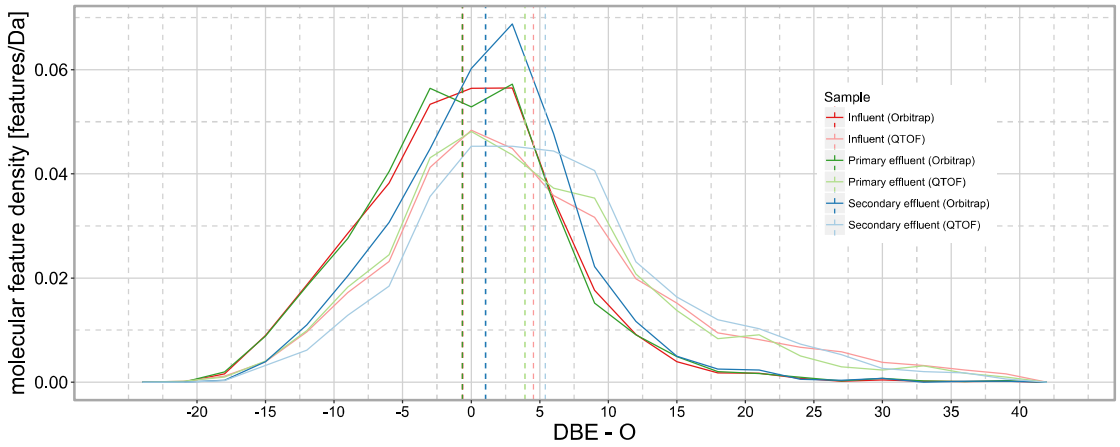



**Figure 4.11 Density distribution of DBE-O of Orbitrap-MS and QTOF-MS molecular features. The dashed lines show means of distributions.**

The difference of chemical profiles of Orbitrap and QTOF data was observed in the van Krevelen plot. Orbitrap data displayed more saturated molecular features at H/C > 2.0 than the QTOF data with the majority these molecular features having mass > 400 Da (Figure 4.12), and being identified as CHON and CHONS matter (Figure 4.13).

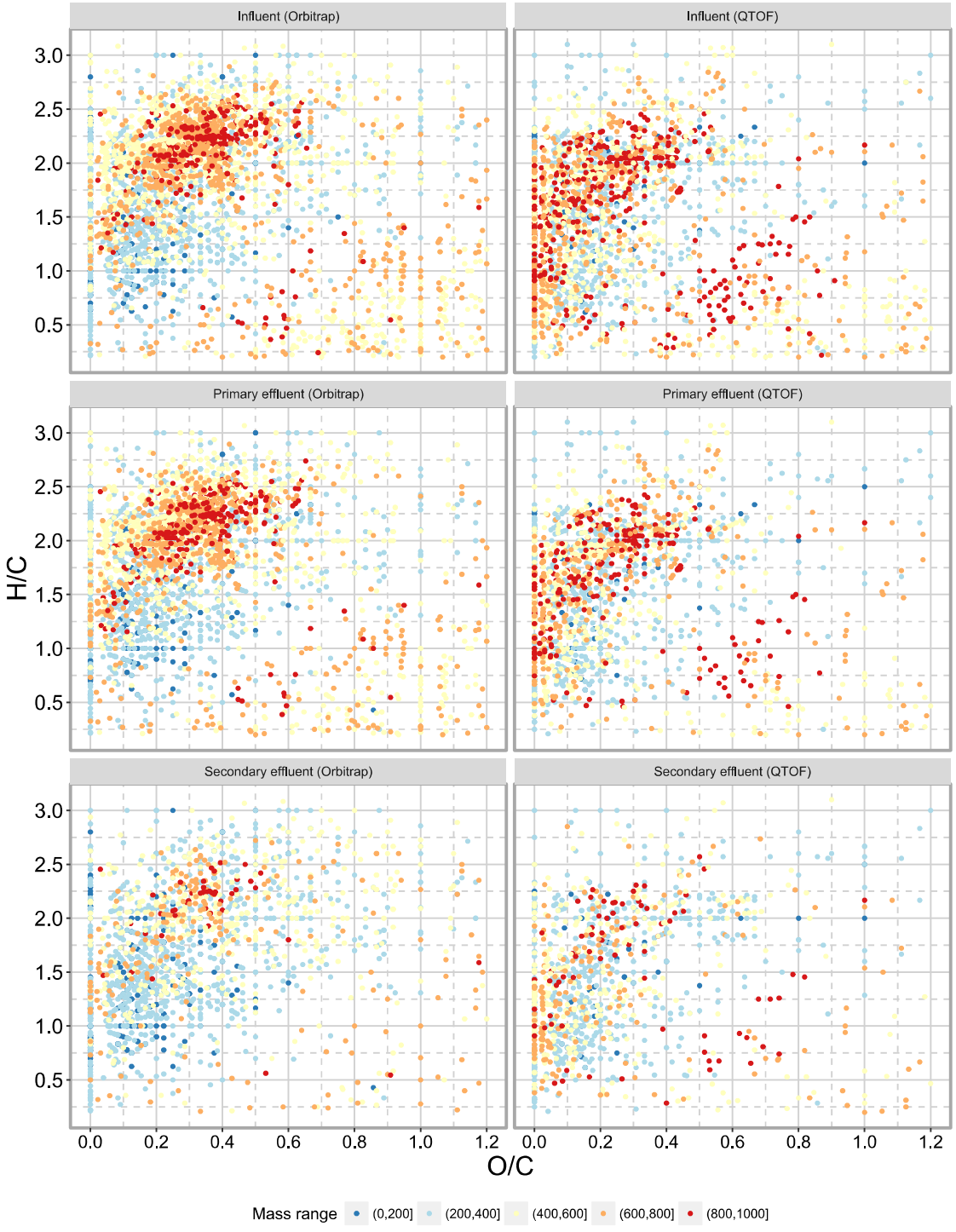



**Figure 4.12 Van Krevelen plot of Orbitrap-MS and QTOF-MS molecular features. Color indicates mass ranges of features.**

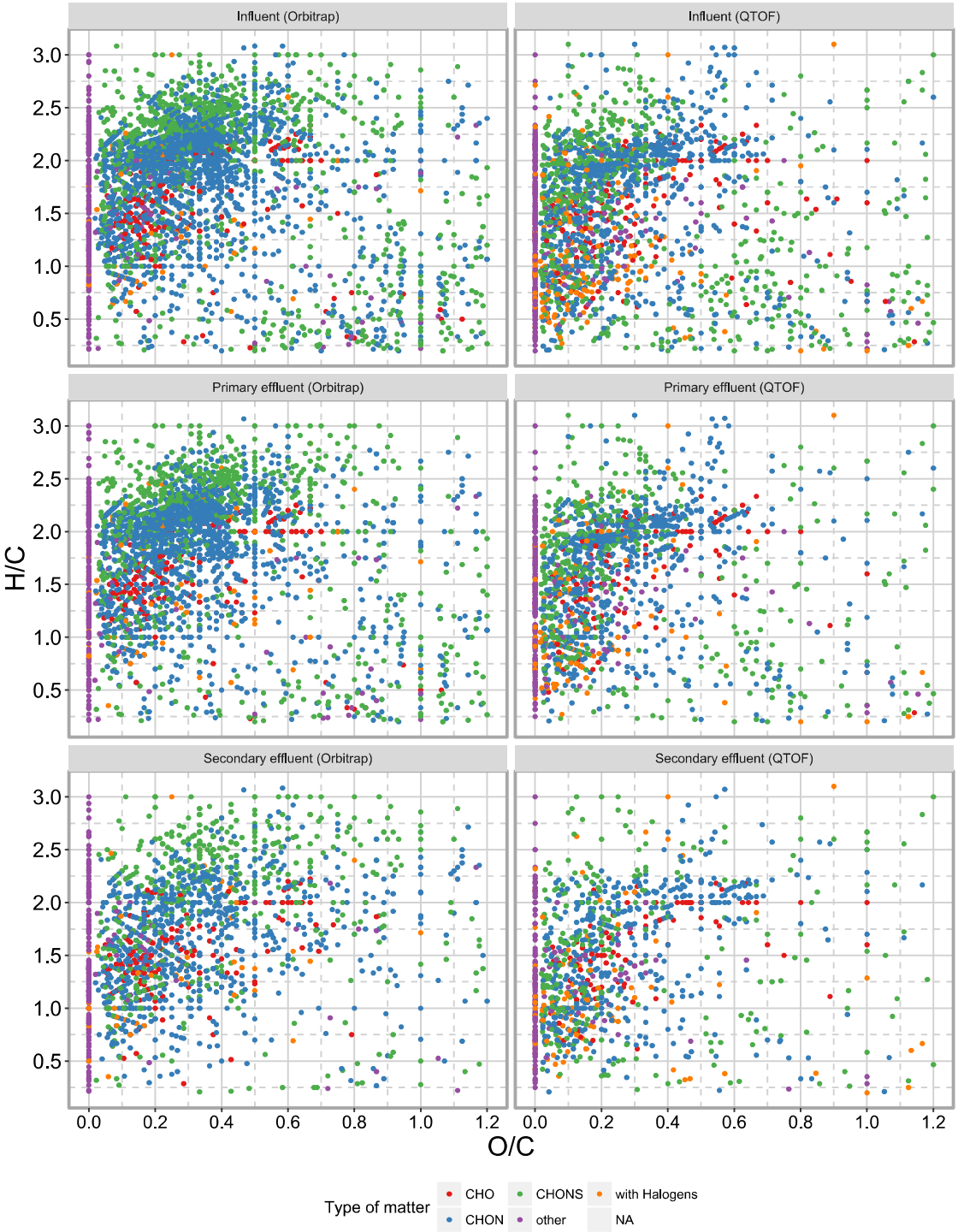



**Figure 4.13 Van Krevelen plot of Orbitrap-MS and QTOF-MS molecular features. The color indicates the elemental composition, where "other" signifies the not explicitly mentioned combinations of elements C, H, O, N, S.**

The van Krevelen plot revealed more halogenated compounds in Orbitrap data. The percentage of detected halogenated substances was not independent of LC-MS setup according to a Chi-squared independence test (for example, comparing influent $X^2 = 202.93$, degrees of freedom = 4, $p < 0.001$) indicating that the fraction was indeed more detectable by the QTOF setup.

QTOF data have shown a lower resolution but a higher intensity of *m/z* signals compared to Orbitrap in the given experimental setup. The detection of chlorinated substances depends on the detection of the M+1 isotopologue which in small molecules has a lower intensity than the monoisotopic M signal. Such signals in Orbitrap were disappearing more often under the limit of detection compared to QTOF where the stronger signal lead to a better molecular feature recognition, under the applied experimental conditions.

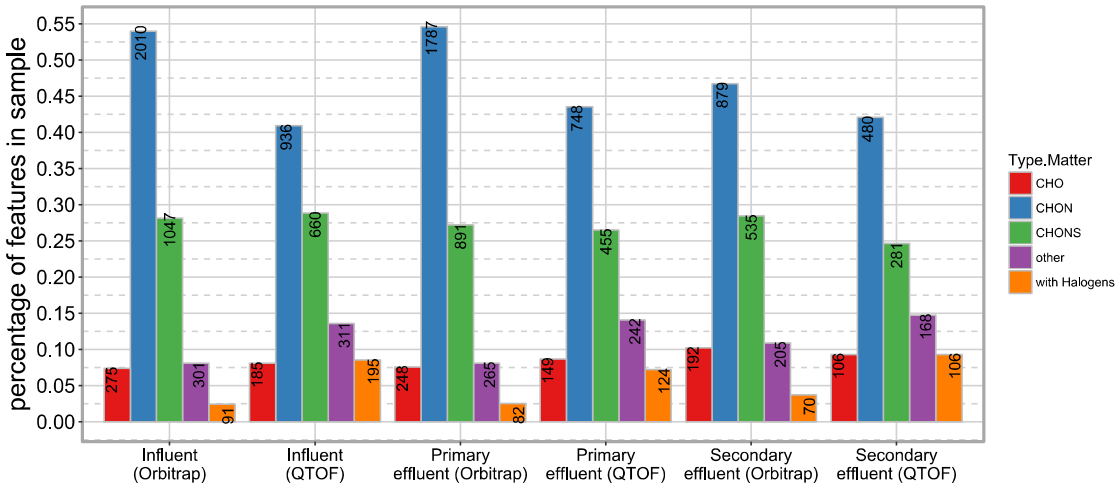



**Figure 4.14 Type of matter for detected features with identified elemental composition in Orbitrap-MS and QTOF-MS data. Absolute number of features is indicated within the bars.**

## 4.3 Suspect screening

### 4.3.1 Transformation of micro-contaminants supporting non-targeted analysis

The confirmation with $MS^2$ yielded 40 compounds in 11 elimination series of parents and TPs (Figure 4.15 and the remaining profiles in Appendix 2 Figure 7.2 - Figure 7.4). The series contain 8 tentative TPs without a confirmed structure but included into the analysis due to a set of restrictive selection rules described in . For the remainder of substances, the structures were compared by either available IS or $MS^2$ fragments (Appendix 2 Figure 7.5, Figure 7.6).

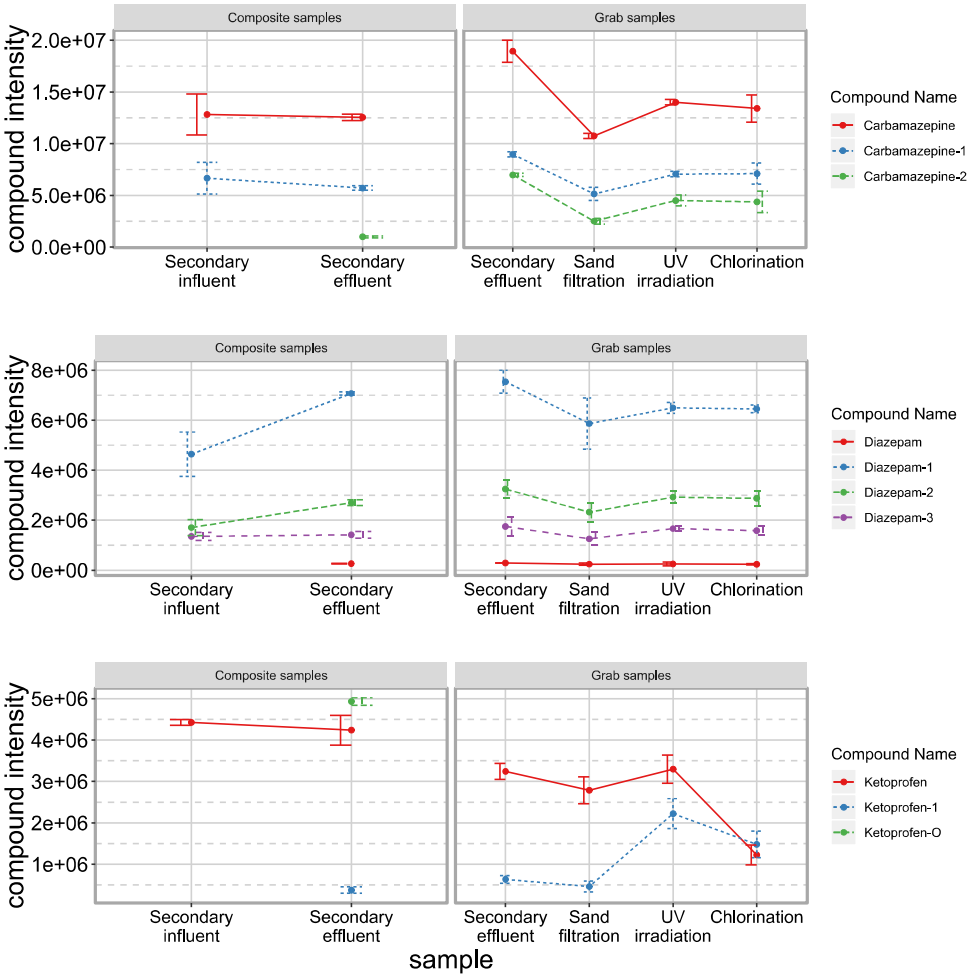



**Figure 4.15 Removal profiles of corrected intensity for Carbamazepine, Diazepam, Ketoprofen, and their confirmed and tentative TPs in a multi-stage wastewater treatment system. Each series depicts two profiles divided due to the sampling mode.**

Secondary treatment reduced the intensity of 6 parents, while 2 parent compound increased in intensity. Some parents, e.g. 17β-estradiol, have shown a lower intensity in the secondary influent samples compared to secondary effluent. The higher intensity in effluent due to a matrix suppression is not the responsible factor regarding this trend after an intensity correction using IS and considering that the intensity increase was not observed for most parent compounds. The lower influent intensity can indicate a release of substances during secondary treatment, for example from colloids or suspended particles. A release of parent compounds in the effluent from conjugates could also be responsible for the intensity increase in secondary effluent, as was previously shown for Carbamazepine [241]. At the same time, the matrix effect cannot be neglected for samples with low concentration approaching the limit of detection, since these will be simply not detected and are not subject to a posterior intensity correction. Here, 2 parents were not detected in the secondary influent, but appeared in the secondary effluent.

At the same time, the secondary treatment reduced the intensity of 10 confirmed TPs which were found already in the influent. Additionally, 8 TPs appeared in the secondary effluent. In case of the 8 tentative TP suspects, 7 were present already in the influent and 4 of them were abated in course of secondary treatment. Ketoprofen-O appeared in secondary treatment. It was not observed in the grab sample study.

7 parent compounds were abated in the tertiary treatment, while 3 showed an intensity increase. 17 TPs have shown an intensity decrease in the tertiary treatment, while the intensity of 6 TPs increased after the treatment. Tertiary treatment decreased the intensity of 5 tentative TPs while 2 TPs have shown an intensity increase. Overall, the tertiary treatment rather removes micro-contaminants with 73 % of observed compounds showing an intensity reduction. As described previously, a removal for micro-contaminants was not necessarily expected for tertiary treatment, since the task is rather to remove pathogens and various biological agents. Despite that, many compounds have shown significant intensity changes in course of the tertiary treatment. A pattern is observed for the majority of compounds within the steps secondary effluent – sand filtration – UV irradiation. Here, an intensity drop is observed in sand filtration effluent which recovers in UV effluent. This behavior can be caused by treatment processes for TPs, yet it is peculiar to observe it for parent compounds.

Possible reasons were explored of the intensity drop in sand filtration. A manual check of surrogate standard intensities for sample triplicates confirmed a successful SPE. A bad integration or omission of signals during data extraction in Exact Finder was avoided by a manual check of all peaks. Also, the intensity correction with IS did not introduce the trend observed in elimination series.

An addition of too much or too little IS is possible, but the comparison of IS intensities in sample replicates and exclusion of replicates which introduce a high CV reduce the possibility. The matrix effect for effluent samples is comparable and mediated by IS intensity correction so the intensity drop in sand filtration is not caused by ion suppression.

The intensity drop might be explained by variations introduced by grab sampling. Yet, another grab sampling series 24 h after the analyzed one shows comparable intensity profiles, thus reducing the possibility of strong intensity deviations.

The average intensities of DOM features were compared among the samples to see if the intensity shifts of micro-contaminants described above are systematic on DOM level. A non-targeted signal extraction was performed for the DOM observed in the analysis using the same LC-HRMS data. Due to the non-targeted analysis of the intensity profile of DOM, there is a systematic intensity shift for DOM features.

The protocol followed the extraction described in section 3.3.1.4.1 with exception of allowing signals without an isotopic pattern, since the main purpose was to extract molecular feature intensity and not chemical information. The non-targeted DOM transformation profile shows a pattern similar to the suspect screening with an intensity drop in sand filtration effluent, a strong increase of intensity in UV irradiation effluent and then an intensity reduction in the chlorination step (Figure 4.16).

To confirm the significance of these results for observed DOM, a one-way ANOVA test was performed checking differences between the average logarithmic sample intensities ($anova_{log10(Intensity)}$: $F(3, 111436) = 27.13$, $p < 0.001$). The test confirmed significant intensity differences except for the sample pair of secondary effluent and sand filtration effluent (Figure 4.17).

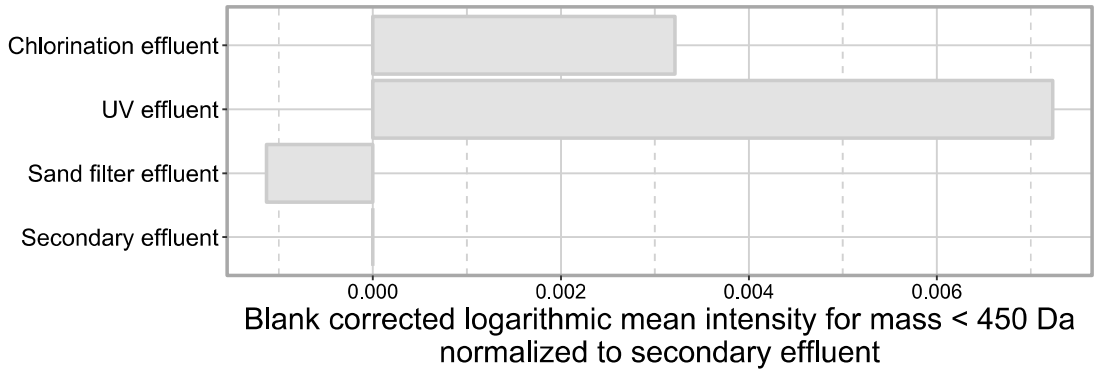



**Figure 4.16 Comparison of average $log_{10}$-transformed intensities of DOM molecular features from the same LC-MS measurement as micro-contaminant suspect screening. The intensities were normalized to the secondary effluent intensity for better comparison.**

However, there is a significant increase in DOM intensity in UV and chlorination effluents compared to the sand filtration effluent which is comparable with the trend in suspect screening. No correction of signal intensity was performed for the molecular features in the non-targeted analysis. Therefore, some aberrations due to the matrix-assisted ion suppression were expected. Caution has to be exercised when using non-corrected data, as shown in section 4.1.6, yet they still provide conclusive results on a statistical level.

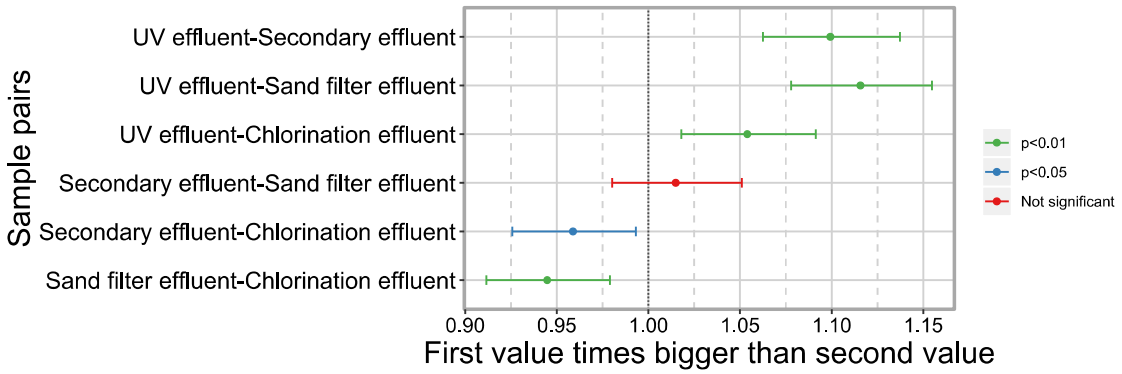



**Figure 4.17 Tukey honest significant differences of average logarithmic intensity with a confidence interval of 95 % for the pairs of samples' used in the ANOVA testing. The x-axis describes the by how many times the mean intensity of the first sample is larger than the mean intensity of the second sample. The color shows the significance level of difference.**

Reasons for the systematic behavior of elimination profiles were explored on the level of WWTP operation and sampling. There are some uncertainties in the sampling before and after the sand filter due to the backwashing applied at this stage of WWTP. The backwashing is executed with sand filter effluent water which reduces the risk of wrong intensity measurement. Regardless, backwashing does not decrease contaminants' concentrations but keeps them at a similar level. However, backwashing might affect the hydraulic retention time in the system. The sampling was according to the theoretical residence time yet a tracer test was not run.

Another possible reason for intensity decrease after sand filtration can be the excessive age of the sand filter. Long-time operation sand filters develop a biofilm able to influence the transformation of DOM (compare to the chapter 4.3.2 on micro-contaminant removal in a BAC filter). With a bioactive sand filter, there is a possibility of conjugate composition with decomposition following a change in pH, temperature, redox conditions in subsequent steps. In addition, an adsorption on colloids in the suspended solids of a saturated sand filter and a release in subsequent steps is possible.

Low uncertainty was foreseen for the grab sampling before and after the UV unit, due to the short hydraulic retention time with an average wastewater flow of 312 $m^3 \cdot h^{-1}$ at a distance of 3 m between the sampling points. The UV-lamps in the UV irradiation treatment step exceeded their lifetime leading to a reduced performance, yet not to an intensity increase. Additionally, the observed concentration increase for micro-contaminants at this treatment stage might be explained by the addition of 5 mL 3 % $HNO_3$.

It is dosed every 15 minutes for 70 s with a flow rate of 0.74 $m^3 \cdot h^{-1}$ and using two pumps after the sand filtration sampling point and before the UV sampling point. A change of pH can change and affect the concentrations of micro-contaminants, for example releasing substances attached to the lamp. This possibility cannot be neglected while keeping in mind that the flow of acid is small compared to the wastewater flow.

Some uncertainties exist for the sampling of chlorination effluent. Yet, the variability at the end of such WWT system is low compared, for example to the WWTP influent, so uncertainties are low as well [242].

### 4.3.2 Test of application in a novel BAC-UF treatment reactor

Elimination of pharmaceuticals and their TPs was explored in a novel BAC-UF pilot wastewater treatment to examine the value of the presented suspect screening for emerging treatment technologies. The current setup had two challenges for micro-contaminant analytics. The influent of BAC had already been treated by secondary biological treatment making the elucidation of TPs appearing through metabolisms in BAC challenging. Also, the recirculation of UF-centrate into the BAC influent complicated the balancing of micro-contaminants in the system.

25 Individual compounds in 11 series showed varying elimination profiles and no strong correlation of removal in BAC-UF on molecular properties was found (Figure 4.18 and Figure 4.19). Additional removal profiles and the structural formulae of confirmed compounds can be found in Appendix 3.

The positive charge of Atenolol at wastewater pH and strong adsorption to carbon granules is responsible for the strong rejection on the negatively charged UF membrane at Day 0 and Day 229 by means of adsorption to the membrane [98]. At Day 269 an increase of intensity after UF treatment indicates a breakthrough of micro-contaminants and its TP.

Carbamazepine shows a behavior comparable to Atenolol throughout the samples due to the similar adsorption capacity and the positive charge [88]. The low intensity of Carbamazepine-1 hints at its removal already in the preceding secondary biological treatment.

Iopromide and Iopromide-1 displayed an intensity decrease after BAC-UF treatment. In case of Iopromide-1, it was difficult to make estimations due to the low intensity of the compound probably metabolized already in secondary treatment.

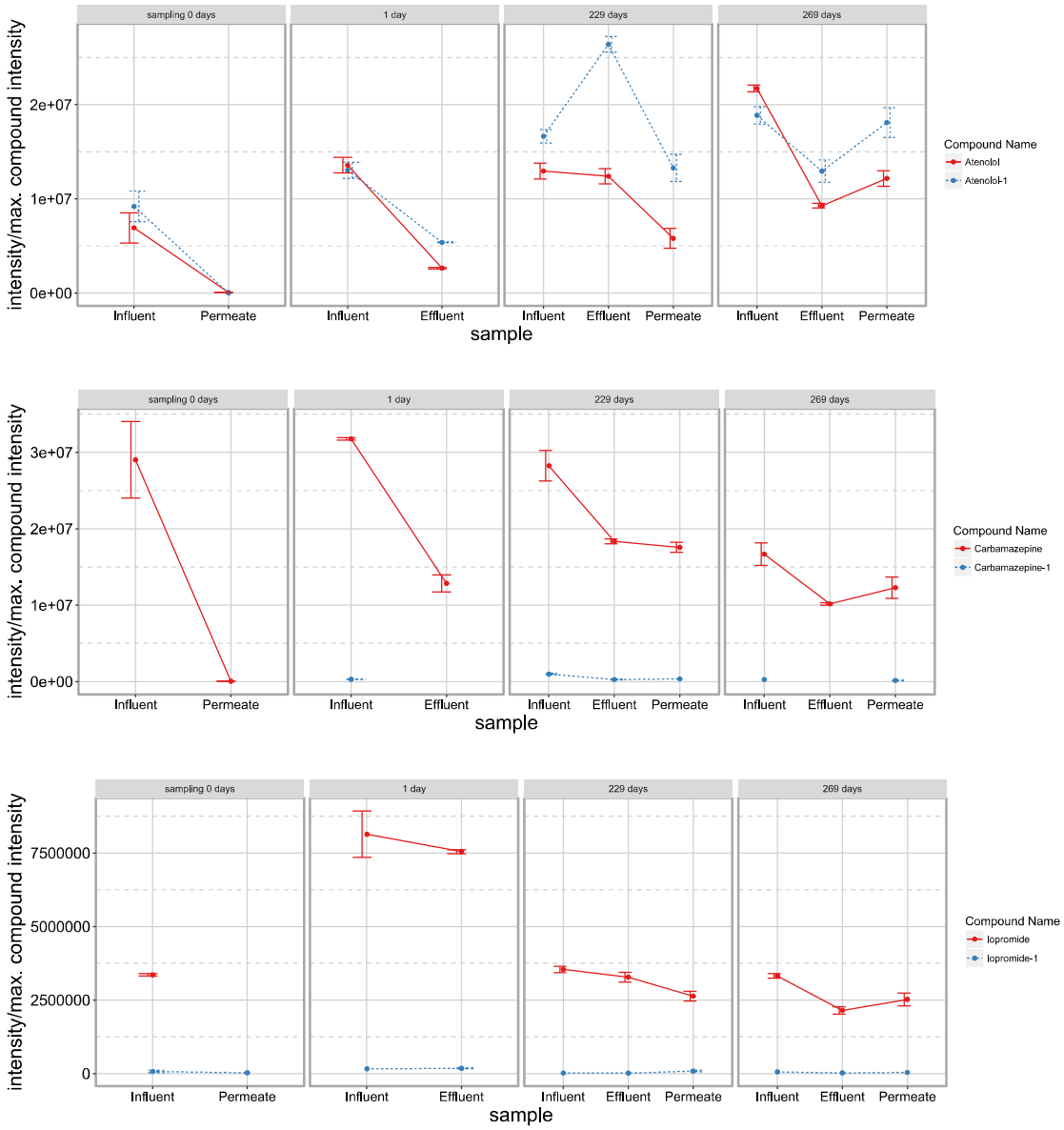



**Figure 4.18 Elimination profiles of corrected intensity for parents and TPs of Atenolol, Azithromycin, Bezafibrate, and Iopromide at 4 sampling dates and 3 sampling sites of the BAC-UF pilot treatment system.**

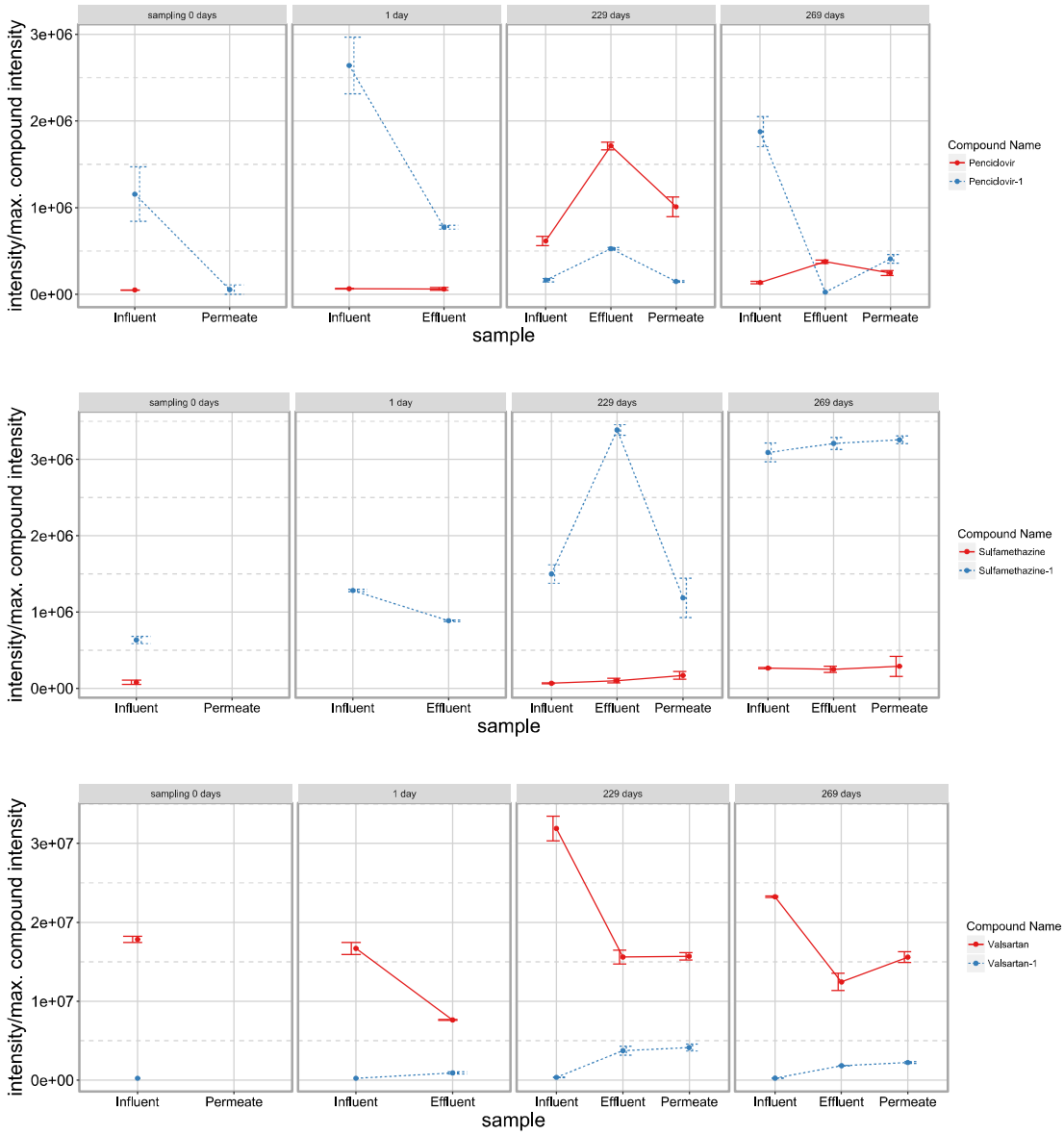



**Figure 4.19 Elimination profiles of corrected intensity for parents and TPs of Penciclovir, Sulfamethazine, and Valsartan at 4 sampling dates and 3 sampling sites of the BAC-UF pilot treatment system.**

Sulfamethoxazole showed a low removal in BAC-UF reactor, expectedly due to its negative charge and low log P which decrease its adsorptive capacity and an intensity drop in BAC effluent. However, Sulfamethazine-1 with one hydroxy group added compared to the parent displayed an inverse removal profile on Day 229 which demonstrates the complexity of micro-contaminant analysis even for structurally similar compounds.

Valsartan-1 intensity increased after BAC treatment at Day 1, Day 229, and Day 269. Such consistent increase is an indication of removal of Valsartan in BAC filter by metabolic activity. However, given that the effect was not observed for other TPs this behavior has to be explored in the future.

The analysis uncovered a more efficient removal of parents and TPs in BAC-UF at Day 0 and Day1 compared to Day 229 and Day 269 when the BAC filter bed was saturated. A clogging of saturated filtering systems leads to breakthrough of micro-contaminants or leaching of micro-contaminants from a saturated filter bed. BAC-UF decreased the intensity of most compounds at Day 229 and Day 269 when BAC influent and UF effluent are compared. However, compounds in groups of Atenolol, Bezafibrate, Penciclovir, and Iopromide show an intensity increase after BAC treatment at Day 229 and a drop at Day 269 which reverses back to the initial intensity after UF treatment.

The pore size of UF membrane allows micro-contaminants to pass through unhindered. However, it was previously speculated that membranes can grow a biofilm that influences the rejection of organic matter by steric and electrostatic interactions [243]. A negative UF membrane covered by a biofilm might explain the found intensity recovery after UF treatment in this study. The increase of intensity in MS due to a decrease of the matrix effect in UF effluent samples cannot be neglected either, despite an intensity correction by IS.

Despite this possibility, such intensity changes have to be caused by the treatment itself, considering that the diametric transformation profiles occur in sample sets with the comparable matrix at Day 229 and Day 269. No recorded special procedure, as backwashing, etc., was applied to the reactor between the two sampling dates. In lack of special events, the difference in behavior is explained by two factors: 40 days between sampling dates led to additional accumulation of OM in BAC filter and raised suspended solids in BAC effluent. The suspended solids could then act as carriers for adsorbed chemicals to be released downstream. A conjugation of compounds due to metabolic activity of BAC biofilm and latter release downstream might have accounted for the difference between profiles at two dates as well.

## 4.4 Value of pre-concentration in non-targeted analysis

The influence of pre-concentration was explored by comparing the results of non-targeted analysis of directly injected samples from Section 4.1 and the multiphase pre-concentrated samples in Section 4.2. Hereby the influent and effluent of secondary treatment were compared.

The number of detected features was higher in the analysis of pre-concentrated samples. However, here the rate of successful formula assignment was lower. Without pre-concentration, the data extraction is able to recognize signals with higher intensity while omitting weak signals disappearing into the instrumental noise. Another indication of this effect is the distribution of absolute intensity ratio when comparing intensities of secondary effluent to WWTP influent.

Recalcitrant features of secondary treatment of pre-concentrated samples show effluent-to-influent intensity ratio distribution with more density at low values compared to directly injected DOM. Having a higher number of lower values corresponds to having more features displaying a stronger before-after difference in intensity of a chemical.

This implies that the LC-MS acquisition and data extraction of pre-concentrated samples unveiled low-intensity molecular features in the influent which then were well removed in secondary treatment.

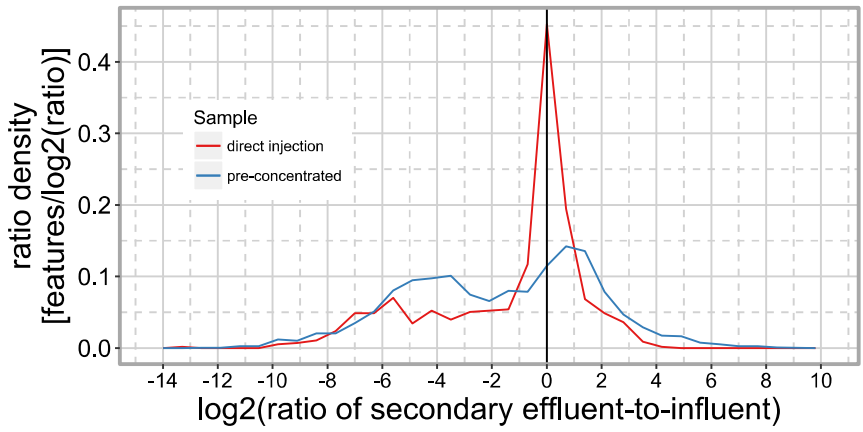



**Figure 4.20 Distribution of the logarithmic ratio secondary effluent-to-influent of absolute intensity for recalcitrant DOM features with directly injected and multiphase pre-concentrated samples.**

This underestimation of low-intensity features in directly injected samples can be both an advantage and a drawback. The detected features are high-impact chemicals with a strong signature in case of direct injection, hence a lower chance of error when it comes to formula assignment and intensity calculation. On the other hand, omitting weak signals neglects micro-contaminants and TPs with low concentration which were the focus of the methodology.

The author is inclined toward direct injection due to the simpler and cheaper preparative procedure. Sample pre-concentration has shown previously to reduce the matrix-directed ion suppression. A matrix correction algorithm was not applied in this study due to the difficulty to predict the physiochemical properties of individual DOM constituents in the non-targeted analysis.

This leads to the assumption that the observed intensity increase of recalcitrant features after secondary treatment was at least partially accounted for by the matrix effect. Since the effect was observed in both directly injected and pre-concentrated data, a better matrix-suppression reduction could not be assumed. For this reason, the author is in favor of direct injection until a more robust matrix effect correction is developed.

The trends in changes of mass fractions as well as the increase of unsaturation by observing DBE-O was detected within both sets of samples. The van Krevelen plot has shown a similar distribution of features in both cases with a large part of CHONS matter residing at $H/C > 2$ and CHON matter at $2.2 > H/C > 1.0$.

Caution has to be exercised when comparing the two analyses due to the different location, grab sampling, a different chromatographic method, and acquisition in PI mode only in the case of pre-concentrated samples. Despite these differences, it is remarkable that many outputs coincided with each other, giving promise to development of a generic fingerprinting mechanism for WWT and transformation of DOM.

# 5 GENERAL CONCLUSIONS

## 5.1 Summary

### 5.1.1 Non-targeted analysis of multi-stage WWTP

The first chapter of this thesis (Chapter 4.1) dealt with the building, testing, and application of non-targeted LC-MS data extraction and analysis of directly injected samples to fingerprint a multi-stage wastewater treatment system.

Various quality control tests were run to assure meaningful data in the non-targeted analysis. The precision and chromatographic range of DOM signals were evaluated using a set of standards with varying physio-chemical properties and *m/z* of all detected signals were corrected using a linear model based on *m/z* of standards. Ion chromatogram extraction and molecular formula prediction within the non-targeted data extraction protocol are complex steps. They were therefore evaluated prior to analyzing the data. Artifact and noise signals were removed in a series of filtering steps in MZmine and R, such as constricting the CV among replicates to values lower than 30 %. The applied molecular formula prediction settings were evaluated with a set of pharmaceuticals and resulted in a yield of 86 % correct formulae. The formula prediction in HRMS relies heavily on isotopic pattern prediction. Therefore, only molecular features with a found isotopic pattern were evaluated in this workflow. This step removes a large complement of molecular features, but safeguards against a wrong formula assignment and noise peaks.

Significant changes of DOM were identified. Secondary treatment removed 1617 of 2409 (67%) detected molecular features while 255 of 1047 (24%) features appeared after secondary treatment. Subsets of molecular features within the treatment stages exhibited significant changes. The unsaturation increased in secondary treatment by 1.5 DBE-O in the bulk DOM and by 4.0 DBE-O comparing the subsets of disappeared and appeared molecular features.

The van Krevelen plots indicated highly oxygenated, unsaturated matter and highly aliphatic, nitrogen-rich matter. The distribution of features did not follow the established regions for the NOM and was attributed to anthropogenic compounds found in wastewater. Biotransformation reaction pairs were explored for simple structural moieties allowing investigation of the properties of individual molecular features. The comparison with literature for analogous moieties has shown that the fractions of discovered reaction pairs are correlating with the previous findings, despite the total number of pairs being lower in our analysis.

KMD plots uncovered 31 prioritized, abundant homologous series for the -$CH_2$- moiety. The fraction of -$CH_2$- series exhibited a preferential removal of homologs at high masses. On the other hand, 38 -$C_2H_2O$- series were largely removed, suggesting complex behavior of heterogeneous wastewater DOM. Altogether, the presented not-targeted analysis combined various existing techniques to analyze wastewater. It explored the bulk properties of wastewater LC-HRMS data and the subsets of features within the samples.

The approach described a chemical fingerprint of DOM transformation in wastewater. However, the analysis is still limited by the scope of LC-MS detection, signal suppression in the samples and the precision of data extraction.

### 5.1.2 LC-MS setup influence in non-targeted analysis

The second chapter (Section 4.2) explored the resilience of the developed non-targeted analysis of wastewater DOM with respect to the choice of the HRMS-setup used to acquire data. Hereby, LC-QTOF and LC-Orbitrap, which are both widely used setups in environmental analytics, were compared. Both LC-MS setups delivered robust results with respect to the developed non-targeted methodology. Orbitrap has shown higher accuracy and resolving power. For example, the non-targeted analysis with the Orbitrap data showed a higher amount of detected molecular formulae and a higher rate of molecular formula assignment. QTOF was prolific in including a subset of higher molecular weight features > 400 Da due it to the stronger stability of resolving power of QTOF at high *m/z* of a signal. In addition, QTOF encompassed a larger mass range for the detected complement of molecular features.

Both Orbitrap and QTOF data were able to identify a reduction of average mass after secondary treatment as well as the preferential disappearance of heavy DOM constituents and appearance of DOM with mass 100-400 Da.

The increase of unsaturation expressed by DBE-O was detected in both sets of data. However, Orbitrap data demonstrated a lower DBE-O in all samples compared to QTOF implying the detection of a slightly different complement of DOM. This peculiarity was also observed in the van Krevelen plot.

In the van Krevelen plot, Orbitrap samples showed more molecular features in the region of saturated DOM at H/C > 2 attributed to features with CHON and CHONS composition and molecular weight > 600 Da. QTOF outperformed Orbitrap in the detection of chlorinated compounds due to the better prediction of isotopic pattern with a higher intensity of M+1 isotope in QTOF samples in the given experimental setup.

### 5.1.3 Comparison of non-targeted and suspect analyses

The suspect screening for pharmaceuticals of interest and their TPs in a multi-phase WWTP revealed a strong removal of micro-contaminants in the secondary treatment and a less pronounced one in tertiary treatment. The stages of tertiary treatment did not follow a monotonous removal pattern which was also confirmed for the DOM of the pre-concentrated samples using a non-targeted DOM signal extraction.

Suspect screening remains the procedure to look at micro-contaminants of interest despite the ability of non-targeted analysis to recognize DOM transformations in WWT. Pharmaceuticals and their TPs discovered here in a suspect screening were compared to analogous signals in not pre-concentrated non-targeted data. The main discovered challenge was the low intensity of compounds in non-targeted data. For example, the intensity of Valsartan in secondary effluent was > 100 times lower in the non-targeted analysis that did not apply a pre-concentration step.

The suspect screening was performed with HLB SPE pre-concentrated samples which narrows down the extracted complement of DOM leading to the reduction of the matrix effect. The reduction of the matrix improved the response in MS increasing the micro-

contaminant intensity and allowing better peak recognition compared to signals in the non-targeted analysis.

In addition, the correction of the matrix effect in suspect screening data in comparison to injected IS improved the assessment of intensity shifts among samples. At the same time, intensity correction is much more straightforward in a suspect screening than in non-targeted screening.

It is easy to pick a proxy for a structurally related compound as IS in suspect screening, but it is difficult to find a set of proxies to cover the properties of all DOM constituents in non-targeted screening. While not described here, the author explored intensity correction in the non-targeted analysis by using native MZmine normalization algorithms and own scripts to correct spectral ranges corresponding to logarithmic partition coefficient (log P) or retention time range based on multiple IS. A satisfactory model of intensity correction was not found.

An additional test of the suspect screening methodology was performed for a novel BAC-UF pilot WWTP. The suspect screening for pharmaceuticals of interest and their TPs revealed micro-contaminant elimination differences between the phases of not-saturated and saturated BAC filter. BAC-UF showed a higher removal of micro-contaminants with the unsaturated filter. In the latter runtime phase, BAC-UF showed a lower removal and in some cases a release of micro-contaminants explained by desorption of micro-contaminants from the biomass in the reactor and the release of substances from conjugates down the treatment line.

The suspect screening was able to follow the removal patterns throughout time and treatment stages and the results correlated with various expected effects in the system, e.g. clogging of the filter after a long operational period.

## 5.2 Outlook

Non-targeted analysis drives the exploration of DOM transformation in wastewater gradually displacing the interest in targeted methods in the researcher community. This shift is explained by the development of user-friendly yet powerful computational data processing tools for LC-MS analysis of wastewater DOM. Yet, many researchers still shy away from the non-targeted analysis of wastewater DOM, owing to the fact that it remains a complex procedure with various challenges.

For example, the acquisition of LC-MS data in non-targeted screening continues to limit the scope of the analysis. The scope of detectable compounds is limited by the sample preparation, chromatographic column, ionization chamber (in case of LC), and resolution/MS specification of the detector. Multiphase SPE was explored in this thesis and provided satisfactory data for non-targeted analysis. However, a multiphase SPE has to be carefully considered due to the preparative complexity. Novel ways of pre-concentrating and isolating relevant DOM fractions have to be researched. Chromatographic separation has to be explored in the future by adding a separation step with a HILIC column. This will widen the number of detected chemicals and help to uncover polar substances, such as hydrolyzed metabolites and TPs.

In addition, the influence of the MS-setup on the output of non-targeted analysis was detected in this thesis. Exploring these differences by comparing additional spectrometer setups will further limit the error of untargeted analysis.

Another challenge of non-target screening as presented here is the extraction of LC-MS data. In particular, peak recognition and assignment of elemental composition to molecular features pose difficulties. Partially, the challenge of extracting molecular features is limited by the physio-chemical properties of LC-MS analysis itself. For example, the matrix-induced ion suppression leads to the intensity reduction of signals or a challenging chromatography can cause tailing and cleaved peaks.

However, the improvement of algorithms in LC-MS data treatment will improve the yield of molecular features. Therefore, novel LC-MS data toolboxes from both commercial (as

Compound Discoverer® from Thermo) and open license (as Mass Spectrometry Development Kit) sources have to be continuously explored and compared for their capabilities, additionally to the MZmine-based methodology developed in this thesis.

The matrix-assisted ion suppression in MS remains one of the main challenges of the non-targeted methodology compared to the targeted or suspect screening. Hitherto it was not answered suitably in the literature despite many attempts, due to the author's best knowledge. The correction of the matrix effect requires the development of a range of IS that mimics the physio-chemical properties of wastewater DOM in a representative fashion.

The IS applied in this thesis were largely based on pharmaceutical compounds, however in the future, the IS should also include IS related to NOM (for example tryptophan, etc.) and highly polar compounds mimicking the TPs of DOM. Due to the cost of isotopically labeled IS this task will be approached with great care in the future. Moreover, a correction of the intensity relies on intelligent algorithms which have to be adjusted to the analyzed data. Therefore, algorithms, for example with averaged IS intensity vs. chromatographic time-weighted IS intensity, will need to be compared to estimate the best outcome for wastewater DOM.

The non-targeted analysis will become a powerful tool in the analysis of the transformation of wastewater DOM. However, the right application and fitting case studies have to determine the application of the analysis. In this thesis, the non-targeted analysis has shown notable transformations of DOM in secondary biological treatment.

The author's hope is that this type of analysis, assisted by a suspect screening as applied in this thesis as well, will be used to pinpoint changes in small molecules for novel technologies. In particular, technologies, which depend strongly on the transformation of small molecules, will benefit from a non-targeted DOM analysis. For example, membrane-filtering technologies as the reverse osmosis or nano-filtration need a way to discern small molecule fractions while looking at the bulk properties, since these can cast light on the chemical-biofilm-membrane interactions in the system. In the future, such systems shall be explored more intensively with respect to non-targeted analysis and the accompanying suspect screening for transformations.

# 6 LITERATURE

(1) European Chemicals Agency: Registered substances https://echa.europa.eu/information-on-chemicals/registered-substances (accessed Jan 9, 2018).

(2) Michael-Kordatou, I.; Michael, C.; Duan, X.; He, X.; Dionysiou, D. D.; Mills, M. A.; Fatta-Kassinos, D. Dissolved Effluent Organic Matter: Characteristics and Potential Implications in Wastewater Treatment and Reuse Applications. *Water Res.* **2015**, *77*, 213–248.

(3) Richardson, S. D.; Postigo, C. Discovery of New Emerging Dbps by High-Resolution Mass Spectrometry. In *Comprehensive Analytical Chemistry*; Elsevier, 2016; Vol. 71, pp 335–356.

(4) Daughton, C. G. Non-Regulated Water Contaminants: Emerging Research. *Environ. Impact Assess. Rev.* **2004**, *24* (7–8), 711–732.

(5) Le-Minh, N.; Khan, S. J.; Drewes, J. E.; Stuetz, R. M. Fate of Antibiotics during Municipal Water Recycling Treatment Processes. *Water Res.* **2010**, *44* (15), 4295–4323.

(6) Brieudes, V.; Lardy-Fontan, S.; Vaslin-Reimann, S.; Budzinski, H.; Lalere, B. Development of a Multi-Residue Method for Scrutinizing Psychotropic Compounds in Natural Waters. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **2017**, *1047*, 160–172.

(7) Stuart, M. E.; Lapworth, D. J.; Thomas, J.; Edwards, L. Fingerprinting Groundwater Pollution in Catchments with Contrasting Contaminant Sources Using Microorganic Compounds. *Sci. Total Environ.* **2014**, *468–469*, 564–577.

(8) Besha, A. T.; Gebreyohannes, A. Y.; Tufa, R. A.; Bekele, D. N.; Curcio, E.; Giorno, L. Removal of Emerging Micropollutants by Activated Sludge Process and Membrane Bioreactors and the Effects of Micropollutants on Membrane Fouling: A Review. *J. Environ. Chem. Eng.* **2017**, *5* (3), 2395–2414.

(9) Siegrist, H.; Joss, A. Review on the Fate of Organic Micropollutants in Wastewater Treatment and Water Reuse with Membranes. *Water Sci. Technol.* **2012**, *66* (6), 1369–1376.

(10) Jones, V.; Gardner, M.; Ellor, B. Concentrations of Trace Substances in Sewage Sludge from 28 Wastewater Treatment Works in the UK. *Chemosphere* **2014**, *111*, 478–484.

(11) Deblonde, T.; Cossu-Leguille, C.; Hartemann, P. Emerging Pollutants in Wastewater: A Review of the Literature. *Int. J. Hyg. Environ. Health* **2011**, *214* (6), 442–448.

(12) Comber, S.; Gardner, M.; Jones, V.; Ellor, B. Source Apportionment of Trace Contaminants in Urban Sewer Catchments. *Environ. Technol.* **2015**, *36* (5), 573–587.

(13) Díaz, R.; Ibáñez, M.; Sancho, J. V.; Hernández, F. Target and Non-Target Screening Strategies for Organic Contaminants, Residues and Illicit Substances in Food, Environmental and Human Biological Samples by UHPLC-QTOF-MS. *Anal Methods* **2012**, *4* (1), 196–209.

(14) Hernández, F.; Ibáñez, M.; Botero-Coy, A.-M.; Bade, R.; Bustos-López, M. C.; Rincón, J.; Moncayo, A.; Bijlsma, L. LC-QTOF MS Screening of More than 1,000 Licit and Illicit Drugs and Their Metabolites in Wastewater and Surface Waters from the Area of Bogotá, Colombia. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6405–6416.

(15) Jobst, K. J.; Shen, L.; Reiner, E. J.; Taguchi, V. Y.; Helm, P. A.; McCrindle, R.; Backus, S. The Use of Mass Defect Plots for the Identification of (Novel) Halogenated Contaminants in the Environment. *Anal. Bioanal. Chem.* **2013**, *405* (10), 3289–3297.

(16) Llorca, M.; Lucas, D.; Ferrando-Climent, L.; Badia-Fabregat, M.; Cruz-Morató, C.; Barceló, D.; Rodríguez-Mozaz, S. Suspect Screening of Emerging Pollutants and Their Major Transformation Products in Wastewaters Treated with Fungi by Liquid Chromatography Coupled to a High Resolution Mass Spectrometry. *J. Chromatogr. A* **2016**, *1439*, 124–136.

(17) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098.

(18) Moschet, C.; Piazzoli, A.; Singer, H.; Hollender, J. Alleviating the Reference Standard Dilemma Using a Systematic Exact Mass Suspect Screening Approach with Liquid Chromatography-High Resolution Mass Spectrometry. *Anal. Chem.* **2013**, *85* (21), 10312–10320.

(19) Krauss, M.; Singer, H.; Hollender, J. LC–high Resolution MS in Environmental Analysis: From Target Screening to the Identification of Unknowns. *Anal. Bioanal. Chem.* **2010**, *397* (3), 943–951.

(20) Richardson, S. D.; Ternes, T. A. Water Analysis: Emerging Contaminants and Current Issues. *Anal. Chem.* **2014**, *86* (6), 2813–2848.

(21) Brezina, E.; Prasse, C.; Meyer, J.; Mückter, H.; Ternes, T. A. Investigation and Risk Evaluation of the Occurrence of Carbamazepine, Oxcarbazepine, Their Human Metabolites and Transformation Products in the Urban Water Cycle. *Environ. Pollut.* **2017**, *225*, 261–269.

(22) Causanilles, A.; Kinyua, J.; Ruttkies, C.; van Nuijs, A. L. N.; Emke, E.; Covaci, A.; de Voogt, P. Qualitative Screening for New Psychoactive Substances in Wastewater Collected during a City Festival Using Liquid Chromatography Coupled to High-Resolution Mass Spectrometry. *Chemosphere* **2017**, *184*, 1186–1193.

(23) Wode, F.; van Baar, P.; Dünnbier, U.; Hecht, F.; Taute, T.; Jekel, M.; Reemtsma, T. Search for over 2000 Current and Legacy Micropollutants on a Wastewater Infiltration Site with a UPLC-High Resolution MS Target Screening Method. *Water Res.* **2015**, *69*, 274–283.

(24) Nürenberg, G.; Schulz, M.; Kunkel, U.; Ternes, T. A. Development and Validation of a Generic Nontarget Method Based on Liquid Chromatography – High Resolution Mass Spectrometry Analysis for the Evaluation of Different Wastewater Treatment Options. *J. Chromatogr. A* **2015**, *1426*, 77–90.

(25) Maizel, A. C.; Remucal, C. K. The Effect of Advanced Secondary Municipal Wastewater Treatment on the Molecular Composition of Dissolved Organic Matter. *Water Res.* **2017**, *122*, 42–52.

(26) Wang, B.; Wan, Y.; Zheng, G.; Hu, J. Evaluating a Tap Water Contamination Incident Attributed to Oil Contamination by Nontargeted Screening Strategies. *Environ. Sci. Technol.* **2016**, *50* (6), 2956–2963.

(27) Bade, R.; Rousis, N. I.; Bijlsma, L.; Gracia-Lor, E.; Castiglioni, S.; Sancho, J. V.; Hernandez, F. Screening of Pharmaceuticals and Illicit Drugs in Wastewater and Surface Waters of Spain and Italy by High Resolution Mass Spectrometry Using UHPLC-QTOF MS and LC-LTQ-Orbitrap MS. *Anal. Bioanal. Chem.* **2015**, *407* (30), 8979–8988.

(28) Remucal, C. K.; Cory, R. M.; Sander, M.; McNeill, K. Low Molecular Weight Components in an Aquatic Humic Substance as Characterized by Membrane Dialysis and Orbitrap Mass Spectrometry. *Environ. Sci. Technol.* **2012**, *46* (17), 9350–9359.

(29) Alvarino, T.; Suarez, S.; Lema, J.; Omil, F. Understanding the Sorption and Biotransformation of Organic Micropollutants in Innovative Biological Wastewater Treatment Technologies. *Sci. Total Environ.* **2018**, *615*, 297–306.

(30) Brack, W.; Ait-Aissa, S.; Burgess, R. M.; Busch, W.; Creusot, N.; Di Paolo, C.; Escher, B. I.; Mark Hewitt, L.; Hilscherova, K.; Hollender, J.; et al. Effect-Directed Analysis Supporting Monitoring of Aquatic Environments — An in-Depth Overview. *Sci. Total Environ.* **2016**, *544*, 1073–1118.

(31) Daughton, C. G. Pharmaceuticals and the Environment (PiE): Evolution and Impact of the Published Literature Revealed by Bibliometric Analysis. *Sci. Total Environ.* **2016**, *562*, 391–426.

(32) Evgenidou, E. N.; Konstantinou, I. K.; Lambropoulou, D. A. Occurrence and Removal of Transformation Products of PPCPs and Illicit Drugs in Wastewaters: A Review. *Sci. Total Environ.* **2015**, *505*, 905–926.

(33) Hollender, J.; Schymanski, E. L.; Singer, H. P.; Ferguson, P. L. Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? *Environ. Sci. Technol.* **2017**, *51* (20), 11505–11512.

(34) Hufsky, F.; Scheubert, K.; Böcker, S. Computational Mass Spectrometry for Small-Molecule Fragmentation. *TrAC Trends Anal. Chem.* **2014**, *53*, 41–48.

(35) Schymanski, E. L.; Singer, H. P.; Slobodnik, J.; Ipolyi, I. M.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund, P.; Letzel, T.; Grosse, S.; et al. Non-Target Screening with High-Resolution Mass Spectrometry: Critical Review Using a Collaborative Trial on Water Analysis. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6237–6255.

(36) Verlicchi, P.; Al Aukidy, M.; Zambello, E. Occurrence of Pharmaceutical Compounds in Urban Wastewater: Removal, Mass Load and Environmental Risk after a Secondary treatment—A Review. *Sci. Total Environ.* **2012**, *429*, 123–155.

(37) Zedda, M.; Zwiener, C. Is Nontarget Screening of Emerging Contaminants by LC-HRMS Successful? A Plea for Compound Libraries and Computer Tools. *Anal. Bioanal. Chem.* **2012**, *403* (9), 2493–2502.

(38) Qian, F.; He, M.; Song, Y.; Tysklind, M.; Wu, J. A Bibliometric Analysis of Global Research Progress on Pharmaceutical Wastewater Treatment during 1994–2013. *Environ. Earth Sci.* **2015**, *73* (9), 4995–5005.

(39) Nelson, E. D.; Do, H.; Lewis, R. S.; Carr, S. A. Diurnal Variability of Pharmaceutical, Personal Care Product, Estrogen and Alkylphenol Concentrations in Effluent from a Tertiary Wastewater Treatment Facility. *Environ. Sci. Technol.* **2011**, *45* (4), 1228–1234.

(40) Papageorgiou, M.; Kosma, C.; Lambropoulou, D. Seasonal Occurrence, Removal, Mass Loading and Environmental Risk Assessment of 55 Pharmaceuticals and Personal Care Products in a Municipal Wastewater Treatment Plant in Central Greece. *Sci. Total Environ.* **2016**, *543*, 547–569.

(41) Dai, G.; Huang, J.; Chen, W.; Wang, B.; Yu, G.; Deng, S. Major Pharmaceuticals and Personal Care Products (PPCPs) in Wastewater Treatment Plant and Receiving Water in Beijing, China, and Associated Ecological Risks. *Bull. Environ. Contam. Toxicol.* **2014**, *92* (6), 655–661.

(42) Eide, I.; Neverdal, G.; Thorvaldsen, B.; Grung, B.; Kvalheim, O. M. Toxicological Evaluation of Complex Mixtures by Pattern Recognition: Correlating Chemical Fingerprints to Mutagenicity. *Environ. Health Perspect.* **2002**, *110* (Suppl 6), 985.

(43) Eide, I.; Neverdal, G.; Thorvaldsen, B.; Arneberg, R.; Grung, B.; Kvalheim, O. M. Toxicological Evaluation of Complex Mixtures: Fingerprinting and Multivariate Analysis. *Environ. Toxicol. Pharmacol.* **2004**, *18* (2), 127–133.

(44) Hug, C.; Sievers, M.; Ottermanns, R.; Hollert, H.; Brack, W.; Krauss, M. Linking Mutagenic Activity to Micropollutant Concentrations in Wastewater Samples by Partial Least Square Regression and Subsequent Identification of Variables. *Chemosphere* **2015**, *138*, 176–182.

(45) Schelli, K.; Rutowski, J.; Roubidoux, J.; Zhu, J. Staphylococcus Aureus Methicillin Resistance Detected by HPLC-MS/MS Targeted Metabolic Profiling. *J. Chromatogr. B* **2016**.

(46) Brack, W. Effect-Directed Analysis: A Promising Tool for the Identification of Organic Toxicants in Complex Mixtures? *Anal. Bioanal. Chem.* **2003**, *377* (3), 397–407.

(47) Brack, W.; Schmitt-Jansen, M.; Machala, M.; Brix, R.; Barceló, D.; Schymanski, E.; Streck, G.; Schulze, T. How to Confirm Identified Toxicants in Effect-Directed Analysis. *Anal. Bioanal. Chem.* **2008**, *390* (8), 1959–1973.

(48) Burgess, R. M.; Ho, K. T.; Brack, W.; Lamoree, M. Effects-Directed Analysis (EDA) and Toxicity Identification Evaluation (TIE): Complementary but Different Approaches for Diagnosing Causes of Environmental Toxicity: Environmental Diagnostics: EDA and TIE. *Environ. Toxicol. Chem.* **2013**, *32* (9), 1935–1945.

(49) Padhye, L. P. Fate of Environmental Pollutants. *Water Environ. Res.* **2015**, *87* (10), 1595–1610.

(50) Luo, Y.; Guo, W.; Ngo, H. H.; Nghiem, L. D.; Hai, F. I.; Zhang, J.; Liang, S.; Wang, X. C. A Review on the Occurrence of Micropollutants in the Aquatic Environment and Their Fate and Removal during Wastewater Treatment. *Sci. Total Environ.* **2014**, *473–474*, 619–641.

(51) Cetecioglu, Z.; Atasoy, M. Biodegradation and Inhibitory Effects of Antibiotics on Biological Wastewater Treatment Systems. *Methods Pharmacol. Toxicol.* **2018**, No. 9781493974245, 29–55.

(52) Howell, S. R.; Husbands, G. E. M.; Scatina, J. A.; Sisenwine, S. F. Metabolic Disposition of 14c-Venlafaxine in Mouse, Rat, Dog, Rhesus Monkey and Man. *Xenobiotica* **1993**, *23* (4), 349–359.

(53) Miao, X.-S.; Yang, J.-J.; Metcalfe, C. D. Carbamazepine and Its Metabolites in Wastewater and in Biosolids in a Municipal Wastewater Treatment Plant. *Environ. Sci. Technol.* **2005**, *39* (19), 7469–7475.

(54) Sunkara, M.; Wells, M. J. M. Phase II Pharmaceutical Metabolites Acetaminophen Glucuronide and Acetaminophen Sulfate in Wastewater. *Environ. Chem.* **2010**, *7* (1), 111–122.

(55) Wang, J.; Gardinali, P. R. Identification of Phase II Pharmaceutical Metabolites in Reclaimed Water Using High Resolution Benchtop Orbitrap Mass Spectrometry. *Chemosphere* **2014**, *107*, 65–73.

(56) Kovalova, L.; Siegrist, H.; Singer, H.; Wittmer, A.; McArdell, C. S. Hospital Wastewater Treatment by Membrane Bioreactor: Performance and Efficiency for Organic Micropollutant Elimination. *Environ. Sci. Technol.* **2012**, *46* (3), 1536–1545.

(57) Zheng, W.; Zou, Y.; Li, X.; Machesky, M. L. Fate of Estrogen Conjugate 17α-Estradiol-3-Sulfate in Dairy Wastewater: Comparison of Aerobic and Anaerobic Degradation and Metabolite Formation. *J. Hazard. Mater.* **2013**, *258–259*, 109–115.

(58) Gao, J.; Banks, A.; Li, J.; Jiang, G.; Lai, F. Y.; Mueller, J. F.; Thai, P. K. Evaluation of in-Sewer Transformation of Selected Illicit Drugs and Pharmaceutical Biomarkers. *Sci. Total Environ.* **2017**, *609*, 1172–1181.

(59) Jelic, A.; Rodriguez-Mozaz, S.; Barceló, D.; Gutierrez, O. Impact of in-Sewer Transformation on 43 Pharmaceuticals in a Pressurized Sewer under Anaerobic Conditions. *Water Res.* **2015**, *68*, 98–108.

(60) Jelic, A.; Fatone, F.; Di Fabio, S.; Petrovic, M.; Cecchi, F.; Barcelo, D. Tracing Pharmaceuticals in a Municipal Plant for Integrated Wastewater and Organic Solid Waste Treatment. *Sci. Total Environ.* **2012**, *433*, 352–361.

(61) Prasse, C.; Wagner, M.; Schulz, R.; Ternes, T. A. Oxidation of the Antiviral Drug Acyclovir and Its Biodegradation Product Carboxy-Acyclovir with Ozone: Kinetics and Identification of Oxidation Products. *Environ. Sci. Technol.* **2012**, *46* (4), 2169–2178.

(62) Boxall, A. B.; Sinclair, C. J.; Fenner, K.; Kolpin, D.; Maund, S. J. *Peer Reviewed: When Synthetic Chemicals Degrade in the Environment*; ACS Publications, 2004.

(63) Picó, Y.; Barceló, D. Transformation Products of Emerging Contaminants in the Environment and High-Resolution Mass Spectrometry: A New Horizon. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6257–6273.

(64) Yang, W.; Zhou, H.; Cicek, N. Treatment of Organic Micropollutants in Water and Wastewater by UV-Based Processes: A Literature Review. *Crit. Rev. Environ. Sci. Technol.* **2014**, *44* (13), 1443–1476.

(65) Michael, I.; Hapeshi, E.; Aceña, J.; Perez, S.; Petrović, M.; Zapata, A.; Barceló, D.; Malato, S.; Fatta-Kassinos, D. Light-Induced Catalytic Transformation of Ofloxacin by Solar Fenton in Various Water Matrices at a Pilot Plant: Mineralization and Characterization of Major Intermediate Products. *Sci. Total Environ.* **2013**, *461–462*, 39–48.

(66) Michael, I.; Hapeshi, E.; Osorio, V.; Perez, S.; Petrovic, M.; Zapata, A.; Malato, S.; Barceló, D.; Fatta-Kassinos, D. Solar Photocatalytic Treatment of Trimethoprim in Four Environmental Matrices at a Pilot Scale: Transformation Products and Ecotoxicity Evaluation. *Sci. Total Environ.* **2012**, *430*, 167–173.

(67) Radjenović, J.; Sirtori, C.; Petrović, M.; Barceló, D.; Malato, S. Characterization of Intermediate Products of Solar Photocatalytic Degradation of Ranitidine at Pilot-Scale. *Chemosphere* **2010**, *79* (4), 368–376.

(68) Barritaud, L.; Leroy, G.; Chachignon, M.; Ingrand, V.; Roche, P.; Bichon, E.; Gervais, G.; Bourgin, M.; Antignac, J. P.; Monteau, F.; et al. Identification of Treatment by-Products of the Ozonation of Estrone Sulfate. *Water Sci. Technol. Water Supply* **2013**, *13* (5), 1302.

(69) Bila, D.; Montalvão, A. F.; Azevedo, D. de A.; Dezotti, M. Estrogenic Activity Removal of 17β-Estradiol by Ozonation and Identification of by-Products. *Chemosphere* **2007**, *69* (5), 736–746.

(70) Müller, A.; Weiss, S. C.; Beißwenger, J.; Leukhardt, H. G.; Schulz, W.; Seitz, W.; Ruck, W. K. L.; Weber, W. H. Identification of Ozonation by-Products of 4- and 5-Methyl-1H-Benzotriazole during the Treatment of Surface Water to Drinking Water. *Water Res.* **2012**, *46* (3), 679–690.

(71) Pereira, R. de O.; de Alda, M. L.; Joglar, J.; Daniel, L. A.; Barceló, D. Identification of New Ozonation Disinfection Byproducts of 17β-Estradiol and Estrone in Water. *Chemosphere* **2011**, *84* (11), 1535–1541.

(72) Tay, K. S.; Madehi, N. Ozonation of Acebutolol in Aqueous Solution: Ozonation by-Products and Degradation Pathway. *Sep. Purif. Technol.* **2014**, *135*, 48–63.

(73) González-Mariño, I.; Rodríguez, I.; Quintana, J. B.; Cela, R. Investigation of the Transformation of 11-nor-9-Carboxy-Δ9-Tetrahydrocannabinol during Water Chlorination by Liquid Chromatography–quadrupole-Time-of-Flight-Mass Spectrometry. *J. Hazard. Mater.* **2013**, *261*, 628–636.

(74) Duirk, S. E.; Lindell, C.; Cornelison, C. C.; Kormos, J.; Ternes, T. A.; Attene-Ramos, M.; Osiol, J.; Wagner, E. D.; Plewa, M. J.; Richardson, S. D. Formation of Toxic Iodinated Disinfection by-Products from Compounds Used in Medical Imaging. *Environ. Sci. Technol.* **2011**, *45* (16), 6845–6854.

(75) Hanigan, D.; Thurman, E. M.; Ferrer, I.; Zhao, Y.; Andrews, S.; Zhang, J.; Herckes, P.; Westerhoff, P. Methadone Contributes to N-Nitrosodimethylamine Formation in Surface Waters and Wastewaters during Chloramination. *Environ. Sci. Technol. Lett.* **2015**, *2* (6), 151–157.

(76) Mascolo, G.; López, A.; Passino, R.; Ricco, G.; Tiravanti, G. Degradation of Sulphur Containing S-Triazines during Water Chlorination. *Water Res.* **1994**, *28* (12), 2499–2506.

(77) Matsushita, T.; Kobayashi, N.; Hashizuka, M.; Sakuma, H.; Kondo, T.; Matsui, Y.; Shirasaki, N. Changes in Mutagenicity and Acute Toxicity of Solutions of Iodinated X-Ray Contrast Media during Chlorination. *Chemosphere* **2015**, *135*, 101–107.

(78) Merel, S.; LeBot, B.; Clément, M.; Seux, R.; Thomas, O. Ms Identification of Microcystin-LR Chlorination by-Products. *Chemosphere* **2009**, *74* (6), 832–839.

(79) Merel, S.; Clément, M.; Mourot, A.; Fessard, V.; Thomas, O. Characterization of Cylindrospermopsin Chlorination. *Sci. Total Environ.* **2010**, *408* (16), 3433–3442.

(80) Negreira, N.; López de Alda, M.; Barceló, D. Degradation of the Cytostatic Etoposide in Chlorinated Water by Liquid Chromatography Coupled to Quadrupole-Orbitrap Mass Spectrometry: Identification and Quantification of by-Products in Real Water Samples. *Sci. Total Environ.* **2015**, *506–507*, 36–45.

(81) Negreira, N.; Regueiro, J.; López de Alda, M.; Barceló, D. Degradation of the Anticancer Drug Erlotinib during Water Chlorination: Non-Targeted Approach for the Identification of Transformation Products. *Water Res.* **2015**, *85*, 103–113.

(82) Negreira, N.; Regueiro, J.; López de Alda, M.; Barceló, D. Transformation of Tamoxifen and Its Major Metabolites during Water Chlorination: Identification and in Silico Toxicity Assessment of Their Disinfection Byproducts. *Water Res.* **2015**, *85*, 199–207.

(83) Quintana, J. B.; Rodil, R.; Cela, R. Reaction of β-Blockers and β-Agonist Pharmaceuticals with Aqueous Chlorine. Investigation of Kinetics and by-Products by Liquid Chromatography Quadrupole Time-of-Flight Mass Spectrometry. *Anal. Bioanal. Chem.* **2012**, *403* (8), 2385–2395.

(84) Rodil, R.; Quintana, J. B.; Cela, R. Transformation of Phenazone-Type Drugs during Chlorination. *Water Res.* **2012**, *46* (7), 2457–2468.

(85) Wendel, F. M.; Lütke Eversloh, C.; Machek, E. J.; Duirk, S. E.; Plewa, M. J.; Richardson, S. D.; Ternes, T. A. Transformation of Iopamidol during Chlorination. *Environ. Sci. Technol.* **2014**, *48* (21), 12689–12697.

(86) Bourgin, M.; Beck, B.; Boehler, M.; Borowska, E.; Fleiner, J.; Salhi, E.; Teichler, R.; von Gunten, U.; Siegrist, H.; McArdell, C. S. Evaluation of a Full-Scale Wastewater Treatment Plant Upgraded with Ozonation and Biological Post-Treatments: Abatement of Micropollutants, Formation of Transformation Products and Oxidation by-Products. *Water Res.* **2018**, *129*, 486–498.

(87) Delgado, L. F.; Charles, P.; Glucina, K.; Morlay, C. The Removal of Endocrine Disrupting Compounds, Pharmaceutically Activated Compounds and Cyanobacterial Toxins during Drinking Water Preparation Using Activated carbon—A Review. *Sci. Total Environ.* **2012**, *435–436*, 509–525.

(88) Paredes, L.; Fernandez-Fontaina, E.; Lema, J. M.; Omil, F.; Carballa, M. Understanding the Fate of Organic Micropollutants in Sand and Granular Activated Carbon Biofiltration Systems. *Sci. Total Environ.* **2016**, *551–552*, 640–648.

(89) Sun, B.; Ma, J.; Sedlak, D. L. Chemisorption of Perfluorooctanoic Acid on Powdered Activated Carbon Initiated by Persulfate in Aqueous Solution. *Environ. Sci. Technol.* **2016**, *50* (14), 7618–7624.

(90) Xiao, Y.; Yaohari, H.; De Araujo, C.; Sze, C. C.; Stuckey, D. C. Removal of Selected Pharmaceuticals in an Anaerobic Membrane Bioreactor (AnMBR) With/Without Powdered Activated Carbon (PAC). *Chem. Eng. J.* **2017**, *321*, 335–345.

(91) Janssens, R.; Mandal, M. K.; Dubey, K. K.; Luis, P. Slurry Photocatalytic Membrane Reactor Technology for Removal of Pharmaceutical Compounds from Wastewater: Towards Cytostatic Drug Elimination. *Sci. Total Environ.* **2017**, *599–600*, 612–626.

(92) Brix, R.; Bahi, N.; Lopez de Alda, M. J.; Farré, M.; Fernandez, J.-M.; Barceló, D. Identification of Disinfection by-Products of Selected Triazines in Drinking Water by LC-Q-ToF-MS/MS and Evaluation of Their Toxicity. *J. Mass Spectrom.* **2009**, *44* (3), 330–337.

(93) Chys, M.; Demeestere, K.; Ingabire, A. S.; Dries, J.; Van Langenhove, H.; Van Hulle, S. W. H. Enhanced Treatment of Secondary Municipal Wastewater Effluent: Comparing (Biological) Filtration and Ozonation in View of Micropollutant Removal, Unselective Effluent Toxicity, and the Potential for Real-Time Control. *Water Sci. Technol.* **2017**, *76* (1), 236–246.

(94) Alvarino, T.; Komesli, O.; Suarez, S.; Lema, J. M.; Omil, F. The Potential of the Innovative SeMPAC Process for Enhancing the Removal of Recalcitrant Organic Micropollutants. *J. Hazard. Mater.* **2016**, *308*, 29–36.

(95) Boehler, M.; Zwickenpflug, B.; Hollender, J.; Ternes, T.; Joss, A.; Siegrist, H. Removal of Micropollutants in Municipal Wastewater Treatment Plants by Powder-Activated Carbon. *Water Sci. Technol.* **2012**, *66* (10), 2115–2121.

(96) Hu, J.; Shang, R.; Heijman, B.; Rietveld, L. Influence of Activated Carbon Preloading by EfOM Fractions from Treated Wastewater on Adsorption of Pharmaceutically Active Compounds. *Chemosphere* **2016**, *150*, 49–56.

(97) Mailler, R.; Gasperi, J.; Coquet, Y.; Derome, C.; Buleté, A.; Vulliet, E.; Bressy, A.; Varrault, G.; Chebbo, G.; Rocher, V. Removal of Emerging Micropollutants from Wastewater by Activated Carbon Adsorption: Experimental Study of Different Activated Carbons and Factors Influencing the Adsorption of Micropollutants in Wastewater. *J. Environ. Chem. Eng.* **2016**, *4* (1), 1102–1109.

(98) Kårelid, V.; Larsson, G.; Björlenius, B. Pilot-Scale Removal of Pharmaceuticals in Municipal Wastewater: Comparison of Granular and Powdered Activated Carbon Treatment at Three Wastewater Treatment Plants. *J. Environ. Manage.* **2017**, *193*, 491–502.

(99) Zhang, D.; Li, W.; Wang, K.; Gong, H. Bacterial Community Dynamics and Its Effects during Biological Activited Carbon Filter Process for Drinking Water Treatment. **2010**, No. Icbee, 0–5.

(100) Kalkan, Ç.; Yapsakli, K.; Mertoglu, B.; Tufan, D.; Saatci, A. Evaluation of Biological Activated Carbon (BAC) Process in Wastewater Treatment Secondary Effluent for Reclamation Purposes. *Desalination* **2011**, *265* (1–3), 266–273.

(101) Velten, S.; Boller, M.; Köster, O.; Helbing, J.; Weilenmann, H. U.; Hammes, F. Development of Biomass in a Drinking Water Granular Active Carbon (GAC) Filter. *Water Res.* **2011**, *45* (19), 6347–6354.

(102) Gerrity, D.; Gamage, S.; Holady, J. C.; Mawhinney, D. B.; Quiñones, O.; Trenholm, R. A.; Snyder, S. A. Pilot-Scale Evaluation of Ozone and Biological Activated Carbon for Trace Organic Contaminant Mitigation and Disinfection. *Water Res.* **2011**, *45* (5), 2155–2165.

(103) Justo, A.; González, O.; Sans, C.; Esplugas, S. BAC Filtration to Mitigate Micropollutants and EfOM Content in Reclamation Reverse Osmosis Brines. *Chem. Eng. J.* **2015**, *279*, 589–596.

(104) Zhang, S.; Gitungo, S. W.; Axe, L.; Raczko, R. F.; Dyksen, J. E. Biologically Active filters–An Advanced Water Treatment Process for Contaminants of Emerging Concern. *Water Res.* **2017**, *114*, 31–41.

(105) Snyder, S. A.; Adham, S.; Redding, A. M.; Cannon, F. S.; DeCarolis, J.; Oppenheimer, J.; Wert, E. C.; Yoon, Y. Role of Membranes and Activated Carbon in the Removal of Endocrine Disruptors and Pharmaceuticals. *Desalination* **2007**, *202* (1–3), 156–181.

(106) Knopp, G.; Prasse, C.; Ternes, T. A.; Cornel, P. Elimination of Micropollutants and Transformation Products from a Wastewater Treatment Plant Effluent through Pilot Scale Ozonation Followed by Various Activated Carbon and Biological Filters. *Water Res.* **2016**, *100*, 580–592.

(107) Benstoem, F.; Nahrstedt, A.; Boehler, M.; Knopp, G.; Montag, D.; Siegrist, H.; Pinnekamp, J. Performance of Granular Activated Carbon to Remove Micropollutants from Municipal wastewater—A Meta-Analysis of Pilot- and Large-Scale Studies. *Chemosphere* **2017**, *185*, 105–118.

(108) De Ridder, D.; Verliefde, A.; Heijman, S.; Verberk, J.; Rietveld, L.; Van Der Aa, L.; Amy, G.; Van Dijk, J. Influence of Natural Organic Matter on Equilibrium Adsorption of Neutral and Charged Pharmaceuticals onto Activated Carbon. *Water Sci. Technol.* **2011**, *63* (3), 416–423.

(109) Gardner, M.; Jones, V.; Comber, S.; Scrimshaw, M. D.; Coello - Garcia, T.; Cartmell, E.; Lester, J.; Ellor, B. Performance of UK Wastewater Treatment Works with Respect to Trace Contaminants. *Sci. Total Environ.* **2013**, *456–457*, 359–369.

(110) Cirja, M.; Ivashechkin, P.; Schäffer, A.; Corvini, P. F. X. Factors Affecting the Removal of Organic Micropollutants from Wastewater in Conventional Treatment Plants (CTP) and Membrane Bioreactors (MBR). *Rev. Environ. Sci. Biotechnol.* **2008**, *7* (1), 61–78.

(111) Quesnel, D. M.; Oldenburg, T. B. P.; Larter, S. R.; Gieg, L. M.; Chua, G. Biostimulation of Oil Sands Process-Affected Water with Phosphate Yields Removal of Sulfur-Containing Organics and Detoxification. *Environ. Sci. Technol.* **2015**, *49* (21), 13012–13020.

(112) Clesceri, L. S.; Eaton, A. D.; Greenberg, A. E. *Standard Methods for the Examination of Water and Wastewater*; American Public Health Association, 1998.

(113) Huber, S. A.; Balz, A.; Abert, M.; Pronk, W. Characterisation of Aquatic Humic and Non-Humic Matter with Size-Exclusion Chromatography – Organic Carbon Detection – Organic Nitrogen Detection (LC-OCD-OND). *Water Res.* **2011**, *45* (2), 879–885.

(114) Chen, W.; Westerhoff, P.; Leenheer, J. A.; Booksh, K. Fluorescence Excitation−Emission Matrix Regional Integration to Quantify Spectra for Dissolved Organic Matter. *Environ. Sci. Technol.* **2003**, *37* (24), 5701–5710.

(115) Prasse, C.; Stalter, D.; Schulte-Oehlmann, U.; Oehlmann, J.; Ternes, T. A. Spoilt for Choice: A Critical Review on the Chemical and Biological Assessment of Current Wastewater Treatment Technologies. *Water Res.* **2015**, *87*, 237–270.

(116) Pérez-Fernández, V.; Mainero Rocca, L.; Tomai, P.; Fanali, S.; Gentili, A. Recent Advancements and Future Trends in Environmental Analysis: Sample Preparation, Liquid Chromatography and Mass Spectrometry. *Anal. Chim. Acta* **2017**, *983*, 9–41.

(117) Andrade-Eiroa, A.; Canle, M.; Leroy-Cancellieri, V.; Cerdà, V. Solid-Phase Extraction of Organic Compounds: A Critical Review (Part I). *TrAC Trends Anal. Chem.* **2016**, *80*, 641–654.

(118) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A Simple and Efficient Method for the Solid-Phase Extraction of Dissolved Organic Matter (SPE-DOM) from Seawater. *Limnol. Oceanogr. Methods* **2008**, *6* (6), 230–235.

(119) Kern, S.; Fenner, K.; Singer, H. P.; Schwarzenbach, R. P.; Hollender, J. Identification of Transformation Products of Organic Contaminants in Natural Waters by Computer-Aided Prediction and High-Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2009**, *43* (18), 7039–7046.

(120) Li, Y.; Harir, M.; Lucio, M.; Kanawati, B.; Smirnov, K.; Flerus, R.; Koch, B. P.; Schmitt-Kopplin, P.; Hertkorn, N. Proposed Guidelines for Solid Phase Extraction of Suwannee River Dissolved Organic Matter. *Anal. Chem.* **2016**, *88* (13), 6680–6688.

(121) Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. An Overview of Matrix Effects in Liquid Chromatography-Mass Spectrometry. *Mass Spectrom. Rev.* **2011**, *30* (3), 491–509.

(122) Tyrkkö, E.; Pelander, A.; Ojanperä, I. Prediction of Liquid Chromatographic Retention for Differentiation of Structural Isomers. *Anal. Chim. Acta* **2012**, *720*, 142–148.

(123) Iparraguirre, A.; Navarro, P.; Rodil, R.; Prieto, A.; Olivares, M.; Etxebarria, N.; Zuloaga, O. Matrix Effect during the Membrane-Assisted Solvent Extraction Coupled to Liquid Chromatography Tandem Mass Spectrometry for the Determination of a Variety of Endocrine Disrupting Compounds in Wastewater. *J. Chromatogr. A* **2014**, *1356*, 163–170.

(124) Taylor, P. J. Matrix Effects: The Achilles Heel of Quantitative High-Performance Liquid Chromatography–electrospray–tandem Mass Spectrometry. *Clin. Biochem.* **2005**, *38* (4), 328–334.

(125) Gago-Ferrero, P.; Schymanski, E. L.; Bletsou, A. A.; Aalizadeh, R.; Hollender, J.; Thomaidis, N. S. Extended Suspect and Non-Target Strategies to Characterize Emerging Polar Organic Contaminants in Raw Wastewater with Lc-Hrms/Ms. *Environ. Sci. Technol.* **2015**, *49* (20), 12333–12341.

(126) Hühnerfuss, H.; Shah, M. R. Enantioselective chromatography—A Powerful Tool for the Discrimination of Biotic and Abiotic Transformation Processes of Chiral Environmental Pollutants. *J. Chromatogr. A* **2009**, *1216* (3), 481–502.

(127) Bletsou, A. A.; Jeon, J.; Hollender, J.; Archontaki, E.; Thomaidis, N. S. Targeted and Non-Targeted Liquid Chromatography-Mass Spectrometric Workflows for Identification of Transformation Products of Emerging Pollutants in the Aquatic Environment. *TrAC Trends Anal. Chem.* **2015**, *66*, 32–44.

(128) Milman, B. L. General Principles of Identification by Mass Spectrometry. *TrAC Trends Anal. Chem.* **2015**, *69*, 24–33.

(129) Pitarch, E.; Portolés, T.; Marín, J. M.; Ibáñez, M.; Albarrán, F.; Hernández, F. Analytical Strategy Based on the Use of Liquid Chromatography and Gas Chromatography with Triple-

Quadrupole and Time-of-Flight MS Analyzers for Investigating Organic Contaminants in Wastewater. *Anal. Bioanal. Chem.* **2010**, *397* (7), 2763–2776.

(130) Gros, M.; Rodríguez-Mozaz, S.; Barceló, D. Fast and Comprehensive Multi-Residue Analysis of a Broad Range of Human and Veterinary Pharmaceuticals and Some of Their Metabolites in Surface and Treated Waters by Ultra-High-Performance Liquid Chromatography Coupled to Quadrupole-Linear Ion Trap Tandem Mass Spectrometry. *J. Chromatogr. A* **2012**, *1248*, 104–121.

(131) Huntscha, S.; Hofstetter, T. B.; Schymanski, E. L.; Spahr, S.; Hollender, J. Biotransformation of Benzotriazoles: Insights from Transformation Product Identification and Compound-Specific Isotope Analysis. *Environ. Sci. Technol.* **2014**, *48* (8), 4435–4443.

(132) Hug, C.; Ulrich, N.; Schulze, T.; Brack, W.; Krauss, M. Identification of Novel Micropollutants in Wastewater by a Combination of Suspect and Nontarget Screening. *Environ. Pollut.* **2014**, *184*, 25–32.

(133) Singer, H. P.; Wössner, A. E.; McArdell, C. S.; Fenner, K. Rapid Screening for Exposure to "Non-Target" Pharmaceuticals from Wastewater Effluents by Combining HRMS-Based Suspect Screening and Exposure Modeling. *Environ. Sci. Technol.* **2016**, *50* (13), 6698–6707.

(134) Radjenović, J.; Petrović, M.; Barceló, D. Fate and Distribution of Pharmaceuticals in Wastewater and Sewage Sludge of the Conventional Activated Sludge (CAS) and Advanced Membrane Bioreactor (MBR) Treatment. *Water Res.* **2009**, *43* (3), 831–841.

(135) Martínez Bueno, M. J.; Ulaszewska, M. M.; Gomez, M. J.; Hernando, M. D.; Fernández-Alba, A. R. Simultaneous Measurement in Mass and Mass/Mass Mode for Accurate Qualitative and Quantitative Screening Analysis of Pharmaceuticals in River Water. *J. Chromatogr. A* **2012**, *1256*, 80–88.

(136) Little, J. L.; Williams, A. J.; Pshenichnov, A.; Tkachenko, V. Identification of "Known Unknowns" Utilizing Accurate Mass Data and ChemSpider. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (1), 179–185.

(137) Gómez-Pérez, M. L.; Romero-González, R.; Vidal, J. L. M.; Frenich, A. G. Identification of Transformation Products of Pesticides and Veterinary Drugs in Food and Related Matrices: Use of Retrospective Analysis. *J. Chromatogr. A* **2015**, *1389*, 133–138.

(138) Hernández, F.; Ibáñez, M.; Portolés, T.; Cervera, M. I.; Sancho, J. V.; López, F. J. Advancing towards Universal Screening for Organic Pollutants in Waters. *J. Hazard. Mater.* **2015**, *282*, 86–95.

(139) Scheubert, K.; Hufsky, F.; Böcker, S. Computational Mass Spectrometry for Small Molecules. *J Cheminformatics* **2013**, *5*, 12.

(140) Vergeynst, L.; Van Langenhove, H.; Demeestere, K. Balancing the False Negative and Positive Rates in Suspect Screening with High-Resolution Orbitrap Mass Spectrometry Using Multivariate Statistics. *Anal. Chem.* **2015**, *87* (4), 2170–2177.

(141) Sjerps, R. M. A.; Vughs, D.; van Leerdam, J. A.; ter Laak, T. L.; van Wezel, A. P. Data-Driven Prioritization of Chemicals for Various Water Types Using Suspect Screening LC-HRMS. *Water Res.* **2016**, *93*, 254–264.

(142) Leendert, V.; Van Langenhove, H.; Demeestere, K. Trends in Liquid Chromatography Coupled to High-Resolution Mass Spectrometry for Multi-Residue Analysis of Organic Micropollutants in Aquatic Environments. *TrAC Trends Anal. Chem.* **2015**, *67*, 192–208.

(143) Zhang, Y.; Zhang, N.; Zhao, P.; Niu, Z. Characteristics of Molecular Weight Distribution of Dissolved Organic Matter in Bromide-Containing Water and Disinfection by-Product Formation Properties during Treatment Processes. *J. Environ. Sci.* **2017**.

(144) Zhang, H.; Zhang, Y.; Shi, Q.; Hu, J.; Chu, M.; Yu, J.; Yang, M. Study on Transformation of Natural Organic Matter in Source Water during Chlorination and Its Chlorinated Products Using Ultrahigh Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2012**, *46* (8), 4396–4402.

(145) Lavonen, E. E.; Gonsior, M.; Tranvik, L. J.; Schmitt-Kopplin, P.; Köhler, S. J. Selective Chlorination of Natural Organic Matter: Identification of Previously Unknown Disinfection Byproducts. *Environ. Sci. Technol.* **2013**, *47* (5), 2264–2271.

(146) Gong, T.; Tao, Y.; Xian, Q. Selection and Applicability of Quenching Agents for the Analysis of Polar Iodinated Disinfection Byproducts. *Chemosphere* **2016**, *163*, 359–365.

(147) Ruff, M.; Mueller, M. S.; Loos, M.; Singer, H. P. Quantitative Target and Systematic Non-Target Analysis of Polar Organic Micro-Pollutants along the River Rhine Using High-Resolution Mass-Spectrometry – Identification of Unknown Sources and Compounds. *Water Res.* **2015**, *87*, 145–154.

(148) Plassmann, M. M.; Tengstrand, E.; Åberg, K. M.; Benskin, J. P. Non-Target Time Trend Screening: A Data Reduction Strategy for Detecting Emerging Contaminants in Biological Samples. *Anal. Bioanal. Chem.* **2016**, *408* (16), 4203–4208.

(149) Wu, Z.; Rodgers, R. P.; Marshall, A. G. Two- and Three-Dimensional van Krevelen Diagrams: A Graphical Analysis Complementary to the Kendrick Mass Plot for Sorting Elemental Compositions of Complex Organic Mixtures Based on Ultrahigh-Resolution Broadband Fourier Transform Ion Cyclotron Resonance Mass Measurements. *Anal. Chem.* **2004**, *76* (9), 2511–2516.

(150) Artemenko, K. A.; Zubarev, A. R.; Samgina, T. Y.; Lebedev, A. T.; Savitski, M. M.; Zubarev, R. A. Two Dimensional Mass Mapping as a General Method of Data Representation in Comprehensive Analysis of Complex Molecular Mixtures. *Anal. Chem.* **2009**, *81* (10), 3738–3745.

(151) D'Andrilli, J.; Cooper, W. T.; Foreman, C. M.; Marshall, A. G. An Ultrahigh-Resolution Mass Spectrometry Index to Estimate Natural Organic Matter Lability: FTICRMS Organic Matter Molecular Lability Index. *Rapid Commun. Mass Spectrom.* **2015**, *29* (24), 2385–2401.

(152) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra. *Anal. Chem.* **2001**, *73* (19), 4676–4681.

(153) Lavonen, E. E.; Kothawala, D. N.; Tranvik, L. J.; Gonsior, M.; Schmitt-Kopplin, P.; Köhler, S. J. Tracking Changes in the Optical Properties and Molecular Composition of Dissolved Organic Matter during Drinking Water Production. *Water Res.* **2015**, *85*, 286–294.

(154) Sleighter, R. L.; Hatcher, P. G. Molecular Characterization of Dissolved Organic Matter (DOM) along a River to Ocean Transect of the Lower Chesapeake Bay by Ultrahigh Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Mar. Chem.* **2008**, *110* (3–4), 140–152.

(155) Fang, Z.; He, C.; Li, Y.; Chung, K. H.; Xu, C.; Shi, Q. Fractionation and Characterization of Dissolved Organic Matter (DOM) in Refinery Wastewater by Revised Phase Retention and Ion-Exchange Adsorption Solid Phase Extraction Followed by ESI FT-ICR MS. *Talanta* **2017**, *162*, 466–473.

(156) Cortés-Francisco, N.; Caixach, J. Molecular Characterization of Dissolved Organic Matter through a Desalination Process by High Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2013**, *47* (17), 9619–9627.

(157) Pereira, A. S.; Islam, M. S.; Gamal El-Din, M.; Martin, J. W. Ozonation Degrades All Detectable Organic Compound Classes in Oil Sands Process-Affected Water; an Application of High-Performance Liquid Chromatography/Obitrap Mass Spectrometry: Ozonation Degrades All Detectable Organic Compound Classes in OSPW. *Rapid Commun. Mass Spectrom.* **2013**, *27* (21), 2317–2326.

(158) Phungsai, P.; Kurisu, F.; Kasuga, I.; Furumai, H. Molecular Characterization of Low Molecular Weight Dissolved Organic Matter in Water Reclamation Processes Using Orbitrap Mass Spectrometry. *Water Res.* **2016**, *100*, 526–536.

(159) Verkh, Y.; Rozman, M.; Petrovic, M. A Non-Targeted High-Resolution Mass Spectrometry Data Analysis of Dissolved Organic Matter in Wastewater Treatment. *Chemosphere* **2018**, *200*, 397–404.

(160) Mesfioui, R.; Love, N. G.; Bronk, D. A.; Mulholland, M. R.; Hatcher, P. G. Reactivity and Chemical Characterization of Effluent Organic Nitrogen from Wastewater Treatment Plants Determined by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Water Res.* **2012**, *46* (3), 622–634.

(161) Gonsior, M.; Zwartjes, M.; Cooper, W. J.; Song, W.; Ishida, K. P.; Tseng, L. Y.; Jeung, M. K.; Rosso, D.; Hertkorn, N.; Schmitt-Kopplin, P. Molecular Characterization of Effluent Organic Matter Identified by Ultrahigh Resolution Mass Spectrometry. *Water Res.* **2011**, *45* (9), 2943–2953.

(162) Müller, A.; Schulz, W.; Ruck, W. K. L.; Weber, W. H. A New Approach to Data Evaluation in the Non-Target Screening of Organic Trace Substances in Water Analysis. *Chemosphere* **2011**, *85* (8), 1211–1219.

(163) Sleno, L. The Use of Mass Defect in Modern Mass Spectrometry: Mass Defect in Mass Spectrometry. *J. Mass Spectrom.* **2012**, *47* (2), 226–236.

(164) Loos, M.; Singer, H. Nontargeted Homologue Series Extraction from Hyphenated High Resolution Mass Spectrometry Data. *J. Cheminformatics* **2017**, *9* (1).

(165) Zhang, H.; Zhang, Y.; Shi, Q.; Ren, S.; Yu, J.; Ji, F.; Luo, W.; Yang, M. Characterization of Low Molecular Weight Dissolved Natural Organic Matter along the Treatment Trait of a Waterworks Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Water Res.* **2012**, *46* (16), 5197–5204.

(166) Taguchi, V. Y.; Nieckarz, R. J.; Clement, R. E.; Krolik, S.; Williams, R. Dioxin Analysis by Gas Chromatography-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (GC-FTICRMS). *J. Am. Soc. Mass Spectrom.* **2010**.

(167) Minor, E. C.; Steinbring, C. J.; Longnecker, K.; Kujawinski, E. B. Characterization of Dissolved Organic Matter in Lake Superior and Its Watershed Using Ultrahigh Resolution Mass Spectrometry. *Org. Geochem.* **2012**, *43*, 1–11.

(168) Lu, M.; Zhang, Z.; Qiao, W.; Wei, X.; Guan, Y.; Ma, Q.; Guan, Y. Remediation of Petroleum-Contaminated Soil after Composting by Sequential Treatment with Fenton-like Oxidation and Biodegradation. *Bioresour. Technol.* **2010**, *101* (7), 2106–2113.

(169) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **2003**, *75* (20), 5336–5344.

(170) Herzsprung, P.; von Tümpling, W.; Hertkorn, N.; Harir, M.; Büttner, O.; Bravidor, J.; Friese, K.; Schmitt-Kopplin, P. Variations of DOM Quality in Inflows of a Drinking Water Reservoir: Linking of van Krevelen Diagrams with EEMF Spectra by Rank Correlation. *Environ. Sci. Technol.* **2012**, *46* (10), 5511–5518.

(171) Tseng, L. Y.; Gonsior, M.; Schmitt-Kopplin, P.; Cooper, W. J.; Pitt, P.; Rosso, D. Molecular Characteristics and Differences of Effluent Organic Matter from Parallel Activated Sludge and Integrated Fixed-Film Activated Sludge (IFAS) Processes. *Environ. Sci. Technol.* **2013**, *18* (47), 10277–10284.

(172) Marshall, J. W.; Schmitt-Kopplin, P.; Schuetz, N.; Moritz, F.; Roullier-Gall, C.; Uhl, J.; Colyer, A.; Jones, L. L.; Rychlik, M.; Taylor, A. J. Monitoring Chemical Changes during Food Sterilisation Using Ultrahigh Resolution Mass Spectrometry. *Food Chem.* **2018**, *242*, 316–322.

(173) Cortés-Francisco, N.; Harir, M.; Lucio, M.; Ribera, G.; Martínez-Lladó, X.; Rovira, M.; Schmitt-Kopplin, P.; Hertkorn, N.; Caixach, J. High-Field FT-ICR Mass Spectrometry and NMR Spectroscopy to Characterize DOM Removal through a Nanofiltration Pilot Plant. *Water Res.* **2014**, *67*, 154–165.

(174) Korsten, H. Characterization of Hydrocarbon Systems by DBE Concept. *AIChE J.* **1997**, *43* (6), 1559–1568.

(175) Badertscher, M.; Bischofberger, K.; Munk, M. E.; Pretsch, E. A Novel Formalism To Characterize the Degree of Unsaturation of Organic Molecules. *J. Chem. Inf. Model.* **2001**, *41* (4), 889–893.

(176) Koch, B. P.; Dittmar, T. From Mass to Structure: An Aromaticity Index for High-Resolution Mass Data of Natural Organic Matter. *Rapid Commun. Mass Spectrom.* **2006**, *20* (5), 926–932.

(177) Sleighter, R. L.; Liu, Z.; Xue, J.; Hatcher, P. G. Multivariate Statistical Approaches for the Characterization of Dissolved Organic Matter Analyzed by Ultrahigh Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2010**, *44* (19), 7576–7582.

(178) Chen, H.; Stubbins, A.; Perdue, E. M.; Green, N. W.; Helms, J. R.; Mopper, K.; Hatcher, P. G. Ultrahigh Resolution Mass Spectrometric Differentiation of Dissolved Organic Matter Isolated by Coupled Reverse Osmosis-Electrodialysis from Various Major Oceanic Water Masses. *Mar. Chem.* **2014**, *164*, 48–59.

(179) Schollée, J. E.; Schymanski, E. L.; Avak, S. E.; Loos, M.; Hollender, J. Prioritizing Unknown Transformation Products from Biologically-Treated Wastewater Using High-Resolution Mass Spectrometry, Multivariate Statistics, and Metabolic Logic. *Anal. Chem.* **2015**, *87* (24), 12121–12129.

(180) Kalogiouri, N. P.; Alygizakis, N. A.; Aalizadeh, R.; Thomaidis, N. S. Olive Oil Authenticity Studies by Target and Nontarget LC–QTOF-MS Combined with Advanced Chemometric Techniques. *Anal. Bioanal. Chem.* **2016**, *408* (28), 7955–7970.

(181) Glauser, G.; Veyrat, N.; Rochat, B.; Wolfender, J.-L.; Turlings, T. C. J. Ultra-High Pressure Liquid Chromatography–mass Spectrometry for Plant Metabolomics: A Systematic Comparison of High-Resolution Quadrupole-Time-of-Flight and Single Stage Orbitrap Mass Spectrometers. *J. Chromatogr. A* **2013**, *1292*, 151–159.

(182) Marshall, A. G.; Hendrickson, C. L. High-Resolution Mass Spectrometers. *Annu. Rev. Anal. Chem.* **2008**, *1* (1), 579–599.

(183) Makarov, A.; Denisov, E.; Lange, O. Performance Evaluation of a High-Field Orbitrap Mass Analyzer. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1391–1396.

(184) Niessen, W.; Falck, D. Introduction to Mass Spectrometry, a Tutorial. *Anal. Biomol. Interact. Mass Spectrom.* **2015**, 1–54.

(185) McLafferty, F. W.; Tureček, F. *Interpretation of Mass Spectra*, 4th ed.; University Science Books: Mill Valley, Calif, 1993.

(186) Pleil, J. D.; Isaacs, K. K. High-Resolution Mass Spectrometry: Basic Principles for Using Exact Mass and Mass Defect for Discovery Analysis of Organic Molecules in Blood, Breath, Urine and Environmental Media. *J. Breath Res.* **2016**, *10*, 012001.

(187) Erve, J. C. L.; Gu, M.; Wang, Y.; DeMaio, W.; Talaat, R. E. Spectral Accuracy of Molecular Ions in an LTQ/Orbitrap Mass Spectrometer and Implications for Elemental Composition Determination. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (11), 2058–2069.

(188) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. Dynamic Range of Mass Accuracy in LTQ Orbitrap Hybrid Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (7), 977–982.

(189) Valkenborg, D.; Mertens, I.; Lemière, F.; Witters, E.; Burzykowski, T. The Isotopic Distribution Conundrum. *Mass Spectrom. Rev.* **2012**, *31* (1), 96–109.

(190) Urban, J.; Afseth, N. K.; Štys, D. Fundamental Definitions and Confusions in Mass Spectrometry about Mass Assignment, Centroiding and Resolution. *TrAC Trends Anal. Chem.* **2014**, *53*, 126–136.

(191) Huang, N.; Siegel, M. M.; Kruppa, G. H.; Laukien, F. H. Automation of a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer for Acquisition, Analysis, and E-Mailing of High-Resolution Exact-Mass Electrospray Ionization Mass Spectral Data. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (11), 1166–1173.

(192) Mihaleva, V. V.; Verhoeven, H. A.; de Vos, R. C. H.; Hall, R. D.; van Ham, R. C. H. J. Automated Procedure for Candidate Compound Selection in GC-MS Metabolomics Based on Prediction of Kovats Retention Index. *Bioinformatics* **2009**, *25* (6), 787–794.

(193) Hawkes, J. A.; Dittmar, T.; Patriarca, C.; Tranvik, L. J.; Bergquist, J. Evaluation of the Orbitrap Mass Spectrometer for the Molecular Fingerprinting Analysis of Natural Dissolved Organic Matter (DOM). *Anal. Chem.* **2016**.

(194) Wang, C.; Huang, R.; Klamerth, N.; Chelme-Ayala, P.; Gamal El-Din, M. Positive and Negative Electrospray Ionization Analyses of the Organic Fractions in Raw and Oxidized Oil Sands Process-Affected Water. *Chemosphere* **2016**, *165*, 239–247.

(195) Katajamaa, M.; Orešič, M. Data Processing for Mass Spectrometry-Based Metabolomics. *J. Chromatogr. A* **2007**, *1158* (1–2), 318–328.

(196) Katajamaa, M.; Miettinen, J.; Oresic, M. MZmine: Toolbox for Processing and Visualization of Mass Spectrometry Based Molecular Profile Data. *Bioinformatics* **2006**, *22* (5), 634–636.

(197) Katajamaa, M.; Orešič, M. Processing Methods for Differential Analysis of LC/MS Profile Data. *BMC Bioinformatics* **2005**, *6* (1), 179.

(198) Pluskal, T.; Uehara, T.; Yanagida, M. Highly Accurate Chemical Formula Prediction Tool Utilizing High-Resolution Mass Spectra, Ms/Ms Fragmentation, Heuristic Rules, and Isotope Pattern Matching. *Anal. Chem.* **2012**, *84* (10), 4396–4403.

(199) Röst, H. L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weisser, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.-C.; Gutenbrunner, P.; Kenar, E.; et al. OpenMS: A Flexible Open-Source Software Platform for Mass Spectrometry Data Analysis. *Nat. Methods* **2016**, *13* (9), 741–748.

(200) Chiaia-Hernandez, A. C.; Schymanski, E. L.; Kumar, P.; Singer, H. P.; Hollender, J. Suspect and Nontarget Screening Approaches to Identify Organic Contaminant Records in Lake Sediments. *Anal. Bioanal. Chem.* **2014**, *406* (28), 7323–7335.

(201) Lange, E.; Tautenhahn, R.; Neumann, S.; Gröpl, C. Critical Assessment of Alignment Procedures for LC-MS Proteomics and Metabolomics Measurements. *BMC Bioinformatics* **2008**, *9* (1), 375.

(202) Aral, H.; Çelik, K. S.; Altındağ, R.; Aral, T. Synthesis, Characterization, and Application of a Novel Multifunctional Stationary Phase for Hydrophilic Interaction/Reversed Phase Mixed-Mode Chromatography. *Talanta* **2017**, *174*, 703–714.

(203) Eichhorn, P.; Pérez, S.; Barceló, D. Time-of-Flight Mass Spectrometry Versus Orbitrap-Based Mass Spectrometry for the Screening and Identification of Drugs and Metabolites. In *Comprehensive Analytical Chemistry*; Elsevier, 2012; Vol. 58, pp 217–272.

(204) Dhungana, B.; Becker, C.; Zekavat, B.; Solouki, T.; Hockaday, W. C.; Chambliss, C. K. Characterization of Slow-Pyrolysis Bio-Oils by High-Resolution Mass Spectrometry and Ion Mobility Spectrometry. *Energy Fuels* **2015**, *29* (2), 744–753.

(205) Saito-Shida, S.; Hamasaka, T.; Nemoto, S.; Akiyama, H. Multiresidue Determination of Pesticides in Tea by Liquid Chromatography-High-Resolution Mass Spectrometry: Comparison between Orbitrap and Time-of-Flight Mass Analyzers. *Food Chem.* **2018**, *256*, 140–148.

(206) Wick, A.; Wagner, M.; Ternes, T. A. Elucidation of the Transformation Pathway of the Opium Alkaloid Codeine in Biological Wastewater Treatment. *Environ. Sci. Technol.* **2011**, *45* (8), 3374–3385.

(207) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010**, *11* (1), 395.

(208) Yaroslav Verkh; Marko Rozman; Mira Petrovic. Extraction and Cleansing of Data for a Non-Targeted Analysis of High-Resolution Mass Spectrometry Data of Wastewater. *MethodsX* **2018**, *5*, 395–402.

(209) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.

(210) Verkh Yaroslav. *MzMineR R Package*; Girona, Spain, 2016.

(211) Cao, Y.; Charisi, A.; Cheng, L.-C.; Jiang, T.; Girke, T. ChemmineR: A Compound Mining Framework for R. *Bioinformatics* **2008**, *24* (15), 1733–1734.

(212) Horan, K.; Girke, T. *ChemmineOB: R Interface to a Subset of OpenBabel Functionalities*; 2017.

(213) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminformatics* **2011**, *3* (1), 33.

(214) Wickham, H. *Tidyverse: Easily Install and Load the "Tidyverse";* 2017.

(215) Bocker, S.; Liptak, Z. A Fast and Simple Algorithm for the Money Changing Problem. *Algorithmica* **2007**, *48* (4), 413–432.

(216) Bocker, S.; Letzel, M.; Liptak, Z.; Pervukhin, A. SIRIUS: Decomposing Isotope Patterns for Metabolite Identification. *Bioinformatics* **2009**, *25* (2), 218–224.

(217) Bocker, S.; Liptak, Z.; Martin, M.; Pervukhin, A.; Sudek, H. DECOMP—from Interpreting Mass Spectrometry Peaks to Solving the Money Changing Problem. *Bioinformatics* **2008**, *24* (4), 591–593.

(218) Bocker, S.; Letzel, M.; Liptak, Z.; Pervukhin, A. Decomposing Metabolomic Isotope Patterns. In *Proc. of Workshop on Algorithms in Bioinformatics (WABI 2006)*; Lect. Notes Comput. Sci.; Springer, Berlin, 2006; Vol. 4175, pp 12–23.

(219) Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag New York, 2016.

(220) *Inkscape*.

(221) Herman Bergwerf. MolView http://molview.org/ (accessed Apr 23, 2018).

(222) Zachary Charlop-Powers. *Depict*.

(223) Papageorgiou, A.; Stylianou, S. K.; Kaffes, P.; Zouboulis, A. I.; Voutsa, D. Effects of Ozonation Pretreatment on Natural Organic Matter and Wastewater Derived Organic Matter – Possible Implications on the Formation of Ozonation by-Products. *Chemosphere* **2017**, *170*, 33–40.

(224) Zubarev, R.; Mann, M. On the Proper Use of Mass Accuracy in Proteomics. *Mol. Cell. Proteomics* **2006**, *6* (3), 377–381.

(225) Keller, B. O.; Sui, J.; Young, A. B.; Whittal, R. M. Interferences and Contaminants Encountered in Modern Mass Spectrometry. *Anal. Chim. Acta* **2008**, *627* (1), 71–81.

(226) Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Orešič, M. Normalization Method for Metabolomics Data Using Optimal Selection of Multiple Internal Standards. *BMC Bioinformatics* **2007**, *8* (1), 93.

(227) Kind, T.; Fiehn, O. Seven Golden Rules for Heuristic Filtering of Molecular Formulas Obtained by Accurate Mass Spectrometry. *BMC Bioinformatics* **2007**, *8* (1), 105.

(228) Aalizadeh, R.; Schymanski, E.; Thomaidis, N.; Williams, A.; Glowacka, N.; Cirka, L. Merged norman suspect list "SusDat" http://www.norman-network.com/sites/default/files/files/suspectListExchange/NORMAN_SusDat_MergedSuspects24052017.xlsx (accessed Jul 20, 2017).

(229) Schymanski, E. L.; Singer, H. P.; Longrée, P.; Loos, M.; Ruff, M.; Stravs, M. A.; Ripollés Vidal, C.; Hollender, J. Strategies to Characterize Polar Organic Contamination in Wastewater: Exploring the Capability of High Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2014**, *48* (3), 1811–1818.

(230) Trier Xenia; David Lunderberg; Graham Peaslee; Zhanyun Wang. PFAS Suspect List (fluorinated substances) http://www.norman-network.com/sites/default/files/files/suspectListExchange/PFAS%20csv%20database%2011.26.15.csv (accessed Mar 2, 2017).

(231) Grabowska, I. Polychlorinated Biphenyls (Pcbs) in Poland: Occurrence, Determination and Degradation. *Pol. J. Environ. Stud.* **2010**, *19* (1).

(232) Gallert, C.; Winter, J. Bacterial Metabolism in Wastewater Treatment Systems. In *Environmental Biotechnology*; Wiley-VCH Verlag GmbH & Co. KGaA, 2004; pp 1–48.

(233) Bozo Zonja. Unpublished Results: Database of Parents and Biodegradation Transformation Product of Wastewater Treatment. 2017.

(234) Jarusutthirak, C.; Amy, G. Role of Soluble Microbial Products (SMP) in Membrane Fouling and Flux Decline. *Environ. Sci. Technol.* **2006**, *40* (3), 969–974.

(235) Park, M.-H.; Lee, T.-H.; Lee, B.-M.; Hur, J.; Park, D.-H. Spectroscopic and Chromatographic Characterization of Wastewater Organic Matter from a Biological Treatment Plant. *Sensors* **2009**, *10* (1), 254–265.

(236) Federle, T. W.; Itrich, N. R. Fate of Free and Linear Alcohol-Ethoxylate-Derived Fatty Alcohols in Activated Sludge. *Ecotoxicol. Environ. Saf.* **2006**, *64* (1), 30–41.

(237) Vega Morales, T.; Torres Padrón, M. E.; Sosa Ferrera, Z.; Santana Rodríguez, J. J. Determination of Alkylphenol Ethoxylates and Their Degradation Products in Liquid and Solid Samples. *TrAC Trends Anal. Chem.* **2009**, *28* (10), 1186–1200.

(238) Kind, T.; Fiehn, O. Metabolomic Database Annotations via Query of Elemental Compositions: Mass Accuracy Is Insufficient Even at Less than 1 Ppm. *BMC Bioinformatics* **2006**, *7* (1), 234.

(239) Carballa, M.; Omil, F.; Lema, J. M. Removal of Cosmetic Ingredients and Pharmaceuticals in Sewage Primary Treatment. *Water Res.* **2005**, *39* (19), 4790–4796.

(240) Katsoyiannis, A.; Samara, C. The Fate of Dissolved Organic Carbon (DOC) in the Wastewater Treatment Process and Its Importance in the Removal of Wastewater Contaminants. *Environ. Sci. Pollut. Res. - Int.* **2007**, *14* (5), 284–292.

(241) Maggs, J. L.; Pirmohamed, M.; Kitteringham, N. R.; Park, B. K. Characterization of the Metabolites of Carbamazepine in Patient Urine by Liquid Chromatography/Mass Spectrometry. *Drug Metab. Dispos.* **1997**, *25* (3), 275–280.

(242) Aymerich, I.; Acuña, V.; Ort, C.; Rodríguez-Roda, I.; Corominas, L. Fate of Organic Microcontaminants in Wastewater Treatment and River Systems: An Uncertainty Assessment in View of Sampling Strategy, and Compound Consumption Rate and Degradability. *Water Res.* **2017**, *125*, 152–161.

(243) Bellona, C.; Drewes, J. E.; Xu, P.; Amy, G. Factors Affecting the Rejection of Organic Solutes during NF/RO Treatment—a Literature Review. *Water Res.* **2004**, *38* (12), 2795–2809.

(244) Brenton, A. G.; Godfrey, A. R. Accurate Mass Measurement: Terminology and Treatment of Data. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1821–1835.

# 7 APPENDICES

# 1 DETECTED IS AND TESTS IN NON-TARGETED ANALYSIS OF BIOLOGICAL TREATMENT

## Detected IS

**Table 7.1 Properties of measured IS as prevalent ionization mode, conventional name, molecular formula, average retention time, intensity measured in the blank and the CV of three replicates, the monoisotopic mass, the mass deviation from the theoretical in mDa and in ppm after the *m/z* correction and also the mass deviation in ppm before the *m/z* correction in R.**

| ESI mode | Name | Formula | RT | Intensity | CV | Mass [Da] | Mass error, corr [mDa] | Mass error, corr [ppm] | Mass error, not corr [ppm] |
|---|---|---|---|---|---|---|---|---|---|
| PI | Acetaminophen-$d_4$ | $D_4C_8H_5NO_2$ | 2.19 | 9.64E+05 | 7.86 | 155.0884 | -0.1174 | -0.75 | -1.13 |
| | Antipyrine-$d_3$ | $D_3C_{11}H_9N_2O$ | 3.59 | 1.64E+07 | 8.44 | 191.1138 | 0.0427 | 0.22 | -0.67 |
| | Atenolol-$d_7$ | $D_7C_{14}H_{15}N_2O_3$ | 2.24 | 5.79E+06 | 2.70 | 273.2070 | -0.0796 | -0.29 | -0.64 |
| | azaperone-$d_4$ | $D_4C_{19}FH_{18}N_3O$ | 3.51 | 2.10E+06 | 14.16 | 331.1998 | 0.0106 | 0.03 | -0.04 |
| | Bezafibrate-$d_4$ | $D_4C_{19}ClH_{16}NO_4$ | 6.72 | 1.55E+06 | 15.53 | 365.1332 | 0.0687 | 0.19 | -0.16 |
| | Carbamazepine-$d_{10}$ | $D_{10}C_{15}H_2N_2O$ | 5.46 | 9.95E+06 | 12.70 | 246.1577 | -0.0184 | -0.07 | -0.93 |
| | cimetidine-$d_3$ | $D_3C_{10}H_{13}N_6S$ | 2.24 | 4.20E+06 | 2.41 | 255.1345 | -0.0414 | -0.16 | -0.50 |
| | Citalopram-$d_4$ | $D_4C_{20}FH_{17}N_2O$ | 5.26 | 1.02E+07 | 13.76 | 328.1889 | 0.0388 | 0.12 | -0.40 |
| | Cprofloxacin-$d_8$ | $D_8C_{17}FH_{10}N_3O_3$ | 3.36 | 1.88E+04 | 18.38 | 339.1834 | -0.0052 | -0.02 | -1.19 |
| | Diclofenac-$d_4$ | $D_4C_{14}Cl_2H_7NO_2$ | 7.87 | 4.35E+05 | 25.29 | 299.0418 | -0.0582 | -0.19 | -0.19 |
| | Dexamethasone-$d_4$ | $D_4C_{22}FH_{25}O_5$ | 5.58 | 3.13E+05 | 22.02 | 396.225 | -0.1089 | -0.27 | -0.27 |
| | diazepam-$d_5$ | $D_5C_{16}ClH_8N_2O$ | 6.67 | 1.69E+07 | 11.83 | 289.103 | 0.0630 | 0.22 | -0.04 |
| | Diltiazem-$d_3$ | $D_3C_{22}H_{23}N_2O_4S$ | 5.34 | 9.38E+06 | 15.29 | 417.1802 | 0.0657 | 0.16 | -0.87 |
| | Fluoxetine-$d_5$ | $D_5C_{17}F_3H_{13}NO$ | 6.13 | 3.46E+06 | 22.22 | 314.1654 | -0.0077 | -0.02 | -0.13 |
| | Glyburide-$d_3$ | $D_3C_{23}ClH_{25}N_3O_5S$ | 7.77 | 6.04E+05 | 22.52 | 496.1627 | 0.0396 | 0.08 | 0.17 |
| | Indomethacin-$d_4$ | $D_4C_{19}ClH_{12}NO_4$ | 7.85 | 4.29E+05 | 13.27 | 361.1019 | -0.0308 | -0.09 | -0.02 |
| | Ketoprofen-$d_3$ | $D_3C_{16}H_{11}O_3$ | 6.61 | 9.64E+05 | 11.95 | 257.1131 | -0.0452 | -0.17 | -0.75 |
| | Lincomycin-$d_3$ | $D_3C_{18}H_{31}N_2O_6S$ | 2.91 | 1.84E+06 | 8.39 | 409.2326 | 0.001 | 0 | -0.29 |

| ESI mode | Name | Formula | RT | Intensity | CV | Mass [Da] | Mass error, corr [mDa] | Mass error, corr [ppm] | Mass error, not corr [ppm] |
|---|---|---|---|---|---|---|---|---|---|
| | meloxicam-$d_3$ | $D_3C_{14}H_{10}N_3O_4S_2$ | 6.72 | 5.52E+06 | 15.86 | 354.0536 | 0.0758 | 0.21 | 0.17 |
| | Ofloxacin-$d_3$ | $D_3C_{18}FH_{17}N_3O_4$ | 3.31 | 1.58E+05 | 16.84 | 364.1626 | 0.0397 | 0.11 | -0.25 |
| | Ronidazole-$d_3$ | $D_3C_6H_5N_4O_4$ | 2.62 | 9.60E+04 | 29.13 | 203.0734 | -0.0395 | -0.19 | -0.09 |
| | Sertraline-$d_3$ | $D_3C_{17}Cl_2H_{14}N$ | 6.21 | 1.04E+06 | 24.71 | 308.0926 | 0.0564 | 0.18 | 0.11 |
| | Simvastatin-$d_6$ | $D_6C_{25}H_{32}O_5$ | 9.84 | 7.50E+04 | 57.74 | 424.3096 | -0.3364 | -0.79 | 1.15 |
| | sulfamethoxazole-$d_4$ | $D_4C_{10}H_7N_3O_3S$ | 4.37 | 1.92E+06 | 11.56 | 257.0772 | -0.0080 | -0.03 | -0.92 |
| | Sulfapyridine-N-Acetyl-$d_4$ | $D_4C_{13}H_9N_3O_3S$ | 3.26 | 1.03E+05 | 17.18 | 295.0929 | -0.3205 | -1.08 | -1.19 |
| | Trimethoprim-$d_3$ | $D_3C_{14}H_{15}N_4O_3$ | 3.22 | 1.75E+07 | 11.73 | 293.1567 | 0.0264 | 0.09 | -0.19 |
| | Valsartan-$d_8$ | $D_8C_{24}H_{21}N_5O_3$ | 7.01 | 1.37E+06 | 17.82 | 443.2773 | -0.0406 | -0.09 | 0.03 |
| | venlafaxine-$d_6$ | $D_6C_{17}H_{21}NO_2$ | 4.51 | 1.05E+07 | 15.86 | 283.2418 | 0.0474 | 0.17 | -0.65 |
| | Verapamil-$d_6$ | $D_6C_{27}H_{32}N_2O_4$ | 5.85 | 2.14E+06 | 26.23 | 460.3208 | 0.0168 | 0.04 | -0.13 |
| | Warfarin-$d_5$ | $D_5C_{19}H_{11}O_4$ | 7.14 | 8.49E+06 | 13.89 | 313.1362 | -0.1319 | -0.42 | -0.18 |
| | Xylazine-$d_6$ | $D_6C_{12}H_{10}N_2S$ | 3.83 | 1.42E+07 | 16.12 | 226.1411 | -0.0107 | -0.05 | -0.65 |
| NI | Bezafibrate-$d_4$ | $D_4C_{19}ClH_{16}NO_4$ | 5.70 | 4.88E+05 | 9.98 | 365.1332 | 0.0056 | 0.02 | 0.92 |
| | Chloramphenicol-$d_5$ | $C_{11}D_5H_7Cl_2N_2O_5$ | 3.76 | 8.18E+05 | 12.64 | 327.0432 | 0.6003 | 1.84 | 2.60 |
| | Citalopram-$d_4$ | $D_4C_{20}FH_{17}N_2O$ | 8.88 | 1.99E+06 | 10.76 | 328.1889 | -1.0516 | -3.21 | -3.86 |
| | Glyburide-$d_3$ | $D_3C_{23}ClH_{25}N_3O_5S$ | 6.69 | 9.81E+05 | 12.33 | 496.1627 | 0.7229 | 1.46 | 0.35 |
| | meloxicam-$d_3$ | $D_3C_{14}H_{10}N_3O_4S_2$ | 5.60 | 5.33E+05 | 1.72 | 354.0536 | 0.3688 | 1.04 | 0.50 |
| | Valsartan-$d_8$ | $D_8C_{24}H_{21}N_5O_3$ | 5.99 | 1.19E+06 | 2.14 | 443.2773 | 0.7001 | 1.58 | 0.57 |
| | Warfarin-$d_5$ | $D_5C_{19}H_{11}O_4$ | 5.98 | 5.42E+06 | 6.67 | 313.1362 | -0.0315 | -0.1 | 0.19 |

## Testing of formula prediction

**Table 7.2 Test of formula prediction using pharmaceuticals measured at the same spectrometric settings as wastewater samples showing the pharmaceuticals' ionization mode, trivial name, molecular formula of neutral compound, *m/z* of $[M+H]^+$ or $[M-H]^-$ ion depending on ionization, status whether the formula was recognized correctly and the corresponding isotopic pattern score, mass deviation from theoretical mass, and normalized intensity of the monoisotopic peak in respective ionization mode.**

| Polarity | Name | Molecular formula | $[M+H]^{\pm}$ *m/z* [Da] | Prediction status | Isotopic pattern score% | Mass Error ppm | Signal intensity, normalized |
|---|---|---|---|---|---|---|---|
| PI | Propyphenazone | $C_{14}H_{18}N_2O$ | 231.1489 | correct | 0.98 | 1.3 | 3.27 |
| | Verapamil | $C_{27}H_{38}N_2O_4$ | 455.2902 | correct | 0.97 | 0.5 | 2.42 |
| | Irbesartan | $C_{25}H_{28}N_6O$ | 429.2396 | correct | 0.96 | 0.4 | 2.37 |
| | Loratadine | $C_{22}ClH_{23}N_2O_2$ | 383.1518 | correct | 0.93 | 0.7 | 2.37 |
| | Diazepam | $C_{16}ClH_{13}N_2O$ | 285.0787 | correct | 0.94 | 0.9 | 2.30 |
| | Diltiazem | $C_{22}H_{26}N_2O_4S$ | 415.1684 | correct | 0.94 | 0.4 | 2.26 |
| | Norverapamil | $C_{26}H_{36}N_2O_4$ | 441.2746 | correct | 0.98 | 0.5 | 1.94 |
| | Xylazine | $C_{12}H_{16}N_2S$ | 221.1103 | correct | 0.96 | 1.7 | 1.85 |
| | Carbamazepine | $C_{15}H_{12}N_2O$ | 237.1020 | correct | 0.98 | 1.1 | 1.81 |
| | Clopidogrel | $C_{16}ClH_{16}NO_2S$ | 322.0660 | correct | 0.93 | 0.8 | 1.72 |
| | Alprazolam | $C_{17}ClH_{13}N_4$ | 309.0900 | correct | 0.95 | 0.6 | 1.50 |
| | Albendazole | $C_{12}H_{15}N_3O_2S$ | 266.0955 | correct | 0.96 | 1.1 | 1.50 |
| | Trimethoprim | $C_{14}H_{18}N_4O_3$ | 291.1447 | correct | 0.96 | 1.5 | 1.43 |
| | Propranolol | $C_{16}H_{21}NO_2$ | 260.1642 | correct | 0.98 | 1.1 | 1.41 |
| | 10,11-epoxyCBZ | $C_{15}H_{12}N_2O_2$ | 253.0968 | correct | 0.97 | 1.3 | 1.35 |
| | Thiabendazole | $C_{10}H_7N_3S$ | 202.0430 | correct | 0.92 | 1.7 | 1.29 |
| | Levamisole | $C_{11}H_{12}N_2S$ | 205.0790 | correct | 0.92 | 1.8 | 1.27 |
| | Acridone | $C_{13}H_9NO$ | 196.0754 | correct | 0.98 | 1.3 | 1.17 |
| | Tamsulosin | $C_{20}H_{28}N_2O_5S$ | 409.1787 | correct | 0.94 | 1.2 | 1.08 |

| Polarity | Name | Molecular formula | $[M+H]^{\pm}$ *m/z* [Da] | Prediction status | Isotopic pattern score% | Mass Error ppm | Signal intensity, normalized |
|---|---|---|---|---|---|---|---|
| | Piroxicam | $C_{15}H_{13}N_3O_4S$ | 332.0696 | correct | 0.94 | 1.0 | 1.06 |
| | Carazolol | $C_{18}H_{22}N_2O_2$ | 299.1750 | correct | 0.97 | 1.2 | 1.04 |
| | Desloratadine | $C_{19}ClH_{19}N_2$ | 311.1308 | correct | 0.95 | 0.5 | 0.93 |
| | Meloxicam | $C_{14}H_{13}N_3O_4S_2$ | 352.0418 | correct | 0.80 | 0.6 | 0.71 |
| | Codeine | $C_{18}H_{21}NO_3$ | 300.1590 | correct | 0.98 | 1.3 | 0.63 |
| | Oxycodone | $C_{18}H_{21}NO_4$ | 316.1537 | correct | 0.97 | 1.9 | 0.58 |
| | Nadolol | $C_{17}H_{27}NO_4$ | 310.2009 | correct | 0.97 | 1.2 | 0.39 |
| | Torasemide | $C_{16}H_{20}N_4O_3S$ | 349.1324 | correct | 0.94 | 1.3 | 0.37 |
| | Cimetidine | $C_{10}H_{16}N_6S$ | 253.1225 | correct | 0.83 | 1.8 | 0.32 |
| | Warfarin | $C_{19}H_{16}O_4$ | 309.1122 | correct | 0.96 | 0.2 | 0.28 |
| | Losartan | $C_{22}ClH_{23}N_6O$ | 423.1693 | correct | 0.91 | 0.4 | 0.18 |
| | Sulfamethoxazole | $C_{10}H_{11}N_3O_3S$ | 254.0589 | correct | 0.95 | 1.7 | -0.01 |
| | Sotalol | $C_{12}H_{20}N_2O_3S$ | 273.1263 | correct | 0.96 | 1.6 | -0.08 |
| | Sertraline | $C_{17}Cl_2H_{17}N$ | 306.0812 | correct | 0.83 | 0.3 | -0.20 |
| | Eryhromycin | $C_{37}H_{67}NO_{13}$ | 734.4688 | wrong | 0.93 | 0.9 | -0.22 |
| | Ketoprofen | $C_{16}H_{14}O_3$ | 255.1013 | correct | 0.98 | 1.0 | -0.24 |
| | Salbutamol | $C_{13}H_{21}NO_3$ | 240.1589 | correct | 0.98 | 2.1 | -0.28 |
| | Clarithromycin | $C_{38}H_{69}NO_{13}$ | 748.4843 | wrong | 0.94 | 1.9 | -0.29 |
| | Glibenclamide | $C_{23}ClH_{28}N_3O_5S$ | 494.1512 | correct | 0.89 | 0.1 | -0.34 |
| | Lorazepam | $C_{15}Cl_2H_{10}N_2O_2$ | 321.0190 | correct | 0.84 | 0.8 | -0.35 |
| | Famotidine | $C_8H_{15}N_7O_2S_3$ | 338.0517 | correct | 0.92 | 1.4 | -0.47 |
| | Bezafibrate | $C_{19}ClH_{20}NO_4$ | 362.1149 | correct | 0.83 | 1.3 | -0.51 |
| | Valsartan | $C_{24}H_{29}N_5O_3$ | 436.2339 | correct | 0.93 | 0.8 | -0.59 |
| | Indomethacin | $C_{19}ClH_{16}NO_4$ | 358.0839 | correct | 0.84 | 0.4 | -0.60 |
| | Metronidazole | $C_6H_9N_3O_3$ | 172.0713 | correct | 0.97 | 1.9 | -0.61 |
| | Dimetridazole | $C_5H_7N_3O_2$ | 142.0609 | correct | 0.98 | 1.7 | -0.63 |
| | Tenoxicam | $C_{13}H_{11}N_3O_4S_2$ | 338.0259 | correct | 0.91 | 1.3 | -0.68 |

| Polarity | Name | Molecular formula | $[M+H]^{\pm}$ *m/z* [Da] | Prediction status | Isotopic pattern score% | Mass Error ppm | Signal intensity, normalized |
|---|---|---|---|---|---|---|---|
| | Ronidazole | $C_6H_8N_4O_4$ | 201.0614 | correct | 0.97 | 1.9 | -0.71 |
| | Gemfibrozil | $C_{15}H_{22}O_3$ | 251.1638 | correct | 0.93 | 1.5 | -0.78 |
| | Venlafaxine | $C_{17}H_{27}NO_2$ | 278.2113 | correct | 0.91 | 0.7 | -0.81 |
| | Metronidazole-OH | $C_6H_9N_3O_4$ | 188.0663 | correct | 0.95 | 1.3 | -0.83 |
| | Cefalexin | $C_{16}H_{17}N_3O_4S$ | 348.1007 | wrong | 0.91 | 1.0 | -0.84 |
| | Azithromycin | $C_{38}H_{72}N_2O_{12}$ | 749.5159 | wrong | 0.90 | 2.0 | -0.84 |
| | Metoprolol | $C_{15}H_{25}NO_3$ | 268.1903 | wrong | NA | NA | -0.85 |
| NI | Irbesartan | $C_{25}H_{28}N_6O$ | 427.2254 | correct | 0.93 | 0.6 | 0.05 |
| | Acridone | $C_{13}H_9NO$ | 194.0613 | correct | 0.99 | 0.8 | 0.17 |
| | Warfarin | $C_{19}H_{16}O_4$ | 307.0980 | correct | 0.96 | 1.2 | -0.06 |
| | Losartan | $C_{22}ClH_{23}N_6O$ | 421.1555 | correct | 0.89 | 1.3 | -0.21 |
| | Valsartan | $C_{24}H_{29}N_5O_3$ | 434.2201 | correct | 0.94 | 0.7 | -0.34 |
| | Meloxicam | $C_{14}H_{13}N_3O_4S_2$ | 350.0279 | correct | 0.82 | 1.3 | -0.57 |
| | Glibenclamide | $C_{23}ClH_{28}N_3O_5S$ | 492.1371 | wrong | 0.84 | 1.5 | -0.59 |
| | Torasemide | $C_{16}H_{20}N_4O_3S$ | 347.1188 | wrong | 0.94 | 1.9 | -0.61 |
| | Tamsulosin | $C_{20}H_{28}N_2O_5S$ | 407.1649 | wrong | 0.90 | 2.1 | -0.63 |
| | Bezafibrate | $C_{19}ClH_{20}NO_4$ | 360.1012 | correct | 0.80 | 1.2 | -0.65 |
| | 10,11-epoxyCBZ | $C_{15}H_{12}N_2O_2$ | 251.0828 | correct | 0.96 | 0.9 | -0.67 |
| | Pravastatin | $C_{23}H_{36}O_7$ | 423.2393 | correct | 0.92 | 1.2 | -0.68 |
| | Amlodipine | $C_{20}ClH_{25}N_2O_5$ | 407.1384 | correct | 0.77 | 1.2 | -0.69 |
| | Furosemide | $C_{12}ClH_{11}N_2O_5S$ | 329.0009 | correct | 0.82 | 1.4 | -0.73 |
| | Lorazepam | $C_{15}Cl_2H_{10}N_2O_2$ | 319.0051 | correct | 0.82 | 1.4 | -0.70 |
| | Piroxicam | $C_{15}H_{13}N_3O_4S$ | 330.0558 | wrong | 0.89 | 1.9 | -0.73 |
| | Carazolol | $C_{18}H_{22}N_2O_2$ | 297.1610 | correct | 0.94 | 0.6 | -0.72 |
| | HCTZ | $C_7ClH_8N_3O_4S_2$ | 295.9575 | correct | 0.74 | 0.9 | -0.71 |
| | Thiabendazole | $C_{10}H_7N_3S$ | 200.0290 | correct | 0.92 | 0.8 | -0.70 |
| | Gemfibrozil | $C_{15}H_{22}O_3$ | 249.1500 | correct | 0.96 | 1.4 | -0.68 |
| | Ranitidine | $C_{13}H_{22}N_4O_3S$ | 313.1345 | correct | 0.90 | 1.5 | -0.76 |

| Polarity | Name | Molecular formula | $[M+H]^{\pm}$ *m/z* [Da] | Prediction status | Isotopic pattern score% | Mass Error ppm | Signal intensity, normalized |
|---|---|---|---|---|---|---|---|
| | Salicylic acid | $C_7H_6O_3$ | 137.0244 | correct | 0.98 | 0.5 | -0.73 |
| | Clarithromycin | $C_{38}H_{69}NO_{13}$ | 746.4698 | correct | 0.95 | 0.3 | -0.78 |
| | Salbutamol | $C_{13}H_{21}NO_3$ | 238.1450 | correct | 0.96 | 0.7 | -0.80 |
| | Sulfamethoxazole | $C_{10}H_{11}N_3O_3S$ | 252.0450 | correct | 0.91 | 0.8 | -0.81 |
| | Azithromycin | $C_{38}H_{72}N_2O_{12}$ | 747.5013 | wrong | 0.77 | 0.7 | -0.84 |

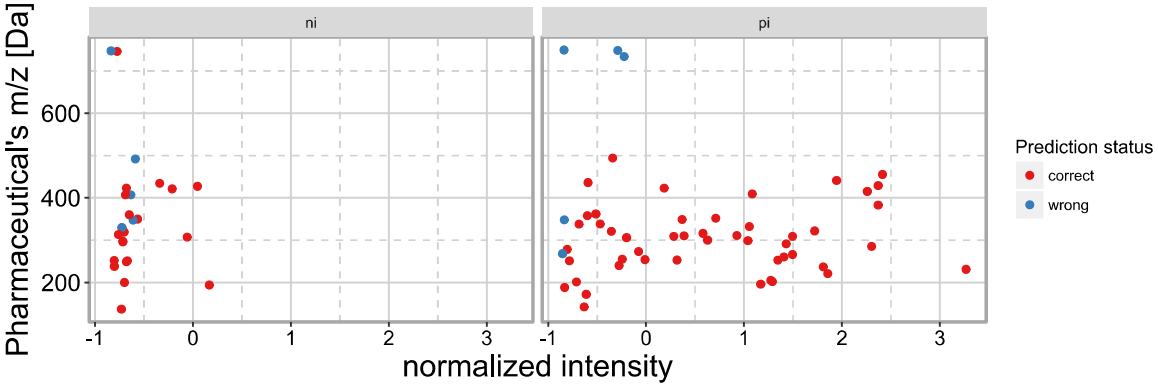



**Figure 7.1 The test of formula prediction for pharmaceuticals depending on *m/z* and normalized intensity. The signals are shown in NI (left) or PI (right) modes.**

# 2 IS IN SUSPECT SCREENING OF BIOLOGICAL TREATMENT AND DETECTED FRAGMENTS

**Table 7.3 IS used for intensity correction of biodegradation and disinfection TPs. Log P values predicted with OpenBabel [212,213], and median RT extracted from all processed samples.**

| Compound Name | Formula | Monoisotopic mass [Da] | Log P | RT [min] |
|---|---|---|---|---|
| Acetaminophen-$d_4$ | $D_4C_8H_5NO_2$ | 155.09 | 1.42 | 1.29 |
| Amlodipine-$d_4$ maleic acid | $D_4C_{20}ClH_{21}N_2O_5$ | 412.17 | 3.01 | 5.04 |
| Antipyrine-$d_3$ | $D_3C_{11}H_9N_2O$ | 191.11 | 1.48 | 2.83 |
| Atenolol-$d_7$ | $D_7C_{14}H_{15}N_2O_3$ | 273.21 | 1.54 | 1.41 |
| azaperone-$d_4$ | $D_4C_{19}FH_{18}N_3O$ | 331.20 | 3.01 | 2.92 |
| Bezafibrate-$d_4$ | $D_4C_{19}ClH_{16}NO_4$ | 365.13 | 3.95 | 5.72 |
| Carbamazepine-$d_{10}$ | $D_{10}C_{15}H_2N_2O$ | 246.16 | 4.15 | 4.56 |
| cimetidine-$d_3$ (N-methyl-$d_3$) | $D_3C_{10}H_{13}N_6S$ | 255.13 | 1.38 | 1.37 |
| Ciprofloxacin-$d_8$ | $D_8C_{17}FH_{10}N_3O_3$ | 339.18 | 1.98 | 2.73 |
| Diclofenac-$d_4$ | $D_4C_{14}Cl_2H_7NO_2$ | 299.04 | 4.44 | 6.72 |
| diazepam-$d_5$ | $D_5C_{16}ClH_8N_2O$ | 289.10 | 2.65 | 5.64 |
| Diltiazem-$d_3$ | $D_3C_{22}H_{23}N_2O_4S$ | 417.18 | 3.14 | 4.58 |
| Fluoxetine-$d_5$ | $D_5C_{17}F_3H_{13}NO$ | 314.17 | 4.83 | 5.33 |
| Ibuprofen-$d_3$ | $D_3C_{13}H_{15}O_2$ | 209.15 | 3.07 | 1.97 |
| Indomethacin-$d_4$ | $D_4C_{19}ClH_{12}NO_4$ | 361.10 | 3.93 | 6.74 |
| meloxicam-$d_3$ | $D_3C_{14}H_{10}N_3O_4S_2$ | 354.05 | 3.04 | 5.62 |
| Ofloxacin-$d_3$ | $D_3C_{18}FH_{17}N_3O_4$ | 364.16 | 1.55 | 2.66 |
| Ronidazole-$d_3$ | $D_3C_6H_5N_4O_4$ | 203.07 | 1.15 | 1.66 |
| Sertraline-$d_3$ | $D_3C_{17}Cl_2H_{14}N$ | 308.09 | 5.57 | 5.40 |

| Compound Name | Formula | Monoisotopic mass [Da] | Log P | RT [min] |
|---|---|---|---|---|
| Simvastatin-$d_6$ | $D_6C_{25}H_{32}O_5$ | 424.31 | 4.59 | 8.42 |
| sulfamethoxazole-$d_4$ | $D_4C_{10}H_7N_3O_3S$ | 257.08 | 3.10 | 3.39 |
| Sulfapyridine-N-Acetyl-$d_4$ | $D_4C_{13}H_9N_3O_3S$ | 295.09 | 3.07 | 2.54 |
| Trimethoprim-$d_3$ | $D_3C_{14}H_{15}N_4O_3$ | 293.16 | 2.42 | 2.53 |
| Valsartan-$d_8$ | $D_8C_{24}H_{21}N_5O_3$ | 443.28 | 4.16 | 6.01 |
| venlafaxine-$d_6$ | $D_6C_{17}H_{21}NO_2$ | 283.24 | 3.04 | 3.78 |
| Verapamil-$d_6$ | $D_6C_{27}H_{32}N_2O_4$ | 460.32 | 5.09 | 5.07 |
| Warfarin-$d_5$ | $D_5C_{19}H_{11}O_4$ | 313.14 | 3.61 | 5.99 |
| Xylazine-$d_6$ | $D_6C_{12}H_{10}N_2S$ | 226.14 | 2.72 | 3.07 |

**Table 7.4 Detected fragments of parents and TPs in multistage WWTP suspect screening in PI mode. Parent compounds with corresponding IS were confirmed using retention time.**

| Compound Name | Ion m/z [Da] | Retention Time [min] | | | | Fragments |
|---|---|---|---|---|---|---|
| 17beta-estradiol | 273.1848 | 8.4 | 255.174 | | | |
| 17beta-estradiol-1 | 279.1591 | 5.3 | | | | |
| 17beta-estradiol-2 | 289.1795 | 7.1 | 243.1739 | 271.1688 | | |
| 17beta-estradiol-3 | 289.1795 | 7.1 | | 271.1688 | | |
| acebutolol | 337.2118 | 3.2 | 319.2015 | 260.1279 | 116.1071 | |
| acebutolol-15 | 295.165 | 3.3 | 278.1386 | 277.1547 | 194.1175 | 152.0706 |
| acebutolol-16 | 222.1122 | 3.3 | | | | |
| acebutolol-17 | 277.1541 | 3.3 | 277.1545 | 152.0706 | | |
| acebutolol-18 | 309.1806 | 2.3 | 298.1704 | 232.0967 | 116.1071 | |
| acebutolol-7 | 343.1877 | 5.2 | | | | |

| Compound Name | Ion m/z [Da] | Retention Time [min] | | | | Fragments |
|---|---|---|---|---|---|---|
| acebutolol-8 | 368.2176 | 2.8 | 219.1123 | 59.0492 | | |
| acebutolol-9 | 368.2176 | 4.0 | 152.0703 | | | |
| Acetaminophen | 152.0706 | 2.0 | 93.03 | 110.06 | 120.04 | 136.04 |
| Acetaminophen-1 | 110.0603 | 0.67 | 92.05 | 93.03 | 110.06 | |
| atenolol | 267.1704 | 1.48 | 116.107 | 162.0911 | 190.086 | 208.0966 |
| Atenolol | 267.1704 | 1.48 | 116.107 | 162.0911 | 190.086 | 208.0966 |
| atenolol-1 | 268.1541 | 2.6 | 250.1434 | 226.107 | 208.0965 | 196.07 |
| Atenolol-1 | 268.1541 | 2.34 | 98.1 | 116.11 | 121.06 | |
| Bezafibrate | 362.1151 | 5.7 | 276.08 | 316.11 | 344.1 | |
| Carbamazepine | 237.1019 | 4.7 | 194.1 | 220.08 | | |
| Carbamazepine-2 | 271.1074 | 3.37 | 79.054 | | | |
| Carbamazepine-3 | 180.0805 | 2.5 | 91.05 | 122.06 | | |
| Diazepam | 285.0787 | 5.7 | 182.0365 | | | |
| Diazepam-1 | 287.0578 | 5.1 | 257.0475 | 273.0604 | | |
| estradiol | 273.1848 | 8.3 | 255.1745 | 175.1116 | | |
| estradiol-1 | 277.143 | 5.7 | 259.1325 | 231.1376 | 121.1012 | |
| iopamidol | 777.8623 | 1.4 | 759.8512 | 686.7989 | 631.9391 | |
| iopamidol-6 | 777.8623 | 1.4 | 759.8512 | 686.7989 | 631.9391 | |
| Ketoprofen-1 | 242.1017 | 3.3 | 209.1 | | | |
| Ofloxacin | 362.1508 | 2.7 | 316.11 | | | |
| propylphenazone-2 | 281.1495 | 2.1 | 263.1391 | 89.06 | | |
| propylphenazone-3 | 233.1282 | 3.1 | 215.1178 | | | |
| propylphenazone-7 | 251.1387 | 3.2 | 219.1126 | 115.0754 | | |
| terbutryne | 242.1431 | 5 | 186.0806 | | | |
| Valsartan | 436.2343 | 6 | 235.1 | 306.1 | 418.22 | |
| Valsartan-1 | 336.1816 | 4.9 | 235.1 | 252.12 | | |
| venlafaxine | 278.2113 | 3.8 | 260.2 | | | |

| Compound Name | Ion m/z [Da] | Retention Time [min] | | | | Fragments |
|---|---|---|---|---|---|---|
| venlafaxine-1 | 264.1956 | 2.9 | 246.19 | | | |
| venlafaxine-3 | 235.1325 | 2.7 | 95.08575 | 81.07005 | 107.0857 | 125.0962 |
| venlafaxine-4 | 265.1433 | 2.4 | 193.049 | | | |
| venlafaxine-d6 | 284.2488 | 3.8 | 266.24 | | | |
| Verapamil | 455.29 | 5.2 | 84.08087 | 109.1012 | | |
| Verapamil-1 | 291.2064 | 4 | 260.16 | | | |

# 3 ADDITIONAL SUSPECT SCREENING REMOVAL PROFILES AND MOLECULAR STRUCTURES OF PARENTS AND TPS

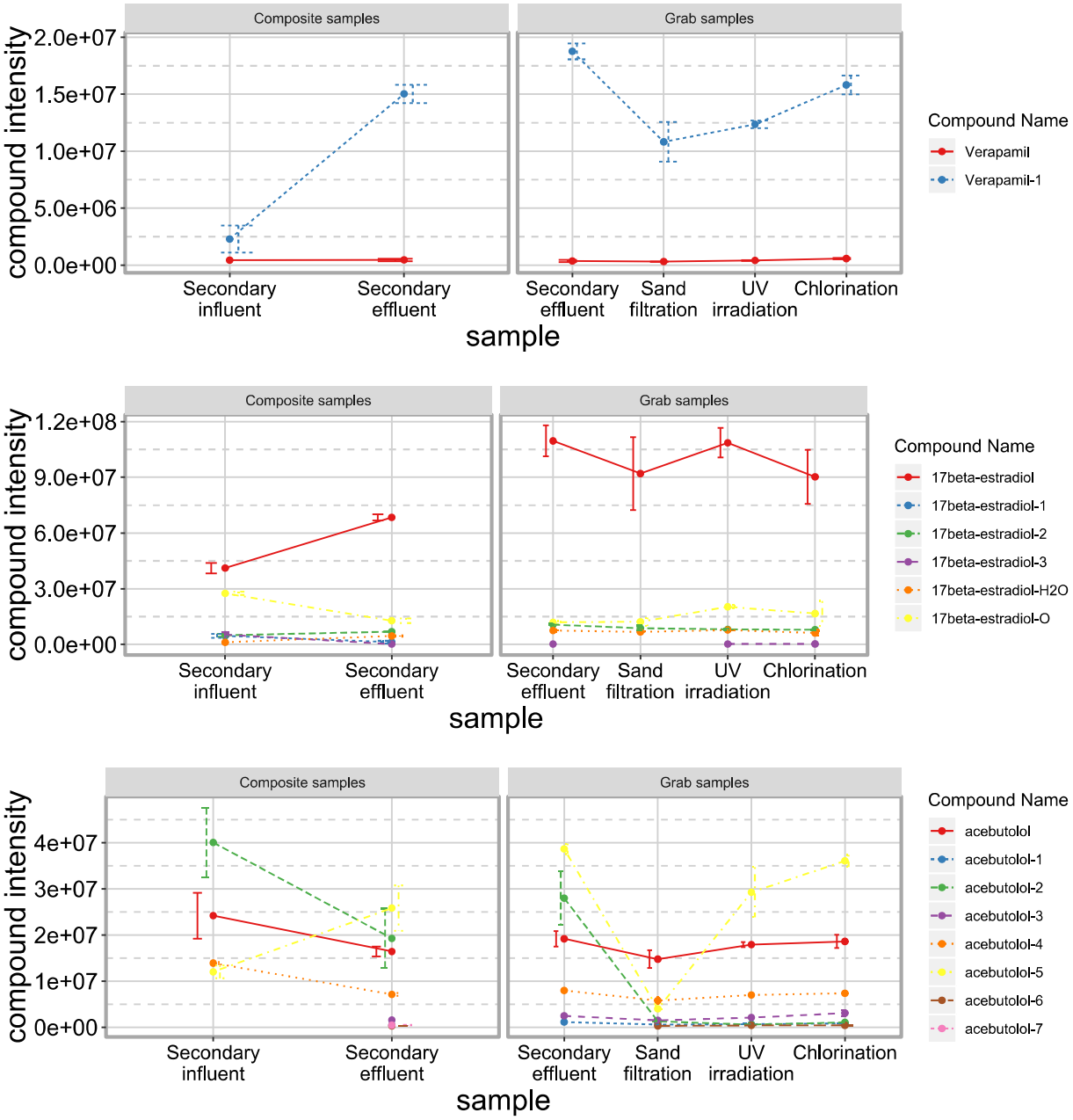



**Figure 7.2 Removal profiles of corrected intensity for Verapamil, 17β-estradiol, Acebutolol, and their confirmed and tentative TPs in a multi-stage wastewater treatment system. Each series depicts two profiles divided due to the sampling mode.**

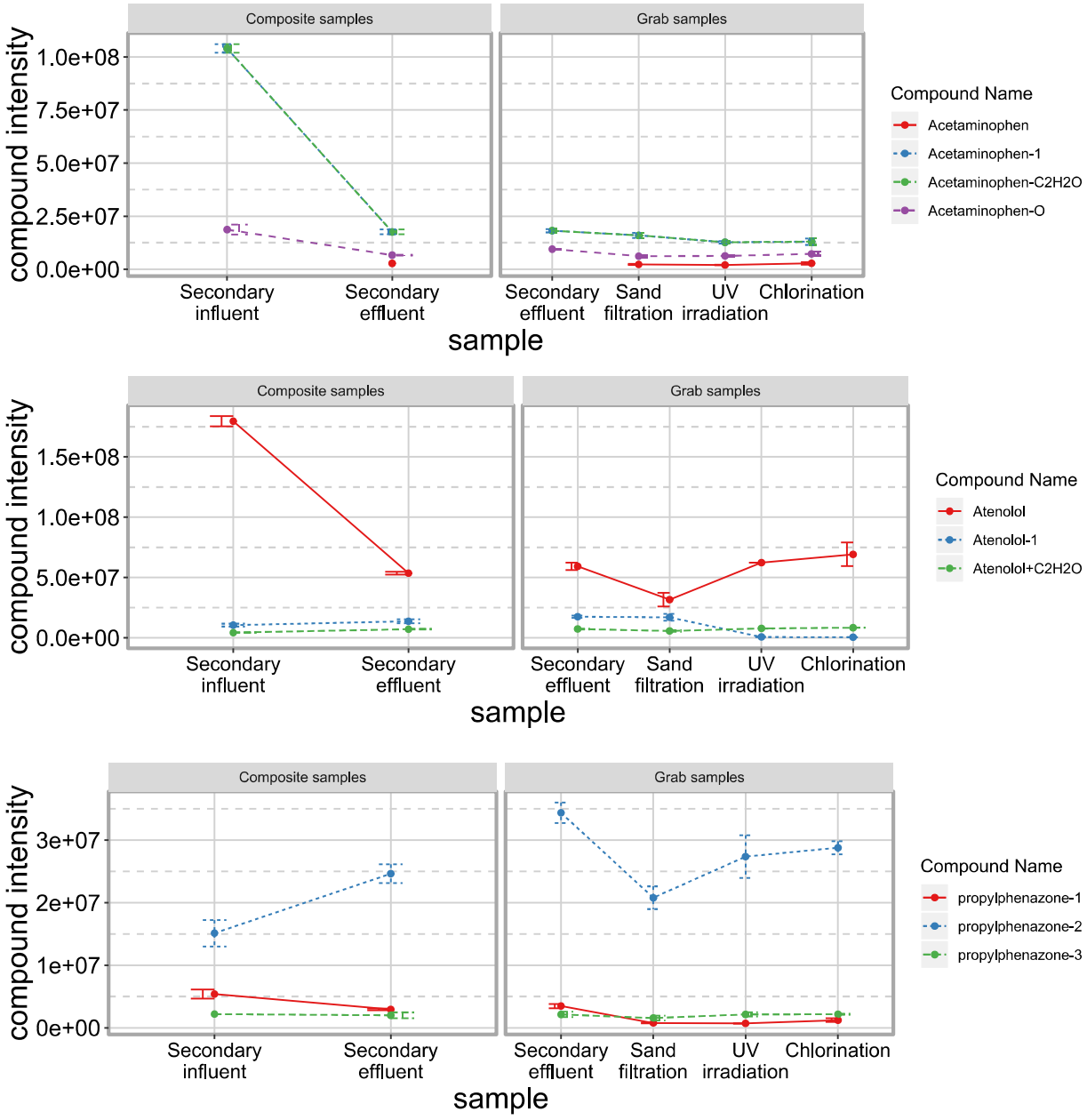



**Figure 7.3 Removal profiles of corrected intensity for Acetaminophen, Atenolol, their confirmed and tentative TPs and TPs of Propylphenazone in a multi-stage wastewater treatment system. Each series depicts two profiles divided due to the sampling mode.**

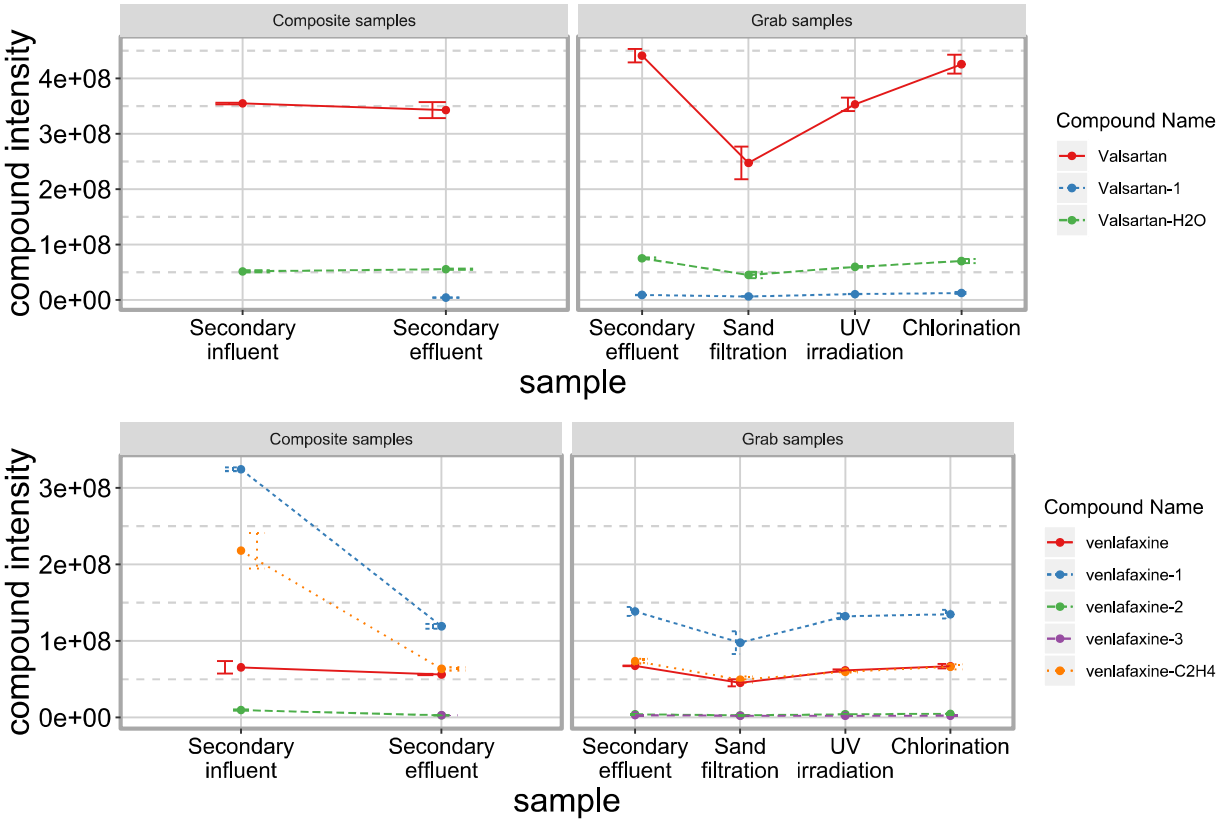



**Figure 7.4 Removal profiles of corrected intensity for Valsartan, Venlafaxine, and their confirmed and tentative TPs in a multi-stage wastewater treatment system. Each series depicts two profiles divided due to the sampling mode.**

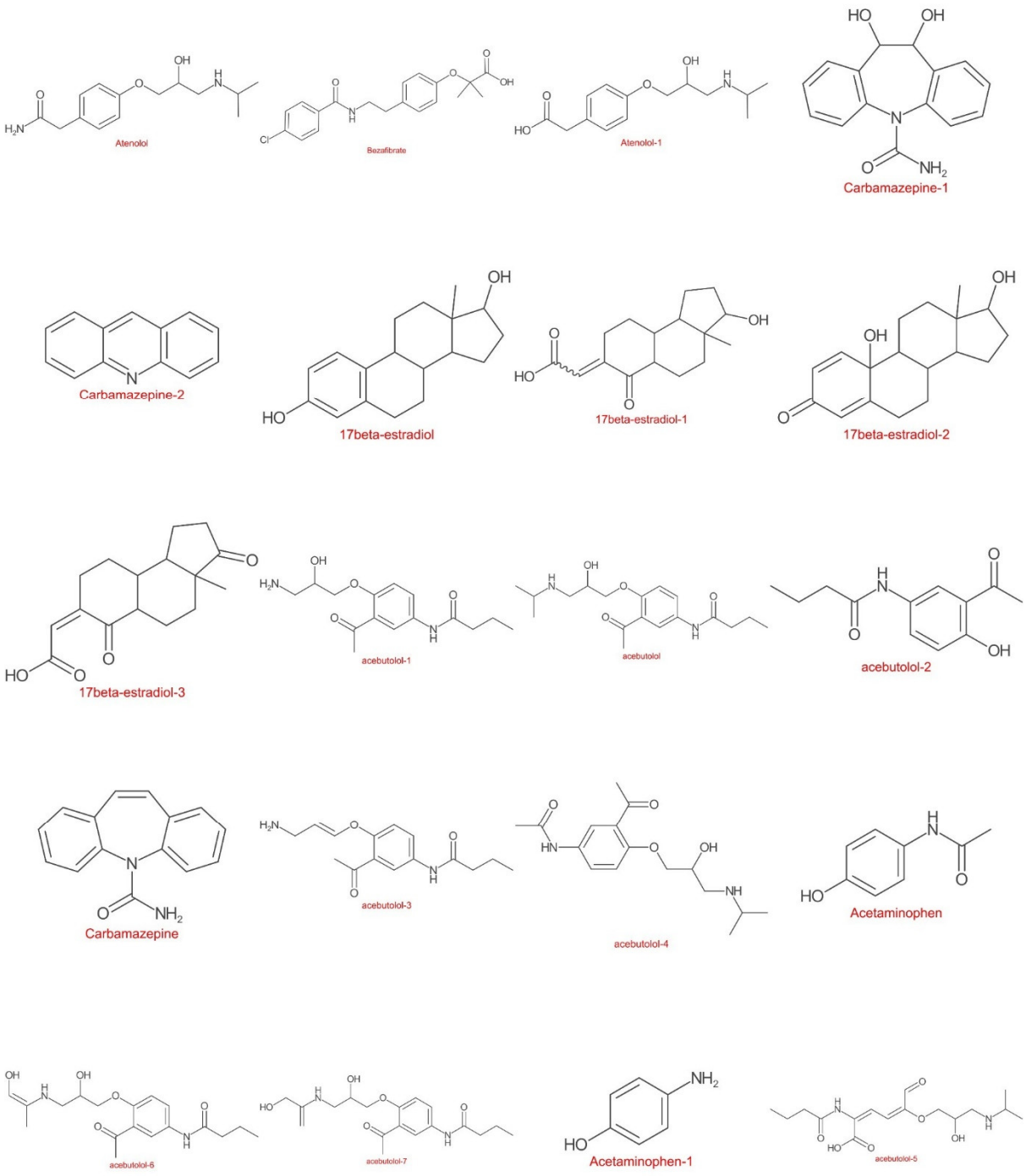



**Figure 7.5 Structures of confirmed parents and TPs in the suspect screening of multi-stage WWTP. Part A**

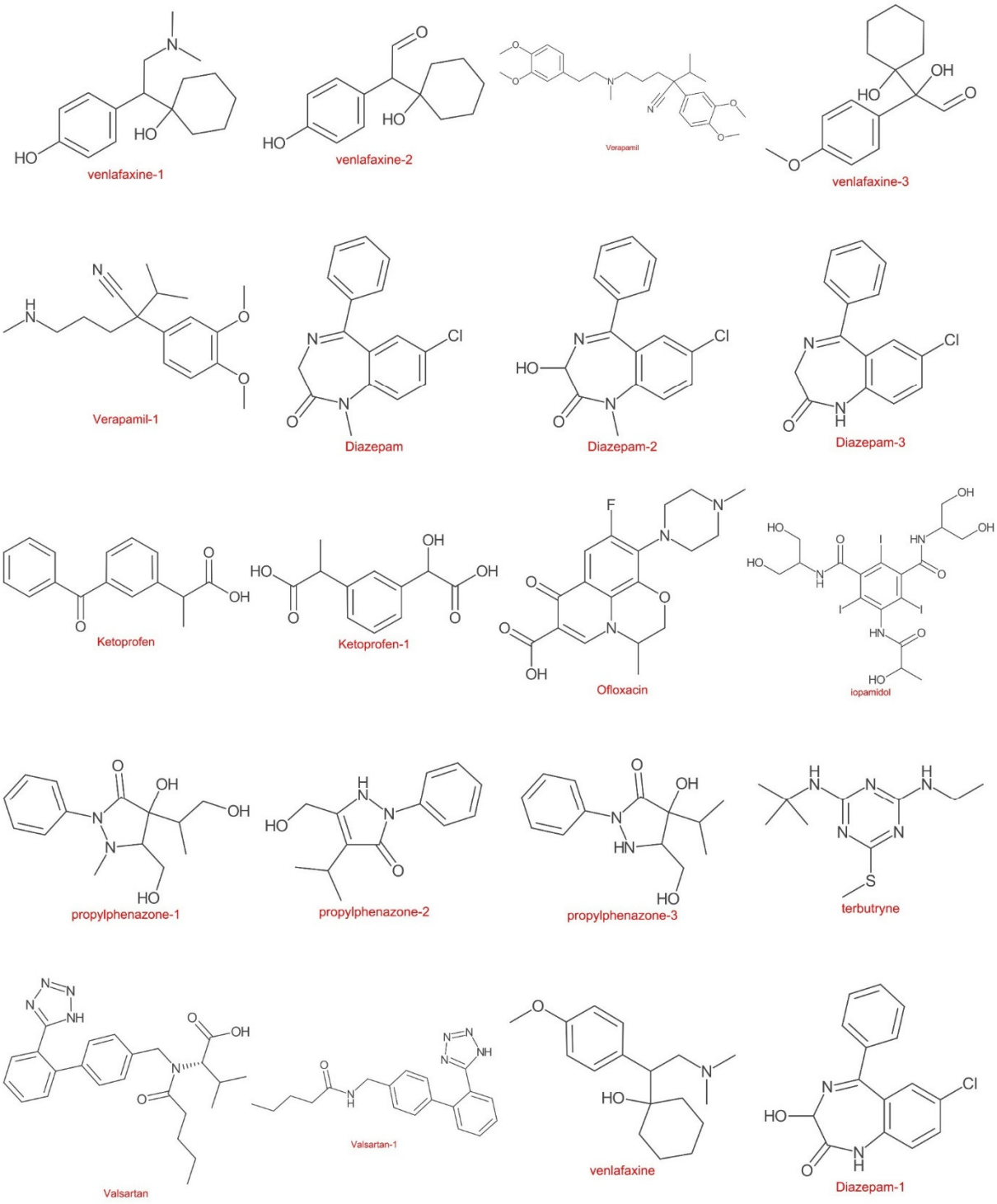



**Figure 7.6 Structures of confirmed parents and TPs in the suspect screening of multi-stage WWTP. Part B**

# 4 BAC-UF REACTOR DESCRIPTION, DETECTED IS, AND ADDITIONAL RESULTS

## BAC-UF reactor description

ORGANOSORB® 10-CO coconut shell-based GAC filled the BAC filter at 2 $m^3$ operating volume and 50 min empty bed contact time. The fixed bed BAC filter received secondary effluent in a down-flow configuration at a flow rate of 48 $m^3 \cdot day^{-1}$, resulting in wastewater treatment of 38.3 bed volumes (BV) per day. BAC outlet fed the ultrafiltration (UF) tank with a total surface of 220 $m^2$, a 5 min filtration mode at 3000 $L \cdot h^{-1}$, a 30 s backwash at 6000 $L \cdot h^{-1}$, a relaxation of 10 s, and a flux of 13 $L \cdot m^{-2} \cdot h^{-1}$. Concentrate of UF was recirculated on top of BAC filter with a flow rate of 60 $m^3 \cdot day$. BAC reactor was continuously oxygenated during the concentrate recycle 10/10 s UF membrane operation mode with scoured coarse bubbles.

## IS detected in BAC-UF system

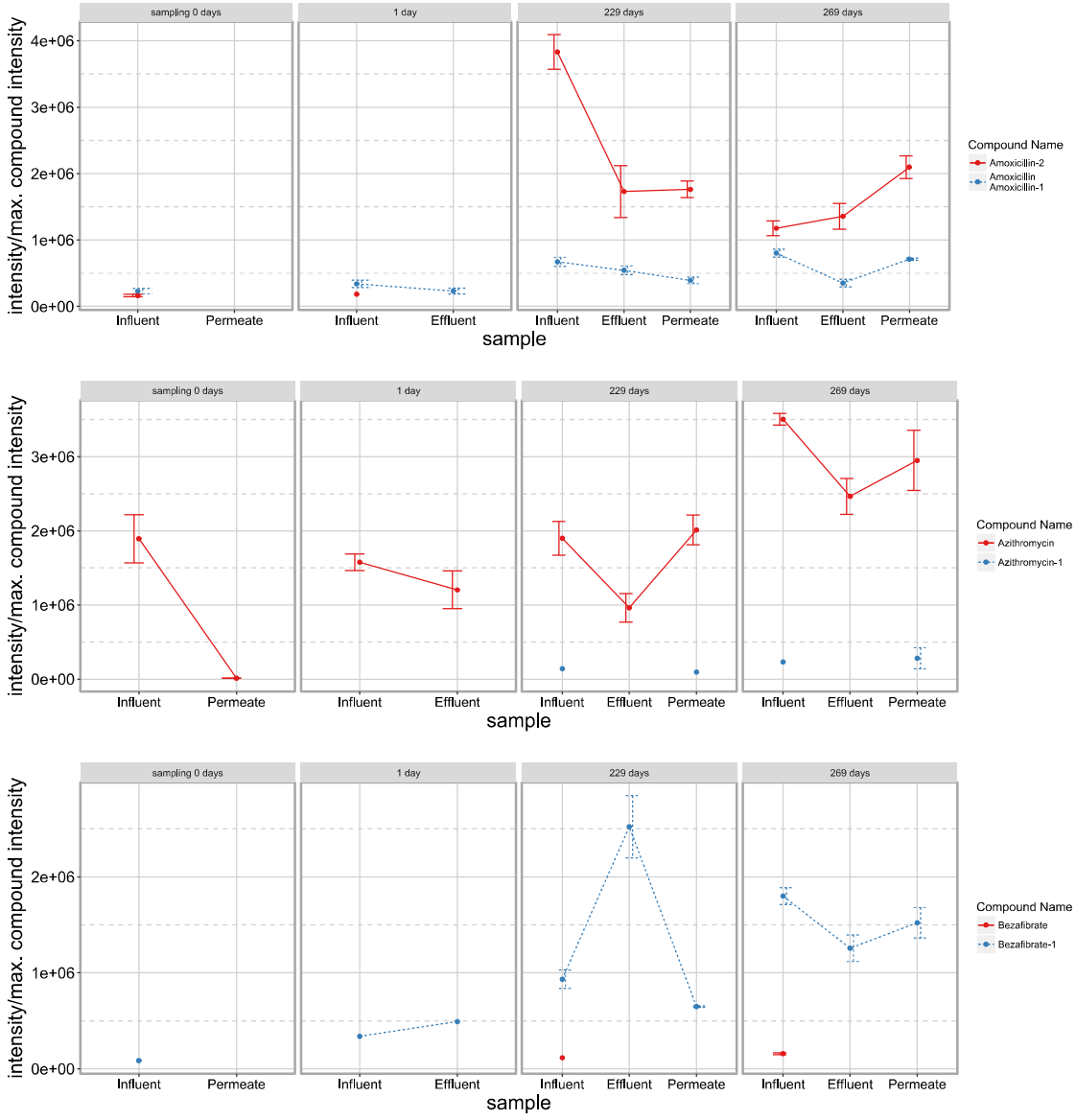
**Table 7.5 IS used for intensity correction of biodegradation TPs. Predicted Log P values and median RT extracted from all detected samples.**

| Compound Name | Formula | Monoisotopic mass [Da] | Log P | RT [min] |
|---|---|---|---|---|
| Acetaminophen-$d_4$ | $D_4C_8H_5NO_2$ | 155.09 | 1.42 | 2.47 |
| Amlodipine-$d_4$ maleic acid | $D_4C_{20}ClH_{21}N_2O_5$ | 412.17 | 3.01 | 4.80 |
| Antipyrine-$d_3$ (Phenazone-$d_3$) | $D_3C_{11}H_9N_2O$ | 191.11 | 1.48 | 3.40 |
| Atenolol-$d_7$ | $D_7C_{14}H_{15}N_2O_3$ | 273.21 | 1.54 | 2.35 |
| azaperone-$d_4$ | $D_4C_{19}FH_{18}N_3O$ | 331.20 | 3.01 | 3.76 |
| Azithromycin-$d_3$ | $D_3C_{38}H_{69}N_2O_{12}$ | 751.53 | 1.84 | 4.02 |
| Bezafibrate-$d_4$ | $D_4C_{19}ClH_{16}NO_4$ | 365.13 | 3.95 | 5.14 |
| Carbamazepine-$d_{10}$ | $D_{10}C_{15}H_2N_2O$ | 246.16 | 4.15 | 4.61 |
| Cimetidine-$d_3$ (N-methyl-$d_3$) | $D_3C_{10}H_{13}N_6S$ | 255.13 | 1.38 | 2.43 |

| Compound Name | Formula | Monoisotopic mass [Da] | Log P | RT [min] |
|---|---|---|---|---|
| Citalopram-$d_4$ Hydrobromide | $D_4C_{20}FH_{17}N_2O$ | 328.19 | 3.81 | 4.23 |
| Cprofloxacin-$d_8$ | $D_8C_{17}FH_{10}N_3O_3$ | 339.18 | 1.98 | 3.28 |
| Diclofenac-$d_4$ | $D_4C_{14}Cl_2H_7NO_2$ | 299.04 | 4.44 | 5.54 |
| Dexamethasone-$d_4$ | $D_4C_{22}FH_{25}O_5$ | 396.23 | 1.89 | 4.96 |
| Diazepam-$d_5$ | $D_5C_{16}ClH_8N_2O$ | 289.10 | 2.65 | 5.22 |
| Diltiazem-$d_3$ | $D_3C_{22}H_{23}N_2O_4S$ | 417.18 | 3.14 | 4.47 |
| Fluoxetine-$d_5$ | $D_5C_{17}F_3H_{13}NO$ | 314.17 | 4.83 | 4.81 |
| Indomethacin-$d_4$ | $D_4C_{19}ClH_{12}NO_4$ | 361.10 | 3.93 | 5.57 |
| Ketoprofen-$d_3$ | $D_3C_{16}H_{11}O_3$ | 257.11 | 3.11 | 5.00 |
| meloxicam-$d_3$ | $D_3C_{14}H_{10}N_3O_4S_2$ | 354.05 | 3.04 | 4.88 |
| Ofloxacin-$d_3$ | $D_3C_{18}FH_{17}N_3O_4$ | 364.16 | 1.55 | 3.14 |
| Ronidazole-$d_3$ | $D_3C_6H_5N_4O_4$ | 203.07 | 1.15 | 2.47 |
| Sertraline-$d_3$ Hydrochloride | $D_3C_{17}Cl_2H_{14}N$ | 308.09 | 5.57 | 4.97 |
| Simvastatin-$d_6$ | $D_6C_{25}H_{32}O_5$ | 424.31 | 4.59 | 6.08 |
| sulfamethoxazole-$d_4$ | $D_4C_{10}H_7N_3O_3S$ | 257.08 | 3.10 | 3.29 |
| Sulfapyridine-N-Acetyl-$d_4$ | $D_4C_{13}H_9N_3O_3S$ | 295.09 | 3.07 | 3.05 |
| Trimethoprim-$d_3$ | $D_3C_{14}H_{15}N_4O_3$ | 293.16 | 2.42 | 2.96 |
| Valsartan-$d_8$ | $D_8C_{24}H_{21}N_5O_3$ | 443.28 | 4.16 | 5.14 |
| venlafaxine-$d_6$ | $D_6C_{17}H_{21}NO_2$ | 283.24 | 3.04 | 4.14 |
| Verapamil-$d_6$ | $D_6C_{27}H_{32}N_2O_4$ | 460.32 | 5.09 | 4.44 |
| Warfarin-$d_5$ | $D_5C_{19}H_{11}O_4$ | 313.14 | 3.61 | 5.22 |
| Xylazine-$d_6$ | $D_6C_{12}H_{10}N_2S$ | 226.14 | 2.72 | 3.40 |

# Additional detected parents and TPs and their structures

Some TPs in Amoxicillin and Codeine series had the same mass as parents despite a different structure. They could not be distinguished by characteristic fragments and all candidates for the corresponding mass were reported here.



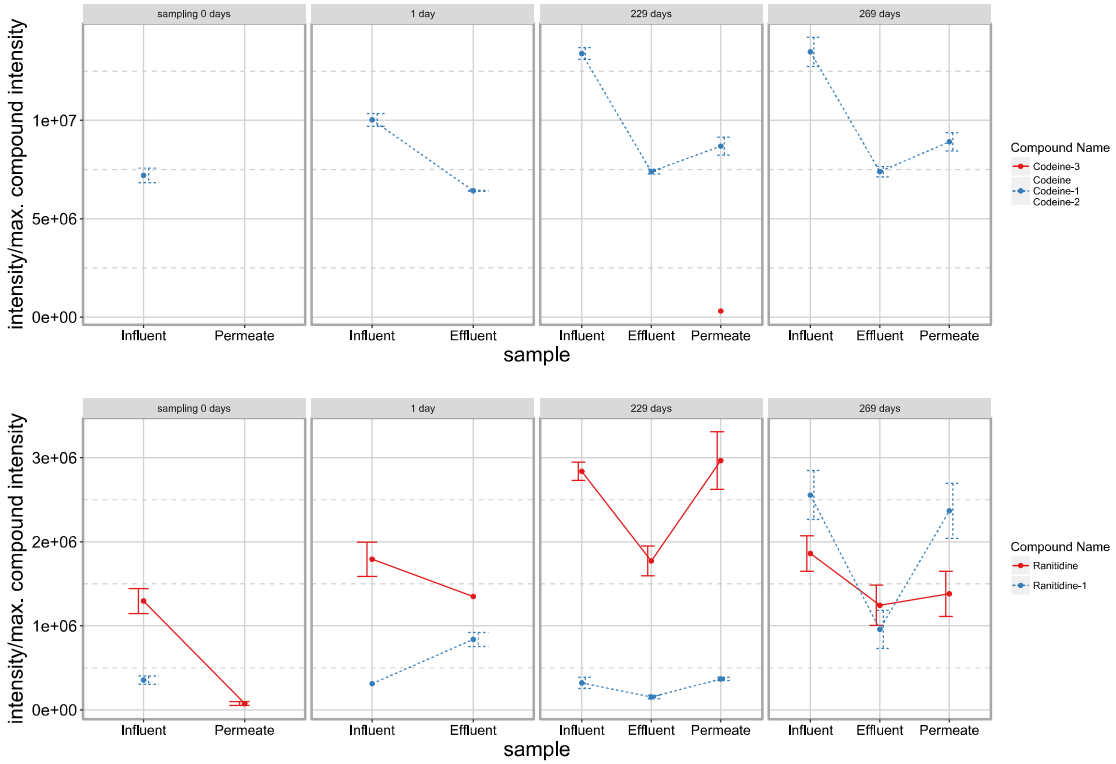



**Figure 7.7 Removal profiles of corrected intensity for parents and TPs of Amoxicillin, Azithromycin, Bezafibrate, Codeine, and Ranitidine at 4 sampling dates and 3 sampling sites of the BAC-UF pilot treatment.**

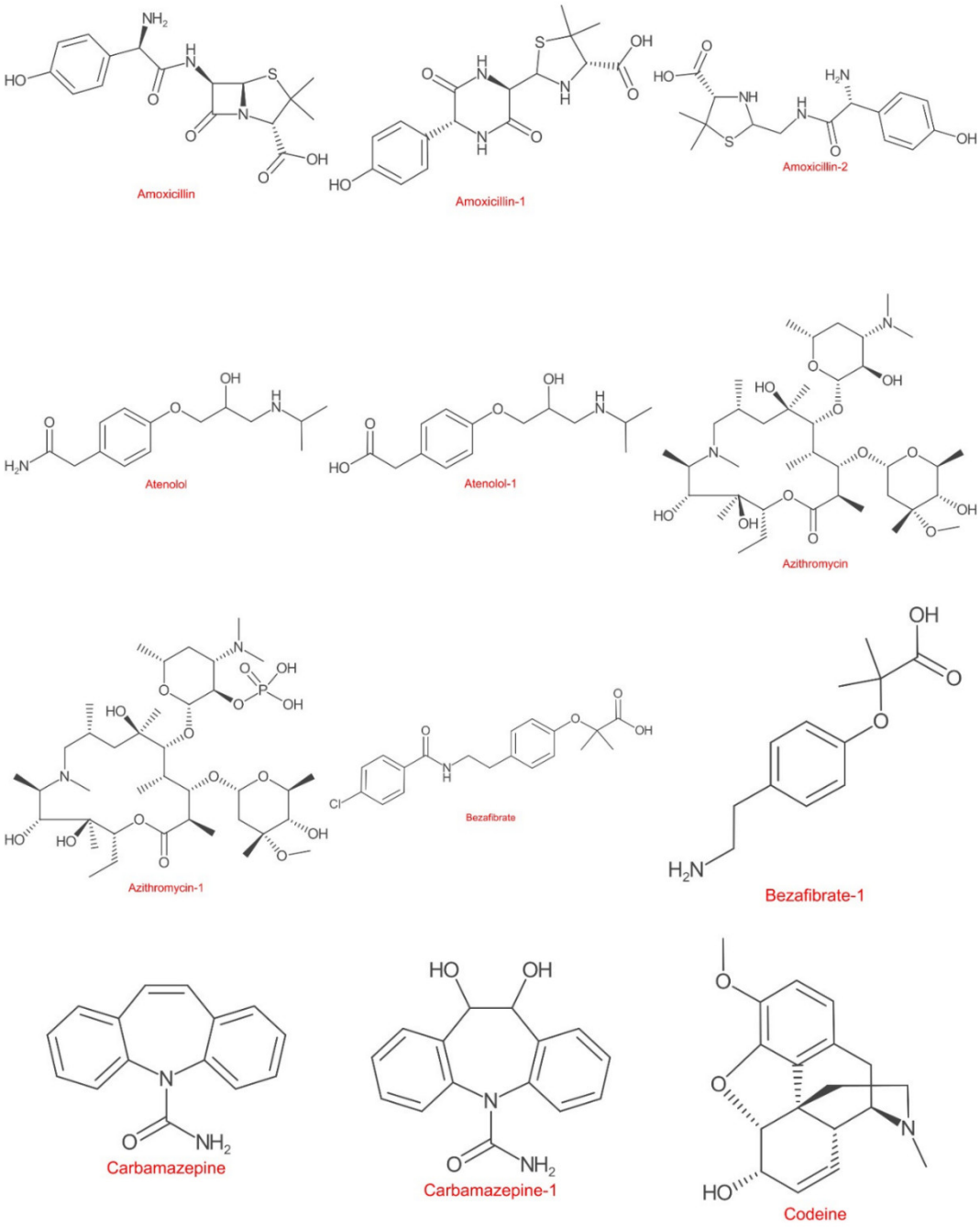



Figure 7.8 Structures of confirmed compounds in BAC-UF treatment. Part A.

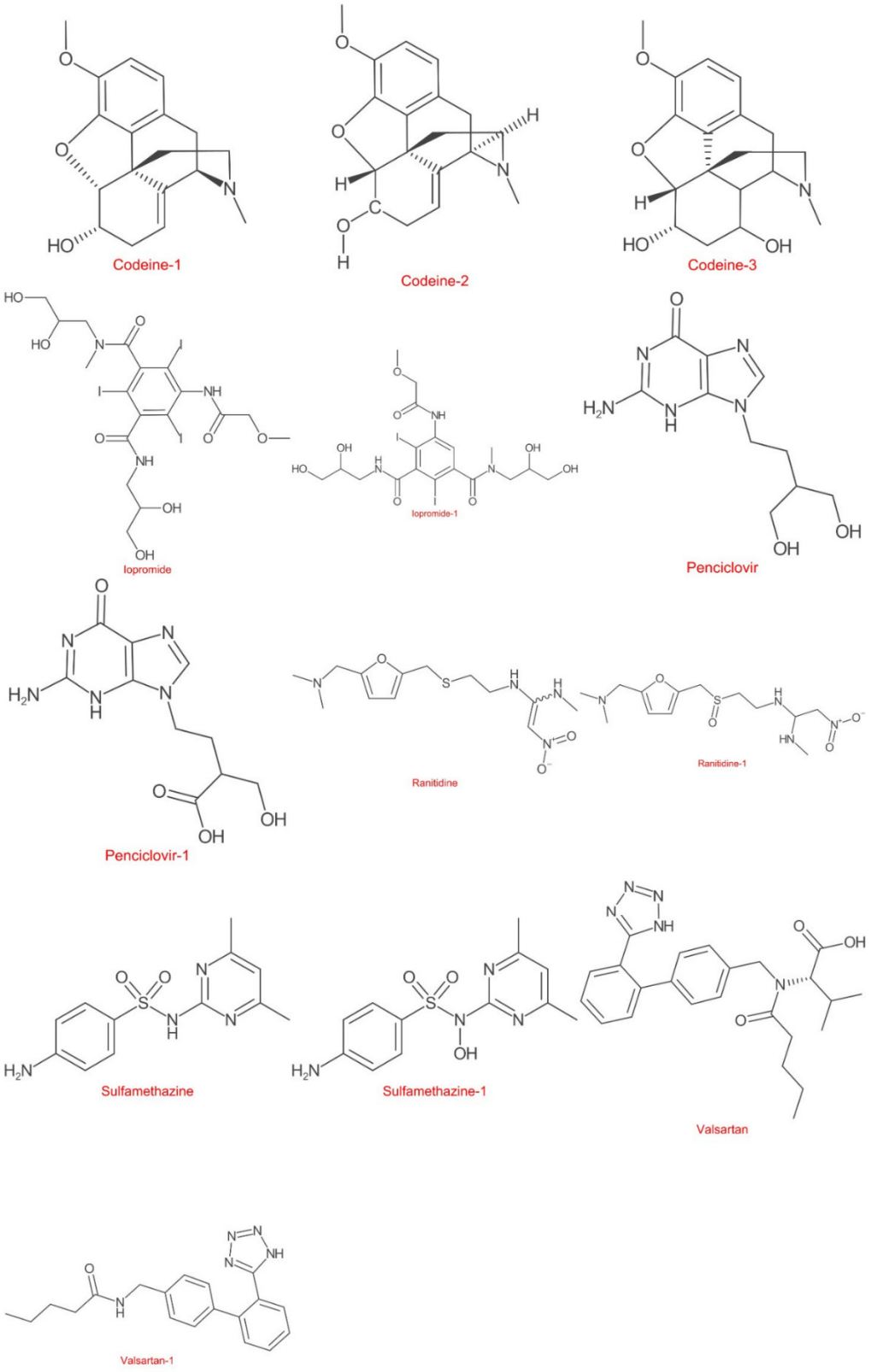



Figure 7.9 Structures of confirmed compounds in BAC-UF treatment. Part B.

Table 7.6 Detected fragments of parents and TPs in BAC-UF samples in PI mode. Parent compounds with corresponding IS were confirmed using retention time.

| Compound Name | Ion m/z [Da] | Retention Time [min] | Fragments | | | | |
|---|---|---|---|---|---|---|---|
| Amoxicillin | 366.1118 | 2.8 | 160.043 | 207.076 | | | |
| Amoxicillin-1 | 366.1118 | 2.8 | 160.043 | 207.076 | | | |
| Amoxicillin-3 | 340.1326 | 3.6 | 165.066 | 102.055 | 116.071 | | |
| Atenolol | 267.1703 | 2.1 | 150.09 | 116.11 | | | |
| Atenolol-1 | 268.1543 | 2.5 | 250.14 | 226.11 | 208.10 | 191.07 | 116.11 |
| Azithromycin-1 | 829.4821 | 3.8 | 212.128 | 189.112 | 131.07 | | |
| Bezafibrate-3 | 224.1281 | 2.6 | 208.1 | 191.07 | | | |
| Carbamazepine-2 | 271.1077 | 3.1 | 255.113 | | | | |
| Codeine | 300.1594 | 2.3 | 282.149 | 243.102 | 225.091 | 187.076 | 253.123 |
| Codeine-2 | 300.1594 | 2.3 | 282.149 | 243.102 | 225.091 | 187.076 | 253.123 |
| Codeine-3 | 300.1594 | 2.3 | 282.149 | 243.102 | 225.091 | 187.076 | 253.123 |
| Codeine-8 | 318.17 | 3.9 | 300.16 | 290.175 | 282.149 | 274.144 | 272.165 |
| Iopromide | 808.9036 | 2.2 | 114.055 | | | | |
| Iopromide-3 | 665.9804 | 2.2 | 114.055 | | | | |
| Penciclovir | 254.1248 | 1.8 | 237.098 | 212.103 | | | |
| Penciclovir-7 | 268.104 | 1.7 | 136.062 | | | | |
| Ranitidine | 315.1485 | 3.7 | 138.091 | 126.091 | | | |
| Ranitidine-4 | 350.1857 | 2.7 | 160.043 | | | | |
| Sulfamethazine | 279.091 | 2.8 | 156.012 | 107.061 | 95.05 | | |
| Sulfamethazine-1 | 295.0859 | 1.8 | 278.059 | | | | |
| Valsartan | 436.2343 | 4.8 | 101.06 | 57.07 | | | |
| Valsartan-1 | 336.1819 | 4.3 | 283.151 | | | | |