Identification of the xenometabolome and novel contaminant markers in fish exposed to a wastewater treatment works effluent

By

Raghad Adnan Ahmed Al-Salhi

A thesis submitted for the degree of doctorate of philosophy of the University of Sussex

School of Life Sciences

University of Sussex

Submitted April 2012

Declaration

I hereby declare that this work has not been submitted in any substance to another University for the award of any other degree or other academic or professional distinction.

Signature ..... Raghad Al-Salhi

Acknowledgments

First of all, I thank GOD for having given me the strength and patience to carry on this project and for blessing me with many great people who have been my greatest support in both my personal and professional life.

There are many people whom I owe a great debt of gratitude and without whose help I could not have completed this work: to all those people I am extremely grateful.

Ministry of Higher Education and Scientific Research in IRAQ is acknowledged for sponsoring my PhD studentship. I hope my work will benefit the country in some way.

A sincere gratitude and appreciation goes to my supervisor, Prof. Elizabeth Hill who gave me the opportunity to work on such an interesting project, for her knowledge, advice and support throughout my Ph.D. I could not have gone this far without your help, and I appreciated greatly your ability to be helpful despite how extremely busy you are.

I would like to thank my second supervisor, Dr Alaa Abdul-Sada for introducing me to the world of mass spectrometry and sharing his expertise and his valuable advice, which without would have made this project impossible.

I would like to express my warmest thanks to Dr Camilla Liscio for being a great friend to me, for putting up with me and being constantly supportive and understanding, thanks for her stimulating discussions, continual advice and patiently providing answers to innumerable and completely indiscriminate questions throughout last two years. A hug thanks to you for helping me to write this piece of work.

I wish to express my enormous gratitude to Dr Pawel Rostkowski for his kindness and dedicated support. I am deeply indebted to him for his help (chemistry and IT knowledge) and continuous encouragement.

A big thank you to all other members of Hill group past and present, whom I was fortunate enough to work with: Dr Paolo, Dr Diana, Dr Elena, Dr Alvine, Lynne, Dr Sumaira, Dr Katy, and Dr Francis. Thanks must go to Julia for her constant advice on the use of laboratory equipment. I wish you all the best with your future endeavours.

A big thank you goes to Mina and her parents for their unwavering support from the first day of my arrival in England until now.

I would like to thank Dr. Asma and her husband for supporting me at times when I needed help during my first year in the UK.

A special thanks must go to my friends in Baghdad (Baida, Maysoon and Niran) for their everlasting friendship.

The gratitude I owe to my host family, to Abida, her husband and her lovely daughters, who have allowed me to come home and enjoy great atmosphere, which has made the hard work of this PhD easier, this will never be forgotten.

Above all I thank my family (father, mother, brothers and sisters) who are the most important part of my life, who have always being there for me in all difficult circumstances, for their love, moral support, encouragement and continuous prayers, for their faith in my ability during my experience in the UK.

This thesis dedicated to those most dear to me- my beloved family.

Many thanks to you all...

Raghad, April 2012

University of Sussex
Raghad Adnan Ahmed Al-Salhi
D.Phil

Identification of the xenometabolome and novel contaminant markers in fish exposed to a wastewater treatment works effluent

Summary

Fish can bioconcentrate complex mixtures of xenobiotics arising from exposure to wastewater effluents discharged into surface waters. Wastewaters contain a complex mixture of organic compounds and little is known about their uptake into fish and their health effects. In this study, a chemical profiling approach was used to characterize the profile of xenobiotics and their metabolites (the xenometabolome) in biofluids (bile and plasma) of juvenile rainbow trout (*Oncorhynchus mykiss*) exposed to a wastewater effluent. Metabolite profiles of effluent-exposed fish were compared with that from control fish exposed to uncontaminated river water. Samples were analysed by ultra performance liquid chromatography-time-of-flight mass spectrometry and data analysed by multivariate statistics. Exposure to effluent resulted in accumulation in trout bile of alkylsulfophenyl and alkylpolyethoxy carboxylates, as well as glucuronide conjugates of nonylphenol ethoxylates, alcohol ethoxylates, naphthols, chlorinated xylenols and phenoxyphenols, chlorophenes, resin acids, mefenamic acid and oxybenzone. Non-conjugated or sulphate conjugates of many of these contaminants were also detected in plasma of effluent-exposed trout. In addition, changes in the concentrations of endogenously derived metabolites were also detected in trout plasma, and these included an increase in blood bile acids, methylbutryolcarnitine and a decrease in sphingosine levels. These observations were verified in a further exposure of sexually mature roach (*Rutilus rutilus*) to concentrations of the same effluent. Exposure to 50% or 100% effluent resulted in dose dependent increases in blood concentrations of xenobiotics, taurocholic acid, syprinol sulphate and lysophospholipids and decreases in sphingosine levels. This work reveals the complex nature of xenobiotics accumulating in effluent-exposed fish together with the identification of changes in concentrations of lipid metabolites associated with hepatotoxicity. These results reveal, for the first time, that metabolite profiling techniques can be used to screen for uptake of complex mixtures of contaminants into fish and also for the detection of natural metabolite pathways in the organism that are disrupted due to exposure to multiple xenobiotics.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

$[M+CH_2O_2]^-$	Formate adduct
$[M+H]^+$	Protonated molecule
$[M+Na]^+$	Sodium adduct
$[M-H]^-$	Deprotonated molecule
+ESI	Positive electrospray ionization
4-OP	4-Octylphenol
AECs	Alkyl ethoxycarboxylates
AEOs	Alkyl polyethoxylates
BCF	Bioconcentration factor
BPA	Bisphenol A
BPI	Base peak intensity
BSTFA	Bis(trimethylsilyl)trifluroacetamide
CDC	Chenodeoxycholic acid
Da	Dalton
EDCs	Endocrine disruption compounds
E_2	17β-estradiol
-ESI	Negative electrospray ionization
ESI-TOFMS	Electrospray-time-of-flight mass spectrometry
eV	Electron volts
FT MS	Fourier transform mass spectrometry
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
HPLC	High performance liquid chromatography
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
LOD	Limit of detection
M	Molar per litre
m/z	Mass to charge ratio
METLIN	Metabolite and tandem MS database
min	Minute
MS	Mass spectrometry
n	Number of samples
ND	Not detected
NIST	National institute of standards and technology
NMR	Nuclear magnetic resonance
NP	Nonylphenol
NPEOs	Nonylphenol ethoxylates
OPLS-DA	Orthogonal partial least squares projection to latent structures discriminant analysis
PAHs	Poly aromatic hydrocarbons
PCA	Principal Component Analysis
PLS-DA	Partial least squares projection to latent structures discriminant analysis
ppm	Part per million
Q^2	Cumulative variation predicted by the PCA or PLS model
R^2	Variation explained by the PCA or PLS model
SD	Standard deviation

SE	Standard error
STWs	Sewage treatment works
T2	Testosterone
TC	Taurocholic acid
TDC	Taurochenodeoxycholic acid
TIC	Total ion chromatogram
TMS	Trimethylsilyl
TOF	Time of flight
UPLC	Ultra-performance liquid chromatography
WwTP	Wastewater treatment plant
WwTWs	Wastewater treatment works

CHAPTER 1: General Introduction

1.1 Overview of emerging contaminants occurring in the aquatic environment

The term 'emerging contaminants' is generally used to describe pollutants either newly developed or newly discovered in the environment which have the potential to cause adverse effects to human and wildlife. Over the last decades scientific interest in the presence of many classes of emerging contaminants in the environment (e.g. personal care product compounds, pharmaceuticals, endocrine disruptors and illicit drugs) has increased significantly. Hundreds of tonnes of these compounds are used by the population every year and many of them are considered potentially hazardous for the environment since they can be ubiquitous, relatively persistent and biologically active, for example with endocrine disruption properties. Furthermore, due to their continuous release into the environment and possible additive or synergistic effects through combined action, even weakly persistent compounds might cause unwanted effects to wildlife (Daughton and Ternes, 1999, Fent et al., 2006). Concentrations of emerging contaminants that can be present in river water can range widely from ng/L to μg/L (Halling-Sorensen et al., 1998, Daughton and Ternes, 1999, Kolpin et al., 2002) and their levels depend mainly on the extent of water dilution resulting from rainfall as well as their concentrations being discharged into the receiving waters.

Personal care product compounds are mainly released into the aquatic environment through wastewater treatment plants (WwTWs). This category of contaminants encompasses compounds such as DEET (N,N-diethyl-meta-toluamide, the most common active ingredient in insect repellents), parabens (alkyl esters of p-hydroxybenzoic acid, used as bacteriostatic agents in drugs, cosmetics, and food), antifungal agents (e.g. triclosan, which is widely used in many household products), polycyclic musks (tonalide and galaxolide used as fragrances in a wide range of cleaning agents) and UV sunscreens (organic filters to UV radiation in sunlight, as benzophenones and methoxycinnamates) (Stuart et al., 2012). Triclosan and its metabolite methyl triclosan have been found in surface water by Lindström *et al.* (2002) whilst tonalide, galaxolide and galaxolide-lactone have been detected in WwTW effluents (Horii et al., 2007). Heberer *et al.* (2002) investigated synthetic musk concentrations in sewage, sewage sludge, surface water, aquatic sediment, and biota samples in relation to bioaccumulation, metabolism in fish, and environmental and human risk assessment. Tonalide and galaxolide were the most relevant compounds in

the characterized matrices and relative concentrations differed depending on the environmental compartment. Another example of personal care products widely distributed in the aquatic environment are sunscreen agents such as benzophenone-type UV, which have also been detected in the the watercourses despite their lipophilic nature. In river water in Korea, the overall concentrations of the UV filters in the sediment samples were between 500 and 18380 ng/kg and in water samples were between 27 and 204 ng/L (Jeon et al., 2006).

Pharmaceuticals are another class of emerging contaminants released extensively in the aquatic environment and are therefore of concern to the environmental community. Although many pharmaceutical compounds are effectively reduced by wastewater treatment, others are sufficiently persistent to occur in the environment. A wide range of pharmaceutical products have been already detected in surface and groundwater, mainly associated with wastewater disposal. Common investigated drugs include antibacterial (e.g. trimethoprim, erythromycin-H_2O and amoxicillin), veterinary and human antibiotics (e.g. ciprofloxacin, erythromycin, lincomycin, sulfamethoxazole and tetracyclines), prescription drugs (e.g. codeine, salbutamol and carbamazepine) anti-inflammatories and analgesics (e.g. paracetamol, tramadol, naproxen, ibuprofen and diclofenac) iodinated X-ray contrast media (e.g. iopromide and iopamidol) and psychotic drugs (e.g. gabapentin) (Ternes and Hirsch, 2000, Nikolaou et al., 2007, Pérez and Barceló, 2007, Barnes et al., 2008, Watkinson et al., 2009, Vulliet and Cren-Olivé, 2011). For instance, antibiotics have been found in hospital effluent at concentrations ranging from 0.01–14.5 μg/L, dominated by the β-lactam, quinolone and sulphonamide groups; and in WwTW influent up to 64 μg/L and in WwTW effluents in the low ng/L range up to a maximum of 3.4 μg/L, and up to 2 μg/L in the surface waters including freshwater, estuarine and marine samples (Watkinson et al., 2009). Several of these pharmaceuticals have been proved to be both ubiquitous and persistent in the aquatic environment (e.g. erythromycin-H_2O, codeine, carbamazepine, gabapentin and valsartan) (Kasprzyk-Hordern et al., 2008). The persistence of the studied compounds was evaluated on the basis of the percent concentration reduction after a certain distance from the effluent discharge. For instance erythromycin-H_2O was reduced by only 12% of its original concentration after 25km of river flow. Other potential threats to surface water are chemotherapy drugs, such as 5-fluorourcil, ifosfamide or cyclophosphamide (Buerge et al., 2006, Zaharie, 2006, Johnson et al., 2008) and illicit drugs as cocaine and

amphetamines (Kasprzyk-Hordern et al., 2008, Zuccato et al., 2008). Illicit drugs have been detected in rivers at levels of ng/L (amphetamine, cocaine and its main metabolite benzoylecgonine). Their occurrence in the environment is primarily associated with their high illegal usage and is strongly related to discharge of insufficiently treated wastewater effluents (Kasprzyk-Hordern et al., 2008). Caffeine and nicotine, and the nicotine metabolite cotinine, have been widely detected in groundwater impacted by sewage effluent (Teijon et al., 2010).

In addition, a wide range of industrial compounds can also be released into the environment affecting wildlife health status. Dioxins, chlorinated solvents, petroleum hydrocarbons (e.g. polyaromatic hydrocarbons and methyl tertiary-butyl ether), and plasticisers (e.g. bisphenols and phthalates) (Moran et al., 2005, Moran et al., 2006, Verliefde et al., 2007) are contaminants which can cause a range of well established health problems. Dioxins are also found in the environment as well as a consequence of degradation of other micropollutants such as triclosan (Mezcua et al., 2004).

Surfactants are widely detected in the environment due to their wide usage in household products. The priority pollutants octyl- and nonyl-phenol (OP and NP) are used in the production of the alkylphenol ethoxylates (APEs). Both the parent ethoxylates and their metabolites, the alkylphenols and the carboxylic degradation products are detected in surface water and effluents (Soares et al., 2008).

In addition, endocrine disruption compounds (EDCs) have become important emerging contaminants, due to their presence in environmental waters and concern about possible harmful effects both to wildlife and humans. Amongst the most common EDCs, vertebrate sex hormones are commonly present in wastewater effluents; they include androgens (e.g. androstenedione and testosterone), estrogens (e.g. estrone, estriol, 17α- and 17β-estradiol), progestins (e.g. progesterone), synthetic androgens such as oxandrolone, nandrolone and more importantly synthetic estrogens such as 17α-ethinyl estradiol and diethylstilbestrol (Johnson et al., 2000, Standley et al., 2008, Vulliet and Cren-Olivé, 2011).

1.2 Sources of contamination in the aquatic environment

Release of contaminants to the aquatic environment occurs via two primary routes: point-source pollution and non-point-source pollution. Point-source pollution derives from discrete sources whose inputs can be defined spatially. Examples of point-source pollution include industrial effluents (pulp and paper mills, steel plants, food

processing plants), municipal waste water treatment plants and storm-water overflows, resource extraction (mining), and land disposal sites (landfill sites, industrial impoundments). On the other hand, non-point-source pollution originates from diffuse sources which occur over broad geographical scales. Examples of nonpoint-source pollution encompass agricultural run-off (pesticides, pathogens, and fertilizers), storm-water and urban run-off, and atmospheric deposition (wet and dry deposition of persistent organic pollutants) (Ritter et al., 2002).

1.2.1 Wastewater treatment works

Often wastewaters contain a wide range of chemicals used in modern life and as a result they can be a significant source of newly emerging contaminants. These chemicals originate from both industrial and domestic sources (including sewage) and depending on the concentration levels may exert toxic effects on wildlife and humans. There are four main stages in wastewater treatment (preliminary, primary, secondary and tertiary treatment) and the number of stages applied depends on the quality of discharge required for a specific environment (Water UK, 2006, EPA Washington, 2004).

1.2.1.1 Preliminary treatment

Sewage undergoes preliminary treatment to make it suitable for the main treatment processes. Preliminary treatment includes screening and removing grit, oil and grease.

Screening

On entering the waste water treatment plant, dirty water passes through screens to remove large articles which could damage machinery or block pipe systems. Screens consist of vertical bars or perforated plates which are regularly cleaned by rakes or water jets. The cleared material (known as screenings) is usually washed and safely disposed of at landfill sites.

Grit removal

Sewage may contain grit and dirt coming from roads or cleaning activities. This inert material cannot be treated and must be removed by a settlement process which allows organic material to remain in suspension for the next treatment stage. The grit is washed and disposed of to landfill as well as the screenings.

Oil and grease removal

This step is exceptional since theoretically oil and grease should not be poured down drains or discharged to a sewer. This process is applied when necessary in order to protect the downstream processes of the wastewater treatment plant.

1.2.1.2 Primary treatment (settlement)

After the preliminary treatment, the waste water flows into large round or rectangular tanks. In these tanks the organic material sinks to the tank floor and is swept by a scraper blade to a submerged outlet. Afterwards, it is pumped to a storage tank for subsequent treatment. Most of the solids in sewage are removed in this process; the concentrated thick slurry deriving from this step is known as sewage sludge and it is dealt with separately. The water after removal of the settled solids flows over a weir to the next stage of treatment.

1.2.1.3 Secondary treatment (biological treatment)

Despite the efficiency of the primary treatment in removing the organic material, the discharge of the settled sewage to a watercourse would still cause problems. Naturally occurring bacteria in the receiving watercourse use organic material as a food source by means of oxidation processes through the oxygen dissolved in the water. Discharges of large quantities of organic matter would result in a rapid decrease of the oxygen levels with consequent harm to fish and wildlife. Wastewater treatment plants use the same process to break down and remove substances harmful to the environment but speed them up within a controlled environment. There are two main types of secondary treatment:

Biological filtration

In this process the settled sewage is distributed via small holes in continuously moving arms over 2 m deep circular or rectangular beds of stones. This step does not involve a proper filtration but the bed of stones provides an ideal place for bacteria to live and grow. The bacteria form a biological film on the stones which oxidises and removes the dissolved organic material as the settled sewage trickles downward. Oxygen from the spaces between the stones allows the microorganisms to respire and multiply. Once the film reaches a maximum thickness, the excess material is continuously washed off. After this step, the water flows to a settlement tank (humus tank) where the excess biological film is separated and removed as humus sludge,

which is normally returned to the primary settlement tanks and removed with the sewage sludge.

Activated sludge

In this process the settled sewage is mixed with a mixture of bacteria (activated sludge) and then aerated by agitators or air blowers in large tanks. The amount of air is regulated according to the respiration requirement, depending on the bacteria concentration and the load of the settled sewage. The excess of bacteria is drawn off as surplus activated sludge which is then mixed with the rest of the sewage sludge. The treated water is then separated from the activated sludge in final settlement tanks and is normally suitable for safe discharge to the environment. The key factor in order to have an effective biological treatment is the efficient separation of the bacteria (activated sludge or biological film) from the treated water. Very fine filter membranes can be used instead of settlement tanks to obtain a very high quality discharge. This process is very expensive though and is normally used for water very sensitive environment. Sewage contains also both nitrogen and phosphorus which can lead to nutrient enrichment of watercourses (eutrophication), resulting in algal blooms potentially harmful to water life. Nitrogen can be removed in specifically designed biological treatment plants whilst phosphorus removal is accomplished adding iron or aluminium salts before the settlement step.

1.2.1.4 Tertiary treatment

Tertiary treatment is applied when very high quality effluents are required. This additional process includes sand filters and natural systems such as wetlands. When discharges are made to bathing waters or shellfish growing areas, disinfection by UV light or removal of bacteria by membrane filtration is required

Over 95% of UK population is connected to wastewater treatment works, which are served by over 300,000 kilometres of sewers. There are 9000 wastewater treatment plants in UK with the majority serving populations of less than 2000. Each person produces approx 150 L of wastewater per day.

1.2.2 Removal efficiency in wastewater treatment plants

The efficiency of the wastewater treatment process plays an important role in minimising the release of contaminants into the aquatic environment (Geoff, 2001). However, a variety of polar contaminants such as pharmaceuticals or personal care product compounds are able to go through the biological wastewater treatment without

being fully degraded (Bernhard et al., 2006, Reemtsma et al., 2006). In order to estimate a pollutant's emission to a receiving watercourse, the removal performance of the treatment plant is usually evaluated by either full-scale balancing or determination of biodegradation rates at lab-scale (Vieno et al., 2007, Wick et al., 2009). Both approaches also rely on the residence time of the water in the treatment plant, normally referred to as the hydraulic retention time (HRT). The HRT can be calculated from flow through and tank volumes and it is therefore an easily accessible parameter. Joss *et al.* (2006) confirmed that the efficiency in contaminant elimination depends on the relative rate of degradation and retention times in the plant. For instance, Maurer *et al.* (2007) proved that β-blockers were not totally removed by the wastewater treatment due to both limited sorption and half lives (typical values ranging from 6h to 14 h) which were similar to the hydraulic retention time in the reactor (6h); the hydraulic retention time is defined as the average length of time that a soluble compound remains in a bioreactor. In fact, many pharmaceuticals may not be in the fully dissolved state in the wastewater or can be often present as the relative more polar glucuronic acid or sulphate conjugates and a combination of these factors can make them harder to remove (Ternes et al., 2004). Rosal *et al.* (2010) reported typical removal efficiencies of 20% for over 70 individual pollutants in a WwTW effluent using biological treatment followed by ozonation. Diclofenac was found at higher concentrations in the effluent than in the influent of a Swedish WwTW (Zorita et al., 2009). Removal efficiencies of >90% have been reported for steroid estrogens and for longer chain nonylphenol ethoxylates (NP4-12EO) (Koh et al., 2009). Horii *et al.* (2007) evaluated removal efficiencies for synthetic musks which ranged from 72% to 98%.

1.3 Bioaccumulation of organic contaminants in fish

Organic contaminants can be released into the aquatic environment by different sources (e.g. effluents, atmospheric deposition, runoff, and groundwater) and then be distributed throughout the water column and underlying sediment. Fish accumulate pollutants both by ingestion of contaminated food (i.e. organisms contaminated via contact with water or sediment and via their own food) and by contact of gills and skin with contaminated water. Accumulated contaminants are distributed throughout the body, and some can reach sites of action to exert toxic effects. However, chemical accumulation and toxicity depend not just on total chemical concentration in the environment but also on the chemical speciation of a contaminant, which can affect how

readily the fish can absorb the chemical at the gill surfaces, across the skin, and within the digestive tract and can also affect how the chemical is distributed throughout the organism (Di Giulio and Hinton, 2008). Chemicals exist in different forms in the environment (chemical species) and their relative concentrations of different chemical species differ among the components of the aquatic ecosystem. A fish can be exposed to a mixture of chemical species via water and food. Some pollutants metabolites have revealed higher toxicity and persistence than the parent compounds. For instance, metabolites deriving from the biodegradation of alkylphenol polyethoxylates (APEOs) in WwTWs are more toxic, more lipophilic, more estrogenic and more persistent than the parent APEOs to fish (Sheahan et al., 2002a, Ying et al., 2002).

Definitions of bioavailability vary markedly but all address how readily chemicals are accumulated under different conditions. Bioavailability can be defined here as the fraction of the bulk amount of the chemical present in soil/sediment and (interstitial) water which can potentially be taken up during the organism's lifetime into the organism's tissues (excluding the digestive tract) (Van der Oost et al., 2003). It is extremely important to correlate the chemical concentration in fish to the real bioavailable concentration otherwise this might lead to underestimation of the bioconcentration potential (Van der Oost et al., 2003).

1.4 Biotransformation in fish

Many xenobiotics are lipophilic and must be biotransformed in order to generate more polar compounds and facilitate renal or biliary excretion. Unfortunately, the process to enhance chemical polarity may create reactive intermediates through bioactivation and the metabolites can be more biologically hazardous than the initial parent compounds (Sarasquete and Segner, 2000). Bioaccumulation of a chemical occurs when uptake rates are significantly higher than metabolic clearance rates. Xenobiotics, taken up by a vertebrate and transported to the liver, can be metabolised or biotransformed by many routes including oxidation, reduction, hydrolysis, hydration, conjugation and condensation reactions. Metabolism of exogenously derived compounds such as xenobiotics is normally divided into two phases: phase I (functionalisation reactions) and phase II (conjugative reactions) (Fil., 2001). In fish, the cytochrome P450 (CYP) (Sarasquete and Segner, 2000) enzymes are known to catalyze oxidative metabolism of enormous number of compounds in the liver, as well as other tissues such as gills, gut epithelium and kidney. The phase I process either adds or

exposes polar atoms within the selected compound. The three main phase I reactions include oxidation, reduction, and hydrolysis. Phase I reactions involve the cytochrome P450 monooxygenases (CYPs) which constitute a superfamily of heme containing proteins which catalyze biological oxidation and reduction reactions (e.g. hydroxylation of aliphatic or aromatic carbons, epoxidation of double bonds, heteroatom oxygenation, hydroxylation or dealkylation, oxidative group transfer, cleavage of esters and dehydrogenation). Phase II reactions normally attempt to further increase polarity through conjugation of the phase I products with a bulky polar endogenous molecule such as glucuronic acid, glutathione, amino acids or peptides, methyl or acetyl groups or inorganic sulphate (Hochachka and Mommsen, 1995). In phase II metabolism, xenobiotics are conjugated to glucuronide, sulphate or glutathione groups, and in fish biliary excretion is considered to be the main pathway for elimination of xenobiotics conjugated to glucuronic acid (Clarke et al., 1991). The respective transferases provide the moieties for the substrates. Alternatively, phase II reactions may protect against bioactivation by masking functional groups (i.e., amines) with groups providing steric hindrance (i.e., methyl, acetyl) rather than augmented polarity. The enzymes responsible for xenobiotic metabolism in fish are located mainly in the liver, gills, intestine, and kidney (Clarke et al., 1991, George et al., 1998, James et al., 1998, Singh et al., 1996).

Glucuronidation involves the transfer of the glucuronic acid component to a substrate by any type of uridine 5'-diphospho(UDP)-glucuronosyltransferases. The UDP-glucuronosyltransferases (UGTs) represent a major group of phase II conjugating enzymes. Sulphation is the enzyme-catalyzed addition of sulphate to a substrate and this process uses its cosubstrate 3'-phosphoadenosine-5'-phosphosulphate (PAPS) to transfer sulphate to a xenobiotic. The sulphotransferase (SULT) family of enzymes catalyzes the transfer of the sulphonate group to hydroxyl and amine groups in endogenous and exogenous substrates (Coughtrie, 2002). The glutathione S-transferases are another family of phase II enzymes that give cellular protection against the toxic effects of a diversity of endogenous and environmental chemicals (Di Giulio and Hinton, 2008).

The liver is the main organ involved in the biotransformation processes, however, when the substrate is present at low levels, biotransformation in the intestine becomes more important. The capacity of the intestine for metabolism is fairly low, because the liver is a larger organ; however if small amounts of xenobiotic are present, the intestine

and other extra hepatic organs such as gills and skin, become important sites of first-pass metabolism (Di Giulio and Hinton, 2008).

1.5 Xenobiotic elimination in fish

Chemical excretion of a compound depends on its physicochemical properties. In fish a toxic compound can be excreted by three principal routes: branchial, urinary and biliary. In branchial excretion, chemicals move from the blood through several diffusion barriers to finally be excreted into the water (Hochachka and Mommsen, 1995). Fish are able to excrete many endogenous and exogenous compounds as well as their metabolites via the biliary route (Kennedy et al., 1989, Larsson et al., 1999). Excretion is partitioned between the kidney and the bile depending on molecular size and lipid solubility of the compounds. The relative importance of biliary excretion as a route of elimination increases with molecular weight. In fish, highly polar chemicals with molecular weights higher than 350 Da are mainly excreted in bile whilst compounds of intermediate molecular weight (<300 Da) and polarity may be eliminated in bile as well as by other routes (Di Giulio and Hinton, 2008). Xenobiotic excretion can be achieved also by the kidney. Small water soluble compounds are excreted by the kidneys while more lipid soluble compounds are excreted by the liver (Hirom et al., 1972, Pritchard et al., 1980). Many lipophilic compounds need metabolic conversion before their excretion via the kidney since they have higher affinity for plasma proteins and they are not filtered by the glomerulus where protein-free filtrate is forced out for further processing to form urine (Hochachka and Mommsen, 1995).

1.6 Exposure to a mixture of contaminants (wastewater treatment works)

Wildlife organisms are seldom exposed to single chemicals but instead are exposed to complex mixtures of contaminants which may act in diverse ways (Thorpe et al., 2001, Silva et al., 2002, Sumpter, 2003, Thorpe et al., 2003) and may induce combination effects (Rajapakse et al., 2002). For instance, wastewater effluents contain a mixture of natural and synthetic xenobiotics (e.g. household and agricultural chemicals, pharmaceuticals, hormones etc) (Stevens et al., 2003). Ecotoxicological data revealed that mixtures might have different effects than single compounds (Cleuvers, 2003, DeLorenzo and Fleming, 2008, Quinn et al., 2009). The behaviour of chemicals in mixtures depends on their mode of toxic action (MOA). When multiple chemicals have different target sites, their effect can be evaluated independently, even if an integral response of the organism is investigated. Mixtures of chemicals with the same

MOA act according to concentration or dose additivity. Concentration Addition (CA) as well as Independent Action (IA) assume a non-interaction between the mixture components. If the mixture components interact with each other, such interactions can lead to either antagonistic (less toxic than expected) or synergistic (more toxic than expected) toxicities. Certain binary pesticide mixtures had recently been proved to have more than additive toxicity toward salmon by Laetz *et. al.,* (2008). Synergistic mixture effects have also been investigated for a mixture of prochloraz (fungicide) and esfenvalerate (insecticide) (Bjergager et al., 2011). Cases of more than additive mixture toxicity are usually specific for mixture (compound types, concentrations and mixture ratios), biological system and end point. Toxicity studies in the literature have shown that mixture of pharmaceuticals at environmentally relevant concentrations and with similar modes of action may exhibit additive effects (DeLorenzo and Fleming, 2008). In this specific study, two different chemical mixtures, simvastatin–clofibric acid mixture and triclosan–fluoxetine mixture, revealed additive effects related to the toxicity threshold for marine phytoplankton. Levels lower than the toxic threshold may lead to harmful effects when in the presence of a mixture of active substances (DeLorenzo and Fleming, 2008). Michael (2003) showed that a mixture of diclofenac and ibuprofen had a stronger toxicity than predicted in cladoceran *Daphnia magna*. The exposure of the freshwater cnidarian *Hydra attenuata* to a pharmaceuticals mixture, revealed sub-lethal effects for environmentally relevant concentrations (µg/L–ng/L) (Quinn et al., 2009).

Wildlife exposure to sewage effluents have been associated in the literature with alterations in sex steroid hormone levels (Folmar et al., 1996, Folmar et al., 2001, Hecker et al., 2002) and impaired gonadal development in adult and juvenile fish (Hemming et al., 2001, Jobling et al., 2002, Sheahan et al., 2002b), altered sexual differentiation in early life periods (Rodgers-Gray et al., 2000), and induction of the egg-yolk precursor protein vitellogenin (VTG) in adult male and juvenile fish of both male and female (Harries et al., 1999, Purdom et al., 1994, Rodgers-Gray et al., 2000). The alteration of the reproductive system of these species have been related to the presence of mixture of estrogenic contaminants in the effluents which can act as endocrine disruptors; these compounds encompass natural and synthetic steroids (Desbrow et al., 1998, Routledge et al., 1998), alkylphenol polyethoxylates (Gimeno et al., 1997, 1998, Jobling et al., 1996, Seki et al., 2003), and phthalates and pesticides (Ankley et al., 2001, Christiansen et al., 2000, Jobling et al., 1995, Sohoni and Sumpter,

1998). These disorders, however, may not necessarily due to estrogenic effects alone. EDCs exhibit multiple mechanisms of toxicity by acting at different sites within the body. Therefore, the overall detected toxic impact of a particular chemical can be a result of the combination of alterations to the endocrine immune and nervous systems and damage to genetic material (Roy and Liehr, 1999, Galloway and Handy, 2003, Choi et al., 2004).

There is now increasing recognition amongst environmental scientists that wildlife and humans are often exposed to chemical mixtures from birth to the end of the life cycle, and this total exposure to chemical contaminants throughout the life course is termed 'the exposome'. The exposome comprises every exposure to which an individual is subjected from conception to death. The concept of exposome was originally developed for a more complete environmental exposure assessment in epidemiological studies. The exposome therefore complements the genome by providing a comprehensive description of lifelong exposure history. However in the broadest sense, three categories of (bio)chemical effects and non-genetic exposures can be considered: internal, specific external and general external (Wild, 2012).

Internal exposure: internal processes influencing the body, such as metabolism, endogenous circulating hormones, body morphology, physical activity, gut microflora, inflammation, lipid peroxidation, oxidative stress and ageing. These internal conditions can affect the cellular environment and have been variously described as host or endogenous factors.

Specific external exposure: such as radiation, infectious agents, chemical contaminants and environmental pollutants, diet etc.

General external exposures (in human case): such as the social, economic and psychological influences on the individual (e.g. social capital, education, financial status, psychological and mental stress, urban–rural environment and climate).

Hypothetically, different components of the exposome will leave their mark or fingerprint, so that it is possible to go not only forward from the molecular characteristics to the clinic but also back to the exposures, epidemiology and public health. The application of omics technologies to biological samples in epidemiological studies has been expressed as a "top–down" approach of measuring the exposome; i.e. the measurement of contaminants or markers of exposure in tissues rather than the "bottom-up" approach of predicting exposure by measuring chemicals in the ambient

environment. The “bottom-up” approach focuses on each category of external exposure in order to quantify contaminants subsequently summed over all categories to estimate individual exposomes. On the other hand, the “top-down” approach adopts untargeted omic methods to characterize features of exposures in a selected biological matrix (e.g. biofluids) (Figure 1.1). The top-down approach is more efficient than a bottom-up approach. The use of ‘omics’ technologies may result for example in the detection of bio(chemical) markers of chemical exposure or toxicity (Wild, 2012).

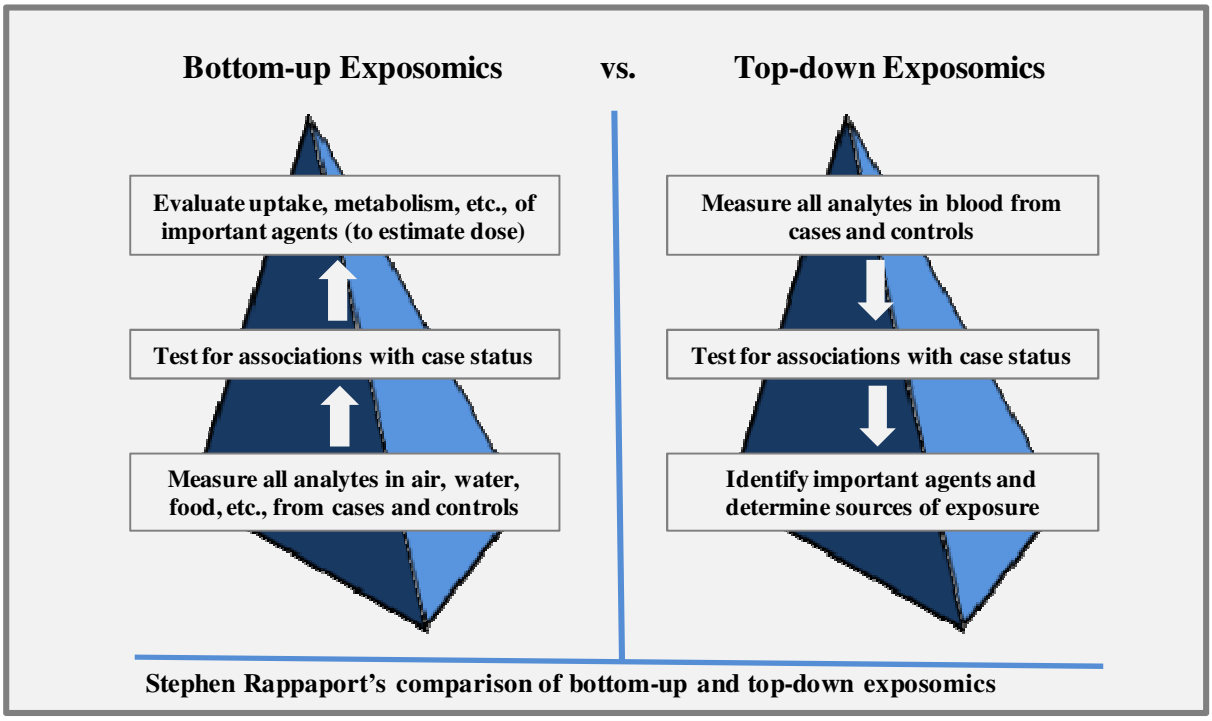


Figure 1.1: Schematic representation of comparison between bottom-up and top-down exposomics approaches. Adapted from ("The exposome: A powerful approach for evaluating environmental exposures and their influences on human disease", 2010).

1.7 Omics-based technologies

Omics-based technologies study changes to genes, transcripts, proteins or metabolites in response to alterations of the cellular environment. “Omics” cover different application fields (Mayer, 2011):

- Genomics (the quantitative study of protein coding genes, regulatory elements and noncoding sequences).
- Transcriptomics (RNA and gene expression).
- Proteomics (protein upregulation and downregulation).
- Metabolomics (metabolites and metabolic networks).

- Pharmacogenomics (the quantitative study of how genetics affects a host response to drugs).
- Physiomics (physiological dynamics and functions of whole organisms).
- Nutrigenomics (identification of the genetic factors influencing the body's response to diet and study of the bioactive constituents of food affecting gene expression).
- Phylogenomics (analysis involving genome data and evolutionary reconstructions).
- Interactomics (molecular interaction networks).

Genomics investigates many genes simultaneously on a genome-wide scale. Structural genomics, comparative genomics and functional genomics are the three main interrelated areas which have been widely described in the literature. Structural genomics determines the genome structure at the sequence level, comparative genomics investigates differences on the molecular basis between organisms at a variety of taxonomic levels whilst functional genomics focuses on the function of genes. Microarrays, either as cDNA or oligonucleotides, and digital transcriptomics are the main techniques employed. Santos et al. (2007) utlized a 17,000 oligonucleotide based microarray to investigate the expression of genes in zebrafish (*Danio rerio*) after exposure to ethinylestradiol (EE2). Garcia-Reyero *et al.* (2008) compared gene expression profile of fish caged upstream and downstream a WwTW effluent with laboratory fish reference using 22,000 oligonucleotide microarray. Genomics approach in environmental studies is still limited by the lack of sequenced and identified genes in ecologically relevant species (Finne et al., 2007).

Proteomics consists in the characterization of the proteome which is the fingerprint of all proteins present in a cell, tissue or organism at a certain physiological stage or as a reaction to a specific treatment. The comparison of proteomes deriving from contaminant treatments versus controls should provide informative results regarding the molecular mechanisms of response to xenobiotics. The main technologies employed in proteomics studies are the separation techniques 2-dimension polyacrylamide gel electrophoresis and the identification techniques LC-MS and fourier transform mass spectrometry FTMS (Miracle and Ankley, 2005). Proteomics has been widely applied in environmental monitoring researches in vertebrate and invertebrate species (Iguchi et al., 2006). Albertsson et al. (2007) characterized the proteome of

rainbow trout (*Oncorhynchus mykiss*) liver after exposure to WwTW effluents. Three proteins (betaine aldehyde dehydrogenase, lactate dehydrogenase and an unidentified protein) were down-regulated after exposure, whilst the only up-regulated protein consisted of both mitochondrial ATP synthase alpha-subunit and carbonyl reductase/20β-hydroxysteroid dehydrogenase (CR/20β-HSD).

Metabolomics gives an overview of the metabolic status of a biological system by the measurement of small molecular size endogenous metabolites. This top-down system biology approach gives insights into the metabolic status of complex living organisms via the non-targeted analysis of metabolites in biological matrices (Nicholson et al., 1999, Lindon et al., 2007). The whole array of metabolites present in the investigated matrix can be estimated at the same time, and metabolic changes induced by external factors (e.g. environment, diet, disease processes, or pharmaceutical interferences) can be screened over time (Nicholson et al., 1999, Lindon et al., 2007). The main limitation in the interpretation of transcriptomics and proteomics results is the difficulty of relating observed gene-expression fold changes or protein-level changes to altered biochemical phenotypes (the metabolome) (Fiehn et al., 2000, Sumner et al., 2003). On the contrary, metabolomics offers the advantage of considering the dynamic metabolic status of the whole living organism and the ability to predict more precisely phenotypical properties (Nicholson et al., 1999, Nicholson and Wilson, 2003, Goodacre et al., 2004). The metabolome gives a snapshot of the final stage in the chain of events from genes to metabolism, and the metabolic phenotype is the most direct reflection of the actual state of a biological system. The terminology relating to metabolomics is still controversial (Ryan and Robards, 2006). The term “metabolome” was first used by Oliver *et al.* (1998) to describe the whole set of metabolites synthesized by an organism. Recently, this description has been limited to “the quantitative complement of all of the low molecular weight molecules present in cells in a particular physiological or developmental state” (Goodacre, 2005). The expression “metabolomics” was proposed by Oliver Fiehn and defined as a comprehensive analysis of all metabolites in a specific biological system (Fiehn, 2001). The confusion in the terminology come from the similar term “metabonomics”, proposed by Nicholson et al. (1999). Metabonomics has been defined as a subset of metabolomics (Fiehn, 2002) which, on the contrary, aims to measure those metabolites which reveal changes in response to a stimulus of one sort. So far, the two expressions have been often used interchangeably. Hence, the

expression metabolomics is usually used to describe a comprehensive, non-targeted analytical approach (universally applicable to identify and quantify the whole array of metabolites characteristic of a selected biological system). Environmental metabolomics is the application of metabolomics to the investigation of living organisms either obtained from the natural environment or from laboratory conditions mimicing natural environment scenarios (Morrison et al., 2007).

The xenometabolome represents the multivariate description of the xenobiotic metabolite profile present in a sample from an individual exposed through any route (either deliberately or accidentally) to drugs, environmental pollutants, or dietary components which cannot be totally catabolized by endogenous metabolic enzyme systems (Holmes et al., 2007). Hence, the xenometabolome can be considered as an exogenous part (not under direct genomic control of an organism) of the metabolic phenotype (metabotype) characteristic of their general environmental exposures to chemicals. Exogenous metabolites represent the biotransformation or metabolism products of xenobiotics resulting from phase I (introduction of a functional group) and/or phase II (conjugation) enzymatic processes. The xenometabolome is useful as it provides potential evaluation of environmental pollutant exposure in an epidemiological context (Roux et al., 2011).

1.7.1 Metabolomic strategies

Metabolite profiling involves the identification and quantitation of a set of unknown metabolites belonging to a selected metabolic pathway (Fiehn, 2001, Dunn and Ellis, 2005). Metabolite fingerprinting aims to the rapid classification of several samples applying multivariate statistics without differentiation of individual metabolites. Target analysis is typically performed for a selected range of metabolite. There are many different strategies which have been employed within the metabolomics field. The following four approaches are currently the most used (Fiehn, 2001, Goodacre et al., 2004, Hollywood et al., 2006):

1. Metabolite target analysis, which is usually limited to metabolites of a specific system that can be directly affected by abiotic or biotic perturbation.
2. Metabolite profiling, which is focused on a particular group of metabolites associated with a specific pathway; in clinical and pharmaceutical field this is often called metabolic profiling, and is used to trace the fate of a drug or metabolite.

3. Metabolomics, which is the comprehensive analysis of the whole metabolome (all measurable metabolites), under a certain set of conditions. This approach is often confused with metabonomics, which on the contrary measures the fingerprint of biochemical perturbations caused by toxins, drugs and disease.
4. Metabolic fingerprinting, which is used to categorize samples according to provenance of either their biological relevance or origin by using a fingerprinting approach; this approach is rapid but does not necessarily give particular metabolite information.

1.7.2 Metabolomics applications

Metabolomics has been applied to study biological systems including microorganisms (MacKenzie et al., 2008, Smedsgaard and Nielsen, 2005), plants (William Allwood et al., 2006, Hall, 2006), mammals (Douglas B, 2006, Dunn et al., 2007, Kenny et al., 2008, Lindon et al., 2007), and environment (Tanaka et al., 2007, Viant, 2007). This omics-based technology has recently demonstrated significant potential in various fields such as responses to environmental stress (Vineis et al., 2009, Workentine et al., 2010), toxicology (Nicholson et al., 1999, Hines et al., 2010, Wang et al., 2009), nutrition (Pexa et al., 2008, Astle et al., 2007, Lee et al., 2010), studying global effects of genetic manipulation (Urakami et al., 2010, Tohge and Fernie, 2010, Pluskal et al., 2010), cancer (Nomura et al., 2010, Wu et al., 2010), natural product discovery (Kim et al., 2010). Metabolomic analysis of biofluids or tissues has been utilized in physiology, diagnostics, functional genomics, pharmacology, toxicology and nutrition fields (Wang et al., 2010).

1.8 Analytical technologies used in metabolomics

Common analytical techniques applied to metabolomics are NMR, GC/MS and LC/MS (Dettmer et al., 2007, Pan and Raftery, 2007, Want et al., 2006). Each technology shows some advantages and they are basically complementary (Dettmer et al., 2007, Pan and Raftery, 2007). Within the field of metabolomics in biofluids, NMR provides an excellent technique for profiling biofluids and is especially adapt for the characterization of complex solutions. ^{1}H and ^{13}C NMR can measure most of the metabolome. However, the extremely large dynamic range which is typical encompassed by biological systems as well as the difficulties due to the coupling of NMR to chromatographic techniques, are significant drawbacks which limit the use of NMR in metabolomics since important aspects of the metabolome composition (e.g.

low abundant signalling metabolites) may not be detected or correctly quantified. Therefore, due to its high sensitivity, its high dynamic range and the capability to analyze extremely complex samples, MS-based approaches have started to complement NMR analyses for metabolomic research. The volatile compounds can be analyzed by GC-MS after derivatization (Cao et al., 2011) whilst LC/MS technique is becoming a useful tool in the study of body fluids, representing a promising micro-separation platform in metabolomics (Johnson et al., 2011). Integrated platforms have been used to provide the sensitive and reproducible detection of the several metabolites present in a biofluid sample. Hence, a combination of different analytical technologies must be used to gain a broad perspective of the metabolome.

Gas chromatography–mass spectrometry (GC–MS) has become more and more popular for metabolite profiling and is often employed for the detection of many metabolic disorders (Chace, 2001). GC-MS can provide high sensitivity, high chromatographic resolution, wide range of detectable metabolite classes and mass spectra can be easily searched in many commercially available electron impact (EI) MS libraries to help in the structural identification (Bino et al., 2004). This technique offers structural information, reasonable quantitative precision and high throughput. So far, most of the metabolic profiling studies using GC–MS have been in the field of plant metabolomics (Fiehn et al., 2000, Bino et al., 2004, Glinski and Weckwerth, 2006). The application of GC–MS for metabolic profiling in pharmacology/toxicology is still underdeveloped compared to NMR and LC-MS. Nevertheless, GC–MS allows the detection and quantification of many metabolite classes including amino and organic acids, fatty acids and some lipids, sugars, sugar alcohols and phosphates, amines, amides and thiols, which are often targets in efficacy and/or toxicity studies (Quinones and Kaddurah-Daouk, 2009). GC–MS can be a powerful tool for metabolic profiling in toxicological evaluations, providing a comprehensive understanding of the response of biological systems to xenobiotic exposure (Chace, 2001, Pasikanti et al., 2008). However, using GC-MS involves some disadvantages for metabolite profiling, including laborious sample preparation often requiring a derivatization step in order to improve volatility of the analyte of interest (e.g. nonvolatile, polar macromolecules are unsuitable).

Liquid chromatography mass spectrometry (LC-MS) as technique of choice for metabolic profiling facilitates metabolite identification by reducing sample complexity

and separating the array of metabolites in the sample prior to detection. High-performance liquid chromatography (HPLC) allows separation of compounds of a wide range of polarity. Furthermore, HPLC coupled to electrospray-ionization mass spectrometry (HPLC-ESI-MS) is becoming the most eligible method for detecting metabolites in complex biological samples (Bowen and Northen, 2010). Reversed-phase HPLC, normally using C18 columns, can efficiently separate many semi-polar compounds. Although normal-phase HPLC allows the separation of very polar compounds, the use of apolar mobile phases makes it more compatible with atmospheric pressure chemical ionization mass spectrometry (APCI-MS) instead of ESI-MS. The coupling of HPLC separation with MS detection improves MS sensitivity and signal reproducibility by easing sample complexity and thus reducing matrix interferences in the ionization process. Additionally, good chromatographic separation gives better quality MS data improving the S/N ratio. Accordingly, recent further developments in LC technology [e.g. capillary monolithic chromatography and ultra-performance LC (UPLC)] have accomplished significant progress in order to improve peak resolution and speeding analysis up (Guillarme et al., 2007, Xiao et al., 2012).

1.9 Mass spectrometry

Mass spectrometry (MS) is a very powerful analytical tool which can be employed in the identification of both unknown organic and inorganic compounds, in the quantification of target compounds, and in the structural elucidation allowing high degree of sensitivity and selectivity. Basically, mass spectrometers convert analyte molecules into gas-phase ions in the ion source and then separate them according to their mass to charge ratio (*m/z*) in the mass analyser. Once separated, ions are guided to the detector where they are counted and the electric signals are recorded by a computer to produce the mass spectrum and to process the data (De Hoffmann and Stroobant, 2007).

A mass spectrometer is mainly constituted by the following elements:

1. A sample inlet to introduce the compound that is analysed (a direct insertion probe or in the case of hyphenated techniques, the GC or LC itself).
2. An ionization source to convert the analyte in gas-phase ions.
3. One or more mass analysers to separate the different ions according to their *m/z*.
4. A detector to ‘count’ the ions emerging from the analyser.
5. A data processing system to produce the mass spectrum.

1.9.1 Ionization techniques

Samples can be introduced into the ion source either directly by direct infusion if the sample is liquid or direct probe if solid, or as eluate from a chromatographic system (as a gas if from a GC or as a liquid if from an LC). The ion source has to convert the sample to gas phase if not already so, to ionize the compounds, and then transfer the produced ions to the mass analyzer. The most common ionization techniques are the electron impact ionization (EI) used in GC chromatography and the electrospray ionization (ESI) for the direct sample infusion or combined with LC. Many ionization parameters can play an important role for a successful detection and in particular, the solvent utilized for sample introduction as the solvent composition is a core part in the ionization process. Some ionization techniques are very energetic and cause extensive fragmentation (hard ionization, e.g. EI, CI) whilst other techniques only produce ions of the molecular species (soft ionization, e.g. ESI, APCI, MALDI). The ion source generates ions mainly ionizing a neutral molecule in the gas phase by different mechanisms: electron ejection, electron capture, protonation, deprotonation, adduct formation or by transfer of a charged species from a condensed phase to the gas phase. Ion production normally implies gas-phase ion–molecule reactions (De Hoffmann and Stroobant, 2007).

1.9.1.1 Electron ionisation

Electron ionisation (El) was one of the first ionisation techniques to be used, developed by Dempster. An electron ionisation source uses a heated filament to produce accelerated electrons with typically energy of 70eV, enabling them to collide with the gaseous analyte molecules introduced into the source. This ionization works well for numerous gas-phase molecules but induces extensive fragmentation, for this reason it is usually difficult to observe the original molecular ions (De Hoffmann and Stroobant, 2007).

19.1.2 Electrospray ionisation

Electrospray ionisation (ESI) was developed by Fenn *et al.* (Fenn et al., 1989, Mann et al., 1989) and has recently become one of the most popular ionisation methods allowing the analysis of big biomolecules directly from the liquid phase. ESI belongs to the soft ionisation techniques category since it generates ions with extremely low internal energies and gives minimal analyte fragmentation. ESI can reach very high sensitivity and is easy to couple to HPLC or capillary electrophoresis. This ionization is

achieved by applying under atmospheric pressure a strong electric field to a liquid flowing through a capillary (flow 1-10 µl/min). The electric field is generated by a potential difference of 3-6 kV between the capillary and a counter-electrode. The generated electric field induces a charge accumulation at the liquid surface located at the end of the capillary leading to the emission of a spray of small highly charged droplets. The solvent contained in the droplets evaporates, passing through a curtain of heated inert nitrogen gas. Excess charges accumulate at the surface of the droplet where the desorption of the charged molecules occurs. For this reason, sensitivity is usually higher for compounds which are more concentrated at the surface. When mixtures of compounds are ionised, those molecules predominant at the surface of the charged droplets can mask even completely the presence of the other compounds more soluble in the bulk (termed ion suppression). The formation of gas phase ions generates protonated $[M+H]^+$, deprotonated $[M-H]^-$ or multiply charged ions, especially from large molecule having many ionisable groups . However, ESI may also generate adducts of the analyte with cations (e. g. Na^+, K^+, NH_4^+) or anions (e. g. CH_3COO^-, Cl^-) (Loo et al., 1989, De Hoffmann and Stroobant, 2007).

Adduct Formation

An adduct is an ion generated by direct combination of a neutral molecule with an ion other than the proton. In +ESI mode the sodium adduct is the most observed adduct, giving rise to a signal higher by 22 mass units than the protonated molecule. It is often found together with the potassium adduct, another 16 mass units higher. If the eluents have not been carefully desalted, the sodium adduct is usually dominant. If an ammonium salt is present in solution, the related adduct $[M+NH_4]^+$ can be formed due to its ability to create hydrogen bonds. In the -ESI mode, the chloride adduct is normally observed with its relative isotopic pattern $[M+35]^-$ and $[M+37]^-$. Whilst the acetate adduct is observed as $[M+59]^-$ (De Hoffmann and Stroobant, 2007).

Formation of Aggregates or Clusters

Dimeric ions $[2M+H]^+$ or of higher order $[nM+H]^+$can also be detected in the ESI mass spectrum. In some cases, the proton can be replaced by another cation (e.g. Na^+). Aggregates are seldom detected in the -ESI mode, because negative charge causes an expansion of the electronic shell, reducing the electric field around the negatively charged ion and this phenomenon can reduce the interactions between the species involved (De Hoffmann and Stroobant, 2007).

1.9.1.3 Atmospheric Pressure Chemical Ionization

APCI is an ionization technique analogous to chemical ionization CI generally used in GC–MS, where primary ions are generated by corona discharges on a solvent spray at atmospheric pressure. APCI is mostly applied to relatively non-polar compounds with moderate molecular weight and generated usually monocharged ions. In the positive ion mode, either proton transfer or adduction of reactant gas ions occurs to produce the molecular ions, depending on the relative proton affinities of the reactant ions and the gaseous analyte molecules. In the negative mode, the molecular ions are generated by either proton abstraction or adduct formation (De Hoffmann and Stroobant, 2007).

1.9.1.4 Matrix-Assisted Laser Desorption Ionization

Matrix-Assisted Laser Desorption Ionization (MALDI) was introduced by Karas and Hillenkamp (Karas et al., 1987, Karas and Hillenkamp, 1988). It is a common and powerful source for the production of intact gas-phase ions from a broad range of large, non-volatile and thermally labile molecules (e.g. proteins, oligonucleotides, synthetic polymers). The choice of a suitable matrix, which provides for both desorption and ionization, is the important factor for a successful ionization method. The method is described by easy sample preparation and has a large tolerance to contamination by salts, buffers, detergents (Chen et al., 1998, Stump et al., 2002). In MALDI, the analyte of interest is dissolved in solution of small organic molecules, called the matrix. The matrix must strongly absorb at the laser wavelength. The obtained mixture must be dried before analysis, resulting in a ‘solid solution’ deposit of analyte-doped matrix crystals. The analyte is then ionized and introduced in the mass spectrometer by ablation of bulk portions of this solid solution by intense laser pulses over a short duration. Since the ablation process is independent of the absorption properties and size of the analyte, MALDI permits desorption and ionization of compounds with very high molecular mass (>100 000 Da) (De Hoffmann and Stroobant, 2007).

1.9.2 Mass Analyzers

Determination of mass-to-charge ratio (*m/z*) is achieved using a combination of electric and/or magnetic fields. Several kinds of mass analyzers are commercially available nowadays and all mass analyzers operate in high vacuum conditions to prevent collision of the generated ions with uncharged molecules. Mass analyzers are often categorised according to their performance (Villas-Boas et al., 2007):

- Nominal mass analyzers, whose mass resolution is unit mass separation. Resolving power is usually 1:1000–2000 and they have integer mass accuracy.
- High resolution mass analyzers, whose mass resolution is higher than 1:7000, presenting mass accuracy below 1 ppm.

High resolution mass analyzer can be used to separate all formulas and isotopic compositions with relevance to metabolomics at approximately below 1000 Da. Resolution (R) or resolving power is the ability to discriminate between close *m/z* values while mass accuracy is the difference between the measured value and the calculated value in relation to the calculated value (m). Mass accuracy is typically expressed as a relative error in parts per million (ppm) (Balogh, 2004, Breitling et al., 2006).

1.9.2.1 Quadrupole

The linear quadrupole mass analyser was developed by Paul and Steinwedel in 1953. It is made of four parallel metal rods positioned along the instrument's z axis. Fixed direct current (dc) and alternating radio frequency (rf) potentials applies to them. The quadrupole analyser separates ions according to their *m/z* ratios using the stability of the trajectories in oscillating electric fields (Villas-Boas et al., 2007).

1.9.2.2 Quadrupole Ion trap

The quadrupole ion trap (QIT) mass analyser was developed in the 1950s by Paul and Steinwedel. The QIT consists of three electrodes (3D trap): two end-cap electrodes and a ring electrode. Ions are injected from the source into the trap through one of the end-caps and trapped by applying an RF-voltage and a DC voltage to the ring electrode and endcaps. In order to control ion motions and lowering ion energy a damping gas is introduced into the trap. By changing the RF-voltage and the DC potentials on one of the end-caps, ions with specific *m/z* values are ejected from the ion trap, and then separated. Ion traps can be classified into two types: the 3D ion trap or the 2D ion trap. 2D ion traps also known as LITs are based on a four rod quadrupole ending in lenses which reflect ions forwards and backwards (Villas-Boas et al., 2007).

1.9.2.3 Sector instruments

High resolution sector analyzers were developed by Mattauch and Herzog (1936) and Johnson and Nier (1953). These analysers are adaptable to many continuous ion sources (e.g. EI, dynamic SIMS, ICP, and ESI). However, they have recently been

replaced by Q-TOFs and FT-ICR instruments, since the latter typically provide good MS/MS data and are much less demanding, being less expensive and much smaller. Nevertheless, sector instruments are still highly performing in high accuracy quantitative measurements, such as isotope ratio determination or analysis of toxic compounds (Ekman et al., 2009b).

1.9.2.4 Time-of-flight

The time-of-flight (TOF) analyser was firstly described by Stephens in 1946 and in 1955, Wiley and McLaren designed a linear TOF mass spectrometer becoming the first commercial instrument (De Hoffmann and Stroobant, 2007). The TOF analyser separates ions initially accelerated by an electric field according to their velocities as they drift in a free-field region (flight tube). Ions reach the analyser as bundles either produced by an intermittent process (e.g. plasma or laser desorption), or expelled by the ionization source focusing lenses. The ions can then be accelerated towards the free-field region by a difference of potential applied between an electrode and the extraction grid. All the ions acquire the same kinetic energy but ions showing a distribution of masses present a distribution of velocities as well. Once left the acceleration region, ions enter into the field-free region where they can separate according to their velocities, before reaching the detector at the other extremity. Mass-to-charge ratios are determined by measuring the time that ions take to move through a field-free region between the source and the detector (De Hoffmann and Stroobant, 2007).

In theory, there is no upper limit of mass range for a TOF analyser and this makes it particularly suitable for soft ionization techniques. Furthermore, TOF analysers have high transmission efficiency which leads to very high sensitivity. TOF analysers have very fast analysis speed and a spectrum over a broad mass range is obtained in micro-seconds. However, the poor number of ions in each individual spectrum does not often grant sufficient precision of mass or abundance measurement; for this reason, it is not possible to record all the individual spectra without exceeding the speed of data transfer and the capacity of data storage. Therefore, recorded spectra are usually obtained by the addition of several individual spectra (De Hoffmann and Stroobant, 2007).

As the mass resolution is proportional to the flight time and the flight path, lengthening the flight tube allows enhancing the resolution of these analysers. However, too long a flight tube decreases TOF performance due to loss of ions by scattering after collisions with gas molecules or by angular dispersion of the ion beam. A possible

solution would be increasing the flight time by means of lowering the acceleration voltage, but lowering this voltage decreases the sensitivity. In order to allow both high resolution and high sensitivity, the use of a long flight tube (length ranging between 1 and 2m) for a higher resolution and a high acceleration voltage (≥20 kV) to grant high sensitivity is required.

If a pulsed source such as MALDI is coupled to the TOF analyser, the quality of its pulsed ion beam is insufficient to obtain high resolution and high mass accuracy. In order to improve this outcome, two techniques have been: the delayed pulsed extraction and the reflectron (De Hoffmann and Stroobant, 2007).

Delayed Pulsed Extraction

Delayed pulsed extraction was initially developed by Wiley and McLaren in the 1950s. The ions formed in the source using the continuous extraction mode are immediately extracted by a continuously applied voltage. Ions with the same *m/z* ratio but with different kinetic energy reach the detector at slightly different times, resulting in peak broadening. To reduce this kinetic energy spread, a time lag between ion formation and extraction can be introduced. Ions are first allowed to expand into the field-free region in the source and after a certain delay (100 ns to few μs) a voltage pulse is applied to extract the ions outside the source. The extraction pulse applied after a certain delay transmits more energy to the ions which kept for a longer time in the source.

Reflectron

Another option in order to improve mass resolution is the use of an electrostatic reflector also known as reflectron. The reflectron solution was proposed initially by Mamyrin (Mamyrin et al., 1973). The reflectron generates a retarding field acting as an ion mirror by deflecting the ions, which are then sent back through the flight tube. The term reflectron time-of-flight (RTOF) analyser is employed to differentiate it from the linear time-of-flight (LTOF) analyser. The simplest type of reflectron, called a single-stage reflectron, consists of a series of equally spaced grid electrodes or ring electrodes connected through a resistive network of equal-value resistors. The reflectron is located behind the field-free region opposite to the ion source and the detector is positioned on the source side of the ion mirror in order to capture ions which have been reflected. The detector can be coaxial with the initial direction of the ion beam or it can be off-axis with respect to the initial direction of the ion beam. The reflectron corrects the kinetic

energy dispersion of the ions leaving the source with the same *m/z* ratio. Ions with more kinetic energy will penetrate the reflectron more deeply than ions with lower kinetic energy, spending therefore more time in the reflectron and reaching in the end the detector at the same time than slower ions having the same *m/z*. However, the higher mass resolution obtained using the reflectron decreases the sensitivity and introduces a mass range limitation. A two-stage reflectron give better performances reducing the size and improving the homogeneity of the electric field by using two successive homogeneous electric fields of different potential gradient. The first stage employs an intense electric field to strongly decelerate the ions while the second stage applies a weaker field. The two-stage reflectrons are usually more compact because of the strong deceleration of the ions at the first stage, but they can suffer from a lower transmission (De Hoffmann and Stroobant, 2007).

1.9.2.5 Fourier transform mass spectrometry

Fourier transform ion cyclotron resonance

Fourier transform mass spectrometry (FTMS) was firstly described by Comisarow and Marshall in 1974 (Comisarow and Marshall, 1974a, Comisarow and Marshall, 1974b) and then reviewed by Amster (1996) and Marshall *et al.* (1998). FTMS excites simultaneously all of the ions present in the cyclotron by scanning rapidly a large frequency range within a short time span (~1 μs). The ions will be forced in a trajectory that comes close to the wall perpendicular to the orbit in phase. This allows transformation of the complex wave detected as a time-dependent function into a frequency-dependent intensity function through a Fourier transform. Ions of each mass have their characteristic cyclotron frequency. The performance of FT-ICR instruments has steadily improved across time and they have recently reached levels of resolution and mass accuracy comparable to the sector analysers (Gross, 2004, De Hoffmann and Stroobant, 2007).

Ions having different masses show characteristic cyclotron frequency. Ions which are excited by an AC irradiation at their own frequency and with the same energy, (i.e. the same V_0 potential), applied during the same time T_{exc}, will have an orbit with the same radius, and with an appropriate radius will all pass close to the detection plate as described by the following equaton:

$$r = \frac{V_0 T_{exc}}{B_0}$$

The ion orbit will have the same radius due to the applied broadband excitation, but at frequencies depending on their *m/z* ratio. The best way to obtain this result is by applying a waveform calculated by the inverse Fourier transform (De Hoffmann and Stroobant, 2007).

Fourier transform orbitrap

The orbitrap is a mass analyser of recent invention. Ions are trapped and stored in a potential well as for quadrupole ion traps, however, ions are not injected for external detection but it is the oscillation frequency of the trapped ions which is measured by the analyzer. This approach allows high mass resolution (100,000 FWHM) and mass accuracy (0.5-1ppm). The orbitrap consists of a thin wire central electrode, a coaxial outer electrode, and two end-cap electrodes. Between the inner and outer electrodes a DC voltage is applied giving a logarithmic potential. Ions injected perpendicular to the wire with an appropriate velocity will circulate in an orbit around the wire. By applying a potential to the end caps, the ions will also be confined axially (Ekman et al., 2009b).

1.9.3 Detectors

The detector measures amount of ions as a function of time. As the *m/z* transmission of the mass analyzer is altered over time, the detector measures the mass as a function of *m/z*. Detection is important for the quality of the data obtained and very sensitive high-speed amplifiers and analog to digital conversions are crucial integrated parts of all detector systems (Villas-Boas et al., 2007).

1.9.4 Tandem mass spectrometry

Generally speaking, tandem mass spectrometry experiments involve the performance of two or more mass spectrometric analysis in sequence after an induced fragmentation process (Figure 1.2). Dissociations of the selected ions can occur either spontaneously in transit through the mass analyzer (metastable ions) or can result from collisions with neutral collision gas (collision-induced dissociation) (Gross, 2004).

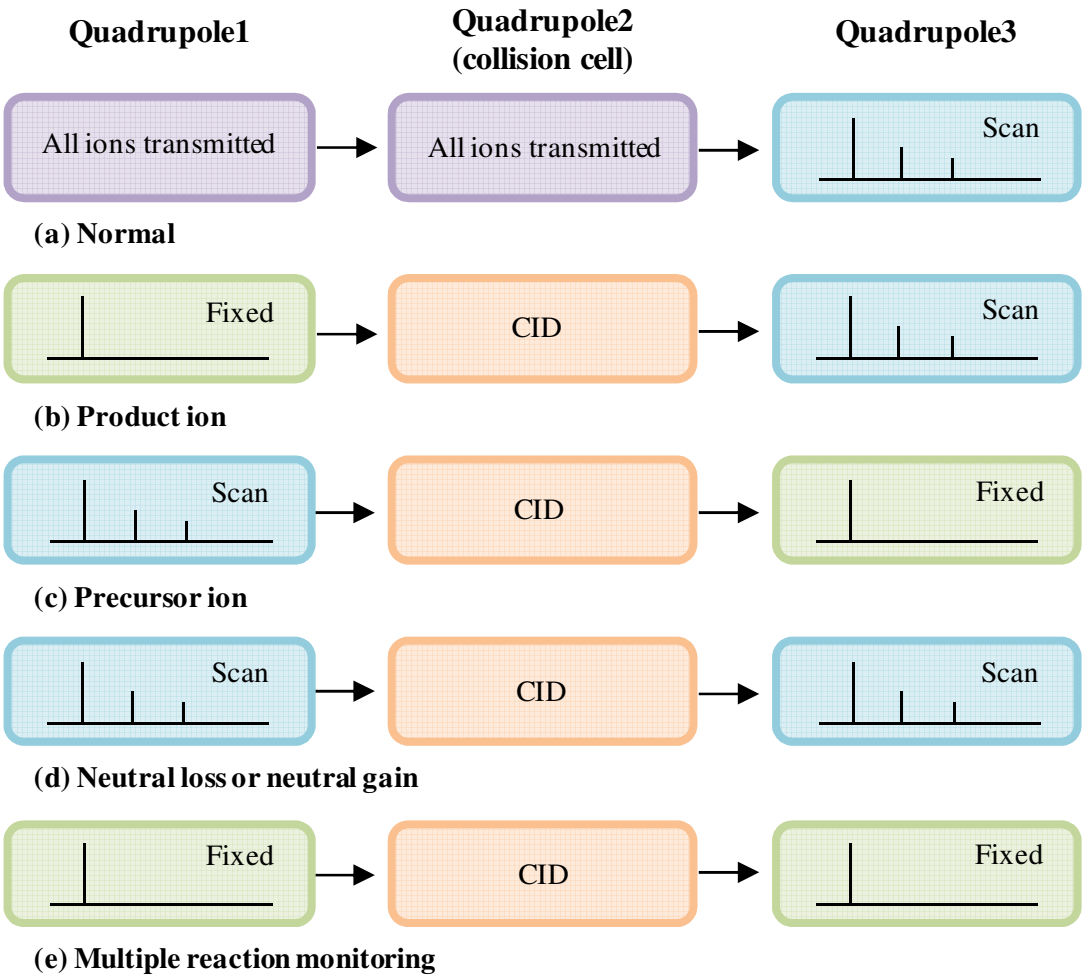


Figure 1.2: Schematic representation of the configuration of a triple quadrupole mass spectrometer for scan types, including (a) normal, (b) product ion, (c) precursor ion, (d) neutral loss or neutral gain, and (e) multiple reaction monitoring. Abbreviation: CID, collision-induced dissociation. Adapted from Blanksby and Mitchell (2010).

Collision-Induced Dissociation

Tandem-MS experiments use activating or reactive collisions within the mass spectrometer to promote ion fragmentation. The most prominent collision technique is collision induced dissociation (CID); known also as collisionally activated dissociation (CAD) or collisional activation (CA). CID allows fragmentation of gaseous ions which were stable before activation. CID is particularly useful for the structure elucidation of ions of low internal energy as those produced by soft ionization methods (Gross, 2004).

1.9.4.1 Tandem-in-space

A tandem-in-space mass spectrometer consists of an ion source, a collision cell to induce fragmentation of the selected precursor ions, and at least two non-trapping mass analyzers. The first analyzer selects precursor ions within a narrow *m/z* range, which are

then introduced in the collision cell to be fragmented. Fragments generated by the CID experiment are then analysed by the second mass analyzer either acquiring a full scan mass fragment spectrum or monitoring a selected *m/z* range (Ekman et al., 2009b).

Quadrupole-Time-of-Flight

Hybrid mass spectrometers are instruments equipped with different kinds of mass analyzers coupled together. A particularly successful hybrid tandem-in-time instrument is the quadrupole mass filter-TOF (Q-TOF). The Q-TOF can be considered as a triple quadrupole where the third quadrupole has been substituted by an orthogonal TOF equipped with a reflector. In MS mode the quadrupole mass filter only acts as transmission element and the mass spectrum can be acquired by the TOF. In MS/MS mode the quadrupole mass filter is set to transmit only the selected precursor ion; fragmentation of the precursor ions occurs in the collision cell through low-energy CID and the fragment mass spectrum can acquire by the TOF analyzer. Q-TOF instruments have high resolution, high mass accuracy, and the chance to record all ions simultaneously without scanning. However, for quantification of targeted compounds triple quadrupole mass spectrometers show higher sensitivity and linear dynamic range. On the other hand, the higher resolution of the TOF can provide a better selectivity, which can be extremely beneficial for structural elucidation (Ekman et al., 2009b).

1.9.4.2 Tandem-in-time

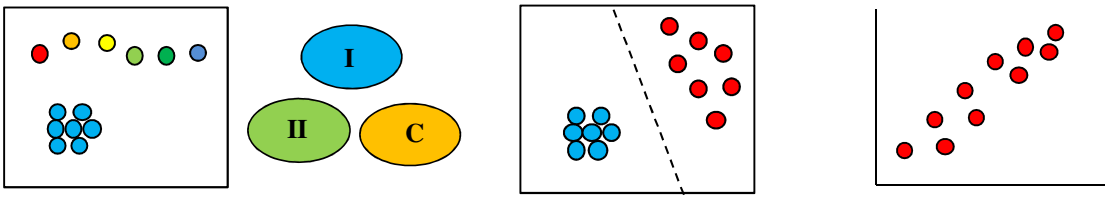
In tandem-in-time mass spectrometer ions generated in the ion source are trapped, isolated, fragmented, and *m/z* separated in the same physical device. This process is only possible in trapping devices (e.g. QIT and FTICR analyzers) (Ekman et al., 2009b).

1.10 Data analysis

Omics technologies usually generate very complex multivariate datasets and multivariate analysis using unsupervised and supervised projection methods is required in order to extract maximum information from complex omic data, to reduce the dimension of the data and to allow the class separating metabolites to be distinguished (Ramadan et al., 2006). The multivariate statistical methods collect relevant information on similarities or differences between the metabolic pathways (Winning et al., 2008, Trygg et al., 2007). Furthermore, a visual inspection of the results after multivariate analyses can quickly reveal errors, so it is often needed to validate the various outputs of the analyses. The methods used in the metabolomics approach include multivariate

projection methods, where principal component analysis (PCA) (Xie et al., 2009, Viant et al., 2008) and partial least squares (PLS) (Mahadevan et al., 2008, Woo et al., 2009, Kim et al., 2008) are the most generally used approaches (Figure 1.3). Multivariate analysis techniques can be categorized as either supervised or unsupervised (Brereton, 2003) (Bernhard Lendl and Karlberg, 2005):

- Unsupervised pattern recognition (cluster analysis): No prior knowledge about the samples to be classified is required. Additionally, when the amount of information available is too large, often cluster analysis is the only suitable tool enabling a search for similarities.
- Supervised pattern recognition (classification): Initial information about the classes (e.g gender, doses, time) is required for the set of samples as a training set. The training set is used to construct a model and assess necessary parameters to be subsequently applied in another set of samples (termed the work set). Supervised pattern recognition is usually used to assign samples to a groups (or classes). It differs from cluster analysis since the relationship between samples is important, there are no predefined groups.



Overview	Classification	Discrimination	Regression
Trends Outliers Quality Control Biological Diversity Patient Monitoring	Pattern Recognition Diagnostics Healthy/Diseased Toxicity Mechanisms Disease Progression	Discriminating between groups Biomarker candidates Comparing studies or instrumentation	Comparing blocks of omics data Metabolomic vs Proteomic vs Genomic Correlation spectroscopy (STOCSY)
PCA	SIMCA	PLS-DA OPLS-DA	O2-PLS

Figure 1.3: Fundamental data analysis objectives using a variety of mulitvariate modelling approaches, "SIMCA" stands for Soft Independent Modeling of Class Analogy. Adapted from Multivariate Data Analysis for Omics, Umetrics workshop, York, UK 2008.

1.10.1 Principal component analysis

Principal component analysis (PCA) is one of the most widespread exploratory techniques in multivariate analysis (Izquierdo-García et al., 2011). PCA is used to simplify complex datasets into interpretable models and present the interrelationship

between the samples (observations) in a low dimensional space (Eriksson et al., 2006). PCA decomposes the data into score vectors and loading vectors, which are used to recreate the original data (Kristian Hovde, 2011). The directions of the loadings are placed in order to maximize the variation spanned by each vector; most variation is in the first component and subsequent orthogonal components show decreasing amounts of variation (Kristian Hovde, 2011). The first principal component accounts for the maximum of the total variance in the original space; the second (non-correlated with the first component) accounts for the maximum of the residual variance, and so on, until the total variance is explained (Izquierdo-García et al., 2011). However, the majority of components are regarded as uninteresting or noise. On the other hand, orthogonality between components in PCA analysis imposes a rigid structure, which is suitable for extracting the highest amount of information from the samples but does not separate natural phenomena well in the components since most naturally occurring phenomena are not orthogonal (Kristian Hovde, 2011).

Aims of PCA

- PCA aims to determine underlying information from multivariate raw data.
- As far as specifically concerns LC-MS data, two pieces of information would be extracted:
    - the scores (i.e. clustering of samples or observations)
    - the loadings (i.e. the retention time X spectra variables) contributing to the score patterns (Brereton, 2003).

How PCA works (concept)

1) It finds a component (dimension vector) which explains most of the x-variation
2) It finds a second component orthogonal to the first (i.e. not correlated with) and explains most of the remaining x-variation (Figure 1.4).

How PCA works (Visually)

PCA visually projects onto a hyperplane defined by the first two components, the multidimensional dataset (Figure 1.4).

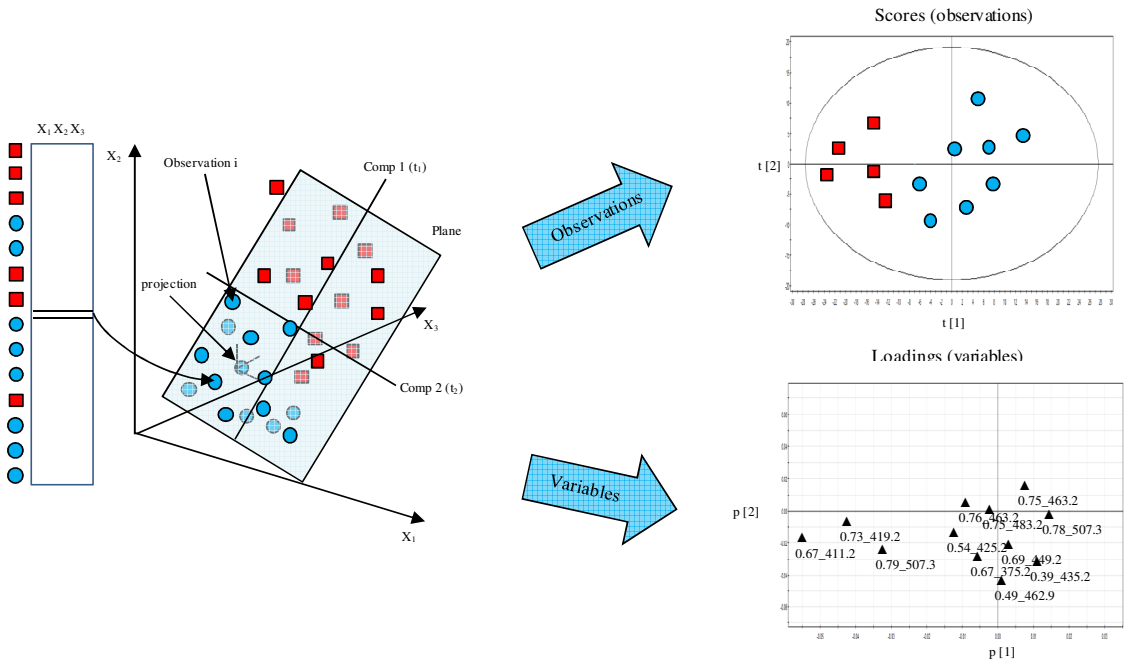


Figure 1.4: A principal component analysis (PCA) model approximates the variation in a data table by a low dimensional model plane. This model plane provides a score plot, where the relation among the observations or samples in the model plane is visualized, for example, if there are any groupings, trends, or outliers. The loading plot describes the influence of the variables in the model plane, and the relation among them. An important feature is that directions in the score correspond to directions in the loading plot, and vice versa. Adapted from Trygg et al. (2007).

Interpretation of the PCA components:

The most difficult part of the PCA approach lies in the interpretation of the components. Generally speaking, PCA takes into account:

- strength and direction of loadings.
- clusters of variables which may be related or have a common origin.

The PCA methodology ensures that components are extracted in decreasing order of explained variance. The first component always explains the majority of the variance, the second component explains the next most significant amount of variance, and so forth. Eventually, the higher-level components represent mainly noise. This is one of main reasons why PCA can result in useful data interpretation: since noise is confined in the higher-level components and it is not present in the first few components. This is due to the fact that all components are orthogonal to each other, therefore statistically independent or uncorrelated.

1.10.2 Partial least squares-discriminant analysis

Partial least squares-discriminant analysis (PLS-DA) is a supervised linear regression method where the latent variables constructed with the multivariate variables (observations or spectral descriptors) are associated with the class membership for each sample.

PLS-DA aims to:

Find those components which can significantly describe relevant variations in the samples and show maximum covariance with the class information vector. PLS-DA can show the goodness of separation between classes and the statistical significance of the obtained results (Izquierdo-García et al., 2011).

Partial least squares methods can do a more efficient and interpretable decomposition than PCA when an informative response (e.g. disease state, sex or other sample-specific information) is available (Kristian Hovde, 2011).

How PLS-DA works:

PLS-DA establishes, through the dataset, an axis as latent variable which attempts to best explain the variance in the X matrix as well as showing good correlation to the Y matrix (Eriksson et al., 2001). The latent variable basically represents high variation within the dataset which can distinguish between sample classes. Similarly to PCA scores and loadings, data can be visualised as a scores plot and discriminatory variables can be identified by searching for the variables with the highest weighting values (Figure 1.5).

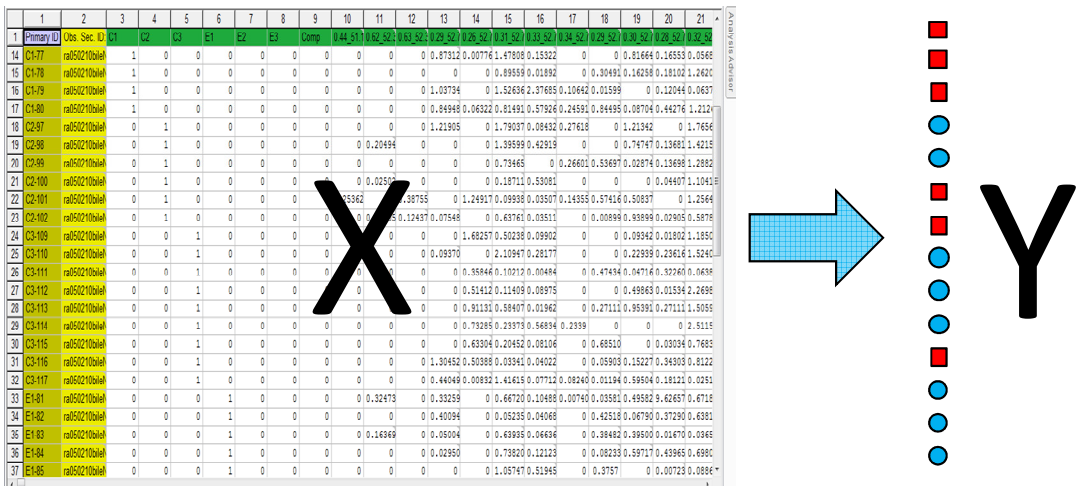


Figure 1.5: Class information can also be used to construct an additional matrix, hereinafter called the Y matrix, consisting of a discrete "Dummy" variable where [1]/[0] indicate belonging to a class. Adapted from Trygg et al. (2007).

PCA and PLS are the two main tools used in regression analysis since they convert a set of highly correlated variables to a set of independent variables by using linear transformations and at the same time they reduce the number of variables in the dataset. The PLS technique is more efficient than PCA for dimension reduction as it specifies a dependent variable for a regression due to the supervised nature of its algorithm (Figure 1.6).

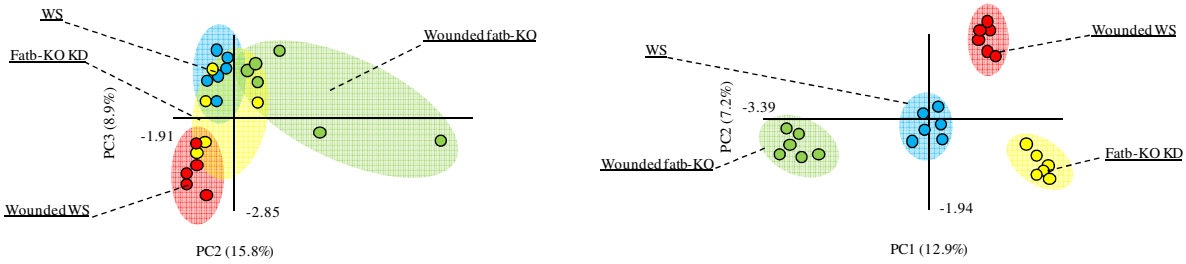


Figure 1.6: Figure to show the potential difference between unsupervised PCA and supervised PLS-DA for class separation. Experiment to show scores plots from 4 different wounding treatments of Arabidopsis plants. Adapted from metabolomics Fiehn Lab webpage ("Statistics: Metabolomics Fiehn Lab").

PLS and PLS-DA models that separate out the X variation as a predictive component due to class differences are called OPLS and OPLS-DA, respectively (Bylesjö et al., 2006). Basically, these models are rotations of the original PLS and PLS-DA models so that they allow easier interpretability, as the rotations are often beneficial with regards to plotting and directions of components. This is one of the main reason for their popularity in metabolomics studies (Kristian Hovde, 2011). OPLS separates the systematic variation in X into two parts: one linearly related to Y and one unrelated (orthogonal) to Y (e.g. Figure 1.7). This approach in turn facilitates model interpretation and model execution on new samples (Trygg and Wold, 2002, Bylesjö et al., 2006).

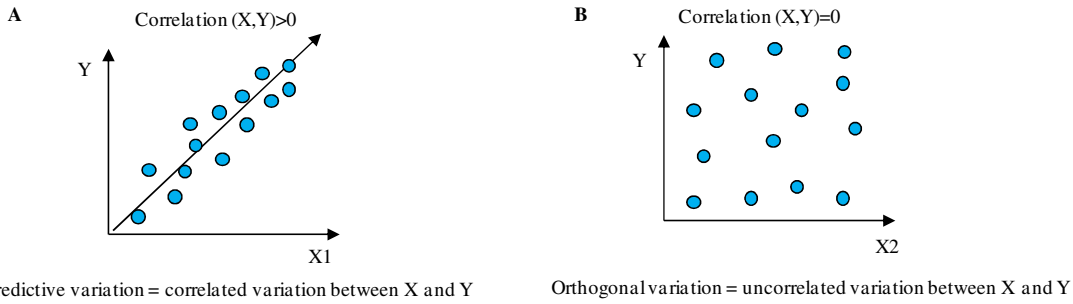


Figure 1.7: Figure to show two types of variation in mulitivariate data. A) predicted variation correlated to Y, B) orthogonal variation not correlated to Y. Adapted from Multivariate Data Analysis for Omics, Umetrics workshop, Your, UK, 2008.

1.10.3 Validation methods

The risk of over-fitting is very high for many methods without proper validation. Cross-validation and test-set validation are the two most widely used technique in multivariate statistics (Kristian Hovde, 2011). Cross-validation splits the calibration data into K equally-sized blocks, ranging from single samples to half the data set, and these blocks can be sampled consecutive, interleaved or randomized. Each of the K blocks is left out once each, and the model is trained on the rest of the blocks using the block held out for validation. Cross-validation is usually employed to compare similar models or to choose the complexity of the model. It can also give an evaluation of the level of error one expected in prediction from new, comparable data.(Kristian Hovde, 2011). On the contrary, test-set validation requires a separate data set, which is used to check the validity and fitting of the results of the statistical analysis. Since the test data is completely separate from the calibration process, the validation should reveal the reproducibility and the accuracy of the models produced. Calibration data and test data must be treated in the same way (data preprocessing). It is important that the validation data span a relevant portion of the sample space (Kristian Hovde, 2011).

1.10.4 Statistical power analysis

Statistical power is defined as the probability of correctly rejecting a null hypothesis which is false. Statistical hypothesis testing is based on 4 interrelated components: power, sample size, significance criterion (α-level), and effect size (Romesburg, 1981, Robert et al., 1997). Each of these components is a function of all the others. Increasing sample size, α, or effect size always increases power (Robert et al., 1997).

Effect size is the difference between the results predicted by the null hypothesis and the actual state of the population being tested and it can therefore measure the biological significance. Power analysis determines the chances for a determined experiment of producing a statistically significant result when there is a biologically significant difference in the population. Values of 0.8 and 0.95 have been suggested in the literature as high power values (Peterman, 1990, Thomas and Juanes, 1996).

Power analysis is used to improve research design (prospective or a priori) and to provide information about results from a previously completed research (retrospective or a posteriori). Prospective power analysis has high probability of detecting biologically significant effects (i.e., high power). Retrospective power analysis can giveinformation about statistical tests in which the null hypothesis was not rejected (Robert et al., 1997). Power analysis is employed to evaluate the sample sizes required to achieve acceptably high power, or to determine the probability for an effect size of interest to be detected with a certain sample size (Peterman, 1990). The determination of prospective power requires the establishment of sample size, α, and a biologically meaningful effect size. Power can be computed with a range of values for each of mentioned parameters and for different experimental designs, yielding a series of power curves indicating the influence of each of the evaluated parameters on the statistical power of the planned research.

The number of samples required for a multivariate approach is still controversial. Pawitan et al. (2005) suggested that the required sample size for a MVA experiment depends on the number and the distribution of the differentiating markers and on the amount of false discovery rate can be tolerated.

1.11 Metabolomics for ecotoxicological investigations

Metabolomics can provide a closer link to functional physiological responses; additionally, environmental metabolomics is an emerging approach for examining metabolic profiles in biological systems exposed to environmental stress. This approach has been employed to suggest biomarkers for the risk assessment of chemicals and for diagnosing diseases in wild animals (Lin et al., 2006). Since metabolomics is a virtually species-independent technique, it results particularly suitable for ecotoxicological investigations (Samuelsson et al., 2006). Knowing the metabolic fate of a xenobiotic aids greatly in understanding its potential toxicity and also its mechanism of toxicity. Metabolomics approaches can determine changes to metabolic pathways which might

not be detected by targeted biochemical assays or might be not present due to the time delay from gene product to metabolic product (Johnson et al., 2012). Employing toxicogenomics to screen the toxicity of chemicals can enable rapid and accurate categorisation into classes of defined mode-of-action (MOA), and prioritises chemicals for further testing. Knowledge of possible toxic modes of action (MOA) of chemicals can provide valuable insights for assessing exposure and effects, diminishing uncertainties related to extrapolation across species, endpoints and chemical structure. However, testing methods based on MOA has rarely been used to assess the ecological risk of chemicals, partly because past regulatory mandates have focused more on adverse effects of chemicals than the pathways through which these effects are elicited.

Metabolomic approaches are increasingly being used in ecotoxicology investigations. For instance, the effects of several contaminants (e.g. endocrine-distrupting chemicals (EDCs), pesticides) on the metabolomes of various aquatic species (e.g. fish, water flea) have been studied. One dimensional (^{1}H) nuclear magnetic resonance (NMR) metabolomics has been used to compare blood plasma and plasma lipid extracts from rainbow trout exposed to the synthetic contraceptive estrogen ethinylestradiol (EE2) with plasma from control fish. The plasma metabolite profile was affected in fish exposed to 10 ng/L but not 0.87 ng/L of EE2, which was in agreement with an induced vitellogenin synthesis in the high dose group only. The main affected metabolites were vitellogenin, alanine, phospholipids and cholesterol (Samuelsson et al., 2006).

^{1}H NMR spectroscopy has been used to profile metabolite changes in the livers of fathead minnows (*Pimephales promelas*) exposed to the EE2. The study revealed greater impact of EE2 in liver metabolism in males compared to females. Biochemical effects observed in the males included changes in metabolites relating to energetics (e.g., glycogen, glucose, and lactate) and liver toxicity (creatine and bile acids). Amino acids associated with vitellogenin (VTG) synthesis increased in livers of EE2-exposed males, a finding consistent with increased plasma concentrations of the lipoprotein in the fish (Ekman et al., 2008). In further work, alterations in hepatic lipid profiles of fathead minnow exposed to EE2 were also determined using ^{1}H NMR spectroscopy-based metabolite profiling. The exposure resulted in a number of sex-specific changes in lipid profiles that were also highly time dependent. Metabolites most affected by

exposure were diglycerides, triglycerides and cholesterol. Changes in the length and degree of unsaturation of hepatic fatty acids observed (Ekman et al., 2009a).

In other work, roach (*Rutilus rutilus*) were exposed to EE2 for 18 days to investigate the effect of estrogen exposure on steroid homeostasis. Exposure to EE2 resulted in a concentration dependent reduction of estrogens and androgens in bile and plasma of both male and female fish. Significant reductions in concentrations of hydroxyprogesterone, androstenedione, 11-hydroxyandrostenedione, and 11-ketotestosterone were detected in the testes metabolome, indicating disruption of steroid biosynthesis upstream of androgen metabolism. Estrogen exposure also resulted in increased biosynthesis of cortisol and cortisone in testes and ovaries, respectively, but did not alter glucocorticoid concentrations in the liver or plasma. This study suggested that both sex steroid and glucocorticoid pathways are one of the primary targets of estrogen exposure in fish gonads (Flores-Valverde et al., 2010).

The molecular responses of male roach (*Rutilus rutilus*) exposed to the antiandrogen fenitrothion were investigated using environmentally realistic concentrations, for a 28 day exposure period and revealed that O-demethylation was confirmed as a major route of pesticide degradation. Fenitrothion significantly depleted acetylcholine, confirming its primary mode of action, and 11-ketotestosterone in plasma and cortisone in testes, showing disruption of steroid metabolism. Metabolomics also revealed significant perturbations to the hepatic phosphagen system and previously undocumented effects on phenylalanine metabolism in liver and testes (Southam et al., 2011).

Changes in metabolism of Japanese medaka (*Oryzias latipes*) embryos exposed to dinoseb (2-sec-butyl-4,6-dinitrophenol) have been determined by *in vivo* ^{31}P NMR, high-pressure liquid chromatography (HPLC)-UV, and ^{1}H NMR metabolomics. Dinoseb exposure at sublethal concentrations resulted in significant declines in (ATP) and phosphocreatine (PCr) at 110 h. Reduced eye growth and diminished heart rate occurred in a concentration-dependent fashion. Metabolic effects measured by in vivo ^{31}P NMR showed a significant increase in orthophosphate levels and significant decreases in (ATP), (PCr). Metabolomics revealed a dose-response relationship between dinoseb and endogenous metabolite changes, with both dinoseb concentrations producing significantly different metabolic profiles from controls, including decreased

concentrations of ATP, PCr, alanine and tyrosine, and increased concentrations of lactate with medaka embryotoxicity (Viant et al., 2006).

^{1}H NMR-based metabolomics has been used to add a suite of metabolic endpoints to the Japanese medaka (*Oryzias latipes*) embryo assay. Medaka exposed throughout embryogenesis to five concentrations of trichloroethylene. While the no-observable-adverse-effect-level for hatching success, metabolic perturbations detected at all exposure concentrations. Twelve metabolites (i.e. amino acids, organic acids, energy related, osmolytes, miscellaneous) exhibited highly significant dose-response relationships. No significant increases in mortality, gross deformity or developmental retardation were observed (Viant et al., 2005).

Metabolomics has also been used to investiage invertibrate response to pollutants. Direct infusion FT-ICR mass spectrometry-based metabolomic has been utilized for toxicity testing in *Daphnia magna*. An OECD (Organisation for Economic Co-operation and Development) 24 h acute toxicity test was conducted with neonates at different copper concentrations. Significant copper-induced changes to the daphnid metabolome, consistent with the documented MOA of copper, was detected thereby validating the approach. In addition, N-acetylspermidine was putatively identified as a novel biomarker of copper toxicity (Taylor et al., 2009).

Hines *et al.,* (2007) have characterized the metabolic variability of the mussel and determined if inherent variability masked the metabolic response to an environmental stressor (hypoxia). Metabolic fingerprints of adductor muscle and mantle have been compared from four groups of *Mytilus galloprovincialis*: animals sampled directly from the field with and without hypoxia and those stabilized in a laboratory, also with and without hypoxia. Laboratory stabilization increased metabolic variability in adductor muscle, therefore completely masking the response to hypoxia. The principal source of metabolic variability in mantle was showen to be gender-based, highlighting the importance of phenotypic anchoring of samples to known life history traits. This study suggested that direct field sampling is recommended for environmental metabolomics since it minimizes metabolic variability and enables stress-induced phenotypic changes to be observed, whilst species and phenotype of the study organism must be known for meaningful interpretation of metabolomics data. In further work, marine mussels (*Mytilus edulis*) have been exposed for 7 days to two different concentrations of copper (12 and 50 μg/L) and pentachlorophenol (50 and 350 μg/L). Metabolic signatures

predictive of scope for growth (SFG) have been sought using optimal variable selection strategies and multivariate regression and then tested upon independently field-sampled mussels from rural and industrialized sites. Metabolic signatures predicted considerably reduced fitness of mussels from the contaminated (SFG = 6.0 J/h/g) versus rural (SFG = 15.2 J/h/g) site (Hines et al., 2010).

1.12 Relevance of research to policy of risk assessment

Most of the commercially used chemicals are not carefully measured in the environment, and this lack of information does not allow the correct evaluation of the potential hazard for human or wildlife exposure. The environmentally relevant chemicals need to be reliably characterized on the basis of both hazard and exposure in a rapid and efficient manner, in order to make possible a prioritization of the mentioned chemicals based on potential risk. The European Water Policy is mainly concerned with the protection, improvement and sustainable use of Europe's water resources. A key piece of legislation across all European states is the European Union Water Framework Directive (WFD) (2000/60/EC) (Commission, 2000). This legislation requires the periodic assessment of all water bodies, including rivers, lakes, estuaries, coastal waters and ground waters. Water bodies are then classified according to a system which grades their deviation from normality (high, good, moderate, poor and bad), with normality defined as a site with no, or very minor, disturbance from human activities (Environment Agency (2002)). The main objective of the WFD is to achieve "good ecological" status for all European water bodies by 2015. It is clear due to the amount of chemicals released in the environment that there is the need to prioritize chemicals for risk assessment and monitoring in the context of the European Union Water Framework Directive (EU WFD). Chemical prioritization can be focused on exposure to a chemical and it's hazard, and takes into account the persistence, bioaccumulation, toxicity and concentration levels in the environment (Muir and Howard, 2006). However, it is extremely difficult to accurately predict bioaccumulation of a chemical in wildlife, such as fish, even by using sophisticated models, and usually, analyses of tissue levels are required. The most promising fish bioaccumulation markers are body loads of persistent organic pollutants such as organochlorines. Easily biodegradable compounds do not accumulate in fish tissues in quantities which can significantly reflect exposure to them. Therefore, measurements of bioaccumulation and biomarker responses in fish exposed to contaminated waters can provide information contributing extensively to

environmental monitoring programs since they may reduce the number of chemicals in the aquatic environment which should be considered for risk assessment.

1.13 Aims and objectives

Wastewater effluents are complex mixtures of a wide variety of chemicals deriving from both domestic and industrial sources.

The main aims of this PhD thesis are to:

1. Fully characterizate the array of organic contaminants and their metabolites which are present in biofluids of fish after exposure to a selected wastewater effluent in order to understand the impact of effluent discharges on natural water bodies.
2. To identify the metabolite pathways in fish which are disrupted by effluent exposure, thereby increasing understanding of the impacts of effluents on these aquatic vertebrates.

In order to achieve these two objectives, chemical profiling and metabolomics approaches have been employed.

This work has been described in detail in the following five chapters.

- Chapter 2 describes the preliminary work performed to optimize the methodology for the xenobiotics profiling in fish bile and includes:
    - Optimization of the UPLC-TOFMS method (e.g. eluents, modifiers, chromatographic separation, ionization modes).
    - Preliminary characterization of the bile matrix in order to ascertain the methods are fit for purpose.
- Chapter 3 details a study of the bioconcentration and depuration of xenobiotics in bile of rainbow trout (*Oncorhynchus mykiss*) exposed to a WwTW for 10 days. The main aim of this study was to investigate the nature of the contaminant chemicals and their mixtures that bioconcentrate in fish exposed to a wastewater effluent. Work included:
    - Chemical profiling of organic contaminants by UPLC-TOFMS and multivariate analysis targeting the exogenous compounds.
- In Chapter 4, the profiles of the xenometabolome in blood samples from the rainbow trout exposure were investigated together with changes in profiles of endogenous metabolites as a result of effluent exposure. Work included:

- Chemical profiling of organic contaminants by UPLC-TOFMS and multivariate analysis targeting the endogenous compounds.

- A follow up study on the effects of two concentrations of the same WwTW effluent on the (xeno) metabolome profiles in plasma of the cyprinid fish, the roach (*Rutilus rutilus*) are described in chapter 5. The main aim of this study was to detect changes in the metabolome obtained from plasma of a different fish species exposed to the same wastewater effluent but for longer time and employing statistically relevant replicates. This included:
  - Investigation of new potential mechanisms and pathways which can be disrupted by effluent exposure.
  - Confirmation of the observed changes in the metabolome of an additional fish species exposed to the same effluent.
  - Determination of gender specific changes in the metabolome.
- A general discussion on the overall findings and suggestions for future work are given in Chapter 6.

CHAPTER 2: Methodology for UPLC-TOFMS Profiling of Xenobiotics and Bile Extracts from Trout

2.1 Introduction

In order to obtain the maximum information from the chemical analyses of samples such as bile, the most appropriate instrumental method should be used to detect as many analytes as possible. Many studies have shown that chromatographic techniques coupled to mass spectrometry can be a suitable tool to successfully characterize the analytical range of metabolites present in complex biological matrices (Fiehn, 2002). Gas chromatography-mass spectrometry (GC-MS) provides a sensitive, specific and reproducible approach for compound detection and quantitation. However, one of the main limits of this technique is its applicability only to volatile molecules, or compounds which can become volatile after suitable derivatization. On the other hand, LC-MS is much more versatile: it encompasses a wider mass range and many compound classes not detectable by GC-MS can be targeted. Thousands of ions in biological matrices have been separated and detected by LC-MS using low-resolution analysers (triple quadrupole or ion trap) (Idborg-Björkman et al., 2003, Lafaye et al., 2003). However, these analyzers do not discriminate between compounds with the same nominal masses; therefore, high resolution mass spectrometers are more suitable for chemical profiling. Ultra performance liquid chromatography (UPLC) coupled to high-resolution mass spectrometers Time-of-Flight (TOF) and Fourier Transform (FT) analyzers offers better options to structurally elucidate unknown metabolites. This cutting edge analytical technique provides (i) accurate mass measurement, which allows to determine the elemental composition of the studied metabolites for further identification, and (ii) structural information after tandem mass experiments, which provides enhanced signal-to-noise ratios (Madalinski et al., 2008). However, significant drawbacks such as ion suppression and isobaric interferences may result from a poor liquid chromatographic separation prior to the MS analysis. The introduction of ultra-performance liquid chromatography (UPLC)-MS has partially solved these limitations providing significantly reduced analysis time and increasing, simultaneously, the sample throughput, resolution and sensitivity. For these reasons it has been considered highly suitable for large-scale untargeted metabolite profiling (Wilson et al., 2005, Nordstrom et al., 2006) and in this study an UPLC system coupled to a high resolution mass spectrometer (TOF) was employed. The coupling between liquid chromatography

and mass spectrometry requires the use of an ion source, which desolvates the eluate and generates ions in gas phase. The number of available ion sources is quite large and the choice of ion source depends mainly on the specific application. Soft ion sources as electrospray ionization (ESI) and atmospheric-pressure chemical ionization (APCI) can produce intact molecular ions of large molecules such as proteins, nucleic acids, or even non-covalently bound complexes. ESI offers several advantages: it allows the analysis of low and high mass compounds, it grants excellent reproducibility and sensitivity, it is a soft ionization technique and it does not require the use of a ionization matrix (as for MALDI) (Siuzdak and Lewis, 1998).

In this chapter preliminary work was focused on the optimization of the analytical approach in order to tailor the methodology for the chemical profiling in fish bile. In order to do that:

- Firstly, the chromatographic separation and the mass spectrometry detection were developed to guarantee the best selectivity and sensitivity.
- Secondly, the real matrix (fish bile) was characterized in order to maximize the valuable information obtained by the analysis of this specific metabolome.

2.2 Chemicals and standards

Cholic acid 98%, deoxycholic acid 99%, taurocholic acid (sodium salt) (>97%), taurodeoxycholic acid (sodium salt) (>98%), β-estradiol 98%, testosterone, hydrocortisone, progesterone (≥97%), bisphenol A (99>%), 4-octylphenol (99>%), 4-nonylphenol (85>%), igepal CO-210 (average of 2EO units), igepal CO-520 (average of 5EO units), igepal CO-720 (average of 12EO units), hexaethylene glycol monohexadecyl ether C16EO6 (≥99%) were purchased from Sigma, UK. Diethyl phthalate (99%) was purchased from May and Baker LTD, Dagenham, UK. Ammonia, formic acid and all solvents (HPLC-grade) were purchased from Rathburn chemicals (Walkerburn, UK).

Individual stock standard solutions were prepared by dissolving 10 mg of standard in 1 mL of methanol and these were stored at -20 °C. Working standard solutions (1 ng/10µL) were obtained by further dilution in methanol/water (1:1 v/v).

2.3 Development of UPLC-TOFMS methodology for analysis of contaminant standards.

2.3.1 Mobile phase optimization

The mobile phase composition has a significant effect on peak shape and also the sensitivity of detection by ESI MS. Therefore, in order to obtain the best results in terms of chromatographic separation and mass spectrometry detection, a preliminary optimization of the mobile phases is usually required. For the organic mobile phase, acetonitrile was selected rather than methanol since acetonitrile, having lower viscosity, gives less pressure fluctuation and less bubble formation when mixed with water. Lower solvent strength (ε^o), the ability of a solvent to elute a particular compound from a column, values correspond to stronger eluents for reverse phase chromatographic separation. Acetonitrile is usually a stronger solvent ($\varepsilon^o = 0.5$) for the elution of non-polar compounds than methanol ($\varepsilon^o = 0.73$). Another parameter which must be taken into account for the selection of the most suitable mobile phase is viscosity (η). In fact, higher viscosity can give higher back pressures, requiring the use of shorter HPLC columns. Methanol is a more viscous solvent (η=0.60) than acetonitrile (η=0.37). These are the main reasons why acetonitrile is usually the most common mobile phase of choice and was used in this study.

The complex nature of bile metabolomes, which encompass a wide variety of ionizable moieties and polarities, suggested that a comprehensive UPLC-TOFMS profiling should employ both +ESI and -ESI mode in order to guarantee a reliable picture of the studied sample. Standard compounds were analysed in both +ESI and –ESI with four different mobile phase eluents. For ionization in positive mode, 0.2% formic acid was used as a modifier with an acetonitrile and water mobile phase. In the -ESI mode, acetonitrile and water mobile phase was used in a number of different compositions: neutral, with 0.2% of formic acid as acid modifier, buffered with 10 mM ammonium acetate as modifier and with ammonia as basic modifier at pH 8, pH 9 and pH 10.

A standard solution mixture 1 ng/μL in MeOH:H_2O (1:1, v/v) of cholic acid (C), taurocholic acid (TC), deoxycholic acid (DC), taurodeoxycholic acid (TDC), hydrocortisone (HCTN), β-estradiol (E_2), testosterone (T_2), progesterone (P), diethyl phthalate (DEPH), bisphenol A (BPA), octylphenol (OP), nonylphenol (NP), hexaethylene glycol monohexadecyl ether (C16EO6), nonylphenol ethoxylates

(NPEOs) was tested in the different mobile phases in –ESI and + ESI to evaluate the effect of the mobile phase on the different classes of compounds.

2.3.2 UPLC-TOFMS analysis

An ACQUITY UPLC system was employed for the chromatographic separation of bile samples. Standards and bile samples were injected onto an Acquity UPLC BEH C18 column (1.0 × 100 mm, particle size 1.7μm, Waters, Elstree, UK) which was maintained at a constant temperature of 30°C. C18 reversed-phase columns provide higher resolution than short chain columns (e.g. C8) which is essential for the analysis of complex samples such as bile and plasma. The mobile phases gradient was 0.0-4.0 min from 0% to 22 % acetonitrile (ACN), 4.0-18.0 min to 50% ACN and then 18.0-28.0 min to the 100% ACN. Re-equilibration to the initial condition was maintained for 6 min.

A Micromass (Waters, Manchester, UK) TOFMS system coupled to UPLC with an electrospray ionisation (ESI) source was used to detect the analytes of interest. The bile samples were profiled twice, once in positive and once in negative ESI mode. All the mass spectrometer parameters were manually tuned to obtain the highest MS signals providing optimum sensitivity and selectivity. The tuning parameters were optimized by direct infusion of 50 pg/μL leucine enkephalin at a flow rate of 10 μL/min using a syringe pump. Previous studies using UPLC-TOFMS indicated that the following parameters provided to highest sensitivity for detecting a ragne of organic compounds including steroidal metabolites (Flores-Valverde and Hill, 2008). The TOF analyzer was used in V mode at 9000 mass resolution. The TOF voltage was set at 9.10 kV and capillary voltage was set at 2.60 kV in +ESI mode and at -2.70 kV in -ESI mode. Argon was used as collision gas at penning pressures of 5.23×10^{-7}mbar. Collision energy was set at 10 eV, the cone and multiplier voltages were 35 and 550 V respectively. The source temperature was 100°C, the desolvation temperature 250°C and the desolvation nitrogen flow was 300 l h^{-1}. Sulphadimethoxine ($C_{12}H_{14}N_4O_4S$) was used as internal lock mass at a concentration of 300 pg μL^{-1} in 1:1 v/v methanol/water in -ESI mode and 100 pg μL^{-1} in 1:1 v/v methanol/water plus 0.1% formic acid in +ESI mode. The calibrant solution was infused at 50μl/min via a lockspray interface in order to ensure accurate mass measurements. The monitored *m/z* for the internal lock mass were 311.0814 (+ESI) and 309.0658 (-ESI). Mass spectra were collected in full scan mode from 50 to 1200 *m/z*.

The relative response of the analytes under different mobile phase conditions and ESI modes was measured by integrating the peak area for the extracted ion chromatogram (XIC) for ion form of each compound. XICs were extracted considering a mass window of 50-100 ppm, and replicate analytical runs were performed for each target analyte.

Typical analytical parameters used in assay validation include: precision, accuracy, linearity, range, ruggedness, limit of detection, limit of quantitation, selectivity and specificity (Chan et al., 2004). The two most important matrices of a chromatographic method are accuracy and precision.

Accuracy: is a measure of the closeness of the experimental value to the actual amount of the substance in the matrix. It determines by performing the method to samples having known amounts of analyte which then analysed against standard and blank solutions to guarantee that no interference exists. The accuracy can be calculated from the test results as a percentage of the analyte recovered by the assay.

Precision: measures how close individual measurements are to each other. It is a measure of the reproducibility of analytical method, including sampling, sample preparation and analysis, under normal performing conditions.

Precision determines by means of the method to assay a sample for a sufficient number of times to achieve statistically valid results. It expresses as the relative standard deviation (RSD):

$$\mathrm{RSD} = \frac{\mathrm{S.D}}{\mathrm{mean}} \times 100$$

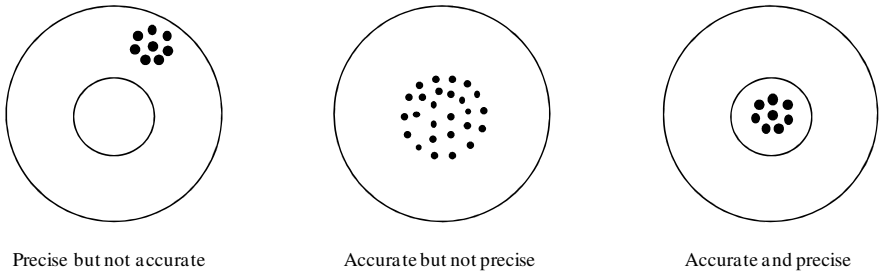


Figure 2.1: The diagram illustrates the differences between the accuracy and precision; a method can have good precision but not be accurate.

To assess assay precision, replicate assays are performed and typically at several concentrations, and the means and RSDs calculated. Most assays are designed to guarantee that the RSD does not exceed 15% irrespective of the concentration. Within-day precisions (intra-batch) determine from a single analytical run and between-day (inter-batch) values by the comparison of values from a number of analytical runs. In general, the inter-batch RSDs is greater than corresponding intra-batch values (add ref).
To order to assess the accuracy, samples having known concentration are required to compare results.
Unfortunately the accuracy and precision of the methods used in this study could not be calculated as the measurements were only based on duplicate analytical runs becuase some of the raw LC-TOFMS data were corrupted.

2.4 Analysis of selected bile samples

Bile samples were obtained from female juvenile rainbow trout (*Oncorhynchus mykiss* n=16) which had been exposed either to final undiluted effluent from a WwTW, or to charcoal filtered river water (the River Blackwater) abstracted upstream of the effluent discharge point (as controls). The details of the wastewater treatment plant and fish exposures are give in Section 3.2.2 Chapter 3. Briefly the trout were exposed for 10 days in $1m^3$ plastic tanks, and were fed commercial trout food, though no food was given 3 days prior to harvesting to maximise bile fluid production. The flow rate of the effluent and river water was 10 L/min through each tank. At the end of the experimental exposure bile samples were taken by terminally anesthetizing the fish and puncturing the gall bladder with a needle and drawing the bile into a syringe. Samples were kept at -80°C prior to further analysis.

Previous work in the Hill laboratory group (Mehinto, 2009) had shown that nonhydrolysed bile samples contained more metabolites than the hydrolysed bile samples. Enzymatic hydrolysis of bile samples is often used to remove glucuronic acid or sulphate groups from conjugated metabolites (Gibson et al., 2005a, Gibson et al., 2005b). The same study has also reported that a solid phase extraction (SPE) of bile sample for clean-up and preconcentration purposes could generate impurities in the samples. Furthermore, a lower number of MS signals were obtained for the bile samples extracted by SPE when compared with samples injected directly onto the LC-MS, which could indicate losses of important metabolites. For this reason, in this study

samples were not hydrolysed or extracted by SPE in order to avoid possible loss of relevant information.

Three composite bile samples were prepared combining aliquots from the 16 control samples and also from the 16 effluent-exposed samples (1 µL of bile from each inidividual fish) in order to provide a representative 'mean' sample containing all the analytes that will be encountered during the analysis. However, it should be recognised that some potential toxicants could be diluted and not detected when using a composite sample analysis. The obtained composite samples (assuming bile matrix as aqueous) were diluted to a final ratio methanol:water (1:1 v/v). Diluted samples were then filtered using a 96-well Strata Protein Precipitation Plates system (0.2µM, Phenomenex, Cheshire, UK). The resulting filtrate was transferred to 1 mL HPLC vials and kept at -20°C before UPLC-TOFMS analysis.

2.5 Results and discussion

2.5.1 Effect of mobile phase composition and ESI mode on detection of contaminant types

The selected compounds for the test standard mixture were: bile acids (C, TC, DC, TDC); surfactants (NP, NPEOs, C16(EO6)); steroids (HCTN, E_2, T_2, P) and phthalates (DEPH). Bile acids were selected since they are the major solute components of bile fluid. NPEOs and AEOs are widely used surfactants, commonly found in wastewaters. Steroids are also likely to be detected in bile and BPA is commonly observed in wastewater due to the daily use of this product. Phthalates could also be expected to be accumulated by fish exposed to wastewater effluent. Furthermore, some of the listed compounds (E_2, NP, NP1EO, NP2EO and BPA) have already been detected in bile from fish exposed to wastewaters (Fenlon et al., 2010).

The mean relative response of duplicate analyses of the standard contaminant compounds at different UPLC-TOFMS conditions are reported in Table 2.1. In the +ESI mode, free and conjugated bile acids were not detected at concentration levels of 1ng/µL and higher concentrations were required to give rise to detectable signals (3ng/µL) (data not shown). E_2, BPA, OP and NP could not be detected at all in positive mode. T_2 was detected predominantly as protonated form $[M+H]^+$ whilst DEPH could only detected as Na-adduct. NPEOs homologues were more abundant as $[M+Na]^+$ and $[M+NH_4]^+$ adducts than the protonated form. The NPEO homologues with more than 4

ethoxylate units showed affinity to NH_4-adducts rather than Na-adducts due to the size of the different metallic ions. This could be explained referring to the destabilization of a crown ether type of complex (a complex between an ethoxylate compound and a metal ion such as Na^+ or NH_4^+) (Cohen et al., 2001). In this process, the oxygen atoms of the polyethoxylate chain can donate their free ion pair electrons to a cation agent; the flexible structure of the chain allows the molecule to wrap itself around the cation. The same phenomenon was observed for C16(EO)6 which showed a high signal as the NH_4-adduct in +ESI mode. In contrast, the use of the acidic mobile phase in -ESI mode allowed only the detection of bile acids (C, TC, DC and TDC) and the steroid HCTN amongst the standards injected at the selected concentration (1ng/μL). Bile acids were detected as deprotonated ions whilst HCTN was detected as $[M\text{-}CH_2O]^-$ after neutral loss of formaldehyde. In the -ESI mode, the neutral mobile phase showed an increase in the responses for most of the analytes compared with an acidic mobile phase but the chromatography was affected and showed large variation in retention time of the standards between repeat injections. Amongst the standards, bile acids (C, TC, DC and TDC) gave rise to the highest signals. An abundant peak was also observed for the HCTN after loss of formaldehyde. E_2, T_2 and P, which could not be detected in -ESI mode using formic acid as modifier, gave rise to detectable signal in neutral condition. BPA showed good response whilst OP, NP, NPEOs and C16(EO)6 could not be detected under these conditions. It is likely that the use of formic acid as a modifier suppressed ionization of the phenolic and ketonic steroids, and to a lesser extent, compounds containing acidic groups such as the bile acid. In negative mode the effect of ammonia as modifier was also investigated, taking into account 3 different pHs (8, 9 and 10). The best responses were obtained for mobile phases added with NH_3 at pH 10, and E_2, BPA, OP and NP showed abundant peaks. Bile acids and HCTN were also detected whilst T_2, P, DEPH, C16(EO)6 and NPEOs could not be detected using this mobile phase.

In summary, in –ESI mode bile acids and HCTN were detected using either acidic, neutral mobile phase conditions or ammonia buffered mobile phases at pH 10. In +ESI mode T_2 and P gave a high response under acidic conditions. The basic mobile phase at pH 10 offered the best sensitivity for the detection of E_2, BPA, OP and NP in -ESI mode. These compounds are ionized more efficiently in basic conditions than in acidic or neutral. In +ESI mode DEPH, C16(EO)6 and NPEOs could be detected more

easily as Na-adducts or NH_4-adducts than in any other form. All the methods resulted in good peak shapes of the analytes. The peak width (at 50% height) was less than 20 seconds (Figure 2.2). From these results it was clear that no one method could be used to detect all the target analytes.

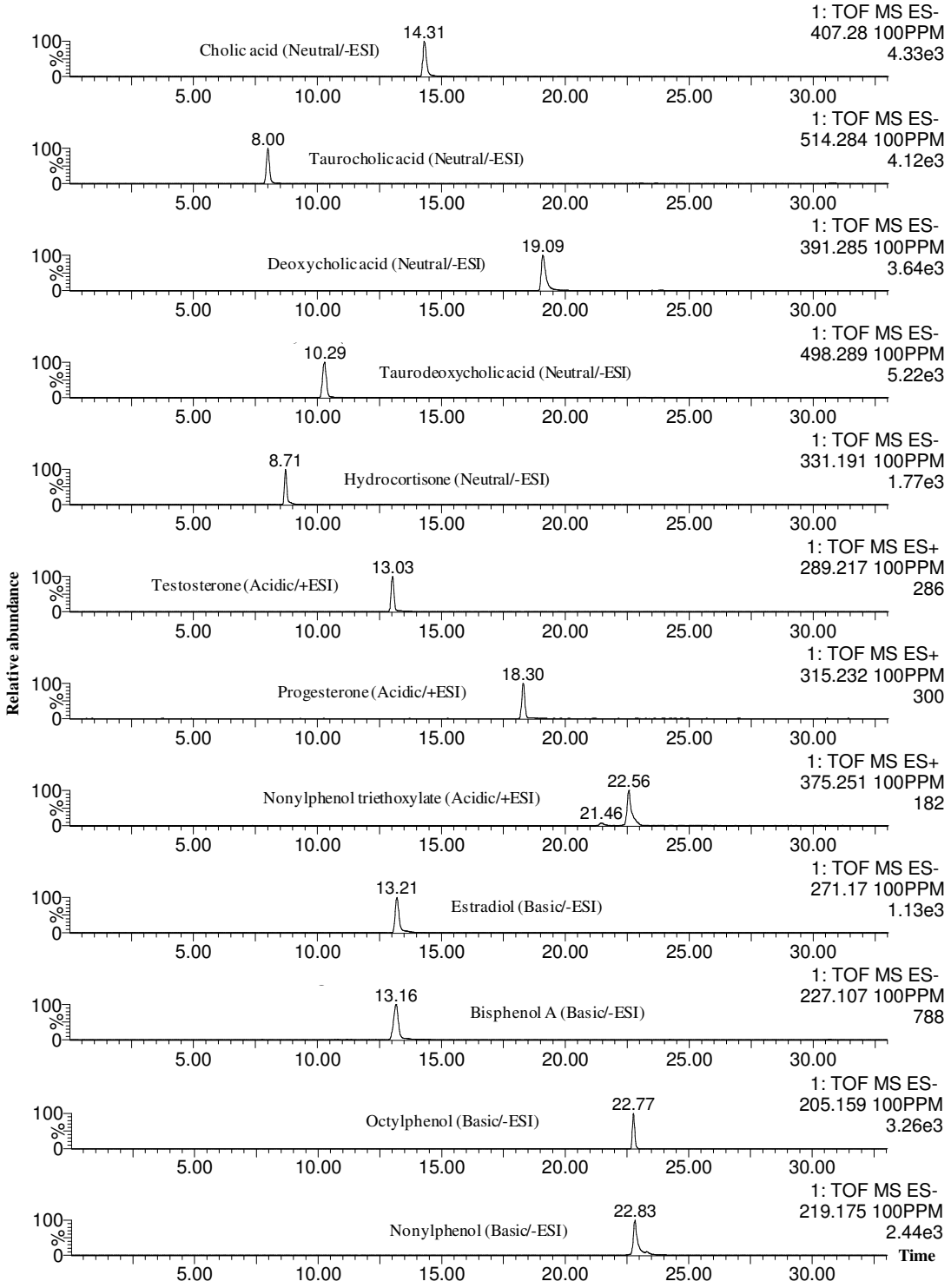


Figure 2.2: UPLC-ESI-TOFMS extracted ion chromatograms of selected standards (1ng/µL) in both ESI modes using three different mobile phase compositions (neutral, acidic and basic).

Table 2.1: UPLC-TOFMS analysis of selected standards (1ng/µL) in aqueous samples: chemical formula, retention time in min; RT, and area as average of duplicate analytical runs.

		+ESI ACN:H_2O buffered 0.2%F.A		-ESI ACN:H_2O buffered 0.2%F.A		-ESI ACN:H_2O Neutral		-ESI ACN:H_2O buffered 10mM Ammonium Acetate		-ESI ACN:H_2O buffered ammonia pH8		-ESI ACN:H_2O buffered ammonia pH9		-ESI ACN:H_2O buffered ammonia pH10	
Compound	**Formula**	**RT**	**Area**	**RT**	**Area**	**RT**	**Area**	**RT**	**Area**	**RT**	**Area**	**RT**	**Area**	**RT**	**Area**
C	$C_{24}H_{40}O_5$	-	nd	14.80	56	14.31	608	10.26	74	7.57	100	7.22	110	6.51	343
TC	$C_{26}H_{45}O_7NS$	-	nd	10.73	53	8.00	738	7.41	121	7.78	170	7.56	187	6.97	295
DC	$C_{24}H_{40}O_4$	-	nd	19.21	96	19.08	1181	15.41	244	8.89	234	8.28	255	7.98	749
TDC	$C_{26}H_{45}O_6NS$	-	nd	13.75	71	10.28	1162	9.67	179	8.87	187	8.61	221	8.40	422
HCTN	$C_{21}H_{30}O_5$	8.67	23	8.65	5[#]	8.72	200[#]	7.58	105[#]	7.60	142[#]	7.59	150[#]	9.31	236[#]
E2	$C_{18}H_{24}O_2$	-	nd	-	nd	12.68	27	-	nd	-	nd	-	nd	13.20	280
T2	$C_{19}H_{28}O_2$	13.02	45	-	nd	13.06	21	-	nd	-	nd	-	nd	-	nd
P	$C_{21}H_{30}O_2$	18.29	55	-	nd	18.32	10	-	nd	-	nd	-	nd	-	nd
DEPH	$C_{12}H_{14}O_4$	14.20	11[*]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
BPA	$C_{15}H_{16}O_2$	-	nd	-	nd	12.97	79	11.36	51	-	nd	-	nd	13.17	171
OP	$C_{14}H_{22}O$	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd	22.76	357
NP	$C_{15}H_{24}O$	-	nd	-	nd	-	nd	22.30	15	-	nd	-	nd	22.83	476
C16(EO)6	$C_{28}H_{58}O_7$	22.80	2203[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP2EO	$C_{19}H_{32}O_3$	22.56	66[*]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP3EO	$C_{21}H_{36}O_4$	22.56	68[*]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP4EO	$C_{23}H_{40}O_5$	22.54	14[*]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP5EO	$C_{25}H_{44}O_6$	22.53	9[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP6EO	$C_{27}H_{48}O_7$	22.53	23[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP7EO	$C_{29}H_{52}O_8$	22.50	43[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP8EO	$C_{31}H_{56}O_9$	22.48	68[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP9EO	$C_{33}H_{60}O_{10}$	22.43	78[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP10EO	$C_{35}H_{64}O_{11}$	22.40	87[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP11EO	$C_{37}H_{68}O_{12}$	22.38	79[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd
NP12EO	$C_{39}H_{72}O_{13}$	22.35	63[**]	-	nd	-	nd	-	nd	-	nd	-	nd	-	nd

+/-ESI: positive/negative electrospray ionization mode; ACN: acetonitrile; FA: formic acid; [*] sodium adduct; [**] ammonium adduct; [#] loss of formaldehyde; nd: not detected; C: cholic acid; DC: deoxycholic acid; TC: taurocholic acid; TDC: taurodeoxycholic acid; HCTN: hydrocortisone; E_2: β-estradiol; T_2: testosterone; P: progesterone; DEPH: diethylphthalate; BPA: bisphenol A; OP: octylphenol; NP: nonylphenol; C16(EO)6: hexaethylene glycol monohexadecyl ether; NP*n*EOs: nonylphenol ethoxylates where *n* represent EO unit. Flow rate 0.075mL/min.

2.4.2 UPLC-TOFMS profiles of bile samples

Bile samples were analysed using the mobile phase compositions tested for the standard mixture in order to select the most suitable conditions for the chemical profiling purpose. Bile samples in both ionization modes (+/-ESI) revealed many saturated peaks in the total ion chromatograms in all tested conditions. It is likely that each of the mobile phase composition tested in this study had advantages as well as disadvantages; for instance the addition of formic acid to the mobile phases would improve the separation efficiency, and facilitated the protonation of the analytes in the +ESI mode (García, 2005). However, the signal intensity of some compounds such as organic acids would be decreased in –ESI mode due to the less efficient deprotonation of the compounds in acidic conditions (Wu et al., 2004). Figure 2.3 shows the comparison between the chromatograms obtained from control and effluent-exposed fish bile samples which were analysed by UPLC-TOFMS in both ESI modes and in the presence of formic acid in the mobile phase. The chromatographic separation with acidic modifier revealed saturated peaks caused by the presence of high amounts of taurocholic acid (TC) in bile samples (control and effluent-exposed fish) in both ESI modes. In -ESI mode, TC produced the ion at *m/z* 514.2839 as $[M-H]^-$ (Figure 2.4). On the other hand, in +ESI mode three signals (different *m/z* but same RT) were found corresponding to TC at *m/z* 516.2989 (molecular ion), *m/z* 498.2883 (loss of H_2O from the molecular ion), *m/z* 480.2780 (loss of $2H_2O$) and *m/z* 462.2672 (loss of $3H_2O$ from the molecular ion) (Figure 2.4). It is interesting to note that TC could not be detected as a standard in positive mode (concentration level 1 ng/μL) whilst in the bile sample its concentration was very high allowing its detection in +ESI despite the use of an unsuitable ionization mode for that specific class of compounds. Other free and conjugated bile acids (C, DC and TDC) were not detected in either ESI modes in the bile. In addition, analyses of the bile samples were also checked for the presence of steroidal and phenolic compounds that were used in the previous analyses of the standard mixture, and they were not detected. However a previous study has already proved the presence of alkylphenols in rainbow bile as glucuronide conjugates (Ferreira-Leach and Hill, 2001). Therefore, UPLC-TOFMS ion chromatograms were carefully examined in order to investigate the presence of the alkylphenols present as conjugated forms (i.e. glucuronidated conjugates). Glucuronide conjugates of NP and it's ethoxylates (i.e. NP1EO-NP5EO, where EO represents the ethoxymer unit), could

be detected as deprotonated ions $[M+Glu-H]^-$ in –ESI mode in the presence of acid modifier. Sodium/ammonium adducts $[M+Glu+Na]^+$/$[M+Glu+NH_4]^+$ were also observed for NP1EO-NP6EO in +ESI mode. An example of the relative abundances of the conjugated compounds obtained for the NPEOs in both +ESI and -ESI are shown in Figure 2.5.

In conclusion, it is extremely likely that ionic or ionisable compounds would be present in a matrix such as bile, therefore the use of buffered mobile phases is strictly recommended in order to ensure reliable results. In the positive mode the best option was certainly a mobile phase added with formic acid since these experimental conditions allowed the detection of the whole series of alkylethoxylates surfactants standards as well as the bile acids and testosterone and progesterone steroids. On the other hand, in the negative ionisation mode both acidic and basic mobile phases (added with formic acid and ammonia, respectively) permitted the detection of bile acids and generally the peak shape was better with the addition of formic acid. However phenolic contaminants such as NP, OP and E_2, were only detected using neutral or basic mobile phases. Analysis of bile samples revealed that NPEO surfactants were present as glucuronide conjugates and it was likely that these and many other conjugated contaminants could be detected as either ammonium, sodium adducts in +ESI and as the deprotonated ion in –ESI even in the presence of formic acid. Hence, samples were profiled with formic acid in the mobile phase for both ESI modes. This would ensure the same retention time in both ESI modes for the same compounds providing an additional confirmation during the identification of unknowns and thus increasing the confidence for correctly assignation of unknown compounds.

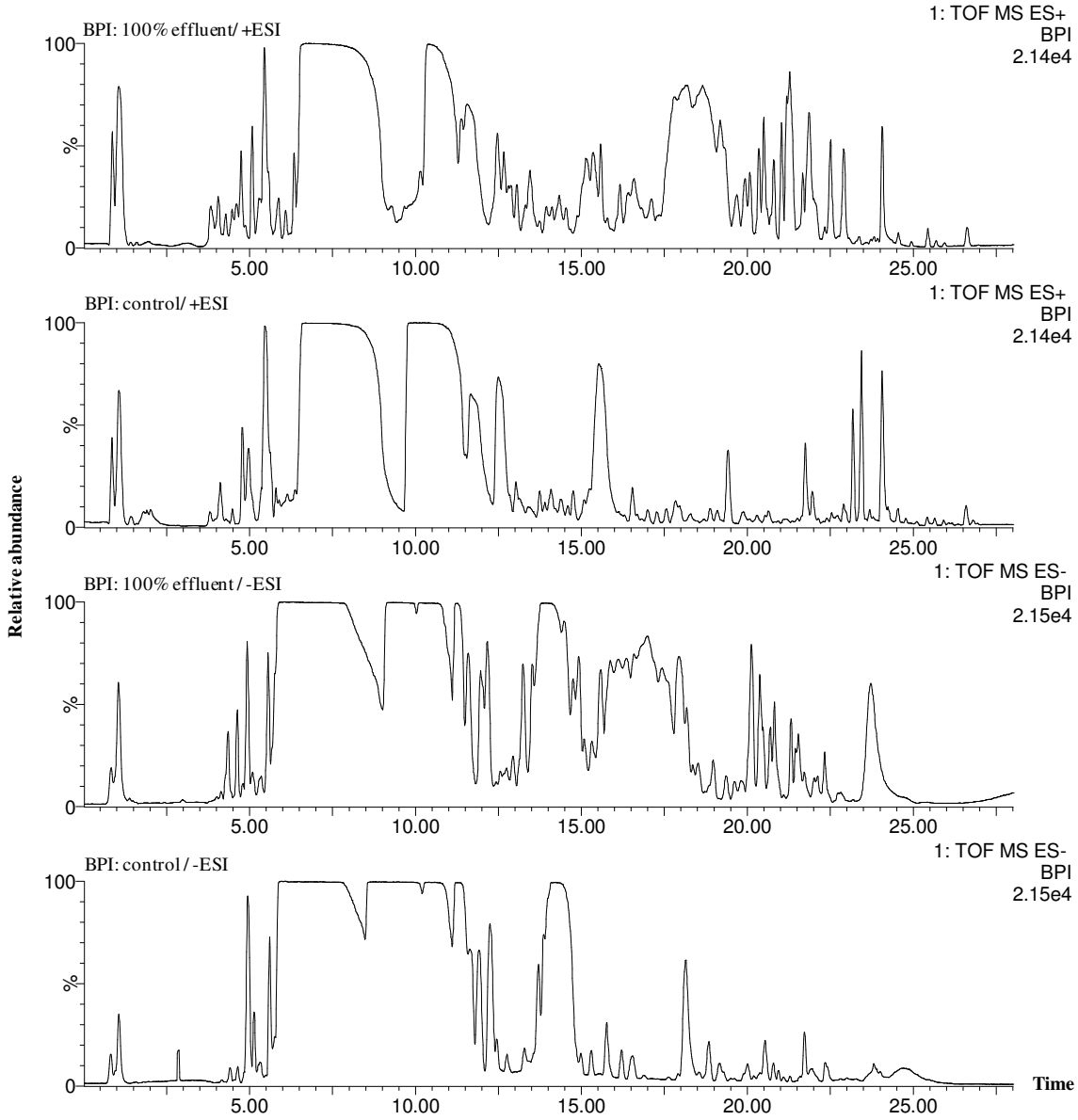


Figure 2.3: Total ion chromatograms (as base peak intensity BPI) of composite bile samples (0.2µL equivalent) for both control (river water exposed fish) and 100% effluent-exposed trout in +/–ESI modes.

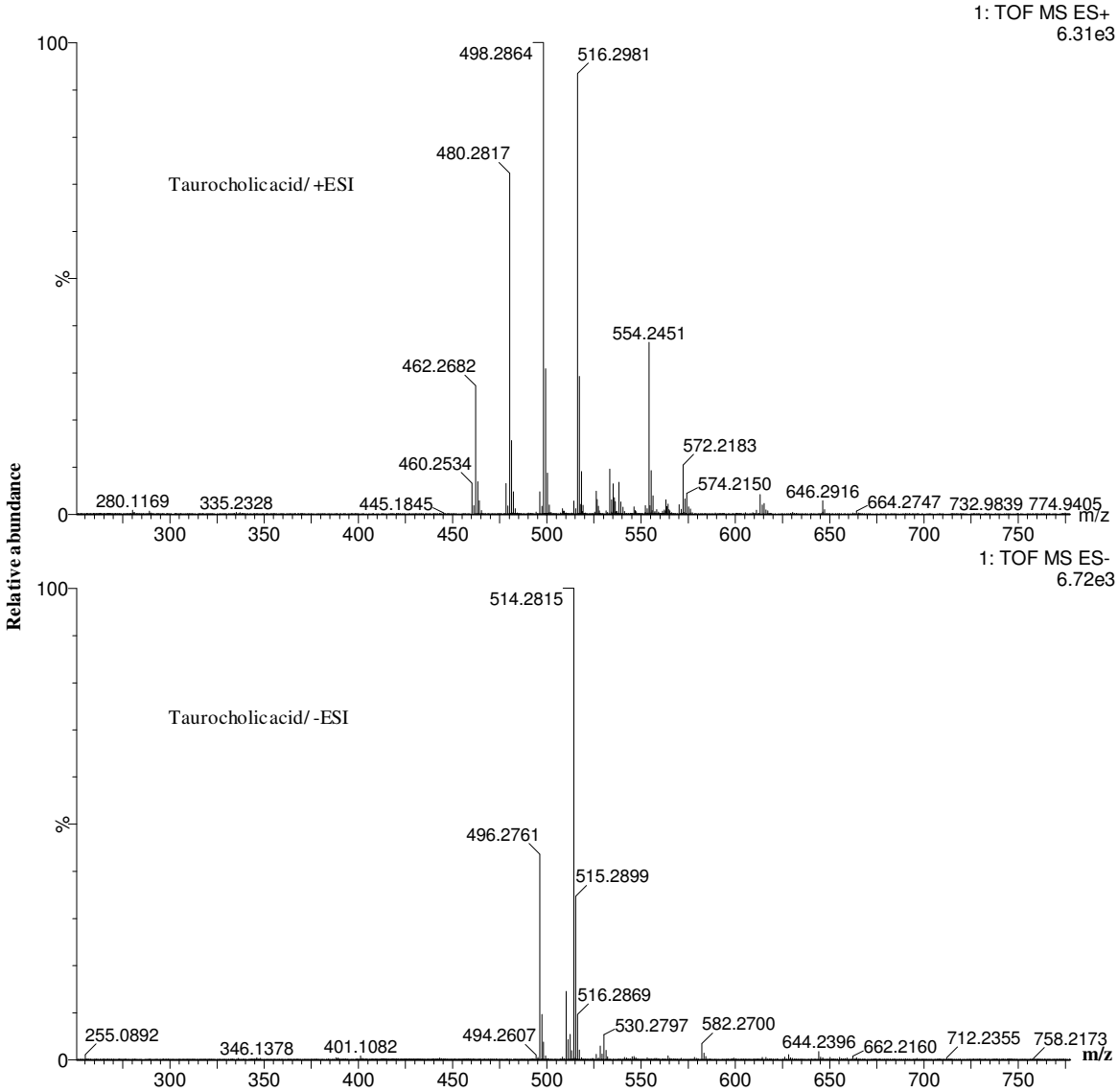


Figure 2.4: TOFMS mass spectra of taurocholic acid found in fish bile (control and effluent-exposed trout) in +/–ESI modes. The spectrum shows peaks for the deprotonated form at *m/z* 514.2815 (molecular ion) and protonated forms at *m/z* 516.2981 (molecular ion), *m/z* 498.2864 (loss of H_2O), *m/z* 480.2817 (loss of $2H_2O$) and *m/z* 462.2682 (loss of $3H_2O$).

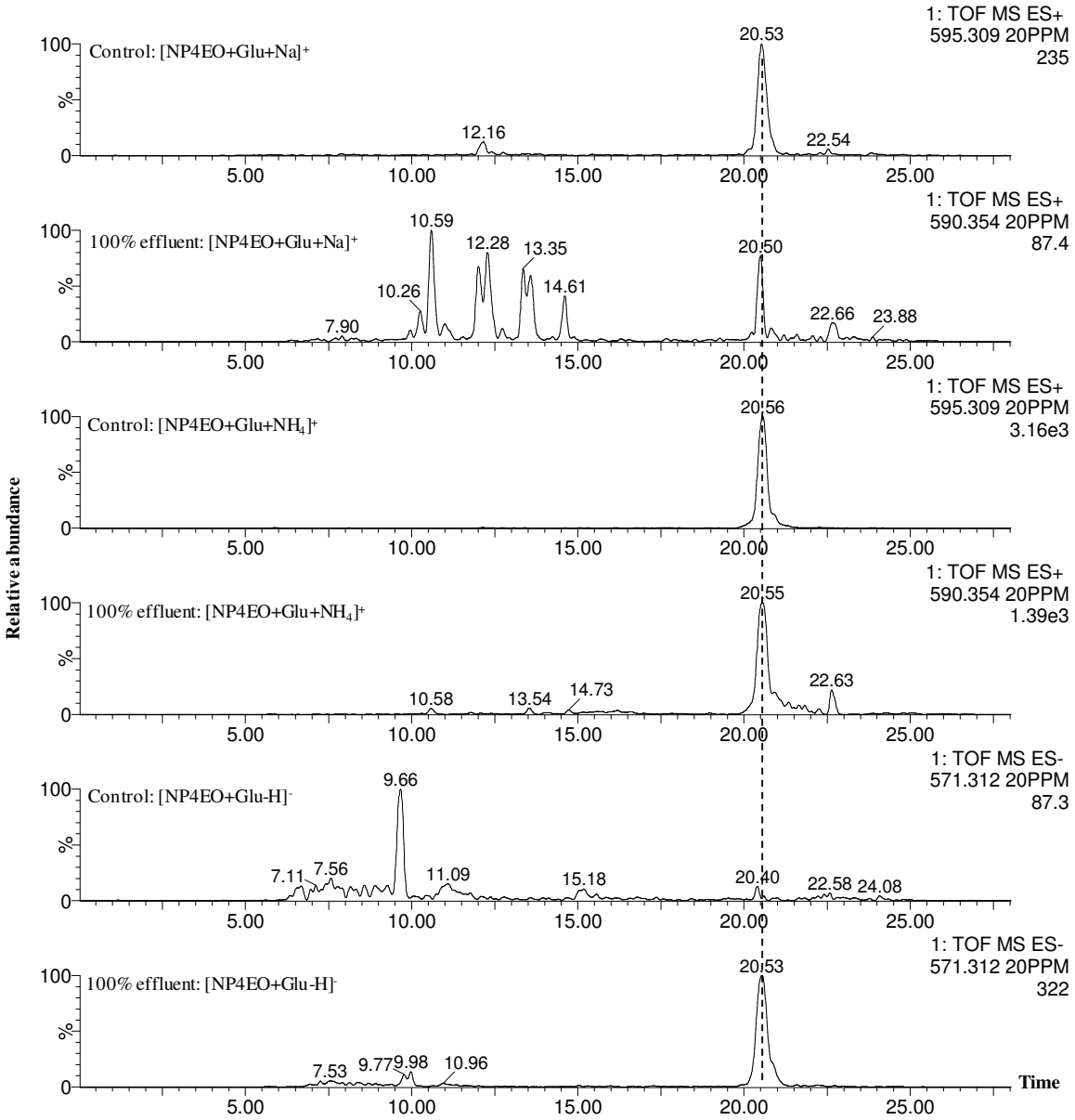


Figure 2.5: UPLC-ESI-TOFMS extracted ion chromatograms in both ESI modes of nonylphenol tetraethoxylate (NP4EO) found in bile samples from control and effluent-exposed fish. This metabolite was detected in the glucuronide conjugate form as sodium $[M+Glu+Na]^+$ and ammonium $[M+Glu+NH_4]^+$ adducts in positive mode and as deprotonated ion$[M-H]^-$ in negative mode, where Glu=glucuronide moiety. The dashed line points out the retention time of the mentioned analyte across the reported chromatograms.

CHAPTER 3: Analysis of the Xenometabolome of Bile from Trout Exposed to Final Effluent from a Wastewater Treatment Works.

3.1 Introduction

The increasing use of chemicals in the pharmaceutical, agricultural, building and manufacturing industries has led to the production of a wide variety of chemical contaminants, many of which may pose a risk to both wildlife and human health. A lot of these contaminants enter the aquatic environment via run off or through discharges from waste effluents or the sewerage systems. Although in the UK, most of the wastewater treatment plants meet UK and EU standards for effluent quality, many organic contaminants are incompletely eliminated (for example steroidal estrogens) and can still effect health of downstream fish (Viant et al., 2003). Numerous studies have documented the presence of different classes of compounds in final effluents such as pharmaceuticals, personal care products, surfactants (and their degradation products), plasticizers, insect repellents, pesticides, and flame retardants (Terzic et al., 2008). Furthermore, in some developing countries the majority of wastewaters are discharged directly into the environment either without any treatment or after being treated only mechanically (Terzic et al., 2008). The concentrations in the waste or surface waters of many of these environmental contaminants are usually extremely low (ng/L) but are still toxicologically relevant. Often the analytical techniques currently available cannot detect them unless efficient preconcentration steps have been applied. Preconcentration can be performed at different stages of the analytical procedure (e.g. during sampling, sample treatment, or preconcentration on-line during instrumental analysis). However, for the analysis of bioavailable waterborne contaminants in fish, bile fluid has been proved to be a useful qualitative and semi-quantitative monitoring aid because it efficiently concentrates xenobiotics >1000 fold from contaminated water (Oikari and Kunnamo-Ojala, 1987, Wachtmeister et al., 1991). Analysis of bile fluid from fish has been performed to detect exposure and uptake of many different substances, e.g. surfactants (Jonsson et al., 2008), polycyclic aromatic hydrocarbons (Mazéas and Budzinski, 2005), resin acids (Meriläinen and Oikari, 2008), bactericides (Adolfsson-Erici et al., 2002), and pharmaceuticals (Kallio et al., 2010). In other studies, bile has also been used as useful matrix for the indication of uptake and exposure of fish to estrogenic substances present in final effluents (Gibson et al., 2005b, Tyler et al., 2005, Fenlon et al., 2010, Pettersson et al., 2006).

Bile is a fluid produced by the liver of most vertebrates which aids the process of digestion of lipids in the small intestine. Bile consists mainly of water (85%) with the addition of other components including bile salts (10%), mucus and pigments (3%), fats (1%), inorganic salts (0.7%) and cholesterol (0.3%). The liver is one of the major organ for detoxification in vertebrates and therefore the prime site for the metabolism and excretion of xenobiotics. Most contaminants undergo hepatic metabolism and are then excreted, depending on their molecular weight and the species, either predominantly in the bile (e.g. molecular weights > 350 Da) or in the urine (< 300 Da) (Di Giulio and Hinton, 2008). The analysis of the bile fluid will therefore give an overview of many, but not all, of the contaminants taken up by the studied organism.

Most of the studies analysing xenobiotics and their metabolites in contaminated organisms have focussed on individual contaminants or a few classes of chemicals. However the introduction of chemical profiling methods, including metabolomics, allows many classes of metabolites to be analysed at one time. The metabolome includes all organic substances naturally occurring from the metabolism of the organism (except biological polymers) and can also include xenobiotics and their biotransformation products. The concept of the xenometabolome has been proposed by Holmes *et al.* (2007) as the multivariate description of the xenobiotic metabolite profile of an individual exposed to drugs, environmental pollutants, or food components that cannot be fully catabolised (the breaking down in living organisms of more complex substances into simpler ones, with the release of energy) by endogenous metabolic enzymes (i.e. non-nutrient compounds) (Wishart, 2008). Metabolomic analysis is semi-quantitative and therefore represents a promising tool for biomarker discovery.

Using available analytical methods for metabolic approach, chemical profiling of biological samples provides information on thousands of metabolites via a single analytical run. Therefore, chemometric and mathematical modelling methods (e.g. principal components analysis (PCA) and partial least-squares discriminant analysis (PLS-DA)) are required to interpret the complex resulting data sets (Lindon et al., 2006, Coen et al., 2008).

The aims of the study in this chapter were to:

1. Determine the profiles of organic contaminants and their metabolites present in bile from trout exposed to a final effluent from a wastewater treatment plant.
2. Investigate how the contaminant profiles in the bile of effluent-exposed trout change in response to a depuration period.
3. Determine the structures of the contaminants in the bile of effluent-exposed trout using a variety of mass spectrometry approaches.

Juvenile trout were exposed either to a WwTW effluent or reference river water for up to 10 days. At this time, some of the effluent-exposed trout were transferred to reference river water to investigate the effect of a depuration period on contaminant profiles in the bile. UPLC-TOFMS technique was used to profile the bile samples from the fish. Statistical methods (multivariate data analysis) based on data clustering, dimensional reduction, and multiple hypothesis testing were performed to prioritize characteristics for successive identification of key metabolites. Orthogonal partial least squares analysis discriminant analysis (OPLS-DA) was used to identify the discriminatory metabolites arising from effluent exposure. The putative identity of the metabolite was obtained after examination of the raw chromatograms. Accurate mass measurements (<2 ppm) were obtained by performing TOFMS analysis in full scan mode (W mode) in order to first examine the possible identity of the detected compounds. However, metabolite identification cannot be uniquely based on evaluation of *m/z* alone since this approach does not meet the Metabolomics Standards Initiative criteria for metabolite identification (Sumner et al., 2007), and any metabolite identifications must be regarded as putative. Kind and Fiehn have highlighted the importance of mass accuracy and how it impacts peak identification. High mass accuracy (<1 ppm) alone is not enough to eliminate enough candidates with complex elemental compositions (C, H, N, S, O, P, F, Cl, Br and Si). Utilization of isotopic abundance patterns as a single further constraint can eliminate >95% of false candidates. The application of this orthogonal filter can reduce several thousand candidates down to only a limited of molecular formulas (Kind and Fiehn, 2006, "Accurate Mass: Metabolomics Fiehn Lab"). In addition the use of the seven rules can lead to automatic exclusion of molecular formulas which are either wrong or which contain an erroneous number of elements. These rules are: (1) restrictions for the number of elements, (2) LEWIS and SENIOR chemical rules, (3)

isotopic patterns, (4) hydrogen/carbon ratios, (5) element ratio of nitrogen, oxygen, phosphor, and sulphur versus carbon, (6) element ratio probabilities and (7) presence of trimethylsilylated compounds (for GC-MS only). Using these rules, the correct molecular formula can be assigned with a probability of 98% if the formula is present in a database. Novel compounds which donot exist in databases, can be assigned with the correct formula in the first three hits of suggested formuals with a probability of 65–81% (Kind and Fiehn, 2007). However, in this study additional confirmation was provided performing Q-TOFMS experiments (product ion scan mode) based on accurate mass measurement of specific product ions. In order to obtain further information useful for successful identification, fractions of bile samples were hydrolysed by glucuronidase or sulphatase enzymes. After deconjugation parent compounds were derivatized and analysed by gas chromatography-mass spectrometry (GC-MS).

3.2 Materials and Methods

3.2.1 Chemicals

Tridecanol-EO_{12} (technical mixture), 1- naphthol (99≥%), 2-naphthol (99%), 1-hydroxypyrene (98>%), 2,2'-dihydroxybiphenyl (99>%), 2,4-dichlorophenol, 2,4,6-trichlorophenol (98%), irgasan (triclosan) (97≥%),4-chloro-3,5-dimethyl-phenol (chloroxylenol) (99>%), dichlorophene, isopimaric acid (%), mefenamic acid, Oxybenzone (2-hydroxy-4-methoxybenzophenone), potassium 4-nitrophenyl sulphate (%), 4-nitrophenol β-D-glucuronide (%), β-glucuronidase (type VII-A extracted from Escherichia coli), sulphatase (VI from Aerobacter aerogenes), bis(trimethylsilyl) triflouroacetamide (BSTFA) containing 1% trimethylchlorosilane (TMCS), and pyridine were purchased from Sigma, UK. Abietic acid (85>%), was purchased from Acros Organics. 2-benzyl-4-chlorophenol (chlorophene) (95>%) was purchased from Tokyo Chemical Industries, UK . Pimaric acid (99%) was purchased from Caltag Medsystems, UK. Deuterated internal standards [2,2,4,6,6,17α-21,21,21-$^2H_9$] progesterone (P-d_9), [2,4,16,16-$^2H_4$] estrone (E1-d_4) and [2,4,16,16-$^2H_4$] 17β-estradiol sodium 3-sulphate (E2-d_4-S) (isotope purity >98%) were obtained from C/D/N Isotopes (Quebec, Canada).

3.2.2 Fish exposure

The wastewater treatment plant (WwTP) in this study received influent from a population equivalent of 142,370: 99.6% was from domestic sources whilst the remaining 0.4% was from commercial vehicle cleaning, laundering, electroplating, plastic manufacturing, photographic development, meat processing and electronic circuit board manufacturing. The influent was treated by fine bubble diffusion activated sludge and trickling filters processes. The flow rate of the effluent discharge ranged between 52,000 and 74,500 m^3/day and the average residence time was 12 hours. Female juvenile rainbow trout (*Oncorhynchus mykiss*) were obtained from Hatchlands fish farm (Rattery, Devon, UK); individuals were measured from the head to the fork-tail and weighed revealing a mean ± SEM length of 22.8±0.2cm and weight 132.9±3.6g.

In order to investigate the complex mixture of chemical xenobiotics and their persistence in effluent exposed fish, twenty-nine juvenile fish were exposed to final undiluted effluent, and further thirty-one fish were used as a reference population. The control population was held in charcoal filtered river water (the River Blackwater) abstracted upstream of the effluent discharge point. The fish were kept in $1m^3$ plastic tanks, and were fed commercial trout food, though no food was given 3 days prior to harvesting. The flow rate of the effluent and river water was 10L/min through each tank. Tanks were continually aerated through the time of exposure to ensure oxygen levels sufficient to support the biomass of the fish. After a period of ten days, sixteen fish from the effluent tank and from the control tank, respectively, were harvested. The remaining 13 fish in the effluent tank were transferred to charcoal-filtered river water to evaluate the possible presence of further levels of contaminants after a period of depuration. Six of these fish (together with a similar number of control trout) were harvested after four days of depuration (referred as day 14 of the study), and seven were harvested after eleven days of depuration (referred as day 21 of the study) (see Figure 3.1). At the end of the experimental exposure, bile samples were taken by terminally anesthetizing the fish and puncturing the gall bladder with a needle and drawing the bile into a syringe. Samples were kept at -80°C prior to further analysis.

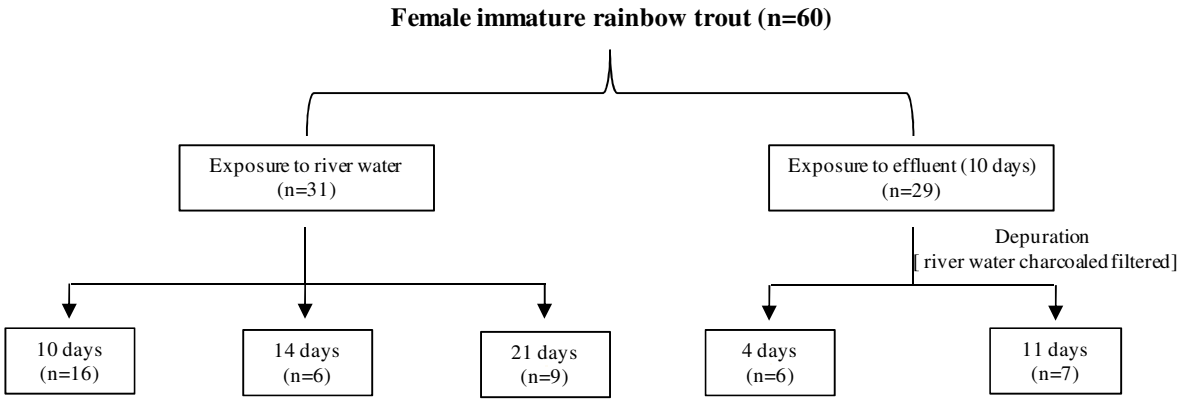


Figure 3.1: Scheme for the exposure of rainbow trout to a wastewater effluent and incorporation of depuration periods.

Based on the hypthesis that most of the contaminants present in effluent can reach the stady state within 10 days of exposure, this study was conducted for 10 days to investage the impact of short term of effluent exposure. For instance, previous studies have reported that the concentration of EE2 in bile of rainbow trout exposed to EE2 dosed water was reached the equilibrium with the EE2 dosed water tank after 10 days of exposure (Skillman et al., 2006, Flores, 2008).

The selection of depuration periods were based on the findings in the Hill laboratory group. Pedersen and Hill (2002) have documented that 4-tert-octylphenol-exposed fish that have been placed in clean water for up to 10 days resulted in a rapid loss of soluble residues from the tissues with half-lives of between 0.7 and 1.0 days (muscle, testis, ovary, gill, blood, kidney), 1.7 days (liver), and 5.9 days (bile). Therefore, 4 days and 11 days of depuration times were designed for the present study to cover depuration rates on non persistent and persistent contaminants that could be present in the effluent.

A statistical power analysis of data obtained in previous studies on the levels of contaminants accumulating in fish (Pedersen and Hill, 2002, Flores, 2008) was calculated using an effect size of 2.0. Analysis of this data revealed a power value >0.9 with a replication of 9. Hence 16 replicates were used in initial exposure in case of mortalitly and in order to investigate potential small effects on the metabolome. However lower replicate numbers wer used in depuration studies where only a higher effect size was needed as the main aim was to investigate depuration of high levels of xenobiotics in the fish biofluids.

3.2.3 Analytical procedure

The schematic summary of the whole analytical procedure applied for the study of the fish bile samples is reported in Figure 3.2. As shown in the flow diagram, sample preparation was followed by UPLC-TOFMS profiling, and the datasets were analysed using a multivariate analysis approach. Selected putative markers of effluent exposure were further characterized analysing samples (either whole bile or fractionated bile) by both GC-MS and UPLC-Q-TOFMS. In order to confirm the structural identity for the different markers, database searches were employed (NIST and various metabolomic databases). The whole analytical procedures are fully described in the following paragraphs.

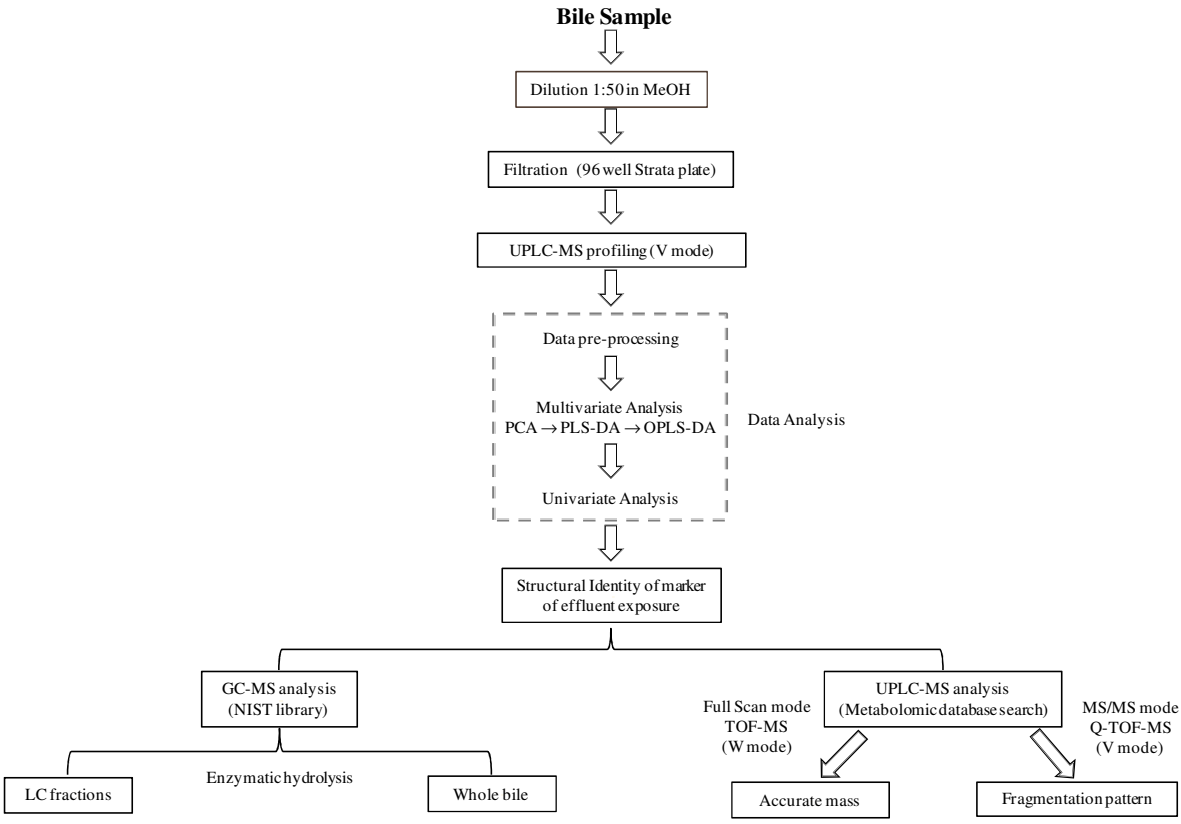


Figure 3.2: Analytical scheme for the chemical profiling and structural identification of compounds accumulating in the bile of effluent-exposed trout.

3.2.3.1 Sample treatment for analyses of chemical profiles

Bile samples (6µL) were obtained from trout exposed to both river water and 100% effluent. Bile was sampled respectively after 10 (n=16), 14 (n=6) and 21 (n=9) days of exposure in river water and after 10 (n=16) days of exposure in 100% effluent;

further samples were obtained after subsequent depuration of the trout exposed to 100% effluent in charcoal-filtered river water for a period of 4 (n=6) and 11 (n=7) days, respectively.

Bile samples were diluted 1:2 in methanol and a known amount of selected internal standards (P-d_9 & E2-d_4-S) (final concentration of 1ng/10µL of sample) was added to the samples before further treatment. Diluted samples were then filtered using a 96-well Strata Protein Precipitation Plates system (0.2µM, Phenomenex, Cheshire, UK). The resulting filtrate was then transferred to HPLC vials and kept at -20°C prior to chemical analysis.

3.2.3.2 UPLC-TOFMS chemical profiling

An ACQUITY UPLC system coupled to a Micromass (Waters, Manchester, UK) TOFMS system with an electrospray ionisation (ESI) source was employed for the chromatographic separation and analysis of bile samples according to the method described in Chapter 2 Section 2.3.2 (injection volume was 7µL for each bile sample).

3.2.3.3 Data analysis

Data pre-processing

MarkerLynx V 4.1 software package (Waters Corporation, Milford, MA, USA) was used to deconvolute and align spectral peaks. Data classification was performed using the following parameters: 0.03 Da as mass unit window, 0.2 min as retention time window, 0.03 Da as tolerance, 20 s as 5% height average peak width, 100 units as the threshold acquisition level, 50 as the number of masses per retention time, 1% of the base peak as minimum intensity allowed for a spectral peak to be considered as the same signal, 0.03 Da as mass window tolerance and 0.2 min as retention time tolerance for a spectral peak to be considered as the same signal. The isotopic peaks and background noise of the data were eliminated and datasets were then normalised (not scaled) to a set maximum total spectral area for each sample prior to exporting to SIMCA-P software for subsequent multivariate analysis (Umetrics UK Ltd, Winkfield, Windsor Berkshire, UK).

Multivariate data analysis

Datasets were Pareto scaled and log-transformed to reduce skewness of the datasets. Pareto scaling method has the advantage of enhancing the contribution of less abundant metabolites without amplification of artefacts and noise (Cloarec et al., 2005).

Principal component analysis (PCA) was performed in order to obtain an overview of the data and to detect any outliers. PCA is a multivariate projection method which represents the data matrix as a low-dimensional plane, typically of 2 to 5 dimensions (components) providing an overview of the whole dataset (Eriksson et al., 2006). The components are uncorrelated variables which are obtained from the original variables of the matrix after transformation and they account for most of the variability of the data (*ibid*). The Hotelling's T^2 and the distance to the model in the X-space (DModX) were applied to identify in the scores plot strong and moderate outliers, respectively. Partial least squares-discriminant analysis (PLS-DA) was then used to identify the metabolites affected by effluent exposure. PLS-DA is a supervised pattern recognition technique that maximises the covariance between the X matrix (LC-MS data) and the Y matrix (treatment class) (Wold et al., 2001). The number of significant components was determined by cross validation (CV). CV consists of dividing the data into a number of groups (SIMCA-P default = 7) and then building a number of parallel models from the data with one of the groups omitted. It then predicts the classification of the excluded data by the different models and compares the predictive values with the real ones. CV assesses the goodness of fit (R^2X for PCA and R^2Y for PLS-DA) and the goodness of prediction (Q^2X or Q^2Y) (Eriksson et al., 2006). Most of the reported metabolomics studies have more variables than observations and this could cause overfitting of the models. Therefore, in order to assess the accuracy of classification of the final PLS-DA models, response permutation testing was used for 60% of the replicates. One replicate was omitted in turn from each treatment class, and the dataset remodelled, and the accuracy of classification of the omitted replicate between the treatment groups was tested.

Finally, to identify the loading variables influencing treatment classification, a control and effluent treatment from the PLS-DA models with high predictive ability (where $Q^2Y > 0.5$) were analysed using orthogonal PLS-DA (OPLS-DA). OPLS-DA is a modification of PLS-DA which is able to separate out the systematic variation in matrix X into Y-related (predictive component) and Y-orthogonal (orthogonal component) (Bylesjö et al., 2006). The discriminative variables (retention time (RT) and x (*m/z*)) between control and treated samples were selected from the S-plot as loading plots of the OPLS models. This plot visualizes the influence of the loading variable in a model (magnitude of the variables or covariance) versus its reliability (correlation)

(Wiklund et al., 2008). The loading plots (S-plots) are used to describe the correlation between the variables of the reference (e.g. control) and the treated (e.g. exposed) samples, thus highlighting the markers responsible for the class separation.

Univariate statistics analysis and false discovery rate

Statistics was performed (e.g. t-test) by using Statistical Package for the Social Sciences (SPSS-V17). After the selection of the potential markers (discriminative variables) of effluent exposure from the S-plots, the p-value for each marker ion between the controls and the treatment groups was calculated. The data were tested for normality and homogeneity of variance using the Kolmogorov-Smirnov and Levene's tests respectively. Normally distributed data were then analysed using a t-test, whilst a Mann-Whitney was used for non-normally distributed data. In order to evaluate the statistical significance of the discriminative markers and to reduce false positives, data were subjected to Bonferroni correction. Bonferroni corrected critical values were calculated dividing 0.05 by the number of variables present in each model.

In addition, for each marker of effluent exposure, the mean of the fold change of the marker between control and effluent exposed fish was calculated. The standard error (confidence interval) of the mean fold change was calculated as following:

$$SE_Q = Q\sqrt{\frac{SEM_A^2}{A^2} + \frac{SEM_B^2}{B^2}}$$

, where Q indicates the fold change (Q = A/B), M the mean, A the mean of the effluent-exposed samples and B the mean of the control (Harvey Motulsky, 1995).

3.2.3.4 Identification of putative markers

The structural identity of significant the markers (discriminative variables) of effluent exposure (detected by means of the 'S' plots and confirmed by univariate analyses) were further characterized by performing different MS experiments. In initial UPLC-TOFMS analyses, the elemental composition of the markers was calculated from their accurate mass and isotopic fit using an elemental composition tool embedded in MassLynx V4.1 software. Further information on structures were obtained from UPLC-Q-TOFMS and GC-MS analyses, described in details below.

UPLC-Q-TOFMS analysis

UPLC-MS using Q-TOF as analyzer was used in order to obtain further structural information for the selected potential markers of effluent exposure. During this analysis, the internal reference standard (lock mass) was continuously infused in order to enable accurate mass measurements. Bile samples were analysed in both +/-ESI modes. The TOF analyzer was used in both V and W mode. V mode set up details has been already reported in Chapter 2 Section 2.3.2. In W mode the instrument gave a higher mass resolution of 18,000.

Therefore two different LC-MS approaches were employed to obtain structural information on the markers of effluent exposure:

1. Determination of accurate mass using high resolution W mode (full scan mode).
2. Product ion scan of selected parent ion in order to give a characteristic fragmentation pattern of the compound of interest using collision energy ranging from 30eV to 50eV in V mode (Q-TOFMS mode).

LC-MS databases used for metabolite identification

The molecular formulae of all the identified markers and their relative fragment ions obtained by means of the different mass spectrometry techniques mentioned above, were searched in a number of metabolomic databases in order to confirm the structural identity of the marker compounds. The databases used in this study were Human Metabolome database (http://www.hmdb.ca/), KEGG LIGAND (http://www.genome.jp/ligand/), LIPID MAPS (http://www.lipidmaps.org/), PubChem (http://pubchem.ncbi.nlm.nih.gov/), ChemSpider (http://www.chemspider.com/), and METLIN (http://metlin.scripps.edu/).

GC-MS analysis

Some markers gave insufficient structural information after UPLC-Q-TOFMS analysis, therefore bile samples needed to be further analysed by gas chromatography coupled to mass spectrometry (GC-MS). For this purpose, bile samples were fractionated using UPLC conditions described in section (2.3.2) and an aliquot of the fractions were subjected to enzymatic hydrolysis (described in details in the following section) before GC-MS analysis. 60µL of the internal standard E1-d_4 (1ng/µL) was added to each fraction before drying down the fraction under nitrogen stream. 30µL of bis(trimethylsilyl) trifluoroacetamide (BSTFA containing trimethylchlorosilane 1%

TMCS) as derivatizing agent and 30µL of pyridine were added then to the dried fractions and the samples were heated at 65°C for 30 min. The silylated samples were evaporated under a stream of nitrogen in order to reduce the volume to about 5µL and then 1µL was injected into the GC-MS.

GC-MS analysis was performed on a gas chromatograph (Trace GC 2000, thermoquest CE Instruments, Texas, USA) coupled with an ion trap mass spectrometer (Polaris Q, Thermoquest Ce Instruments, Texas, USA). Zebron ZB-5MS fused silica capillary column was used (30 m x 0.25 mm x 0.25 µm film thickness plus 5 m guard column) with helium as carrier gas at a flow rate of 1.5 ml/min. The source was operated in positive ionisation mode (electron impact energy: 70eV) and the detection was performed in full-scan mode. The inlet and the transfer line temperatures were both maintained at 280°C while the ion source was kept at 250°C. Samples were injected in splitless mode and separated using a temperature gradient program as follows: 70°C for 4 mins, to 264°C at 8°C/min and then maintained at 264°C for 10 mins; then to 300°C at 10°C/min and maintained at 300 C for further 5 mins. GC-MS spectra were evaluated by Xcalibur v1.2 software (Thermoquest-Finnigan) and searched in the National Institute of Standards and Technology (NIST) MS Search v1.7 and WILEY library browsers. In some cases, GC-MS analysis of fractions revealed very low signals for some of the markers of interest, and in order to overcome this limitation whole bile samples were subjected to enzymatic hydrolysis and derivatized and analysed by GC-MS as described above. GC-MS analyses of structures obtained from both fractionated and unfractionated bile samples were examined and compared with authentic commercially available standards and with the corresponding spectra of standards in the NIST and WILEY libraries.

Bile hydrolysis

Bile fractions obtained after UPLC separation were dried down and hydrolysed by adding a mixture of glucuronidase and sulphatase enzymes as reported previously (Gibson et al., 2005a, Gibson et al., 2005b). The deconjugation of the dried fractions was achieved by adding 20µL of β-glucuronidase type VII from *Escherichia coli* (1000 units/mL), 20 µL sulphatase type VI from *Aerobacter aerogenes* (2 units/mL) and 50µL 0.1M phosphate buffer solution at pH 6.05 (0.2M sodium dihydrogen orthophosphate and 0.2M disodium hydrogen orthophosphate) to the samples. The fractions were vortexed and incubated for 16 hours at 32°C and then 5µL of methanol were added to

stop enzyme activity. The activity and specificity of the individual enzymes was monitored during the hydrolysis as follows: 10µL of standard substrates were deconjugated to monitor β-glucuronidase activity (10mg/mL of standard 4-nitrophenol β-D-glucuronide) and sulphatase activity (10mg/mL of standard potassium 4-nitrophenyl sulphate) in 50µL 0.1M phosphate buffer (pH 6.05) and 20µL of HPLC water. The samples were then incubated for 16 hours at 32°C. A colour change from colourless to yellow was used to indicate the presence of the enzyme activity and the formation of the 4-nitrophenol product. Standards and samples were analysed using an UV spectrophotometer (λ=405nm) and quantification of the product was performed by comparing the absorbance of the samples with a standard curve of 4-nitrophenol concentrations prepared in 0.1M phosphate buffer.

In some cases, for further characterization of some metabolites in whole bile samples after 10 days of exposure a deconjugation step was employed. For this purpose, a composite whole bile sample, of 30 µL was deconjugated by adding 60µL of β-glucuronidase type VII from *Escherichia coli* (1000 units/mL), 60 µL sulphatase type VI from *Aerobacter aerogenes* (2 units/mL) and 150µL 0.1M phosphate buffer.

3.3 Results

3.3.1 Overview of the PCA models of the chemical profiles of bile from effluent-exposed fish

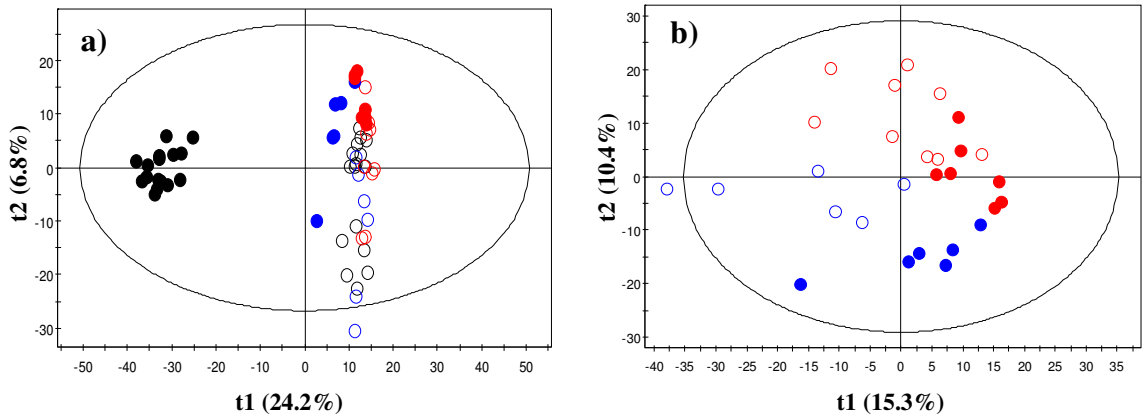
Multivariate analyses (PCA) were applied to bile datasets obtained in +ESI and -ESI TOFMS modes. The PCA models revealed a distinct separation in the metabolome between samples from control trout and those obtained from trout exposed to effluent for 10 days. As shown in Figure 3.3, the first principal component separated the observations (samples) of the effluent-exposed fish (E_{10}: 10 days effluent exposure) from those of all other treatments (C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control), E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration). This result suggests significant changes occurred in the metabolic profile of trout bile due to the effluent exposure. Dataset of control and effluent-exposed trout depurated for 4 or 11 days in river water after the 10 days exposure were reanalyzed separately by PCA. The obtained model showed some discrimination between the C_{14} control samples and the depurated E_{14} samples, but less so between C_{21} and E_{21} samples (Figure 3.3). The percentages of the explained (R^2X)

and predicted variation (Q^2) for all PCA models were low ($R^2X < 0.5$ and $Q^2 < 0.3$) in both positive/negative electrospray ionization (+/-ESI) modes (see Appendix 3.1).

All the depurated groups, whether they are controls or treated (C_{14}, C_{21}, E_{14} and E_{21}) differ significantly from each other by applying univariate statistical analysis (one-way ANOVA) of the score plots and these gave p-values of between 6.4×10^{-8} and 4.3×10^{-4}. Control groups (C_{14} and C_{21}) were considered significantly different from each other according to the first component (PC1) in both ESI modes at p-values of 0.017 (-ESI) and 0.019 (+ESI). The significant difference between C_{14} and C_{21} control groups was likely due to growth of the fish during the experiment altering metabolite profiles, changes in food availability thereby altering dietary metabolite profiles, and due to changes in the river water composition that was used during the depuration period as some contaminants were discovered after analysis of biofluids from the control fish. Effluent groups (E_{14} and E_{21}) were considered different by PC2 in both ESI modes (p-value 0.007 –ESI, 2.2×10^{-4} +ESI) but not by PC1. Significant p-values ranging 0.017 (+ESI) and 0.026 were obtained for C_{14} versus E_{14} in PC1. C_{21} and E_{21} were significantly different only by PC2 in +ESI mode giving a p-value equal to 0.006.

On the basis of the Hotelling's T^2 (critical limit 95%), three observations (samples) were considered as strong outliers: 1 outlier in +ESI dataset and 2 in the –ESI dataset (see Figure 3.3). For subsequent PLS-DA and OPLS-DA analyses, these outliers should be excluded from the datasets. However, excluding the detected outliers did not improve the explained and the predicted variation percentages and also produced new outliers in PCA. Thus, these outliers were not omitted from further analyses and the whole datasets were then reanalysed by PLS/OPLS-DA regardless.

Positive mode



Negative mode

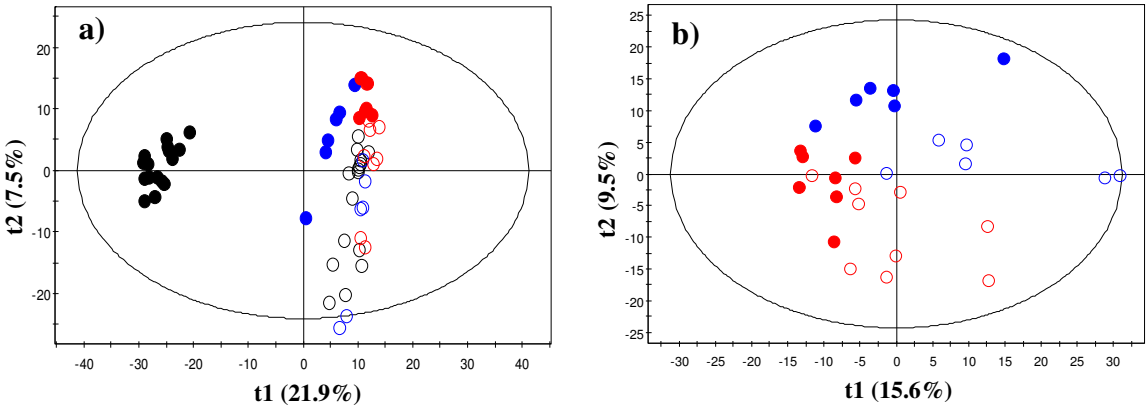


Figure 3.3: Principal component analyses (PCA) scores plots of the chemical profiles of bile samples from trout exposed to either control river water or wastewater effluent in both ESI modes. On the left hand-side, (a) PCA of whole dataset [(○), (○) and (○) symbols represent C_{10}, C_{14} and C_{21} for control trout where n=16, 6, and 9 respectively. (●), (●) and (●) symbols represent E_{10}, E_{14} and E_{11} for effluent-exposed trout where n=16, 6 and 7 respectively] and on the right hand-side (b) PCA of dataset only for the depuration period. The percentages of explained variation (R^2X) for the first two components (t1 and t2) are displayed on the relative axes. One outlier (C_{14} group) and two outliers (C_{14}) were detected in models of the whole datasets from + and – ESI modes respectively.

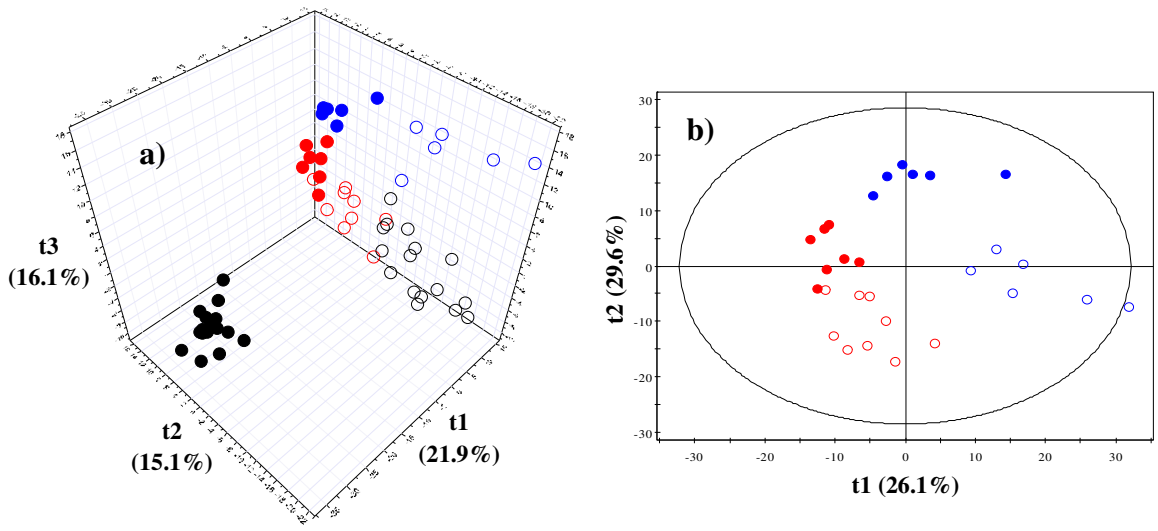
3.3.2 PLS-DA and OPLS-DA analyses of the bile from control and effluent-exposed trout

Due to the inability of PCA models to fully separate all the treatment groups, PLS-DA of the datasets was employed to further investigate class separation between the different treatments. PLS-DA correlates matrix X (variation in the dataset) with matrix Y (class membership). Datasets of bile extracts (C_{10}, C_{14}, C_{21}, E_{10}, E_{14}, E_{21}) resulted in good models in both +/-ESI modes as described by their predicted variation (Q^2 >0.97, calculated by seven-fold cross validation) and the % of explained variation (R^2Y >0.77, which is the total sum of variations in Y explained by the model) (Figure 3.4 and Appendix 3.2). In order to improve the class separation, PLS-DA between the depurated groups (C_{14}, C_{21}, E_{14}, E_{21}) were also modelled. This resulted in a clear separation between the depurated treatments, with a high degree of explained (R^2Y>0.98) and predicted variation (Q^2>0.78) (see Figure 3.4 and Appendix 3.2). This indicates that the elevated or reduced metabolites due to effluent exposure persisted in the bile even after an 11 day depuration period.

Response permutation testing was used to assess the validity of the PLS-DA (Q^2>0.5) models in terms of over-fitting. This resulted in more than 89% accuracy for the classification of treatment groups (control and effluent-exposed fish for 10 days and control and effluent-exposed fish after 4 and 11 days of depuration, respectively) (see Appendix 3.2).

The dataset for each comparison (i.e. C_{10} vs E_{10}, C_{14} vs E_{14}, and C_{21} vs E_{21}) were reanalysed using OPLS-DA to identify the class separating variables. The analyses resulted in models with one significant component (predictive component) and one not significant component (orthogonal component). The OPLS-DA analyses for the control and effluent-exposed fish after 10 days of exposure gave good models as indicated by the explained (R^2Y>99%) and predicted variation (Q^2>98%) in both +/-ESI modes (see Appendix 3.2). The comparison between C_{14} and E_{14} produced acceptable models, whilst, the resulting OPLS-DA models for C_{21} versus E_{21} did not represent accurately the dataset as highlighted by the low values of explained and predicted variation (see Appendix 3.2).

Positive mode



Negative mode

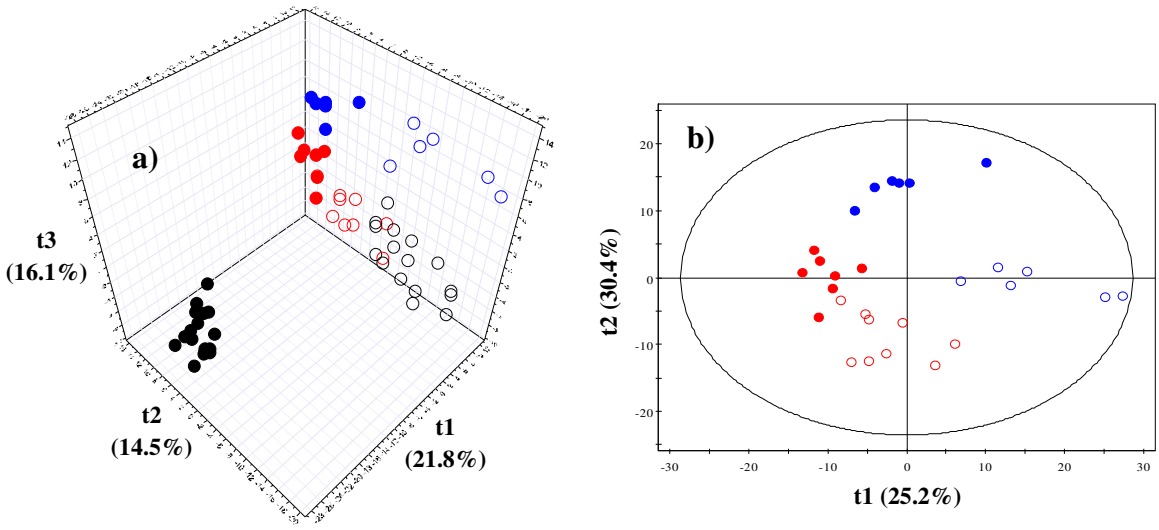


Figure 3.4: Partial least squares-discriminant analyses (PLS-DA) scores plots of the chemical profiles of bile of trout exposed either to 100% wastewater effluent or control river water in both ESI modes. On the left hand-side, (a) PLS-DA of whole dataset [(○), (○) and (○) symbols represent C_{10}, C_{14} and C_{21} for control trout where n=16, 6, and 9 respectively. (●), (●) and (●) symbols represent E_{10}, E_{14} and E_{21} for effluent-exposed trout where n=16, 6 and 7 respectively] and on the right hand-side (b) PLS-DA of dataset only for the depuration period. The percentages of explained variation (R^2Y) modelled for the first two or three latent variables (t1, t2 and/or t3) are displayed on the related axes.

In order to investigate which variables (RT-*m/z*) were influential in class separation, a 'S' plot of the OPLS models of bile samples from the control and effluent-exposed fish (10 day, 14 day, and 21 day exposure) were constructed. 'S' plot analysis of OPLS for C_{10} versus E_{10} model indicated over 150 potential markers (RT-*m/z* signal) in both +/-ESI modes (see Table 3.1a,b). The scores plot of the OPLS-DA model for the first two components is shown in Figure 3.5a,b for + and -ESI modes. An example of a 'S' plot obtained from the OPLS model of the –ESI dataset is given in Figure 3.5c. An example of the profiles of selected markers which were significantly elevated or reduced by effluent exposure is shown in Figure 3.3d,e. This graph shows that the *m/z* 462.9759 and *m/z* 610.2809 markers were increased and decreased, respectively, by the effluent exposure. Fewer markers were obtained from the analysis of the loading plot of OPLS-DA for C_{14} versus E_{14} (10 days effluent exposure followed by 4 days depuration) and C_{21} versus E_{21} (10 days effluent exposure followed by 11 days depuration) of bile samples (see Table 3.1a,b).

The discriminative variables or markers were confirmed as significant markers by comparing their concentrations (intensities from normalised data) in the effluent-exposed fish to intensities in the control fish for each treatment, using parametric (t-test) and non-parametric (Mann-Whitney test) statistical analyses followed by Bonferroni correction. Table 3.1 summarises the *p* values of markers detected in bile extracts from control and effluent-exposed fish after10 days of exposure.

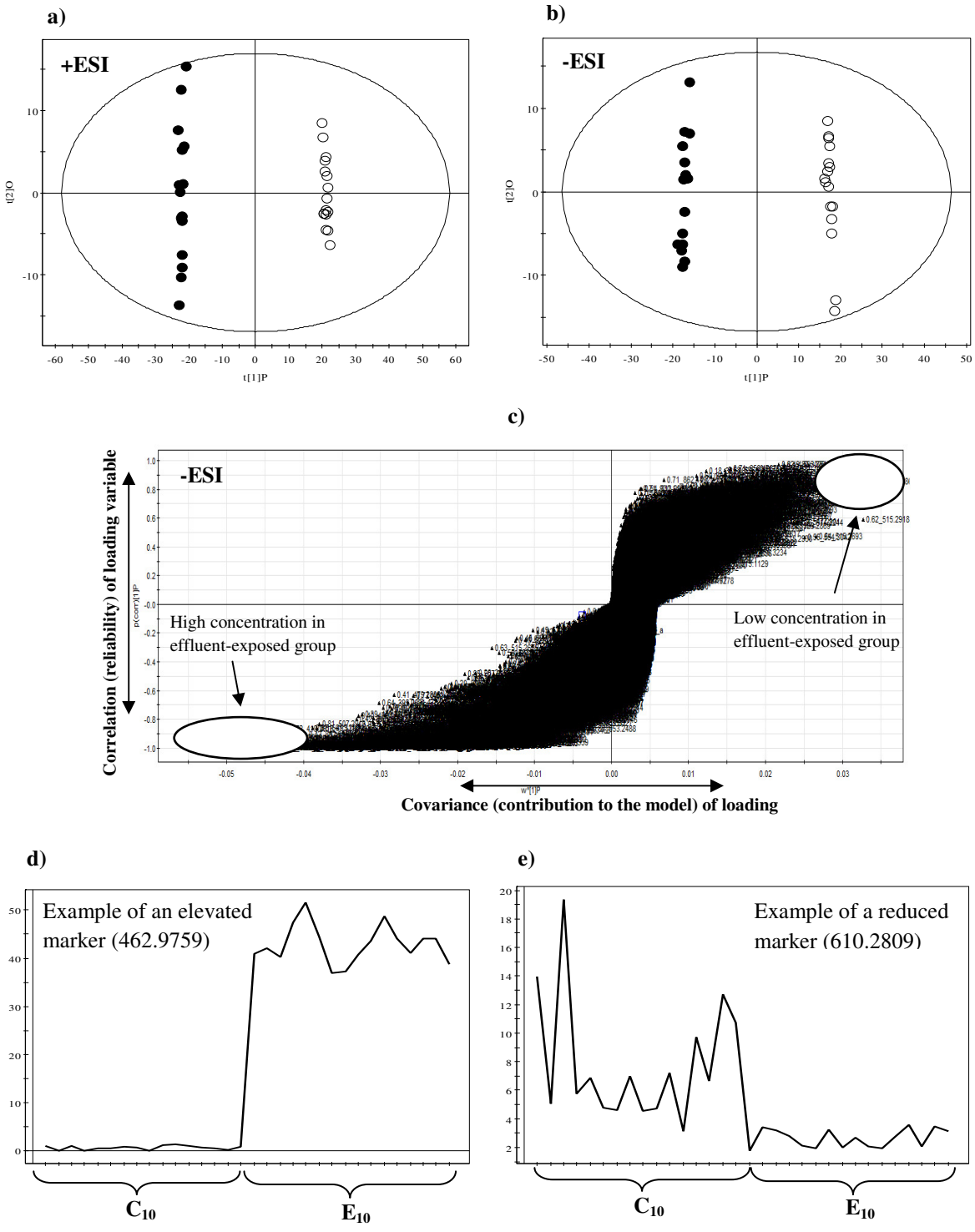


Figure 3.5: a) and b) OPLS-DA scores plots of the chemical profiles of bile of trout exposed either to 100% wastewater effluent or control river water. The samples were profiled in both +/-ESI modes by UPLC-TOFMS; (○) and (●) symbols represent C_{10} for control and E_{10} for effluent-exposed trout for 10 days, respectively, analysed in – or + ESI modes.
c) 'S' plot (contribution of variables versus confidence) from the OPLS-DA model in -ESI.
d) and e) Example of plots of signal intensity of a markers found at the extreme of the 'S' plot, where the marker concentration was either increased or decreased as a response to effluent treatment.

Table 3.1a: Markers of effluent exposure in bile from rainbow trout in -ESI mode.

Marker No.	Observed ion (*m/z*)	RT	*P* value		
			C_{10} vs E_{10}[a]	C_{14} vs E_{14}	C_{21} vs E_{21}
1	257.0482	4.54	3.33E-09	2.00E-03	nd
2	505.1747	4.79	3.33E-09	1.00E-00	nd
3	271.0640	5.16	3.33E-09	1.50E-02	4.37E-01
4	343.1217	5.32	3.33E-09	nd	nd
5	291.0866*	5.83	6.65E-09	5.89E-01	4.20E-02
6	371.1165	5.84	3.33E-09	nd	nd
7	557.0959	5.90	3.33E-09	nd	nd
8	285.0800	5.91	3.33E-09	2.00E-03	4.47E-01
9	319.0815	6.19	3.33E-09	4.00E-03	nd
10	336.9882	6.39	3.33E-09	6.10E-02	nd
11	370.9493	6.56	3.33E-09	nd	nd
12	395.0302	6.58	3.33E-09	2.00E-03	5.00E-03
13	299.0954	6.66	3.33E-09	nd	nd
14	313.1111	6.72	3.33E-09	1.00E-00	nd
15	327.1267	7.32	3.33E-09	nd	nd
16	331.0588	7.41	3.33E-09	2.00E-03	4.23E-01
17	388.1403	7.41	3.33E-09	4.00E-03	4.68E-01
18	361.0694	7.85	3.33E-09	1.82E-01	nd
19	395.0302	8.62	3.33E-09	6.10E-02	4.37E-01
20	365.0197	8.67	3.33E-09	4.55E-01	nd
21	331.1758	9.07	3.33E-09	1.00E-00	nd
22	656.3105	9.23	1.46E-07	1.50E-02	nd
23	423.0853	10.48	3.33E-09	1.00E-00	nd
24	393.0745	10.70	3.33E-09	1.82E-01	4.37E-01
25	429.0146	11.16	3.33E-09	4.55E-01	nd
26	479.2859	11.69	3.33E-09	nd	nd
27	416.1345	12.39	3.33E-09	nd	nd
28	383.2073	12.39	3.33E-09	1.82E-01	nd
29	462.9759	12.49	3.33E-09	2.00E-03	8.00E-03
30	433.1831	13.39	3.33E-09	2.00E-03	1.75E-4
31	403.2329	13.45	3.33E-09	6.10E-02	4.37E-01
32	449.2178	15.64	3.33E-09	3.03E-01	2.12E-01
33	491.3222*	12.77	4.00E-08	9.30E-02	8.37E-01
34	433.2801*	13.3	1.33E-08	1.32E-02	2.99E-01
35	464.2470*	13.68	2.33E-08	2.00E-03	1.20E-02
36	425.2171	13.88	3.33E-09	6.10E-02	nd
37	397.1859	14.14	3.33E-09	1.00E-00	nd
38	381.1922	14.22	3.33E-09	1.82E-02	nd
39	393.1924	14.64	3.33E-09	1.00E-00	4.37E-01
40	407.2071	14.71	3.33E-09	2.00E-03	nd
41	361.2231	14.76	3.33E-09	1.50E-02	nd
42	389.2169	16.13	3.33E-09	4.55E-01	nd
43	395.2072	16.42	3.33E-09	2.00E-03	5.50E-02
44	405.2488	16.79	3.33E-09	6.10E-02	nd
45	375.2383	16.94	3.33E-09	1.82E-01	nd
46	479.3374*	17.5	3.22E-06	2.40E-01	1.14E-01
47	449.2751	17.73	3.33E-09	3.03E-01	2.12E-01
48	439.2331	18.72	3.33E-09	2.00E-03	3.40E-02
49	419.2645	18.82	3.33E-09	2.00E-03	5.74E-01
50	483.2597	19.42	3.33E-09	2.00E-03	1.00E-00
51	463.2908	19.64	3.33E-09	2.00E-03	3.06E-01
52	477.2492	19.80	3.33E-09	7.77E-01	2.35E-01

ESI: electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); C as Control and E as effluent exposure: C_{10}: 10 days river water exposure; C_{14}: 14 days river water exposure; C_{21}: 21 days river water exposure; E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; nd: not detected. All markers signals in bile were increased after effluent exposure except those labelled by *.

[a] The lowest value given for Mann-Whitney test for non-parametric marker data was 3.33×10^{-9}.

Table 3.1a: (continued) Markers of effluent exposure in bile from rainbow trout in -ESI mode.

Marker No.	Observed ion (*m/z*)	*P* value			
		RT	C_{10} vs E_{10}[a]	C_{14} vs E_{14}	C_{21} vs E_{21}
53	527.2860	20.23	3.33E-09	1.43E-01	1.57E-01
54	493.3008	20.31	3.33E-09	nd	nd
55	542.3336	20.31	3.33E-09	4.55E-01	nd
56	507.3171	20.35	3.33E-09	6.10E-02	nd
57	571.3118	20.56	3.33E-09	4.55E-01	2.12E-01
58	389.2541	20.66	3.33E-09	nd	nd
59	537.3271	20.75	3.33E-09	nd	nd
60	615.3391	20.80	7.00E-03	nd	nd
61	581.3547	21.01	3.33E-09	nd	nd
62	610.2809*	21.12	1.50E-07	1.32E-01	2.99E-01
63	433.2801	21.59	3.33E-09	6.10E-02	nd
64	403.2695	21.74	3.22E-06	nd	nd
65	551.3430	21.82	3.33E-09	1.82E-01	2.12E-01
66	595.3699	21.97	3.33E-09	nd	nd
67	639.3967	22.10	3.33E-09	nd	nd
68	477.3068	22.13	3.33E-09	1.00E-00	nd
69	683.4214	22.21	3.33E-09	ND	nd
70	521.3326	22.47	3.33E-09	1.50E-02	nd
71	447.2959	22.47	3.33E-09	4.55E-01	nd
72	565.3592	22.68	3.33E-09	2.20E-02	1.00E-03
73	609.3851	22.79	3.33E-09	6.10E-02	nd
74	653.4116	22.90	3.33E-09	ND	nd
75	491.3223	22.93	2.48E-04	1.00E-00	nd
76	697.4371	22.98	3.33E-09	nd	nd
77	535.3484	23.20	3.33E-09	nd	nd
78	579.3748	23.41	3.33E-09	2.00E-03	8.80E-02
79	623.4008	23.51	3.33E-09	nd	nd
80	667.4275	23.60	3.33E-09	nd	nd
81	711.4532	23.65	3.33E-09	1.00E-00	nd
82	755.4787	23.70	3.33E-09	nd	nd
83	585.3431*	24.13	1.15E-05	8.18E-01	1.14E-01
84	403.3060	24.17	3.33E-09	1.50E-02	nd
85	447.3325	24.13	3.33E-09	6.10E-02	4.37E-01
86	491.3585	24.12	3.33E-09	6.10E-02	6.08E-01
87	535.3850	24.12	3.33E-09	4.55E-01	nd
88	579.4102	24.09	3.33E-09	nd	nd
89	461.3480	24.60	5.66E-08	nd	nd
90	505.3742	24.60	3.33E-09	2.00E-03	nd
91	549.4006	24.60	3.33E-09	4.55E-01	nd
92	593.4269	24.60	3.33E-09	1.00E-00	nd

ESI: electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); C as Control and E as effluent exposure: C_{10}: 10 days river water exposure; C_{14}: 14 days river water exposure; C_{21}: 21 days river water exposure; E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; nd: not detected. All markers signals in bile were increased after effluent exposure except those labelled by *.

[a] The least value given for Mann-Whitney test for non-parametric marker data was 3.33×10^{-9}.

Table 3.1b: Markers of effluent exposure in bile from rainbow trout in +ESI mode.

Marker No.	Observed ion (*m/z*)	RT	*P* value		
			C_{10} vs E_{10}[a]	C_{14} vs E_{14}	C_{21} vs E_{21}
1	341.1816	4.35	3.33E-09	1.82E-01	nd
2	317.1578	5.13	3.33E-09	nd	nd
3	287.1475	5.32	3.33E-09	nd	nd
4	509.1564*	5.85	1.33E-08	4.10E-02	3.51E-01
5	565.1577	6.58	3.33E-09	nd	nd
6	229.0863	7.06	3.33E-09	nd	nd
7	468.1426	8.09	3.33E-09	8.14E-01	8.07E-01
8	233.1905	8.67	3.33E-09	2.00E-03	1.75E-04
9	343.1885	9.16	3.33E-09	2.55E-01	2.00E-02
10	418.1505	12.39	3.33E-09	nd	nd
11	449.2148	13.88	2.48E-04	2.00E-03	4.80E-02
12	378.1705	13.89	3.33E-09	nd	nd
13	473.2175	13.93	3.33E-09	2.00E-03	1.75E-04
14	287.2378	14.17	3.33E-09	nd	nd
15	548.2391	14.96	3.33E-09	1.00E-00	nd
16	257.1905	15.59	3.33E-09	2.00E-03	1.75E-04
17	243.2114	15.95	3.33E-09	2.00E-03	8.74E-05
18	429.2461	16.79	3.33E-09	4.55E-01	1.75E-01
19	473.2727	17.73	3.33E-09	4.55E-01	7.00E-01
20	503.2833	18.14	5.66E-08	nd	nd
21	463.2311	18.72	3.33E-09	1.50E-02	nd
22	443.2622	18.82	3.33E-09	1.50E-02	nd
23	507.2574	19.42	3.22E-06	nd	nd
24	487.2884	19.64	3.33E-09	2.00E-03	nd
25	447.2958	19.7	3.33E-09	1.00E-00	nd
26	551.2836	20.23	3.22E-06	nd	nd
27	517.2990	20.31	3.33E-09	nd	nd
28	531.3151	20.35	3.33E-09	4.55E-01	nd
29	595.3088	20.56	3.33E-09	nd	nd
30	639.3357	20.76	3.33E-09	nd	nd
31	561.3253	20.80	3.33E-09	nd	nd
32	683.3620	20.92	3.33E-09	nd	nd
33	605.3518	21.01	3.33E-09	nd	nd
34	518.3239*	21.42	3.33E-09	2.40E-01	8.37E-01
35	575.3410	21.82	3.33E-09	4.55E-01	nd
36	619.3672	21.97	3.33E-09	1.00E-00	nd
37	663.3934	22.10	2.00E-03	nd	nd
38	501.3036	22.13	3.33E-09	4.55E-01	nd
39	707.4200	22.21	3.33E-09	nd	nd
40	545.3307	22.47	3.33E-09	2.00E-03	nd
41	589.3560	22.68	7.00E-03	nd	nd
42	633.3826	22.79	3.33E-09	6.30E-02	nd
43	677.4083	22.90	3.33E-09	nd	nd
44	443.2983	22.90	3.33E-09	2.00E-03	6.82E-01
45	515.3198	22.93	5.60E-02	1.50E-02	8.00E-03
46	721.4349	22.98	5.36E-07	1.00E-00	9.40E-02
47	765.4606	23.01	3.33E-09	nd	1.00E-00
48	559.3458	23.20	3.33E-09	2.00E-03	nd
49	603.3724	23.41	3.33E-09	nd	nd
50	573.3608	23.41	3.33E-09	nd	nd
51	647.3979	23.51	3.33E-09	nd	nd
52	413.2881	23.56	3.33E-09	2.00E-03	nd
53	457.3139	23.56	3.33E-09	2.00E-03	nd
54	501.3401	23.54	3.33E-09	6.10E-02	nd

ESI: electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); C as Control and E as effluent exposure: C_{10}: 10 days river water exposure; C_{14}: 14 days river water exposure; C_{21}: 21 days river water exposure; E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; nd: not detected. All markers signals in bile were increased after effluent exposure except those labelled by *.

[a] The least value given for Mann-Whitney test for non-parametric marker data was 3.33×10^{-9}.

Table 3.1b: (continued) Markers of effluent exposure in bile from rainbow trout in +ESI mode.

Marker No.	Observed ion (*m/z*)	*P* value			
		RT	C_{10} vs E_{10}[a]	C_{14} vs E_{14}	C_{21} vs E_{21}
55	545.3668	23.52	3.33E-09	nd	nd
56	691.4239	23.60	3.33E-09	nd	nd
57	735.4518	23.65	3.33E-09	1.77E-01	3.11E-01
58	779.4720	23.70	3.33E-09	nd	nd
59	631.4040	23.99	3.33E-09	1.50E-02	4.37E-01
60	427.3042	24.12	4.63E-07	2.00E-03	2.91E-01
61	471.3300	24.13	6.65E-09	2.00E-03	nd
62	515.3560	24.12	3.33E-09	2.00E-03	nd
63	559.3830	24.12	3.33E-09	1.50E-02	nd
64	603.4080	24.09	3.33E-09	nd	nd
65	441.3197	24.59	5.66E-08	2.00E-03	nd
66	485.3453	24.60	4.95E-04	2.00E-03	nd
67	529.3718	24.60	3.33E-09	2.00E-03	nd
68	573.3981	24.60	3.33E-09	4.00E-03	1.72E-01
69	617.4253	24.60	3.33E-09	1.90E-02	1.50E-02
70	661.4506	24.56	3.33E-09	2.00E-03	nd
71	705.4774	24.54	3.33E-09	nd	nd

ESI: electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); C as Control and E as effluent exposure: C_{10}: 10 days river water exposure; C_{14}: 14 days river water exposure; C_{21}: 21 days river water exposure; E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; nd: not detected. All markers signals in bile were increased after effluent exposure except those labelled by *.

[a] The least value given for Mann-Whitney test for non-parametric marker data was 3.33×10^{-9}.

3.3.3 Identification and verification of chemical markers due to wastewater effluent exposure

Data obtained from UPLC-TOFMS analysis can be extremely complex, especially with biological matrices; therefore simple visualization of the chromatograms does not allow full interpretation of the results. For this reason chemometric tools are essential to extract the useful information and assist in the analysis of metabolomics data sets.

Data from UPLC-TOFMS analysis allow the use of accurate mass; however many different chemicals formula can account for the same accurate mass. Therefore, as one of the most crucial step in the identification of unknown, it is necessary to select the correct molecular formula for the selected compound. The next step consists in finding likely structures for the chosen molecular formula by searching different available databases. The candidate structures must be then checked using information coming from fragmentation experiments from the product ion. Finally, the identity of the compound must be confirmed by comparison to relevant commercial standards, verifying both retention time and accurate mass of the characterised compound (Ferrer et al., 2005, Ferrer et al., 2006).

A total of 92 (-ESI mode) and 71 (+ESI mode) of the most abundant markers metabolites were detected after ten days of exposure (Table 3.1a,b). Most of these metabolites were increased in concentration in bile and only nine of them (7 in –ESI and 2 in +ESI) were decreased as a result of effluent exposure. The mass spectra of the decreased markers were examined carefully to ensure they did not result from ion suppression during MS analyses. These down-regulated markers were considered as metabolites of endogenous origin whose concentrations decreased in response to effluent exposure. Further analysis (as described in details in the following sections) revealed that most of the elevated markers were xenobiotics taken up into fish from the effluent water.

The empirical formulae of each marker metabolite were obtained operating the TOF analyser in W mode. The accuracy of the detected masses ranged between 0 and 2.3 ppm. In many cases, a number of signals (RT-*m/z*) were detected for each metabolite marker and these represented adducts of the molecular ions. For instance, markers were detected as positively charged $[M+H]^+$, negatively charged $[M-H]^-$ species, as adducts (i.e. $[M+Na]^+$, $[M+NH_4]^+$, $[M+FA-H]^-$; where FA=formic acid) and/or clusters (i.e. $[M+M+H]^+$, $[M+M+Na]^+$, $[M+M-H]^-$). Markers with multiple ion forms (e.g. adducts) showed similar fold changes between the control and effluent treatments (data not shown) and this result increased the degree of confidence for the correct assignation of studied empirical formulae of the marker metabolite. Additional information coming from the evaluation of the isotopic fit (i-fit) was used to facilitate the identification of the peaks of interest. The fit value is the least square error between the theoretical data and the experimental data. The peak that gives the lowest isotopic fit value is assumed to be the correct monoisotopic peak. Therefore, the closer the i-fit value is to 0 the better is the matching between theoretical and experimental data. I-fit values ranging between 0 and 21.2 were obtained for most of the markers (with an exception of few of them where there was coelution with interfering compounds).

In order to provide additional information besides the accurate mass value, collision induced dissociation (CID) experiments were employed to obtain the fragmentation pattern for the detected markers. The achieved results were then searched against the metabolite formulae within available databases to determine putative metabolite structures.

3.3.4 Identification of metabolite markers of effluent exposure.

Evaluation of the results as described in the previous section (3.3.3) led to the identification of different classes of metabolites present in fish bile after effluent exposure. The identified classes encompassed the most common environmental contaminants and were surfactants, pharmaceuticals, personal care products and disinfectants. In the following sections each of these classes will be described and characterized. Most of the selected markers were present in the sample as conjugated compounds (as either glucuronides or sulphates). In order to characterize the fragmentation pattern of conjugates, two known conjugated standards (nitrophenyl sulphate and nitrophenyl glucuronide) were analysed by Q-TOFMS and the most informative ions were used to confirm analyses of conjugated markers. The Q-TOFMS analyses of the standards in –ESI mode is shown in Appendix 3.3. From the analysis of the nitrophenyl glucuronide ($C_{12}H_{13}NO_9$), a fragment ion was detected at *m/z* 138.0191 ($C_6H_4NO_3$), which corresponds to the aglycone $[M\text{-}Glu\text{-}H]^-$ (where Glu=glucuronide moiety). The fragments at *m/z* 175.0243 ($C_6H_7O_6$), at *m/z* 157.0137 ($C_6H_5O_5$) due to loss of H_2O and at *m/z* 113.0239 ($C_5H_5O_3$) due to loss of CO_2, respectively, were considered as characteristic ions for the glucuronide conjugates (Wen et al., 2007). On the other hand, Q-TOFMS spectrum of the nitrophenyl sulphate ($C_6H_5O_6NS$) showed mainly a deprotonated fragment at *m/z* 138.0191, which corresponds to the deconjugated nitrophenyl $[M\text{-}SO_3\text{-}H]^-$. This fragment ion resulted from the neutral loss of the sulphate moiety (loss of 80 Da) (see Appendix 3.3). The elemental composition tool assisted in the calculation of the theoretical mass formulae of the fragment ions giving the deviation range of 0 to 5ppm between the predicted and the experimental *m/z*.

3.3.4.1 Identification of linear alkylbenzene sulfonic acid (LAS) anionic surfactants and their metabolites in trout bile

A series of LAS anionic surfactants or their metabolites were detected in the bile of effluent-exposed trout. LAS are compounds widely used in the formulation of household detergents and other cleaning products. They are commercially sold as a mixture of different homologues, with alkyl chain lengths ranging from C10-LAS to C13-LAS and each homologue consists of two different isomers which differ on the relative position of the sulfophenyl group respect to the alkyl chain: the LAS homologue is defined as an internal isomer when the sulphophenyl group is linked to the middle of the alkyl chain whilst it is considered as an external isomer when the sulphophenyl group is positioned at the end of the alkyl chain. LAS isomers can be

expressed as mCn-LAS, where m indicate the C atom of the alkyl chain which is linked to the sulfophenyl group whilst n refers to the total length of the alkyl chain. LAS surfactants can be metabolized by ω-oxidation mechanism to the sulfophenyl carboxylic acids (SPCs). SPCs are identified in a similar way to LAS homologues except for the C atom counting which starts from the C atom of the carboxylic group at the end of the alkyl chain. LAS and SPCs isomer examples are shown in Figure 3.6 (HERA, 2009a, Lara-Martín et al., 2007).

A series of peaks (markers 1, 3, and 6-9; Table 3.2) was detected in the bile of effluent-exposed fish in –ESI mode as deprotonated ions $[M-H]^-$. The ions of the series differed from each other by 14 Da, which corresponds to the methylene group. These markers were identified as the LAS metabolite sulfophenyl carboxylic acids (SPCs) with alkyl length ranging from C5 to C10 (Figure 3.7). The elution order of these metabolites was related, as expected, to the respective length of the alkyl chain, due to the interaction between the alkyl chain on the other side of the carboxylic group and the C18 stationary phase.

As briefly mentioned before, the biodegradation pathway of C10-C13LAS homologues in surface waters is via generation of sulfophenyl carboxylic acids (SPCs) due to the initial ω-oxidation of the alkyl chain and its progressive shortening by successive α- and β-oxidations until the completed LAS mineralization (Swisher, 1987). However, detection of both odd and even carbon chain series of SPCs could possibly be explained with the degradation of longer alkyl chain via β-oxidation starting from two different original molecules (i.e. C10-LAS and C11-LAS) (see Figure 3.8).

An ion at *m/z* 343.1271 (Marker 4; Table 3.2) was detected in effluent-exposed fish bile and this signal can correspond to the mono-hydroxylate C10-SPC. The presence of hydroxylated SPC could be due phase I biotransformation of C10-SPC (functionalisation reactions) either within the fish or by bacteria in the environment (see explanation for biotransformation of xenobiotics in section 3.1).

The ion *m/z* 371.1165 (marker 5) was also detected and it can be putatively identified as sulfophenyl dicarboxylic acid (C11-SPdC), assuming it is formed via the ω/β-oxidation mechanism of both alkyl chain free ends of C13-LAS. All LAS metabolites were observed as non-conjugated compounds with exception of the ion *m/z* 505.1747 (marker 2), which was identified as the glucuronide conjugate of the dihydroxylated C10-LAS.

All the above metabolites were identified with high agreement between the theoretical and experimental masses (error <2ppm) whilst fragments generated by Q-TOFMS experiments showed an accuracy within 5 ppm (Table 3.2).

Based on Q-TOFMS fragmentation experiments, the most abundant product ion shared between LAS and SPC metabolites was the ion at *m/z* 183.0116 ($C_8H_7SO_3$), which corresponds to styrene-4-sulfonate (ethenylbenzene-4-sulfonate). According to previous studies, the ion *m/z* 183.0116 was considered as a characteristic ion of LAS and SPC metabolites (González-Mazo et al., 1997). Additional ions at *m/z* 170.0038 ($C_7H_6SO_3$) and/or *m/z* 197.0272 ($C_9H_9SO_3$) were also observed; these two ions correspond to the methyl- and the propenyl-substituted benzenesulfonates, respectively. In Figure 3.9, an example is given for the fragmentation pattern of both internal and external isomers at *m/z* 257.0482 and *m/z* 271.0640 which were positively identified as C5-SPC and C6-SPC, respectively. In both cases, the alkyl chain is fragmented via McLafferty rearrangement, which consists of a hydrogen migration from the C4 of the alkyl chain, considering the carbonyl oxygen in the carboxylic group as C1, followed by a bond cleavage between C2 and C3. As a result, the molecular ions (*m/z* 257.0484 and *m/z* 271.0640) are broken and the fragment ions *m/z* 183.0116 and *m/z* 211.0429 are formed (Figure 3.9) (Lara-Martín et al., 2010).

The metabolites C7-SPC and C9-SPC appeared to be the most abundant amongst the other metabolites detected in the bile extracts (Appendix 3.4). This maybe a result of the rapid chain shortening of the alkyl chain of the longer parent SPCs (C11 and C13) (Lara-Martín et al., 2007).

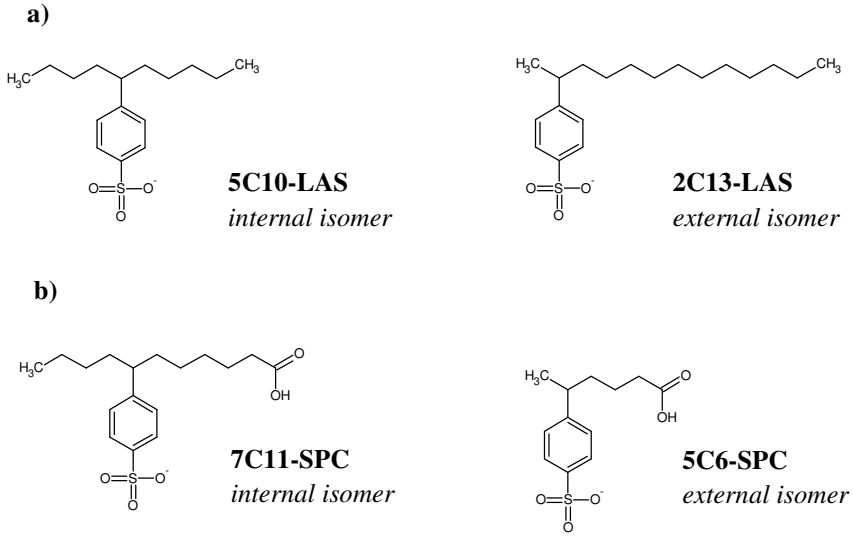


Figure 3.6: Chemical structure of: a) linear alkylbenzene sulfonic acid (LAS) and b) sulfophenyl carboxylic acids (SPCs) isomers. Adapted from Lara-Martin *et al.* (2007)

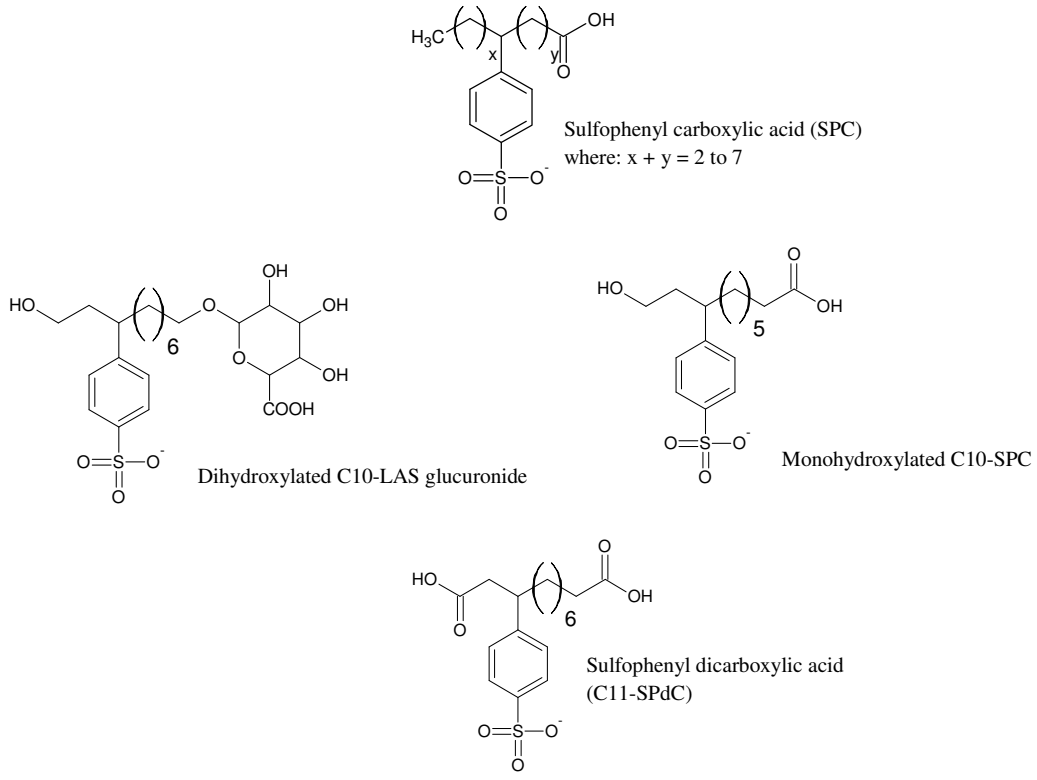


Figure 3.7: Chemical structures of linear alkylbenzene sulfonic acid metabolites identified in trout bile.

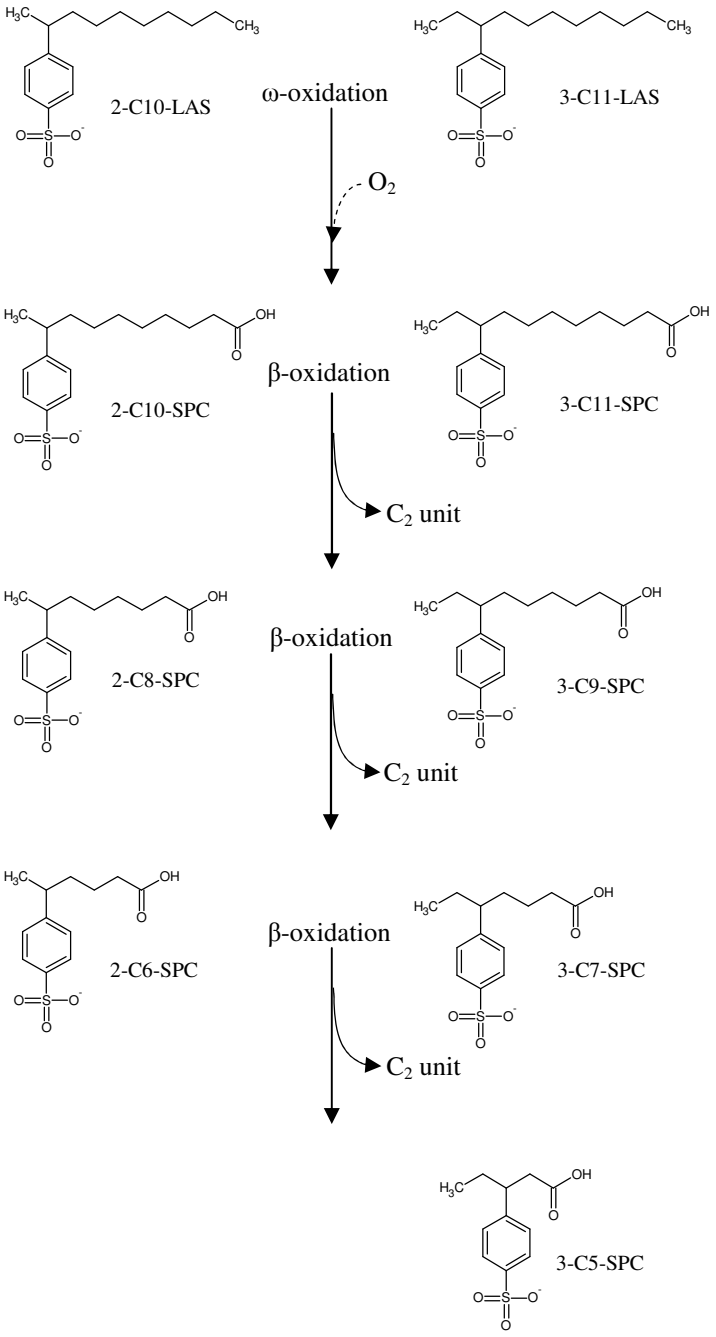
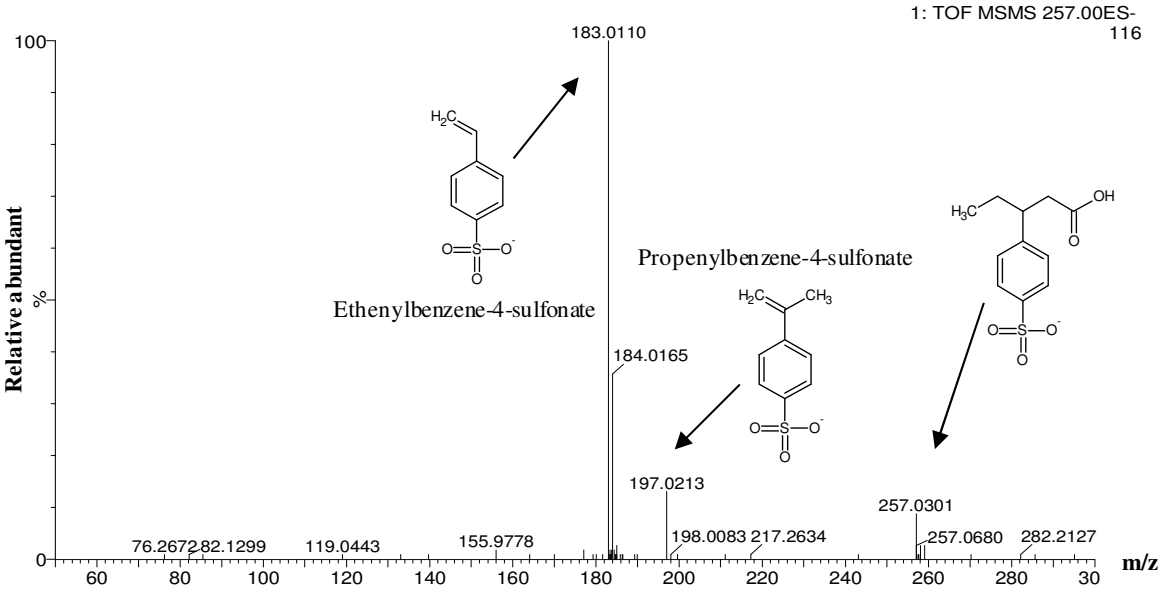


Figure 3.8: Proposed pathway for degradation of linear alkylbenzene sulfonic acid (LAS) in either wastewater effluent or in trout bile.

a) **C5-SPC**



b) **C6-SPC**

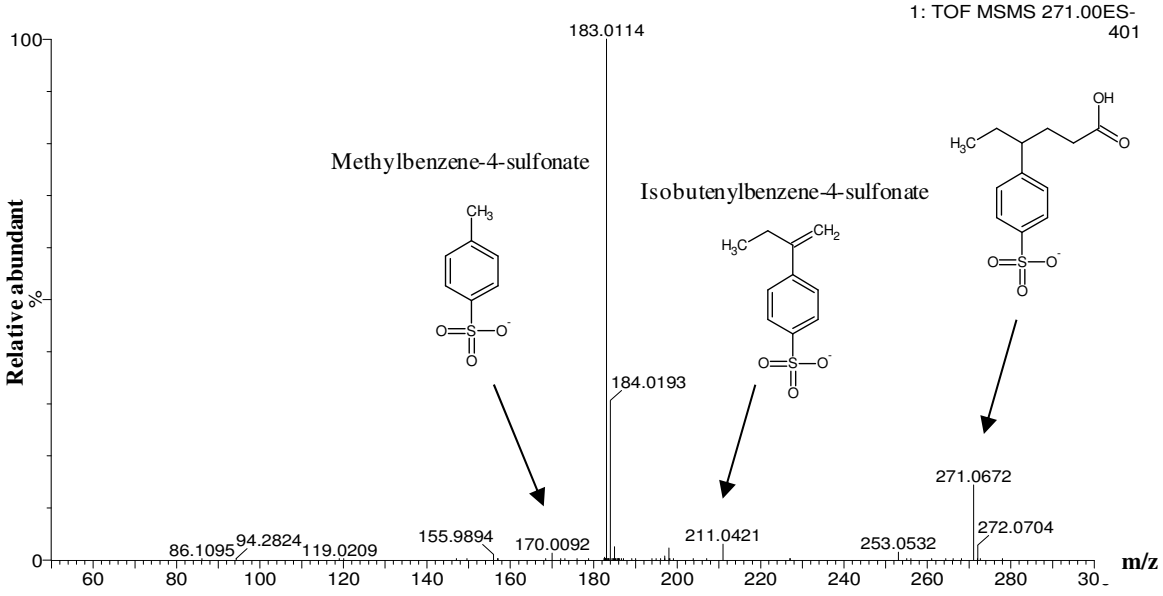


Figure 3.9: Q-TOFMS spectra of two SPCs homologous (C5-SPC and C6-SPC) and tentative structures of their fragment ions using Q-TOF as analyzer and a collision energy of 30eV.

Table 3.2: Linear alkylbenzene sulfonic acids (LAS) and their metabolites identified in trout bile in -ESI mode.

Marker No.	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Ion form	Putative identity
1	257.0482	4.54	$C_{11}H_{13}O_5S$	257.0484	-0.8	1.3	197.0270, 183.011	$C_{11}H_{14}O_5S$	$[M-H]^-$	C5-SPC
2	505.1747	4.79	$C_{22}H_{33}O_{11}S$	505.1744	0.6	0.5	329.1425, 193.0345, 183.0112, 175.0237	$C_{16}H_{26}O_5S$	$[M-H]^-$	Dihydroxylated C10-LAS glucuronide
3	271.0640	5.16	$C_{12}H_{15}O_5S$	271.0640	0.0	1.0	253.0532, 211.0429, 183.0116, 170.0033	$C_{12}H_{16}O_5S$	$[M-H]^-$	C6-SPC
4	343.1217	5.32	$C_{16}H_{23}O_6S$	343.1215	0.6	0.9	325.1122, 197.0273, 183.0120, 170.0043	$C_{16}H_{24}O_6S$	$[M-H]^-$	Putative monohydroxylated C10-SPC
5	371.1165	5.84	$C_{17}H_{23}O_7S$	371.1165	0.0	0.4	353.1052, 183.0117, 170.0036	$C_{17}H_{24}O_7S$	$[M-H]^-$	Putative C11-SPdC
6	285.0800	5.91	$C_{13}H_{17}O_5S$	285.0797	1.1	0.1	225.0583, 183.0118, 170.0036	$C_{13}H_{18}O_5S$	$[M-H]^-$	C7-SPC
7	299.0954	6.66	$C_{14}H_{19}O_5S$	299.0953	0.3	1.1	183.0117, 170.0033	$C_{14}H_{20}O_5S$	$[M-H]^-$	C8-SPC
8	313.1111	6.72	$C_{15}H_{21}O_5S$	313.1110	0.3	1.1	295.1011, 183.0117, 197.0277, 170.0037	$C_{15}H_{22}O_5S$	$[M-H]^-$	C9-SPC
9	327.1267	7.32	$C_{16}H_{23}O_5S$	327.1266	0.3	0.1	183.0111	$C_{16}H_{24}O_5S$	$[M-H]^-$	C10-SPC

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone compound; $[M-H]^-$: deprotonated ion; LAS: linear chain alkylbenzene sulfonate, SPC: sulfophenyl carboxylate; SPdC: sulfophenyl di-carboxylate.

3.3.4.2 Nonylphenol ethoxylates (NPEOs) non-ionic surfactants and their metabolites.

A series of nonylphenol ethoxylates non-ionic surfactants and their metabolites were detected in effluent-exposed trout bile. Amongst non-ionic surfactants, alcohol polyethoxylates (AEOs) and alkyl phenol polyethoxylates (APEOs) are the two major products used in Europe. APEOs are widely used in a variety of applications, for instance as detergents and emulsifiers in household and industry (Ying et al., 2002). APEOs are generally composed of an hydrophobic (mono-alkylphenol) and hydrophilic part (polyethylene glycol chain); the polyethylene glycol chain has a range of 2-25 ethoxylate units (Figure 3.10) (Lee, 1999).

Metabolite structures detected by Q-TOFMS analysis were mainly glucuronide conjugates which were detected in both +/–ESI modes (Figure 3.11 and Table 3.3). The only two exceptions were the glucuronide conjugate of nonylphenol (NP) (marker 1), which gave response only in –ESI mode providing the deprotonated molecule at *m/z* 395.2072, and the NP6EO (marker 7), which was detected only in +ESI mode as both Na and NH_4-adduct at *m/z* 683.3620 and *m/z* 678.4075, respectively. The order of elution in RP-HPLC (C_{18} stationary phase) depended on the polyethoxy chain length (see Table 3.3 and the ion chromatogram of the NPEOs series detected in the bile shown in Appendix 3.5).

Product ion scan experiments (Q-TOFMS) were performed for the whole range of NPEO homologues found in the bile samples after applying collision energies of 20-50 eV. In -ESI mode, a number of different fragments were detected. The molecular ion of the aglycone molecule [M-Glu-H] for each member of the NPEO series was detected, as well as a fragment due to loss of H_2O. The glucuronide moiety gave rise to typical fragments at *m/z* 175.0243, at *m/z* 157.0137 after loss of H_2O and at *m/z* 113.0239 after loss of CO_2 (see previous section 3.3.4). In –ESI mode, Q-TOFMS mass spectra of each ethoxylate homologue showed the characteristic signal at *m/z* 219.1749, which corresponds to the nonylphenoxy ion ($C_{15}H_{23}O$), therefore all their molecular structures could be related to the NPEOs series.

In +ESI mode, the Na-adduct of the NPEO glucuronides, gave only the aglycone fragment (as the Na-adduct) during Q-TOFMS analyses corresponding to loss of 176 Da which was the neutral loss of the glucuronide moiety. NH_4-adducts of glucuronide conjugated NPEOs were also detected, and likewise their fragments showed

predominantly the loss of the glucuronide moiety plus NH_3 (loss of 193 Da). However, unlike the Na-adduct, fragmentation of the NH_4-adduct gave rise to a number of other ions (Table 3.3). The different fragmentation behaviour of Na and NH_4-adducts is due to the strength of the interaction between the molecule and the cation.

In fact, NH_4-adducts are usually more unstable than Na-adducts due to the higher ionic radii (Na: 0.98Å, NH_4: 1.43Å) which destabilizes the interaction between the analyte and the adduct. Furthermore, the alkali metals have a high charge density and therefore they have a stronger interaction with the adduction site resulting in a more stable compound (Cowan et al., 2008). Metallic cations can behave as Lewis acid (empty orbitals in the external shell) and therefore allow unshared pair-empty orbital interactions with molecule having available lone pairs of electrons. On the contrary, the charge in the NH_4^+ ion is delocalised on all its atoms and this allows only the formation of weaker hydrogen-bonds (Cowan et al., 2008). For this reason, NH_4-adducts were chosen for fragmentation studies in order to obtain further structural information for the NPEOs series, since Na-adducts do not give relevant fragmentation in Q-TOFMS mode due to their higher stability.

The aglycone fragment of the NH_4-adducts of the NPEO glucuronides gave rise to a number of further fragments including neutral loss of nonene (loss of 126 Da). Table 3.4 summarises the fragments of NH_4-adducts of the nonylphenol ethoxylates in +ESI mode. Fragment ions at *m/z* 209.1178 ($C_{12}H_{17}O_3$), *m/z* 165.0916 ($C_{10}H_{13}O_2$), and *m/z* 121.0653 (C_8H_9O) were detected after sequential loss of ethoxylated alcohols plus H_2O. Further fragments characteristic of ethoxylated compounds were observed at *m/z* 177.1127 ($C_8H_{17}O_4$) and *m/z* 133.0865 ($C_6H_{13}O_3$): these two ions resulted from the simultaneous loss of nonylphenol and ethoxylated alcohol at both sites of the ethoxylate chain (Plomley et al., 1999).

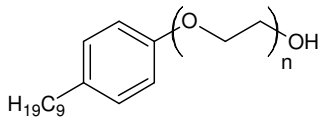
Additional information of the types of fragments obtained from $[M+Na]^+$, $[M+NH_4]^+$ and $[M-H]^-$ species are given in Figure 3.12. Here, an example is given for the fragmentation pattern of the positively identified marker NP4EO glucuronide (marker 5, Table 3.3) in both +/-ESI modes. Fragmentation was performed on three different parent ions: *m/z* 571.3118 $[M-H]^-$, *m/z* 595.3088 $[M+Na]^+$, and *m/z* 590.3544 $[M+NH_4]^+$, respectively.

Further ions at *m/z* 791.4218 ($C_{42}H_{63}O_{14}$) and *m/z* 879.4742 ($C_{46}H_{71}O_{16}$) were also detected in –ESI as potential markers of NPEOs (data not shown). However, their Q-TOFMS fragmentation confirmed their identity as clusters $[M+M-H]^-$ for NP and NP1EO, respectively. The formation of clusters in the spectrum could also be explained by the very high ion abundance of the molecular ions for these metabolites: high ion abundances correspond to high amounts of the ion in the source and therefore this facilitates agglomeration processes. Clusters for the other ethoxylated homologues were not detected either due to the mass cut-off applied on the full scan mode (50-1000 *m/z*) or due to their concentration possibly below the LOD value (minimum required intensity 100 counts as one of the data preprocessing parameters). The structural information obtained from the Q-TOF analysis were sufficient to fully characterized all the compounds belonging to this specific class of contaminants. However, the GC-MS analysis of the whole bile samples for other purposes gave additional information also regarding this class of compounds worth to be mentioned.

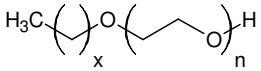
In order to further confirm the identity of the detected compounds, LC fractions of bile samples were subjected to enzymatic hydrolysis and then analysed by GC-MS. However, in some cases the response for the peaks of interest in the LC fractions was too low to allow proper structural characterization. Therefore, whole composite bile samples for both control and effluent exposed fish (10 days exposure) were also subjected to enzymatic hydrolysis in order to overcome the limitations due to poor signals. Analysis of the derivatized samples revealed the characteristic ions and RT of the trimethylsilyl ether (TMS) derivatives of NP, NP1EO and NP2EO (Appendix 3.6). The electron impact mass spectra EI-MS fragmentation pattern of NP, NP1EO and NP2EO presented abundant ions at $[M-29]^+$, $[M-71]^+$, and $[M-99]^+$ which correspond to loss of alkyl radicals from the nonyl chain ($C_2H_5^\bullet$, $C_5H_{11}^\bullet$, and $C_7H_{15}^\bullet$, receptively). The molecular ion was only observed for NP1EO at *m/z* 336 whereas the trimethysilyl group at *m/z* 73 was obtained for all three compounds. The trimethylsiloxy tropylium ion at *m/z* 179 was only detected for NP and NP1EO homologues. Formation of tropylium ion indicates that the hydroxyl group is attached directly to the ring (Wheeler et al., 1997).

As shown in Appendix 3.7 the glucuronide conjugates of NP and the ethoxylate homologue NP1EO appeared to be the most abundant compounds in the bile extracts amongst all the detected homologues.

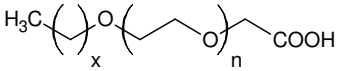
In summary a number of NPEO homologues were detected as marker of effluent exposure in fish bile. These were all glucuronide conjugates and ranged from NP itself, to NP6EO.



Nonylphenol polyethoxylate (NP*n*EO)

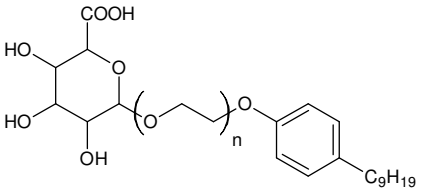


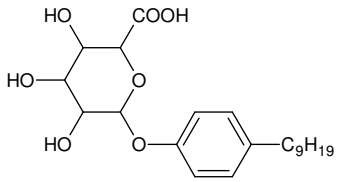
Alcohol polyethoxylate (A*x*EO*n*)



Alkyl polyethoxy carboxylate (A*x*EC*n*)

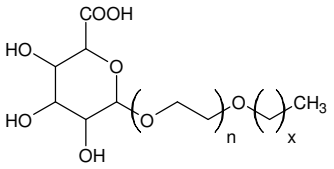
Figure 3.10: The structures of NP*n*EO, A*x*EO*n* and A*x*EC*n* products; *x* represents the number of methyl groups in the alkyl chain; *n* denotes number of ethoxy units.





Glucuronide conjugate of nonylphenol

Glucuronide conjugate of nonylphenol polyethoxylate (NP*n*EO+Glu)



Glucuronide conjugate of alcohol polyethoxylate (A*x*EO*n*+Glu)

Figure 3.11: Glucuronide conjugates of NPEO and AEO series detected in bile from effluent-exposed fish; Glu denotes glucuronide moiety; *x* represents the number of methyl groups in the alkyl chain; *n* denotes number of ethoxylate units.

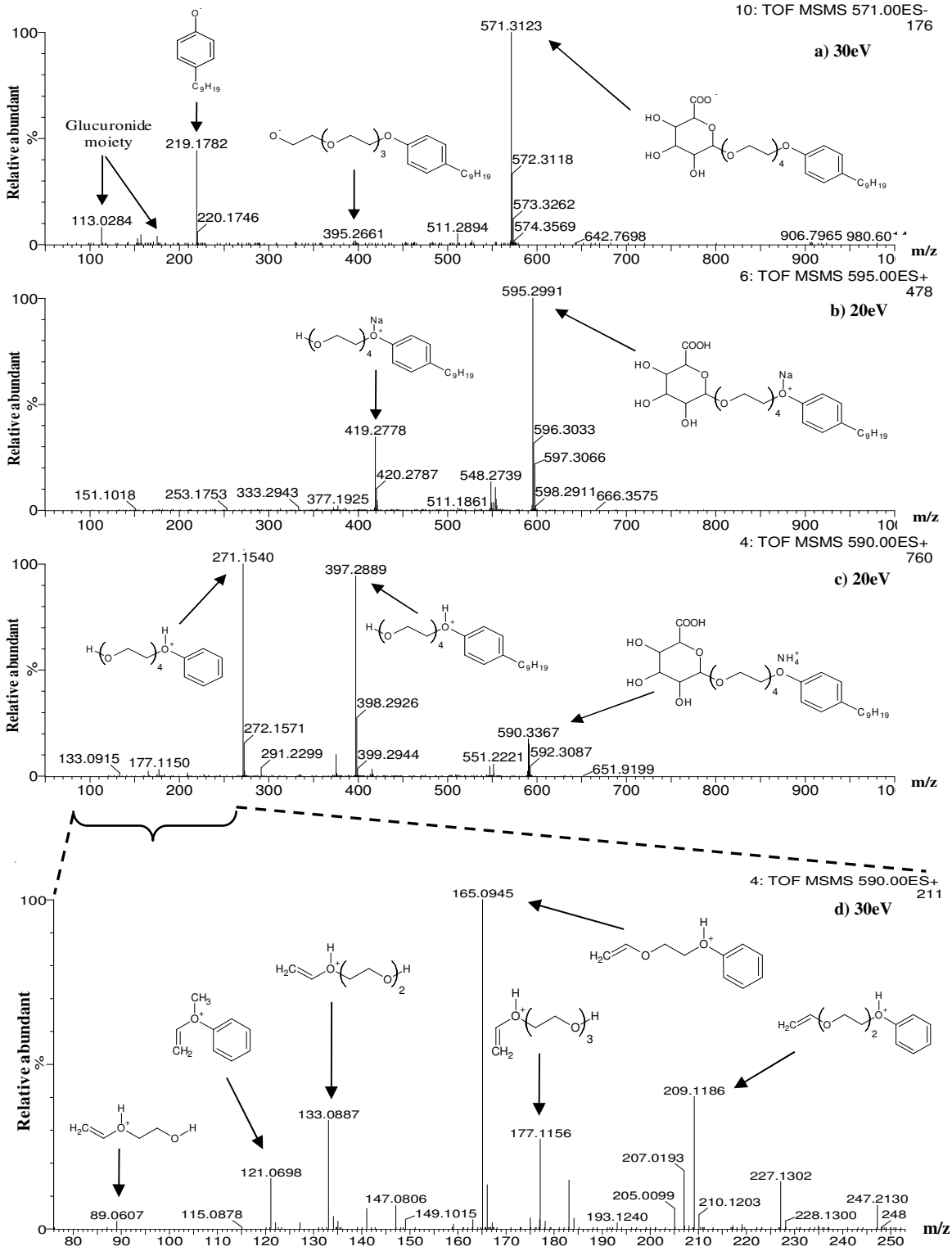


Figure 3.12: Q-TOFMS spectra of nonylphenol-4EO glucuronide and structures of relative fragment ions: a) -ESI mode, b) +ESI mode (Na adduct), c) +ESI mode (NH_4 adduct) and d) +ESI mode magnitude *m/z* range 70-270 (NH_4 adduct).

Table 3.3: Nonylphenol ethoxylates (NPEOs) and their metabolites identified in trout bile in both ESI modes (+/-ESI).

Marker No.	ESI mode	Observed ion (m/z)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments (plus fragments from GCMS analysis of the derivatized aglycone)[#]	Parent compound formula[§]	Ion form	Putative identity
1	-ESI	395.2072	16.42	$C_{21}H_{31}O_7$	395.2070	0.5	7.8	377.1972, 219.1748, 175.0245, 113.0243 (263,221,193,179,73)	$C_{15}H_{24}O$	$[M-H]^-$	Nonylphenol glucuronide
2	+ESI	463.2311	18.72	$C_{23}H_{36}O_8Na$	463.2308	0.6	0.1	287.1986	$C_{17}H_{28}O_2$	$[M+Na]^+$	Nonylphenol-1EO glucuronide
	+ESI	458.2760		$C_{23}H_{40}NO_8$	458.2754	1.3	1.9	265.2171, 139.0761		$[M+NH_4]^+$	
	-ESI	439.2331		$C_{23}H_{35}O_8$	439.2332	-0.2	1.0	421.2224, 263.2012, 219.1758, 175.024 (307,265,251,221,193,179,73)		$[M-H]^-$	
3	+ESI	507.2574	19.42	$C_{25}H_{40}O_9Na$	507.2570	0.8	1.0	331.2241	$C_{19}H_{32}O_3$	$[M+Na]^+$	Nonylphenol-2EO glucuronide
	+ESI	502.3018		$C_{25}H_{44}NO_9$	502.3016	0.4	0.9	309.2426, 183.1026, 165.0915, 121.0655		$[M+NH_4]^+$	
	-ESI	483.2597		$C_{25}H_{39}O_9$	483.2594	0.6	0.1	465.2504, 307.2273, 219.1750, 175.0239 (351,309,281,221,193,73)		$[M-H]^-$	
4	+ESI	551.2836	20.23	$C_{27}H_{44}O_{10}Na$	551.2832	0.7	0.5	375.2498	$C_{21}H_{36}O_4$	$[M+Na]^+$	Nonylphenol-3EO glucuronide
	+ESI	546.3277		$C_{27}H_{48}NO_{10}$	546.3278	-0.2	2.8	353.2689, 227.1281, 165.0923, 133.0871, 121.0647		$[M+NH_4]^+$	
	-ESI	527.2860		$C_{27}H_{43}O_{10}$	527.2856	0.8	0.0	509.2747, 351.2552, 219.1754, 175.0240		$[M-H]^-$	
5	+ESI	595.3088	20.56	$C_{29}H_{48}O_{11}Na$	595.3094	-1.0	0.9	419.2777	$C_{23}H_{40}O_5$	$[M+Na]^+$	Nonylphenol-4EO glucuronide
	+ESI	590.3544		$C_{29}H_{52}NO_{11}$	590.3540	0.7	2.3	397.2951, 271.1538, 209.1180, 177.1119, 165.0924, 133.0863, 121.0655, 89.0607		$[M+NH_4]^+$	
	-ESI	571.3118		$C_{29}H_{47}O_{11}$	571.3118	0.0	0.1	553.2998, 395.2781, 219.1749, 175.0241, 157.0133, 113.0242		$[M-H]^-$	
6	+ESI	639.3357	20.76	$C_{31}H_{52}O_{12}Na$	639.3356	0.2	1.5	463.3029	$C_{25}H_{44}O_6$	$[M+Na]^+$	Nonylphenol-5EO glucuronide
	+ESI	634.3804		$C_{31}H_{56}NO_{12}$	634.3803	0.2	0.6	441.3211, 315.1798, 209.1175, 177.1124, 165.0916, 133.0869, 121.0651		$[M+NH_4]^+$	
	-ESI	615.3391		$C_{31}H_{51}O_{12}$	615.3381	1.6	0.4	439.3059, 219.1739, 175.0235		$[M-H]^-$	
7	+ESI	683.3620	20.92	$C_{33}H_{56}O_{13}Na$	683.3619	0.1	0.9	507.3287	$C_{27}H_{48}O_7$	$[M+Na]^+$	Nonylphenol-6EO glucuronide
	+ESI	678.4075		$C_{33}H_{60}NO_{13}$	678.4065	1.5	2.1	485.3484, 359.2072, 209.1177, 177.1134, 165.0920, 133.0860		$[M+NH_4]^+$	

+/-ESI: positive/negative electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm and additional fragments observed from GC-MS of the aglycone and after derivatization to TMS ether; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; nEO where n: represents number of ethoxy units.

Table 3.4: Q-TOFMS fragments of nonylphenol ethoxylates (ammonium adducts) in +ESI mode.

		Q-TOFMS fragments (*m/z*)							
Metabolite identity	**Observed ion (*m/z*)**	**Aglycone**	**Loss of nonene(C_9H_{18})**	**Ethoxylate residuals**					
NP1EO glucuronide	458.2760	265.2171	139.0761						
NP2EO glucuronide	502.3018	309.2426	183.1026		165.0915	121.0655			
NP3EO glucuronide	546.3277	353.2689	227.1281		165.0923	121.0647		133.0871	
NP4EO glucuronide	590.3544	397.2951	271.1538	209.1180	165.0924	121.0655	177.1119	133.0863	89.0607
NP5EO glucuronide	634.3804	441.3211	315.1798	209.1175	165.0916	121.0651	177.1124	133.0869	
NP6EO glucuronide	678.4075	485.3484	359.2072	209.1177	165.0920		177.1134	133.0860	

m/z: mass to charge ratio; NP: nonylphenol; nEO where n: represents number of ethoxy units.

3.3.4.3 Alcohol ethoxylates (AEOs) non-ionic surfactants and their metabolites

A series of alcohol ethoxylate (AEO) non-ionic surfactants and their metabolites were detected in effluent-exposed trout bile (Figure 3.10). Commercial AEOs are a mixture of homologues with different alkyl chain length (12-18 carbon atoms) and the alkyl chain is connected to an ethylene oxide chain by means of an ether bond (Figure 3.11). As the above mentioned NPEOs, they are mainly used as industrial or household detergents (HERA, 2009b).

AEOs (A*x*EO*n*; where *x*=12-15; *n*=0-8) were detected as glucuronide conjugates in effluent-exposed fish bile in both +/-ESI modes (Table 3.5). These metabolites were detected as deprotonated ions $[M+Glu-H]^-$ in -ESI mode, whereas Na and NH_4 adducts $[M+Glu+Na]^+$/$[M+Glu+NH_4]^+$ were predominant in +ESI mode. Some of these different homologues coeluted on the C_{18} reversed-phase column, probably due to similarities in their chromatographic behaviour. A mixture of branched and linear alkyl chain isomers with same molecular formula was detected for each AEO homologue (see Appendix 3.8 for the extracted ion chromatograms). According to previous studies (Cohen et al., 2001), the major peak in the chromatogram can be assigned to the linear form whilst the broader and smaller peaks correspond to the branched forms.

In order to characterize the fragmentation pattern of ethoxylate compounds, a technical mixture of C13(EO)12 was selected as a standard of known ethoxylates products. The C13(EO)8 homologue was selected as parent ion and analysed by Q-TOFMS and fragmentation experiments gave rise to informative ions which led to the confirmation of the identity for the ethoxylated markers. The Q-TOFMS analysis of the Na-adduct at *m/z* 575.4135 did not give any fragments, even using very high collision energy (50 eV) due to its high stability as an ion (data not shown). On the other hand, the spectrum of the NH_4-adduct was rich of informative fragment ions. Characteristic ions were detected at *m/z* 553.4316 (loss of NH_3), at *m/z* 371.2281 (subsequent loss of NH_3 and tridecanene $C_{13}H_{26}$), and at *m/z* 221.1389, 177.1127, 133.0865 due to simultaneous loss of alkyl group and ethoxylated alcohol at both sides of the ethoxylate chain (see Appendix 3.9).

Analysis of AEOs from the fish bile revealed fragments corresponding to $[M-H_2O-H]^-$ and $[M-CH_3COOH-H]^-$ in –ESI mode as well as fragments associated with the glucuronide moiety (*m/z* 175.0243, 157.0137, 113.0239). An additional ion at *m/z* 193.0348 ($C_6H_9O_7$) was detected, which corresponds to the glucuronic acid anion,

where the glucuronide moiety is attached to an alcohol (Wen et al., 2007). The Na-adducts of the AEO metabolites were too stable to allow substantial fragmentation in Q-TOFMS mode (Houde et al., 2002), therefore the aglycone (Na-adduct) was the only fragment detected in this series (Table 3.5).

Table 3.6 summarises the Q-TOFMS fragments of alcohol ethoxylates (NH_4-adducts) in +ESI mode. Characteristic fragment ions were obtained from the fragmentation of NH_4-adducts (ions at *m/z* ranging from 107.0708 to 371.2281, separated by 44 Da). These ions were generated after neutral loss of an alkene molecule (loss of C_nH_{2n}; where n=12-15) from the aglycone ion of each AEO homologue. Aglycone ions were also detected due to loss of the glucuronide moiety plus NH_3. Ethoxylate series at *m/z* ranging from 89.0603 to 309.1913 were also observed as consequence of the cleavage of the aglycone parent ion at both extremes (simultaneous loss of alkyl group and ethoxylated alcohol). A fragmentation pattern example for a selected AEO [C13(EO)4] is shown in Figure 3.13 in both +/-ESI modes and for Na and NH_4 adducts.

Appendix 3.10 shows the relative distribution for the series of detected AEOs (C12-C15) according to the different number of ethoxylate units. The C13 homologue was the most predominant for ethoxymers with less than 3EO units and detected as the Na adduct. In the series with 4EO units the C13 and C14 homologues were found equally predominant whilst when the number of EO units increased (> 4EO) the C14 and C15 homologues were the most abundant.

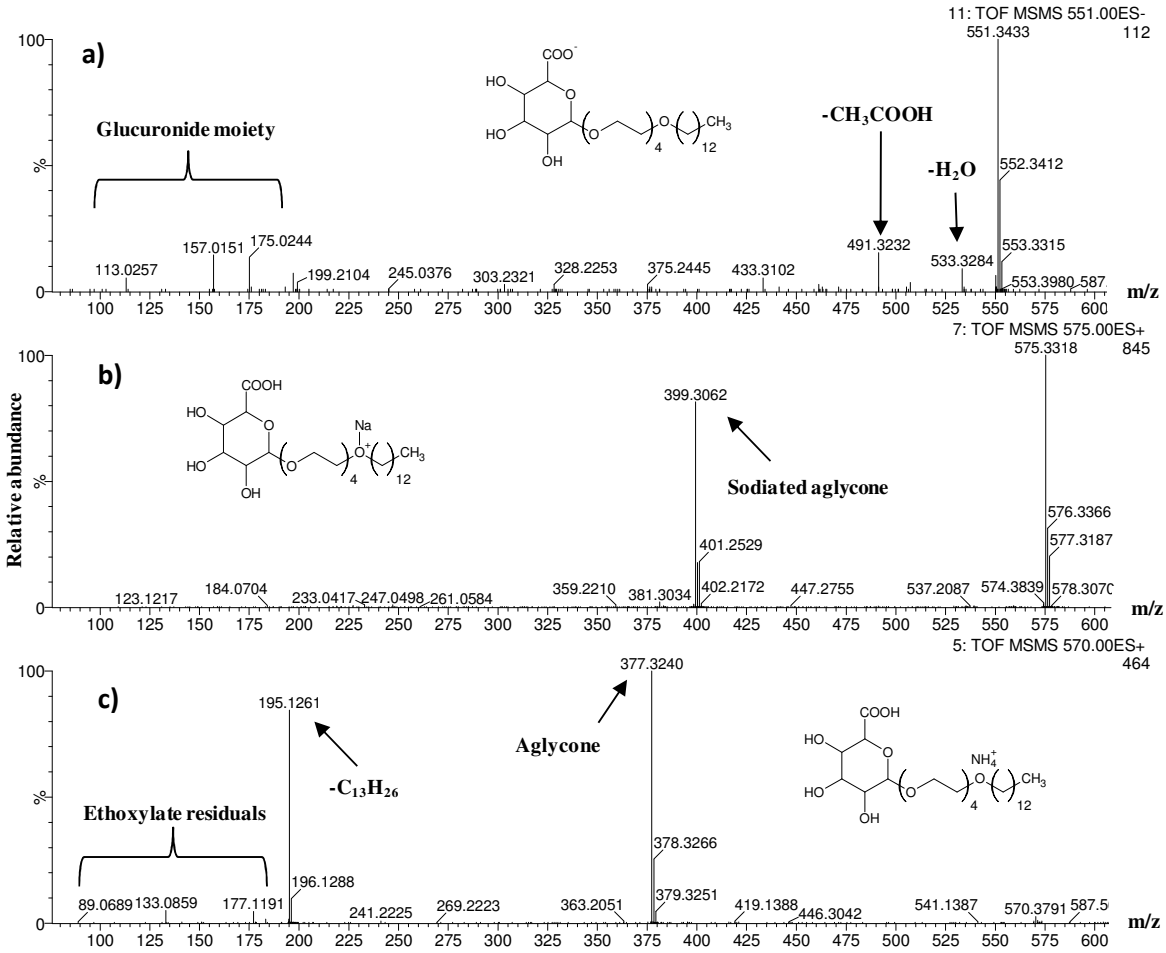


Figure 3.13: Q-TOFMS spectra of tridecanol-4EO [C13(EO)4] glucuronide metabolite using Q-TOF as analyzer and a collision energy of 20 eV: a) -ESI mode $[M-H]^-$, b) +ESI mode $[M+Na]^+$, and c) +ESI mode $[M+NH_4]^+$.

Table 3.5: Alcohol ethoxylates (AEOs) and their metabolites identified in trout bile in both electrospray ionization modes (+/-ESI).

Marker No.	ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments	Parent compound formula[§]	Ion Form	Putative identity
1	-ESI	361.2231	14.76	$C_{18}H_{33}O_7$	361.2226	1.4	0.2	343.2121, 301.2001, 175.0244	$C_{12}H_{26}O$	$[M-H]^-$	Dodecanol glucuronide
2	+ESI	429.2461	16.79	$C_{20}H_{38}O_8Na$	429.2464	-0.7	1.4	253.2142, 199.0217, 181.0112	$C_{14}H_{30}O_2$	$[M+Na]^+$	Dodecanol-1EO glucuronide
	+ESI	424.2904		$C_{20}H_{42}NO_8$	424.2910	-1.4	2.1	231.2327		$[M+NH_4]^+$	
	-ESI	405.2488		$C_{20}H_{37}O_8$	405.2488	0.0	0.4	387.2371, 175.0248		$[M-H]^-$	
3	+ESI	473.2727	17.73	$C_{22}H_{42}O_9Na$	473.2727	0.0	0.6	297.2405	$C_{16}H_{34}O_3$	$[M+Na]^+$	Dodecanol-2EO glucuronide
	+ESI	468.3175		$C_{22}H_{46}NO_9$	468.3173	0.4	1.7	275.2585, 107.0705		$[M+NH_4]^+$	
	-ESI	449.2751		$C_{22}H_{41}O_9$	449.2751	0.0	0.1	431.2652, 389.2538, 175.0246, 157.0134, 113.0240		$[M-H]^-$	
4	+ESI	517.2990	20.31	$C_{24}H_{46}O_{10}Na$	517.2989	0.2	0.7	341.2666	$C_{18}H_{38}O_4$	$[M+Na]^+$	Dodecanol-3EO glucuronide
	+ESI	512.3435		$C_{24}H_{50}NO_{10}$	512.3435	0.0	0.4	319.2852, 151.0966, 133.0861		$[M+NH_4]^+$	
	-ESI	493.3008		$C_{24}H_{45}O_{10}$	493.3013	-1.0	0.5	475.2920, 433.2794, 175.0248, 157.0143		$[M-H]^-$	
5	+ESI	561.3253	20.80	$C_{26}H_{50}O_{11}Na$	561.3251	0.4	0.4	385.2924	$C_{20}H_{42}O_5$	$[M+Na]^+$	Dodecanol-4EO glucuronide
	+ESI	556.3699		$C_{26}H_{54}NO_{11}$	556.3697	0.4	0.2	363.3112, 195.1236, 177.1123, 133.0869, 89.0599		$[M+NH_4]^+$	
	-ESI	537.3271		$C_{26}H_{49}O_{11}$	537.3275	-0.7	3.4	175.0247		$[M-H]^-$	
6	+ESI	605.3518	21.01	$C_{28}H_{54}O_{12}Na$	605.3513	0.8	0.2	429.3177	$C_{22}H_{46}O_6$	$[M+Na]^+$	Dodecanol-5EO glucuronide
	+ESI	600.3966		$C_{28}H_{58}NO_{12}$	600.3959	-1.0	0.9	407.3359, 239.1499,221.1382, 177.1128, 133.0869		$[M+NH_4]^+$	
	-ESI	581.3547		$C_{28}H_{53}O_{12}$	581.3537	1.7	1.0	175.0243		$[M-H]^-$	
7	-ESI	375.2383	16.94	$C_{19}H_{35}O_7$	375.2383	0.0	0.8	357.2283, 315.2153, 199.2068, 175.0240, 157.0130, 113.0237	$C_{13}H_{28}O$	$[M-H]^-$	Tridecanol glucuronide
8	+ESI	443.2622	18.82	$C_{21}H_{40}O_8Na$	443.2621	0.2	1.2	267.2300, 199.0224, 181.0110	$C_{15}H_{32}O_2$	$[M+Na]^+$	Tridecanol-1EO glucuronide
	+ESI	438.3064		$C_{21}H_{44}NO_8$	438.3067	-0.7	2.7	245.2476		$[M+NH_4]^+$	
	-ESI	419.2645		$C_{21}H_{39}O_8$	419.2645	0.0	0.1	401.2540, 359.2431, 175.0246, 157.0140, 113.0238		$[M-H]^-$	

+/-ESI: positive/negative electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; *n*EO where *n*: represents number of ethoxy units.

Table 3.5: (continued) Alcohol ethoxylates (AEOs) and their metabolites identified in trout bile in both electrospray ionization modes (+/-ESI).

Marker No.	ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments	Parent compound formula[§]	Ion Form	Identity of compound
9	+ESI	487.2884	19.64	$C_{23}H_{44}O_9Na$	487.2883	0.2	1.0	311.2552	$C_{17}H_{36}O_3$	$[M+Na]^+$	Tridecanol-2EO
	+ESI	482.3329		$C_{23}H_{48}NO_9$	482.3329	0.0	0.3	289.2740, 107.0712		$[M+NH_4]^+$	glucuronide
	-ESI	463.2908		$C_{23}H_{43}O_9$	463.2907	0.2	0.2	445.2795, 403.2696, 175.0246, 157.0137, 113.0239		$[M-H]^-$	
10	+ESI	531.3151	20.35	$C_{25}H_{48}O_{10}Na$	531.3145	1.1	0.5	355.2821	$C_{19}H_{40}O_4$	$[M+Na]^+$	Tridecanol-3EO
	+ESI	526.3591		$C_{25}H_{52}NO_{10}$	526.3591	0.0	0.6	333.3005, 151.0977, 133.0861		$[M+NH_4]^+$	glucuronide
	-ESI	507.3171		$C_{25}H_{47}O_{10}$	507.3169	0.4	0.2	489.3064, 447.2956, 175.0244, 157.0135, 113.0242		$[M-H]^-$	
11	+ESI	575.341	21.82	$C_{27}H_{52}O_{11}Na$	575.3407	0.5	2.0	399.3088	$C_{21}H_{44}O_5$	$[M+Na]^+$	Tridecanol-4EO
	+ESI	570.3856		$C_{27}H_{56}NO_{11}$	570.3853	0.5	0.9	377.3265, 195.1236, 177.1129, 133.0858, 89.0608		$[M+NH_4]^+$	glucuronide
	-ESI	551.3430		$C_{27}H_{51}O_{11}$	551.3431	-0.2	1.2	533.3345, 491.3204, 175.0241, 113.0240		$[M-H]^-$	
12	+ESI	619.3672	21.97	$C_{29}H_{56}O_{12}Na$	619.3669	0.5	0.9	443.3336	$C_{23}H_{48}O_6$	$[M+Na]^+$	Tridecanol-5EO
	+ESI	614.4119		$C_{29}H_{60}NO_{12}$	614.4116	0.5	2.0	421.324, 239.1497,221.1395, 177.1123, 133.0865		$[M+NH_4]^+$	glucuronide
	-ESI	595.3699		$C_{29}H_{55}O_{12}$	595.3694	0.8	0.7	577.3591, 175.0238		$[M-H]^-$	
13	+ESI	663.3934	22.10	$C_{31}H_{60}O_{13}Na$	663.3932	0.3	1.2	487.3616	$C_{25}H_{52}O_7$	$[M+Na]^+$	Tridecanol-6EO
	+ESI	658.4372		$C_{31}H_{64}NO_{13}$	658.4378	-0.9	0.4	465.3795, 283.1757, 265.1654, 221.1397, 177.1120, 133.0860		$[M+NH_4]^+$	glucuronide
	-ESI	639.3967		$C_{31}H_{59}O_{13}$	639.3956	1.7	0.2	193.0353, 175.0239		$[M-H]^-$	
14	+ESI	707.4200	22.21	$C_{33}H_{64}O_{14}Na$	707.4194	-0.6	1.9	531.3859	$C_{27}H_{56}O_8$	$[M+Na]^+$	Tridecanol-7EO
	+ESI	702.4647		$C_{33}H_{68}NO_{14}$	702.4640	1.0	2.0	509.4032, 327.2017, 309.1901, 221.1388, 177.1127, 133.0871		$[M+NH_4]^+$	glucuronide
	-ESI	683.4214		$C_{33}H_{63}O_{14}$	683.4218	-0.6	1.1	175.0250		$[M-H]^-$	
15	-ESI	389.2541	20.66	$C_{20}H_{37}O_7$	389.2539	0.5	0.4	193.0352, 175.0243, 113.0242	$C_{14}H_{30}O$	$[M-H]^-$	Tetradecanol glucuronide
16	-ESI	433.2801	21.59	$C_{22}H_{41}O_8$	433.2801	0.0	0.9	415.2682, 373.2607, 175.0240, 157.0133, 113.0242	$C_{16}H_{34}O_2$	$[M-H]^-$	Tetradecanol-1EO glucuronide

+/-ESI: positive/negative electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; *n*EO where *n*: represents number of ethoxy units.

Table 3.5: (continued) Alcohol ethoxylates (AEOs) and their metabolites identified in trout bile in both electrospray ionization modes (+/-ESI).

Marker No.	ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments	Parent compound formula[§]	Ion Form	Putative identity
17	+ESI	501.3360	22.13	$C_{24}H_{46}O_9Na$	501.3040	-0.8	0.2	325.2719, 199.0216, 181.0118	$C_{18}H_{38}O_3$	$[M+Na]^+$	Tetradecanol-2EO
	+ESI	496.3487		$C_{24}H_{50}NO_9$	496.3486	0.2	5.1	303.2901, 107.0705		$[M+NH_4]^+$	glucuronide
	-ESI	477.3068		$C_{24}H_{45}O_9$	477.3064	0.8	2.0	459.2952, 417.2859, 175.0242, 157.0138, 113.0236		$[M-H]^-$	
18	+ESI	545.3307	22.47	$C_{26}H_{50}O_{10}Na$	545.3302	0.9	0.4	369.2974	$C_{20}H_{42}O_4$	$[M+Na]^+$	Tetradecanol-3EO
	+ESI	540.3746		$C_{26}H_{54}NO_{10}$	540.3748	-0.4	0.2	347.3161, 151.0974, 133.0868		$[M+NH_4]^+$	glucuronide
	-ESI	521.3326		$C_{26}H_{49}O_{10}$	521.3326	0.0	0.1	503.3214, 461.3107, 175.0246, 113.0238		$[M-H]^-$	
19	+ESI	589.3560	22.68	$C_{28}H_{54}O_{11}Na$	589.3564	-0.7	0.7	413.3248	$C_{22}H_{46}O_5$	$[M+Na]^+$	Tetradecanol-4EO
	+ESI	584.4008		$C_{28}H_{58}NO_{11}$	584.4010	-0.3	0.5	391.3421, 195.1232, 177.1122, 133.0859, 89.0607		$[M+NH_4]^+$	glucuronide
	-ESI	565.3592		$C_{28}H_{53}O_{11}$	656.3588	0.7	0.0	547.3482, 505.3366, 175.0243, 157.0135, 113.0242		$[M-H]^-$	
20	+ESI	633.3826	22.79	$C_{30}H_{58}O_{12}Na$	633.3826	0.0	1.1	457.3497	$C_{24}H_{50}O_6$	$[M+Na]^+$	Tetradecanol-5EO
	+ESI	628.4274		$C_{30}H_{62}NO_{12}$	628.4272	0.3	0.3	435.3681, 239.1495,221.1387, 177.1127, 133.0871		$[M+NH_4]^+$	glucuronide
	-ESI	609.3851		$C_{30}H_{57}O_{12}$	609.3850	0.2	0.7	591.3727, 549.3620, 175.0239, 157.0142, 113.0242		$[M-H]^-$	
21	+ESI	677.4083	22.90	$C_{32}H_{62}O_{13}Na$	677.4088	-0.7	1.4	501.3758	$C_{26}H_{54}O_7$	$[M+Na]^+$	Tetradecanol-6EO
	+ESI	672.4534		$C_{32}H_{66}NO_{13}$	672.4534	0.0	0.6	479.3947, 283.1756, 265.1659, 221.1381, 177.1125, 133.0869		$[M+NH_4]^+$	glucuronide
	-ESI	653.4116		$C_{32}H_{61}O_{13}$	653.4112	0.6	0.0	635.4011, 175.0247, 157.0140, 113.0241		$[M-H]^-$	
22	+ESI	721.4349	22.98	$C_{34}H_{66}O_{14}Na$	721.4350	-0.1	0.4	545.4030	$C_{28}H_{58}O_8$	$[M+Na]^+$	Tetradecanol-7EO
	+ESI	716.4798		$C_{34}H_{70}NO_{14}$	716.4796	0.3	0.5	523.4230, 327.2016, 309.1928, 221.1391, 177.1123, 133.0861		$[M+NH_4]^+$	glucuronide
	-ESI	697.4371		$C_{34}H_{65}O_{14}$	697.4374	-0.4	1.7	193.0343, 175.0241, 113.0234		$[M-H]^-$	
23	+ESI	765.4606	23.01	$C_{36}H_{70}O_{15}Na$	765.4612	-0.8	0.3	589.4287	$C_{30}H_{62}O_9$	$[M+Na]^+$	Tetradecanol-8EO
	+ESI	760.5055		$C_{36}H_{74}NO_{15}$	760.5058	-0.4	1.8	567.4473, 371.2283, 221.1388, 177.1124, 133.0870		$[M+NH_4]^+$	glucuronide
24	-ESI	403.2695	21.74	$C_{21}H_{39}O_7$	403.2696	-0.2	0.3	175.0240	$C_{15}H_{32}O$	$[M-H]^-$	Pentadecanol glucuronide

+/-ESI: positive/negative electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; *n*EO where *n*: represents number of ethoxy units.

Table 3.5: (continued) Alcohol ethoxylates (AEOs) and their metabolites identified in trout bile in both electrospray ionization modes (+/-ESI).

Marker No.	ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments	Parent compound formula[§]	Ion Form	Identity of compound
25	-ESI	447.2959	22.47	$C_{23}H_{43}O_8$	447.2958	0.2	0.8	429.2845, 175.0243, 157.0143	$C_{17}H_{36}O_2$	$[M-H]^-$	Pentadecanol-1EO glucuronide
26	+ESI	515.3198	22.93	$C_{25}H_{48}O_9Na$	515.3196	0.4	2.6	339.2860	$C_{19}H_{40}O_3$	$[M+Na]^+$	Pentadecanol-2EO glucuronide
	+ESI	510.3640		$C_{25}H_{52}NO_9$	510.3642	-0.4	1.3	317.3059, 107.0711		$[M+NH_4]^+$	
	-ESI	491.3223		$C_{25}H_{47}O_9$	491.3220	0.6	1.8	175.024		$[M-H]^-$	
27	+ESI	559.3458	23.20	$C_{27}H_{52}O_{10}Na$	559.3458	0.0	4.3	383.3135	$C_{21}H_{44}O_4$	$[M+Na]^+$	Pentadecanol-3EO glucuronide
	+ESI	554.3906		$C_{27}H_{56}NO_{10}$	554.3904	0.4	0.4	361.3315, 151.0972, 133.0860		$[M+NH_4]^+$	
	-ESI	535.3484		$C_{27}H_{51}O_{10}$	535.3482	0.4	1.4	517.3395, 475.3261, 175.0250, 157.0140, 113.0241		$[M-H]^-$	
28	+ESI	603.3724	23.41	$C_{29}H_{56}O_{11}Na$	603.3720	0.7	0.3	427.3398	$C_{23}H_{48}O_5$	$[M+Na]^+$	Pentadecanol-4EO glucuronide
	+ESI	598.4166		$C_{29}H_{60}NO_{11}$	598.4166	0.0	0.6	405.3581, 195.1239, 177.1124, 133.0860, 89.0604		$[M+NH_4]^+$	
	-ESI	579.3748		$C_{29}H_{55}O_{11}$	579.3744	0.7	0.3	561.3619, 519.3519, 175.0243		$[M-H]^-$	
29	+ESI	647.3979	23.51	$C_{31}H_{60}O_{12}Na$	647.3982	-0.5	0.0	471.3653	$C_{25}H_{52}O_6$	$[M+Na]^+$	Pentadecanol-5EO glucuronide
	+ESI	642.4432		$C_{31}H_{64}NO_{12}$	642.4429	0.5	0.1	449.3824, 239.1495, 221.1397, 177.1127, 133.0868		$[M+NH_4]^+$	
	-ESI	623.4008		$C_{31}H_{59}O_{12}$	623.4007	0.2	3.1	605.3892, 563.3814, 175.0249		$[M-H]^-$	
30	+ESI	691.4239	23.60	$C_{33}H_{64}O_{13}Na$	691.4245	-0.9	1.6	515.3939	$C_{27}H_{56}O_7$	$[M+Na]^+$	Pentadecanol-6EO glucuronide
	+ESI	686.4694		$C_{33}H_{68}NO_{13}$	686.4691	0.4	0.8	493.4114, 283.1760, 265.1647, 221.1390, 177.1122, 133.0859		$[M+NH_4]^+$	
	-ESI	667.4275		$C_{33}H_{63}O_{13}$	667.4269	0.9	0.1	649.4175, 175.0231		$[M-H]^-$	
31	+ESI	735.4518	23.65	$C_{35}H_{68}O_{14}Na$	735.4507	1.5	1.3	559.4166	$C_{29}H_{60}O_8$	$[M+Na]^+$	Pentadecanol-7EO glucuronide
	+ESI	730.4956		$C_{35}H_{72}NO_{14}$	730.4953	0.4	0.6	537.4373, 327.2010, 309.1915, 221.1397, 177.1128, 133.0868		$[M+NH_4]^+$	
	-ESI	711.4532		$C_{35}H_{67}O_{14}$	711.4531	0.1	0.2	693.4417, 175.0239		$[M-H]^-$	
32	+ESI	779.4720	23.70	$C_{37}H_{72}O_{15}Na$	779.4769	0.4	1.5	603.4444	$C_{31}H_{64}O_9$	$[M+Na]^+$	Pentadecanol-8EO glucuronide
	+ESI	774.5209		$C_{37}H_{76}NO_{15}$	774.5215	-0.8	2.2	581.4631, 371.2267, 221.1386, 177.1129, 133.0871		$[M+NH_4]^+$	
	-ESI	755.4787		$C_{37}H_{71}O_{15}$	755.4793	-0.8	3.7	175.0243		$[M-H]^-$	

+/-ESI: positive/negative electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; *n*EO where *n*: represents number of ethoxy units.

Table 3.6: Q-TOFMS fragment ions for alcohol ethoxylates (ammonium adducts) in +ESI mode.

		Q-TOFMS fragments (*m/z*) of NH_4 adduct							
Metabolite identity	**Observed ion (*m/z*)**	**Aglycone**	**Loss of Alkene(C_nH_{2n})**	**Ethoxylate residuals**					
C12(EO)1 glucuronide	424.2904	231.2327							
C12(EO)2 glucuronide	468.3175	275.2585	107.0705						
C12(EO)3 glucuronide	512.3435	319.2852	151.0966					133.0861	
C12(EO)4 glucuronide	556.3699	363.3112	195.1236				177.1123	133.0869	89.0599
C12(EO)5 glucuronide	600.3966	407.3359	239.1499			221.1382	177.1128	133.0869	
C13(EO)1 glucuronide	438.3064	245.2476							
C13(EO)2 glucuronide	482.3329	289.2740	107.0712						
C13(EO)3 glucuronide	526.3591	333.3005	151.0977					133.0861	
C13(EO)4 glucuronide	570.3856	377.3265	195.1236				177.1129	133.0858	89.0608
C13(EO)5 glucuronide	614.4119	421.3224	239.1497			221.1395	177.1123	133.0865	
C13(EO)6 glucuronide	658.4372	465.3795	283.1757		265.1654	221.1397	177.1120	133.0860	
C13(EO)7 glucuronide	702.4647	509.4032	327.2017	309.1901		221.1388	177.1127	133.0871	
C14(EO)2 glucuronide	496.3487	303.2901	107.0705						
C14(EO)3 glucuronide	540.3746	347.3161	151.0974					133.0868	
C14(EO)4 glucuronide	584.4008	391.3421	195.1232				177.1122	133.0859	89.0607
C14(EO)5 glucuronide	628.4274	435.3681	239.1495			221.1387	177.1127	133.0871	
C14(EO)6 glucuronide	672.4534	479.3947	283.1756		265.1659	221.1381	177.1125	133.0869	
C14(EO)7 glucuronide	716.4798	523.4230	327.2016	309.1928		221.1391	177.1123	133.0861	
C14(EO)8 glucuronide	760.5055	567.4473	371.2283			221.1388	177.1124	133.0870	
C15(EO)2 glucuronide	510.3640	317.3059	107.0711						
C15(EO)3 glucuronide	554.3906	361.3315	151.0972					133.0860	
C15(EO)4 glucuronide	598.4166	405.3581	195.1239				177.1124	133.0860	89.0604
C15(EO)5 glucuronide	642.4432	449.3824	239.1495			221.1397	177.1127	133.0868	
C15(EO)6 glucuronide	686.4694	493.4114	283.1760		265.1647	221.1390	177.1122	133.0859	
C15(EO)7 glucuronide	730.4956	537.4373	327.2010	309.1915		221.1397	177.1128	133.0868	
C15(EO)8 glucuronide	774.5209	581.4631	371.2267			221.1386	177.1129	133.0871	

m/z: mass to charge ratio; A*x*EO*n*: *x* represents the number of methyl groups in the alkyl chain; *n* denotes number of ethoxy units.

3.3.4.4 Alkyl polyethoxy carboxylates (AECs) anionic surfactants and their metabolites.

A series of alkyl polyethoxy carboxylates (AEC) anionic surfactants and their metabolites were detected in effluent-exposed trout bile. AECs structure is similar to that of AEOs, except for the terminal carboxylic group instead of the alcoholic group (Figure 3.10). AECs could be formed from ω-oxidation of the ethoxylate group belonging to the corresponding AEO compounds during wastewater treatment (Di Corcia et al., 1998). AECs are also used as anionic surfactants in the textile industry and in household and personal care products and thus could contribute as an additional source to the effluent contamination.

Unlike NPEOs and AEOs, A*x*EC*n* series (where; A=alkyl group and *n*= number of ethoxy units plus a terminal CH_2COOH moiety) were detected as non-conjugated compounds in both +/-ESI modes (Table 3.7). As in the case of the AEOs, no efficient chromatographic separation was observed for most of the analytes and due to coelution, and the MS spectra for each series were overlapping.

The deprotonated ions of AECs revealed fragments at *m/z* 211.2062 and *m/z* 225.2218 in –ESI mode, which corresponded to molecular formula $C_{14}H_{27}O$ and $C_{15}H_{29}O$ respectively. This can be possibly explained by the loss of the ethoxycarboxylate moiety from the molecular ion.

An example of the fragmentation pattern for Na and NH_4 adducts of two selected AEC metabolites is shown in Figure 3.14, and the list of Q-TOFMS fragments of NH_4 and Na adducts are given in Table 3.8. The fragmentation of AE*n*Cs Na-adducts (A=C13-C15, and *n*=4-10) in +ESI highlighted loss of 58 Da which corresponds to the loss of the acetate group and/or central cleavage of the parent ion, giving rise to the ion $[M\text{-}C_nH_{2n}+Na]^+$, where n=13-15. This process leads to the loss of 182, 196 or 210 Da, corresponding to the loss of tridecanene ($C_{13}H_{26}$), tetradecanene ($C_{14}H_{28}$) or pentadecanene ($C_{15}H_{30}$), respectively. Characteristic ions were also detected after fragmentation of the NH_4-adducts, showing loss of alkene plus NH_3 (ethoxylate residual), subsequently followed by loss of H_2O and CO (Ding et al., 1994). Furthermore, a series of fragments corresponding to $[M+H\text{-}2EO\text{-}H_2O]^+$ were obtained from the aglycone.

The relative concentration of the ethoxycarboxylated series (AECs) in bile extracts of effluent-exposed fish was predominated by C14ECs and C15ECs (see Appendix 3.11). Like those ethoxylates previously mentioned (NPEOs and AEOs), AECs demonstrated higher signals for Na-adducts than NH_4-adducts. For AECs the homologues with 6EC and 7EC units had the highest signal amongst their series (e.g. C14(EC)6 and C14(EC)7) (Figure 3.15). However, the lower signal observed for the homologues with fewer ethoxylate units could be caused by ion suppression due to coelution with the more abundant homologues.

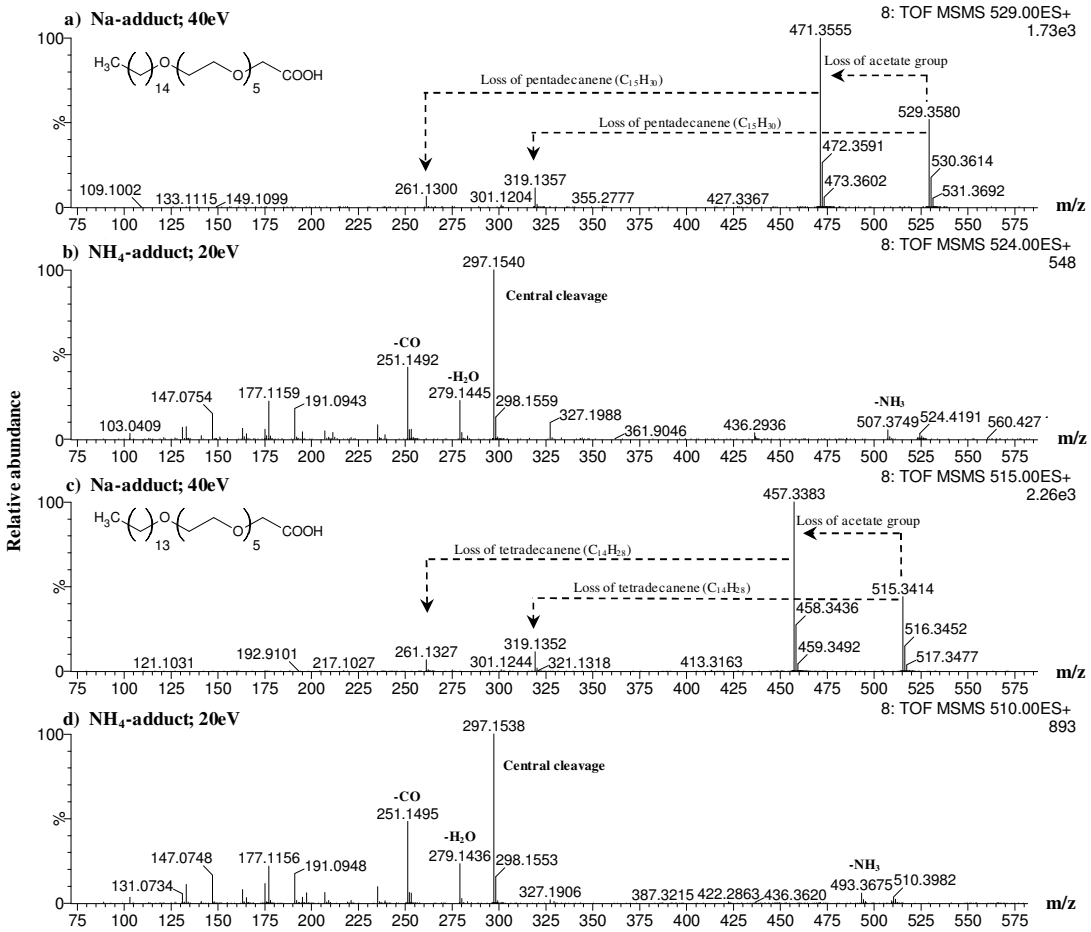


Figure 3.14: Q-TOFMS spectra of two AECs metabolites tetradecanol-6EC [C14(EC)6] (a & b) and pentadecanol-6EC [C15(EC)6] (c & d) using Q-TOF as analyzer in +ESI mode.

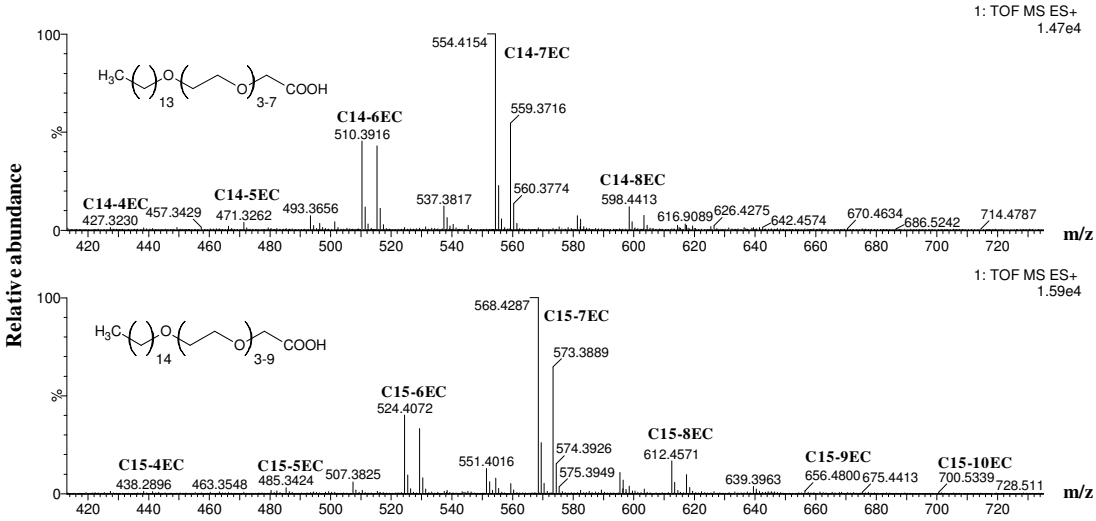


Figure 3.15: TOFMS spectra of tetradecanol ethoxycarboxylates (C14ECs) and pentadecanol ethoxycarboxylates (C15ECs) homologues using Q-TOF as analyzer and a collision energy of 10eV in +ESI mode (full scan mode).

Table 3.7: Alkyl polyethoxycarboxylates (AECs) and their metabolites identified in trout bile in both electrospray ionization modes (+/-ESI).

ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments	Parent compound formula[§]	Ion form	Identity of compound
+ESI	443.2983	22.90	$C_{22}H_{44}O_7Na$	443.2985	-0.5	5.5	385.2911	$C_{22}H_{44}O_7$	$[M+Na]^+$	Dodecanyl-5EC metabolite
+ESI	438.3430		$C_{22}H_{48}NO_7$	438.3431	-0.2	8.9	253.1284, 235.1180, 207.1239, 147.0661		$[M+NH_4]^+$	
+ESI	413.2881	23.56	$C_{21}H_{42}O_6Na$	413.2879	0.5	0.8	355.2822, 231.0846	$C_{21}H_{42}O_6$	$[M+Na]^+$	Tridecanol-4EC metabolite
+ESI	408.3331		$C_{21}H_{46}NO_6$	408.3325	1.5	3.6	209.1033, 191.0925, 163.0965		$[M+NH_4]^+$	
+ESI	457.3139	23.56	$C_{23}H_{46}O_7Na$	457.3141	-0.4	0.5	399.3083, 275.1110, 217.1063	$C_{23}H_{46}O_7$	$[M+Na]^+$	Tridecanol-5EC metabolite
+ESI	452.3589		$C_{23}H_{50}NO_7$	452.3587	0.4	0.1	253.1280, 235.1183, 207.1240, 147.0667		$[M+NH_4]^+$	
+ESI	501.3401	23.54	$C_{25}H_{50}O_8Na$	501.3403	-0.4	0.5	443.3354, 319.1371, 261.1310	$C_{25}H_{50}O_8$	$[M+Na]^+$	Tridecanol-6EC metabolite
+ESI	496.3850		$C_{25}H_{54}NO_8$	496.3849	0.2	1.7	297.1551, 279.1452, 251.1499, 191.0916		$[M+NH_4]^+$	
+ESI	545.3668	23.52	$C_{27}H_{54}O_9Na$	545.3666	0.4	1.4	487.3607, 363.1648, 305.1566	$C_{27}H_{54}O_9$	$[M+Na]^+$	Tridecanol-7EC metabolite
+ESI	540.4111		$C_{27}H_{58}NO_9$	540.4112	-0.2	0.5	341.1814, 323.1707, 295.1755, 235.1181		$[M+NH_4]^+$	
+ESI	427.3042	24.12	$C_{22}H_{44}O_6Na$	427.3036	1.4	1.4	369.2978, 231.0852	$C_{22}H_{44}O_6$	$[M+Na]^+$	Tetradecanol-4EC metabolite
+ESI	422.3484		$C_{22}H_{48}NO_6$	422.3482	0.5	0.6	209.1034, 191.0927, 163.0978		$[M+NH_4]^+$	
-ESI	403.3060		$C_{22}H_{43}O_6$	403.3060	0.0	0.4	211.2071		$[M-H]^-$	
+ESI	471.3300	24.13	$C_{24}H_{48}O_7Na$	471.3298	0.4	0.7	413.3229, 275.1109, 217.1049	$C_{24}H_{48}O_7$	$[M+Na]^+$	Tetradecanol-5EC metabolite
+ESI	466.3741		$C_{24}H_{52}NO_7$	466.3744	-0.6	1.1	253.1284, 235.1182, 207.1228, 147.0651		$[M+NH_4]^+$	
-ESI	447.3325		$C_{24}H_{47}O_7$	447.3322	0.7	0.3	211.2057		$[M-H]^-$	
+ESI	515.3560	24.12	$C_{26}H_{52}O_8Na$	515.3560	0.0	1.0	457.3510, 319.1365, 261.1313	$C_{26}H_{52}O_8$	$[M+Na]^+$	Tetradecanol-6EC metabolite
+ESI	510.4002		$C_{26}H_{56}NO_8$	510.4006	-0.8	1.3	297.1545, 279.1445, 251.1494, 191.0928		$[M+NH_4]^+$	
-ESI	491.3585		$C_{26}H_{51}O_8$	491.3584	0.2	0.4	211.2066		$[M-H]^-$	
+ESI	559.3830	24.12	$C_{28}H_{56}O_9Na$	559.3822	1.4	1.9	501.3774, 363.1620, 305.1581	$C_{28}H_{56}O_9$	$[M+Na]^+$	Tetradecanol-7EC metabolite
+ESI	554.4265		$C_{28}H_{60}NO_9$	554.4268	-0.5	0.6	341.1805, 323.1700, 297.1753, 235.1180		$[M+NH_4]^+$	
-ESI	535.3850		$C_{28}H_{55}O_9$	535.3846	0.7	1.0	211.2064		$[M-H]^-$	
+ESI	603.4080	24.09	$C_{30}H_{60}O_{10}Na$	603.4084	-0.7	2.8	545.4028, 407.1907, 349.1847	$C_{30}H_{60}O_{10}$	$[M+Na]^+$	Tetradecanol-8EC metabolite
+ESI	598.4530		$C_{30}H_{64}NO_{10}$	598.4530	0.0	0.4	385.2066, 367.1965, 339.2015, 279.1446		$[M+NH_4]^+$	
-ESI	579.4102		$C_{30}H_{59}O_{10}$	579.4108	-1.0	0.6	211.2059	$C_{30}H_{60}O_{10}$	$[M-H]^-$	

+/-ESI: positive/negative electrospray ionization; m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; *n*EC where n: represents number of ethoxy units plus a terminal CH_2COOH.

Table 3.7: (continued) Alkyl polyethoxycarboxylates (AECs) and their metabolites identified in trout bile in both electrospray ionization modes (+/-ESI).

ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments	Parent compound formula[§]	Ion form	Identity of compound
+ESI	441.3197	24.59	$C_{23}H_{46}O_6Na$	441.3192	1.1	0.3	383.3127, 231.0856	$C_{23}H_{46}O_6$	$[M+Na]^+$	Pentadecanol-4EC metabolite
+ESI	436.3640		$C_{23}H_{50}NO_6$	436.3638	0.5	1.9	209.1022, 191.0921, 163.0978		$[M+NH_4]^+$	
+ESI	485.3453	24.60	$C_{25}H_{50}O_7Na$	485.3454	-0.2	1.3	427.3387, 275.1116, 217.1062	$C_{25}H_{50}O_7$	$[M+Na]^+$	Pentadecanol-5EC metabolite
+ESI	480.3893		$C_{25}H_{54}NO_7$	480.3900	-1.5	0.6	253.1293, 235.1180, 207.1226, 147.0655		$[M+NH_4]^+$	
-ESI	461.3480		$C_{25}H_{49}O_7$	461.3478	0.4	0.2	225.2234	$C_{25}H_{50}O_7$	$[M-H]^-$	
+ESI	529.3718	24.60	$C_{27}H_{54}O_8Na$	529.3716	0.4	0.1	471.3653, 319.1363, 261.1315	$C_{27}H_{54}O_8$	$[M+Na]^+$	Pentadecanol-6EC metabolite
+ESI	524.4163		$C_{27}H_{58}NO_8$	524.4162	0.2	2.3	297.1552, 279.1445, 251.1492, 191.0922		$[M+NH_4]^+$	
-ESI	505.3742		$C_{27}H_{53}O_8$	505.3740	0.4	0.2	225.2230	$C_{27}H_{54}O_8$	$[M-H]^-$	
+ESI	573.3981	24.60	$C_{29}H_{58}O_9Na$	573.3979	0.3	21.2	515.3925, 363.1637, 305.1564	$C_{29}H_{58}O_9$	$[M+Na]^+$	Pentadecanol-7EC metabolite
+ESI	568.4423		$C_{29}H_{62}NO_9$	568.4425	-0.4	0.1	341.1811, 323.1701, 295.1755, 235.1187		$[M+NH_4]^+$	
-ESI	549.4006		$C_{29}H_{57}O_9$	549.4003	0.5	0.7	225.2226	$C_{29}H_{58}O_9$	$[M-H]^-$	
+ESI	617.4253	24.60	$C_{31}H_{62}O_{10}Na$	617.4241	-1.9	4.5	559.4181, 407.1895, 349.1837	$C_{31}H_{62}O_{10}$	$[M+Na]^+$	Pentadecanol-8EC metabolite
+ESI	612.4689		$C_{31}H_{66}NO_{10}$	612.4687	0.3	3.3	385.2075, 367.1960, 339.2021, 279.1445		$[M+NH_4]^+$	
-ESI	593.4269		$C_{31}H_{61}O_{10}$	593.4265	0.7	0.4	225.2208	$C_{31}H_{62}O_{10}$	$[M-H]^-$	
+ESI	661.4506	24.56	$C_{33}H_{66}O_{11}Na$	661.4503	0.5	2.8	603.4435 , 451.2148, 393.2113	$C_{33}H_{66}O_{11}$	$[M+Na]^+$	Pentadecanol-9EC metabolite
+ESI	656.4947		$C_{33}H_{70}NO_{11}$	656.4949	-0.3	1.7	429.2316, 411.2235, 383.2294, 323.1700		$[M+NH_4]^+$	
+ESI	705.4774	24.54	$C_{35}H_{70}O_{12}Na$	705.4765	1.3	6.3	647.4705 , 495.2407, 437.2348	$C_{35}H_{70}O_{12}$	$[M+Na]^+$	Pentadecanol-10EC metabolite
+ESI	700.5203		$C_{35}H_{74}NO_{12}$	700.5211	-1.1	1.1	473.2620, 455.2514, 427.2547		$[M+NH_4]^+$	

+/-ESI: positive/negative electrospray ionization; m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+NH_4]^+$: ammonium adduct; *n*EC where *n*: represents number of ethoxy units plus a terminal CH_2COOH.

Table 3.8: Q-TOFMS fragment ions for alkyl polyethoxy carboxylate (sodium and ammonium adducts) in +ESI mode.

			MS/MS fragments (*m/z*)					
Identity of compound	**Ion form**	**Observed ion (*m/z*)**	**Loss of acetate group**	**Loss of Alkene(C_nH_{2n})[#]**	**Loss of (alkene plus acetate group)**	**Loss of H_2O**	**Loss of CO**	**Loss of (2EO units plus H_2O)**
C12(EC)5	$[M+Na]^+$		385.2911					
	$[M+NH_4]^+$	438.3431		253.1284		235.1180	207.1239	147.0661
C13(EC)4	$[M+Na]^+$	413.2879	355.2822	231.0846				
	$[M+NH_4]^+$	408.3325		209.1033		191.0925	163.0965	
C13(EC)5	$[M+Na]^+$	457.3141	399.3083	275.1110	217.1063			
	$[M+NH_4]^+$	452.3587		253.1280		235.1183	207.1240	147.0667
C13(EC)6	$[M+Na]^+$	501.3403	443.3354	319.1371	261.1310			
	$[M+NH_4]^+$	496.3849		297.1551		279.1452	251.1499	191.0916
C13(EC)7	$[M+Na]^+$	545.3666	487.3607	363.1648	305.1566			
	$[M+NH_4]^+$	540.4112		341.1814		323.1707	295.1755	235.1181
C14(EC)4	$[M+Na]^+$	427.3036	369.2978	231.0852				
	$[M+NH_4]^+$	422.3482		209.1034		191.0927	163.0978	211.2071
C14(EC)5	$[M+Na]^+$	471.3298	413.3229	275.1109	217.1049			
	$[M+NH_4]^+$	466.3744		253.1284		235.1182	207.1228	147.0651
C14(EC)6	$[M+Na]^+$	515.3560	457.3510	319.1365	261.1313			
	$[M+NH_4]^+$	510.4006		297.1545		279.1445	251.1494	191.0928
C14(EC)7	$[M+Na]^+$	559.3822	501.3774	363.1620	305.1581			
	$[M+NH_4]^+$	554.4268		341.1805		323.1700	297.1753	235.1180
C14(EC)8	$[M+Na]^+$	603.4084	545.4028	407.1907	349.1847			
	$[M+NH_4]^+$	598.4530		385.2066		367.1965	339.2015	279.1446
C15(EC)4	$[M+Na]^+$	441.3197	383.3127	231.0856				
	$[M+NH_4]^+$	436.3640		209.1022		191.0921	163.0978	
C15(EC)5	$[M+Na]^+$	485.3453	427.3387	275.1116	217.1062			
	$[M+NH_4]^+$	480.3893		253.1293		235.1180	207.1226	147.0655
C15(EC)6	$[M+Na]^+$	529.3718	471.3653	319.1363	261.1315			
	$[M+NH_4]^+$	524.4163		297.1552		279.1445	251.1492	191.0922
C15(EC)7	$[M+Na]^+$	573.3981	515.3925	363.1637	305.1564			
	$[M+NH_4]^+$	568.4423		341.1811		323.1701	295.1755	235.1187
C15(EC)8	$[M+Na]^+$	617.4253	559.4181	407.1895	349.1837			
	$[M+NH_4]^+$	612.4689		385.2075		367.1960	339.2021	279.1445
C15(EC)9	$[M+Na]^+$	661.4506	603.4435	451.2148	393.2113			
	$[M+NH_4]^+$	656.4947		429.2316		411.2235	383.2294	323.1700
C15(EC)10	$[M+Na]^+$	705.4774	647.4705	495.2407	437.2348			
	$[M+NH_4]^+$	700.5203		473.2620		455.2514	427.2547	

[#]loss of alkene plus NH_3 from ammonium adduct. Additional fragments were obtained at *m/z* 211.2062 ($C_{14}H_{27}O$) and *m/z* 225.2218 ($C_{15}H_{29}O$) in -ESI mode.

3.3.4.5 Chlorinated phenols and their metabolites

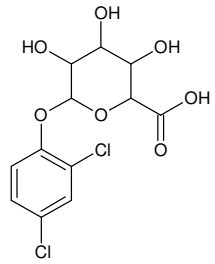
Several chlorinated compounds and their metabolites were detected in effluent-exposed trout bile. Chlorinated phenols were only detected in -ESI mode and were detected as glucuronide conjugates (Table 3.9); the charge site was probably located on the carboxylic group of the glucuronide moiety. Chlorophenol compounds substituted as monochlorine to pentachlorine compounds are commonly used as biocides and as intermediates in the synthesis of dyestuffs and pesticides (Gao et al., 2008).

Elemental compositions were calculated for the deprotonated ions at *m/z* 336.9882 (marker 1, Table 3.9) and 370.9493 (marker 2, Table 3.9) corresponded to the theoretical formulae of $C_{12}H_{11}O_7Cl_2$ and $C_{12}H_{10}O_7Cl_3$, respectively. These two markers were identified as glucuronide conjugates of 2,4-dichlorophenol and an isomer of trichlorophenol (structures shown in Figure 3.16). The typical chlorine isotopic pattern was observed in the mass spectra of both di- and tri-chlorinated compounds, aiding in the confirmation of the identity of these compounds (Figures 3.17a and 3.18a).

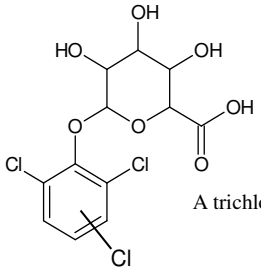
Q-TOFMS spectra of the deprotonated ions $[M-H]^-$ showed mainly two signals at *m/z* 160.9561 and 194.9163 (Figures 3.17b and 3.18b), which correspond to the deprotonated aglycone $[M-Glu-H]^-$ (loss of 176 Da) of dichlorophenol and trichlorophenol, respectively, in addition to other fragments characteristic of the glucuronide moiety (i.e. *m/z* 175.0243, *m/z* 113.0239). Elimination of a neutral glucuronide moiety by cleavage of the glycoside bond led to a very stable product ion. Its stability can probably be due to resonance stabilization and it is responsible of no further fragmentation at higher collision energies. Therefore, GC-MS analysis was needed to obtain further characteristic fragments for these two chlorinated phenols (dichlorophenol and trichlorophenol).

UPLC fractions containing the above mentioned conjugated chlorinated phenols were subjected to enzymatic hydrolysis first and then derivatised for further GC-MS analysis. Analysis of corresponding derivatized UPLC fractions by GC-MS showed the characteristic fragment ions and RT of the trimethylsilyl ether (TMS) derivatives of dichlorophenol whilst trichlorophenol TMS derivatives (RT=15.64) differed in RT when compared to the available commercial standard (2,4,6-trichlorophenol, RT=15.36) (for GC-MS fragmentation see Figure 3.17c and Figure 3.18c). This might indicate that the trichlorinated compound identified in the bile sample is a positional isomer of 2,4,6-trichlorophenol, differeing only in the relative position of the three chlorine atoms on

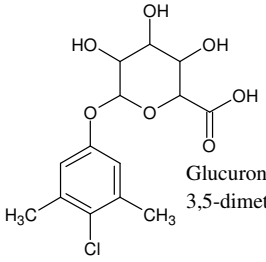
the phenolic ring. The peak at *m/z* 219 and *m/z* 253 in the spectra of dichlorophenol and trichlorophenol were produced by the loss of the methyl group from the molecular ions *m/z* 234 and *m/z* 268, respectively. The ions at *m/z* 219 and *m/z* 253 further fragmented by elimination of hydrogen chloride HCl (loss of 36 Da) and gave signals at *m/z* 183 and *m/z* 217, respectively. Furthermore, the characteristic ions at *m/z* 93 and *m/z* 95 corresponded to the dimethylsilyl chloride fragments $[^{35}Cl\text{-}Si(CH_3)_2]^+$ and $[^{37}Cl\text{-}Si(CH_3)_2]^+$ due to a rearrangement mechanism and elimination of CH_3Cl. This elimination could possibly occur when the chlorine atom is located on the same ring of the phenolic group; ions will show different intensities depending on the number of chlorine atoms and on their relative position respect to the aromatic –OH (ortho-, meta- or para-) (Stalling and Hogan, 1978, Heberer and Stan, 1997). Chlorine isotopic patterns of the derivative molecular ions and their fragments were also observed in the spectra for both phenols compounds, confirming the presence of halogens in the molecule.



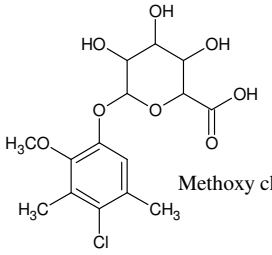
2,4-Dichlorophenol glucuronide



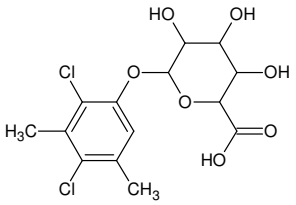
A trichlorophenol glucuronide



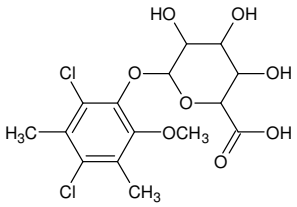
Glucuronide conjugate of 4-Chloro-3,5-dimethyl-phenol (chloroxylenol)



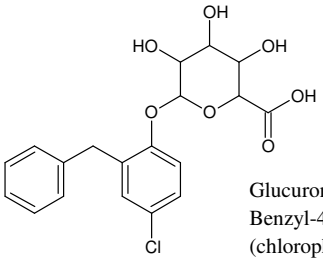
Methoxy chloroxylenol glucuronide



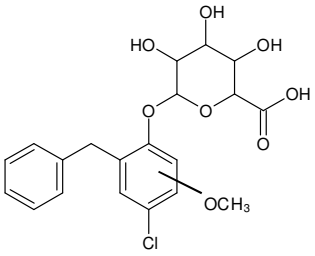
Dichloroxylenol glucuronide



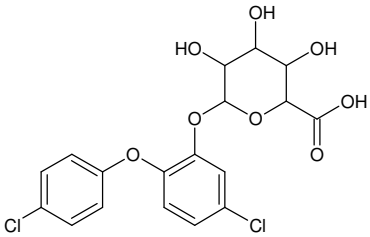
Methoxy dichloroxylenol glucuronide



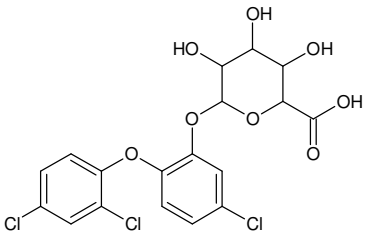
Glucuronide conjugate of 2-Benzyl-4-chlorophenol (chlorophene)



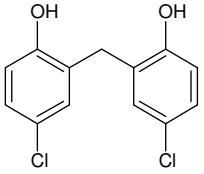
Methoxy chlorophene glucuronide



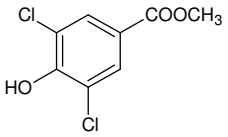
Diclosan glucuronide



Triclosan glucuronide



Dichlorophene



Methyl dichloroparaben

Figure 3.16: Chemical structures of chlorinated compounds and their metabolites found in bile of effluent-exposed fish.

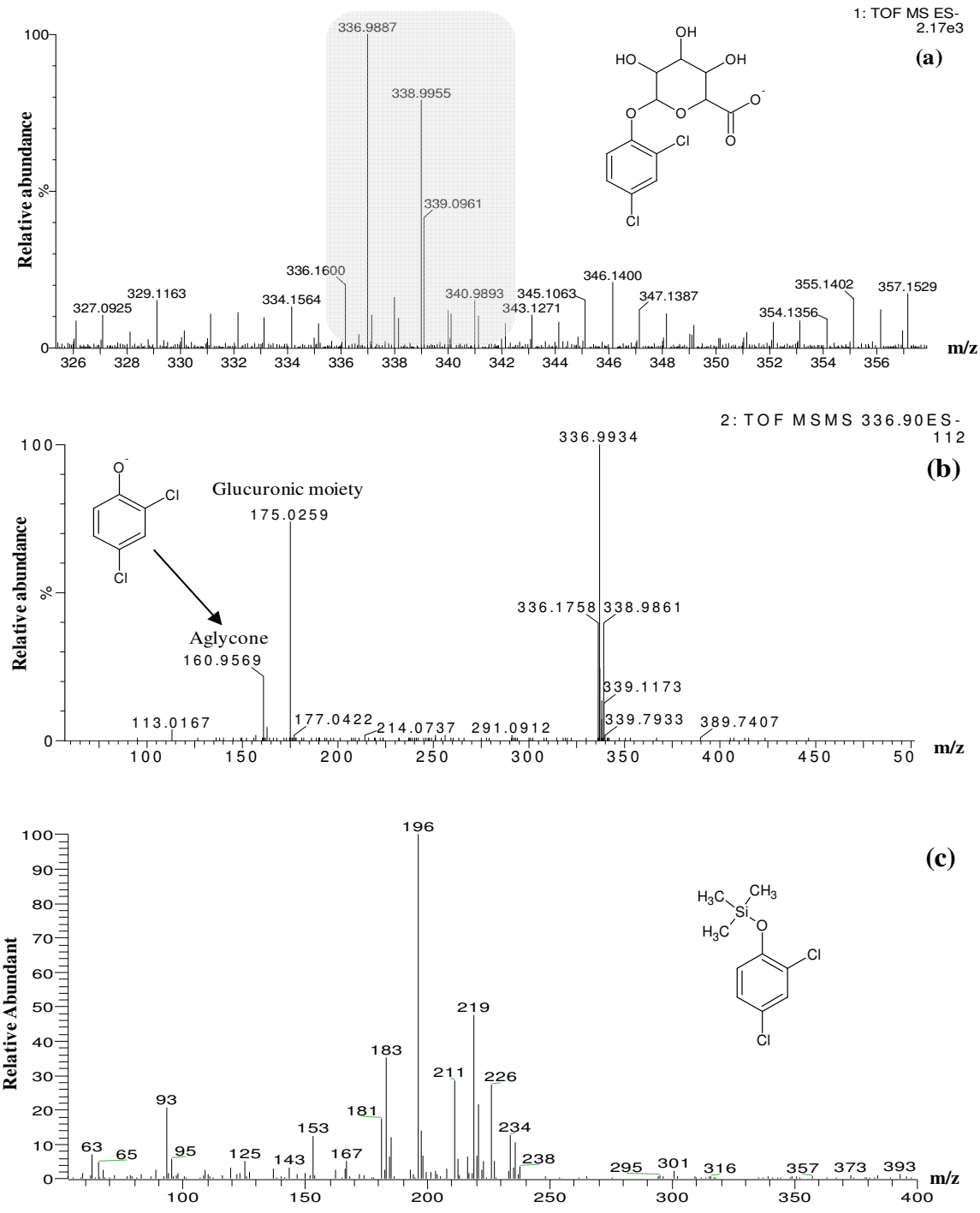


Figure 3.17:
(a) TOFMS mass spectrum of 2,4-dichlorophenol glucuronide detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A, A+2 and A+4.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of 2,4-dichlorophenol obtained applying a collision energy of 10 eV.
(c) EI mass spectra of 2,4-dichlorophenol as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

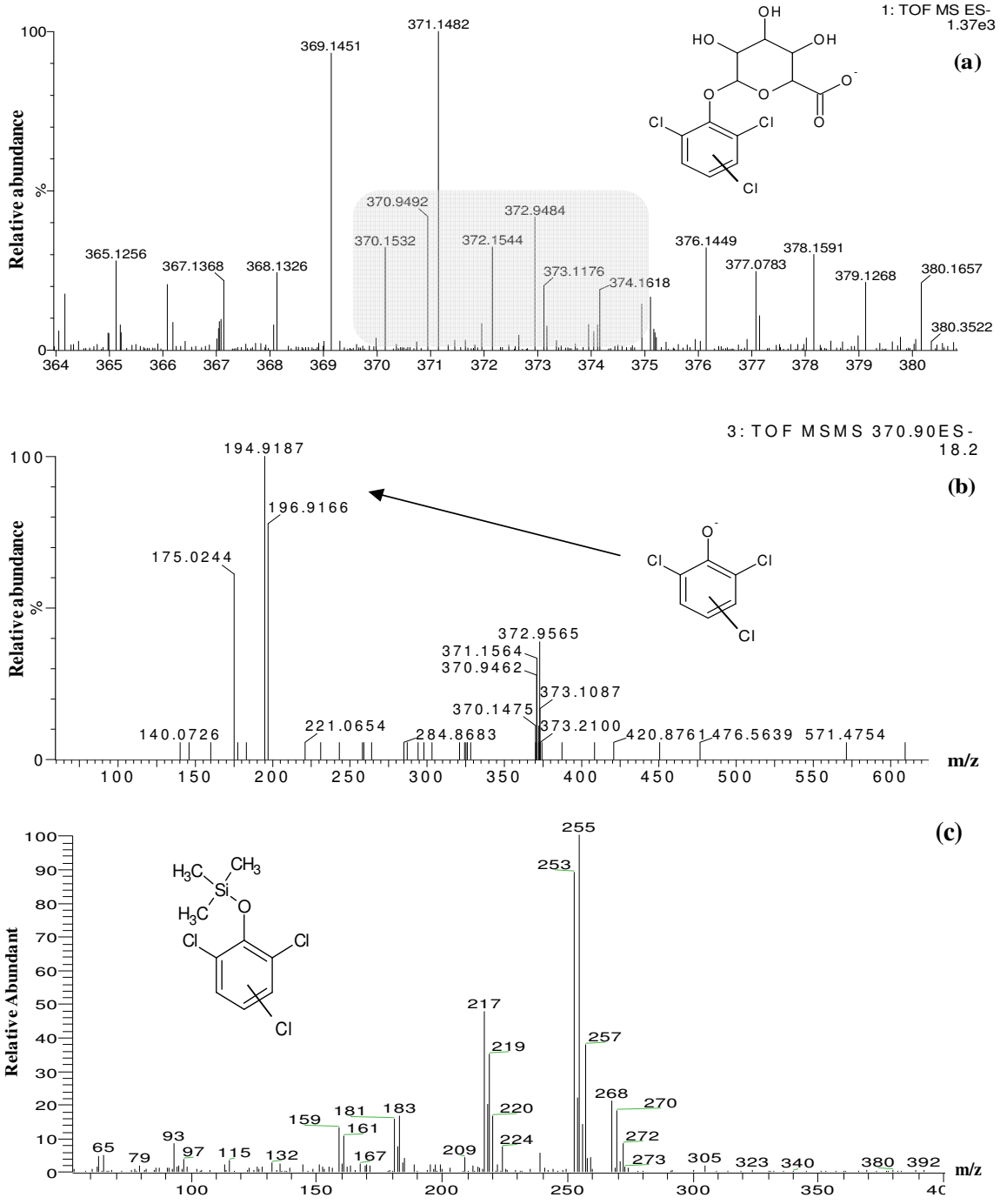


Figure 3.18:
(a) TOFMS mass spectrum of trichlorophenol glucuronide detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A, A+2 and A+4.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of trichlorophenol obtained applying a collision energy of 10 eV.
(c) EI mass spectra of a trichlorophenol as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

Table 3.9: Chlorinated markers and their metabolites identified in trout bile in -ESI mode.

Marker No.	ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments (plus fragments from GCMS analysis of the derivatized aglycone)[#]	Parent compound formula[§]	Ion form	Putative identity
1	-ESI	336.9883	6.39	$C_{12}H_{11}O_7Cl_2$	336.9882	0.3	1.7	175.0248, 160.9561, 113.0125 (219,234,183, 93)	$C_6H_4OCl_2$	$[M-H]^-$	Glucuronide conjugate of 2,4-dichlorophenol
2	-ESI	370.9493	6.56	$C_{12}H_{10}O_7Cl_3$	370.9492	0.3	1.0	194.9163, 175.0237 (268,253,217,183,93)	$C_6H_3OCl_3$	$[M-H]^-$	Glucuronide conjugate of a trichlorophenol
3	-ESI	395.0302	6.58	$C_{15}H_{17}O_8Cl_2$	395.0300	0.5	1.1	175.0224	$C_9H_{10}O_2Cl_2$	$[M-H]^-$	Glucuronide conjugate of putative methoxy metabolite of a dichloroxylenol isomer
4	-ESI	331.0588	7.41	$C_{14}H_{16}O_7Cl$	331.0585	0.9	1.0	175.0243, 155.0271, 113.0241 (228,213,177,163,93)	C_8H_9OCl	$[M-H]^-$	Chloroxylenol glucuronide (4-chloro-3,5-dimethyl-phenol)
5	-ESI	361.0694	7.85	$C_{15}H_{18}O_8Cl$	361.0690	1.1	0.5	185.0372, 175.0245, 170.0136, 113.0240 (258,243,228,193)	$C_9H_{11}O_2Cl$	$[M-H]^-$	Glucuronide conjugate of a methoxy metabolite of chloroxylenol
6	-ESI	395.0302	8.62	$C_{15}H_{17}O_8Cl_2$	395.0300	0.5	0.1	175.0233	$C_9H_{10}O_2Cl_2$	$[M-H]^-$	Glucuronide conjugate of putative methoxy metabolite of a dichloroxylenol isomer
7	-ESI	365.0197	8.67	$C_{14}H_{15}O_7Cl_2$	365.0195	0.5	0.9	188.9879, 175.0249, 157.0144, 113.0240 (262,247,211,197,93)	$C_8H_8OCl_2$	$[M-H]^-$	Dichloroxylenol glucuronide
8	-ESI	423.0853	10.48	$C_{20}H_{20}O_8Cl$	423.0847	1.4	1.7	247.0537, 232.0282, 175.0244 (320,305,290,255)	$C_{14}H_{13}O_2Cl$	$[M-H]^-$	Glucuronide conjugate of a methoxy metabolite of chlorophene
9	-ESI	393.0745	10.70	$C_{19}H_{18}O_7Cl$	393.0741	1.0	0.0	217.0428, 175.0251, 113.0241 (290,275,255,197,165,152,135,91,75)	$C_{13}H_{11}OCl$	$[M-H]^-$	Chlorophene glucuronide (2-benzyl-4-chlorophenol)
10	-ESI	429.0146	11.16	$C_{18}H_{15}O_8Cl_2$	429.0144	0.5	1.7	252.9823, 175.0248, 113.0239 (326,311,276,218,200,185,170)	$C_{12}H_8O_2Cl_2$	$[M-H]^-$	Diclosan glucuronide
11	-ESI	462.9759	12.49	$C_{18}H_{14}O_8Cl_3$	462.9754	1.1	0.0	286.9434, 175.0245, 113.0238 (360,345,310,252,200,185,170)	$C_{12}H_7O_2Cl_3$	$[M-H]^-$	Triclosan glucuronide

+/-ESI: positive/negative electrospray ionization; m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm and additional fragments observed from GC-MS of the aglycone and after derivatization to TMS ether; [§]aglycone formula; $[M-H]^-$: deprotonated ion.

3.3.4.6 Chlorinated xylenols and their metabolites

Two xylenols and their metabolites were detected only in -ESI mode and were detected as glucuronide conjugates in effluent-exposed trout bile (Table 3.9). Chlorinated xylenols are commonly used as antiseptics in household products. Amongst this class of compounds, chloroxylenol is used as disinfectant in the pharmaceutical, industrial, and cosmetic products (Bruch, 1996).

The analysis of the S-plot of markers metabolites included two ions that were significantly increased in bile from effluent-exposed trout; *m/z* 331.0588 (marker 4, Table 3.9) and 365.0197 (marker 7, Table 3.9). Accurate mass measurement, Q-TOFMS fragmentation, and GC-MS analysis led to the identification of the compounds as glucuronide conjugate of chloroxylenol (4-chloro-3,5-xylenol) and dichloroxylenol (structures shown in Figure 3.16). Q-TOFMS fragmentation revealed ions corresponding to the aglycone compounds at (*m/z* 155.0264 chloroxylenol and *m/z* 188.9874 dichloroxylenol), as well as glucuronide moiety fragments. TOFMS spectrum and high collision energy Q-TOFMS spectrum of both xylenols are shown in Figures (3.19a,b and 3.20a,b).

UPLC fractions containing the xylenol conjugates were subjected to enzymatic hydrolysis, and derivatised for GC-MS analysis. In the electron impact mass spectrum (EI-MS) of the mono- and di-chlorinated xylenol TMS derivatives, the presence of the respective molecular ions $[M]^{\bullet+}$ at *m/z* 228 and *m/z* 262 indicated that the hydroxyl groups were successful derivatized (Figure 3.19c and Figure 3.20c). Fragment ions at *m/z* 213 and *m/z* 247 were observed after loss of methyl $[M-15]^{+}$ from the trimethylsilyl group for both derivative compounds. Abundant ions at *m/z* 177 and *m/z* 211 were obtained due to loss of methyl and further loss of HCl from the molecular ions of chloroxylenol and dichloroxylenol, respectively, together with ions at *m/z* 93 and *m/z* 95, which corresponded to dimethylsilyl chloride. Loss of CH_3Cl following loss of methyl from the trimethylsilyl group was observed giving rise to ions at *m/z* 163 and *m/z* 197 for both xylenol compounds (Balmer et al., 2003, Cowan et al., 2008). The observed chlorine isotopic patterns of the fragments and molecular ions were consistent with the structures of the mono- and di-chlorinated compounds and relative fragments.

Methoxy metabolites of glucuronide conjugated chloroxylenols (Table 3.9; marker 3, 5 and 6) were also detected in –ESI mode in effluent-exposed trout bile (Figure 3.21a). Fragment ions corresponding to the aglycone ion of methoxychloroxylenol at *m/z* 185.0369 and ion due to loss of a methyl group from the aglycone itself at *m/z* 170.0135 were obtained by Q-TOFMS experiments (Figure 3.21b) (Thomas and Kotchevar, 2010). However, no sufficient fragmentation was achieved for putative methoxydichloroxylenol metabolite (data not shown). Further GC-MS analysis of the relevant UPLC fractions revealed an ion at *m/z* 258 (Figure 3.21c), which corresponded to the molecular ion of methoxy metabolites of chloroxylenol TMS derivative. The ion at *m/z* 243 resulted from either loss of a methyl from the TMS group or from the elimination of methyl from the methoxy group. The loss of 30 Da from the molecular ion could be due to cleavage of the methoxy group (OCH_3) generating the very stable base peak ion at *m/z* 228 stabilized by resonance. The putative metabolite methoxydichloroxylenol was not detected by analysis of its corresponding fraction by GC-MS, possibly due to its low abundance in the bile samples.

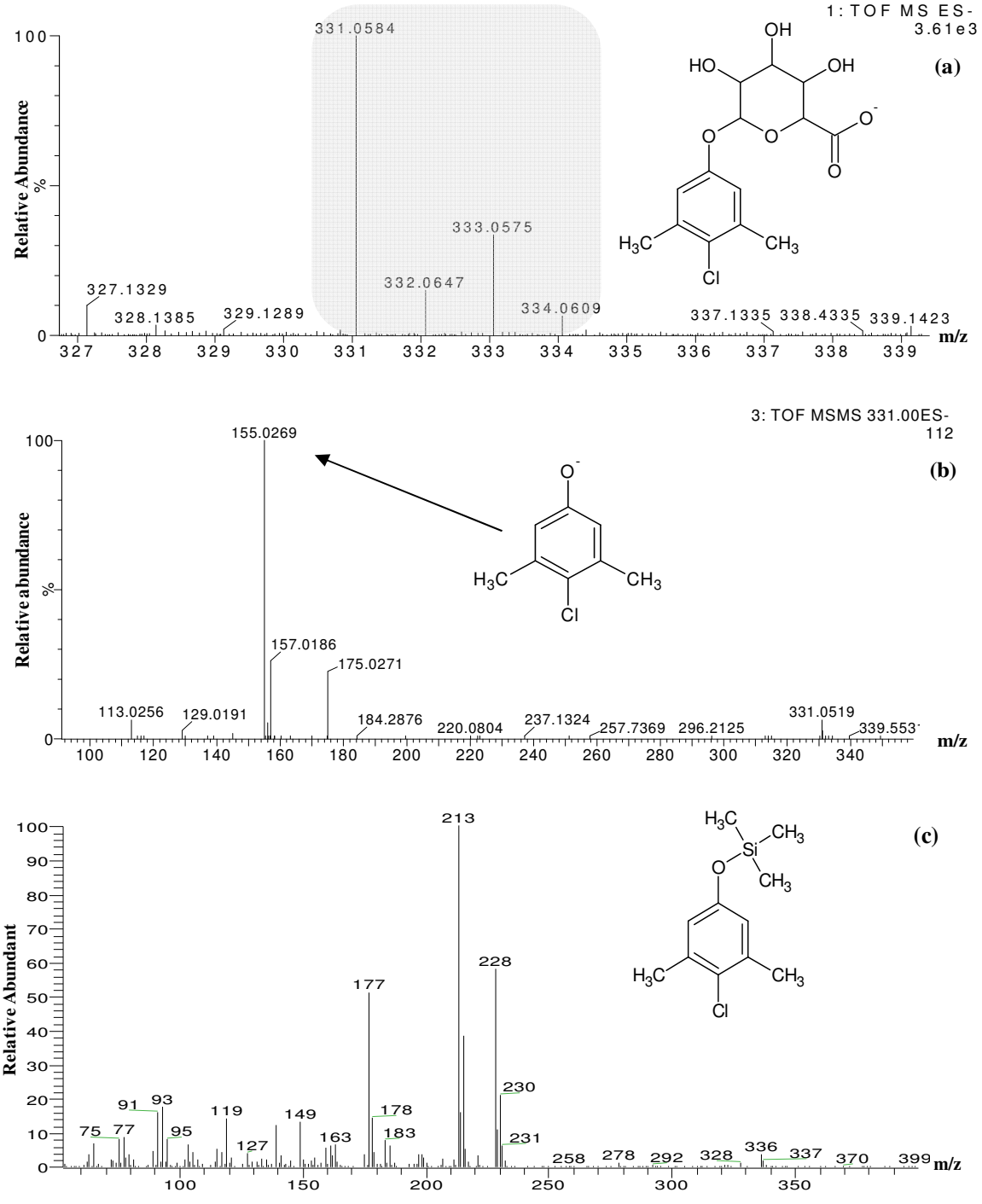


Figure 3.19:
(a) TOFMS mass spectrum of chloroxylenol glucuronide detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A and A+2.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of chloroxylenol obtained applying a collision energy of 20 eV.
(c) EI mass spectrum of chloroxylenol as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

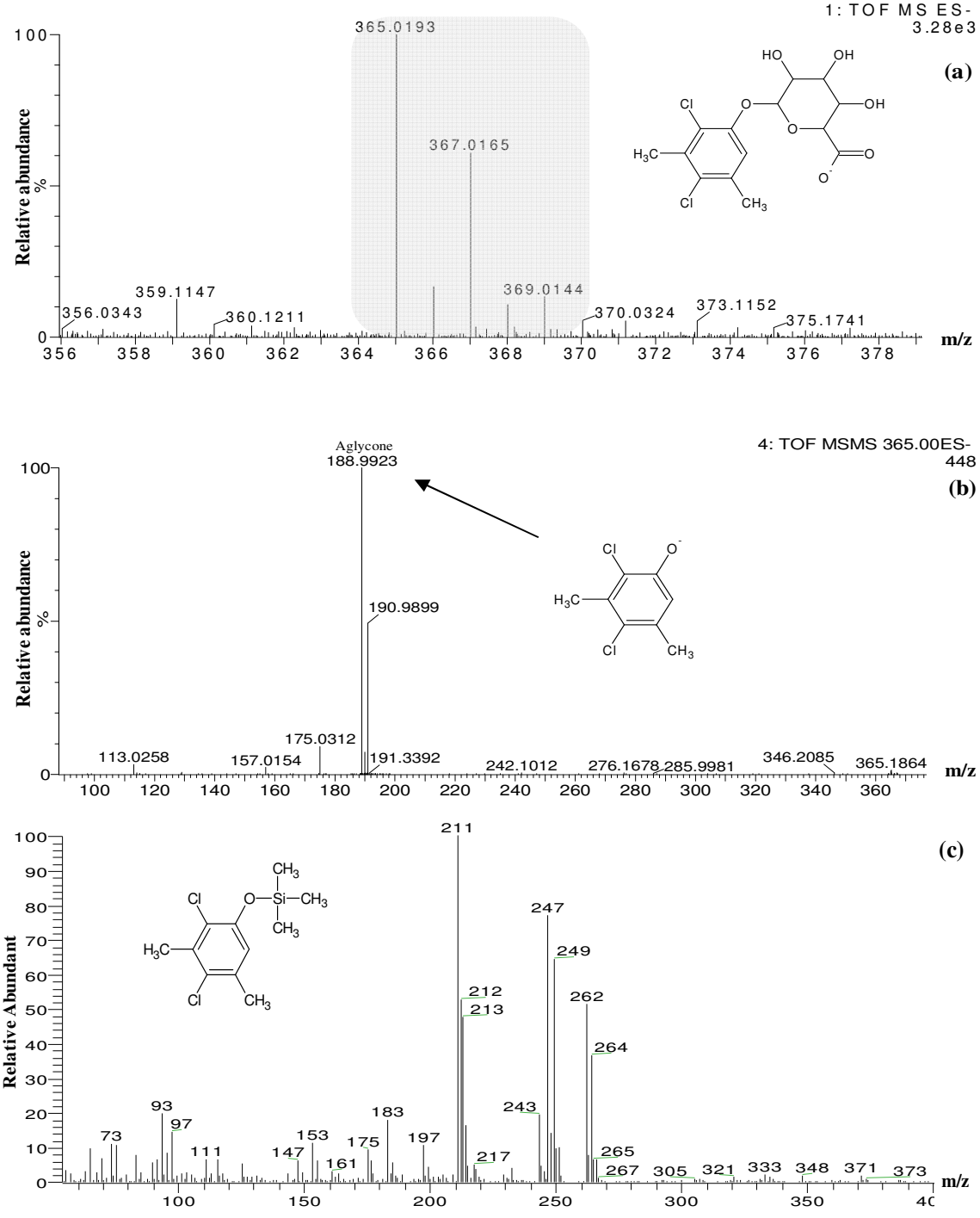


Figure 3.20:
(a) TOFMS mass spectrum of dichloroxylenol glucuronide detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A, A+2 and A+4.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of dichloroxylenol obtained applying a collision energy of 20 eV.
(c) EI mass spectra of dichloroxylenol as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

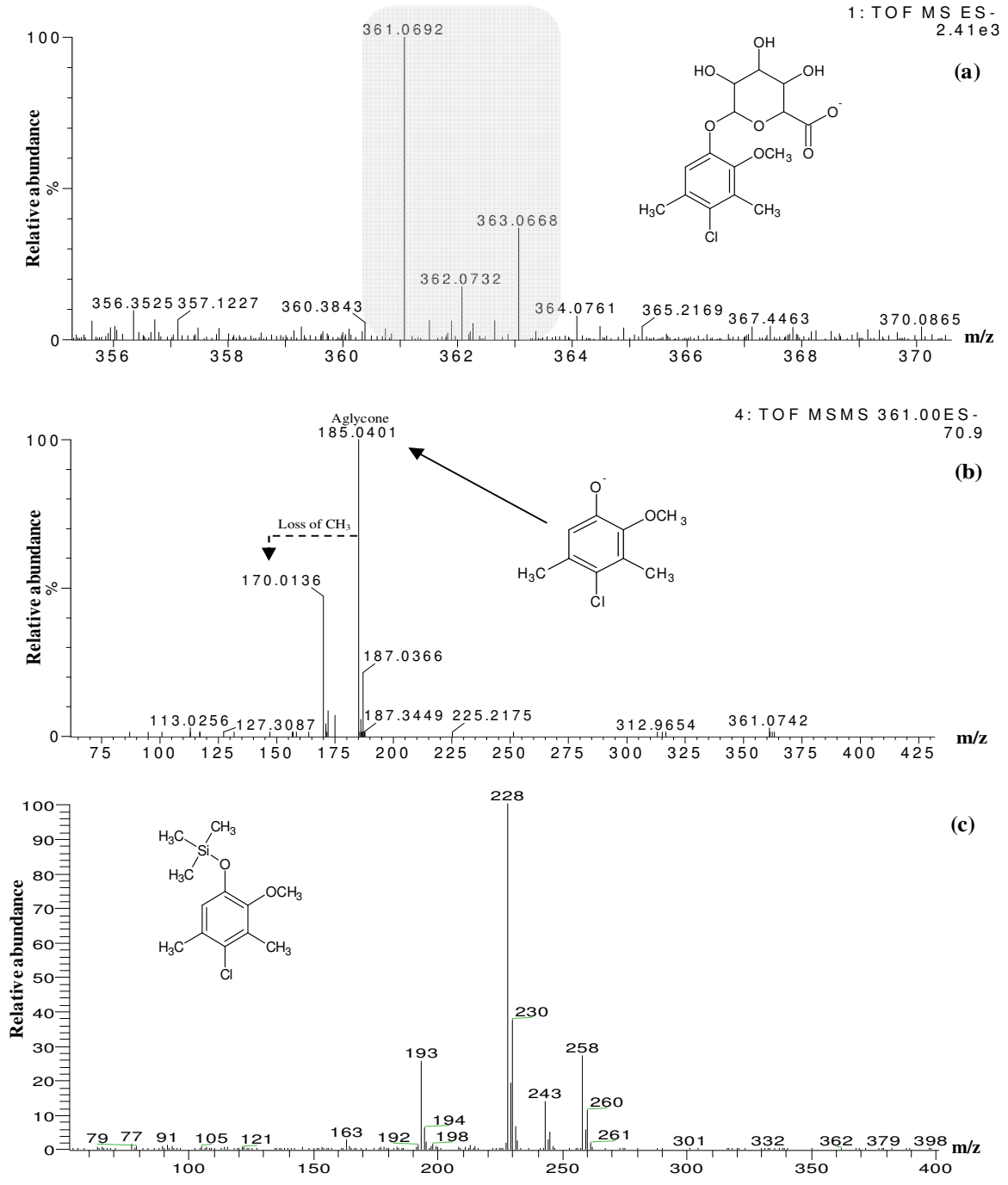


Figure 3.21:
(a) TOFMS mass spectrum of glucuronide conjugate of methoxy chloroxylenol detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A and A+2.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of methoxy chloroxylenol obtained applying collision energy of 20 eV.
(c) EI mass spectra of methoxy chloroxylenol as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

3.3.4.7 Chlorophenes and their metabolites

Chlorophenes and their metabolites were detected amongst the variety of chlorinated compounds detected in bile from effluent-exposed fish. UPLC-TOFMS analysis (-ESI mode) showed peaks, for glucuronide conjugates of chlorophene and for its methoxy metabolite. Prominent ions for these two metabolites were *m/z* 393.0745 (marker 9, Table 3.9) and *m/z* 423.0853 (marker 8, Table 3.9), respectively corresponding to $[M+Glu-H]^-$, (where Glu=glucuronide moiety). Both ions presented the typical isotopic pattern characteristic of mono-chlorinated molecules (Figure 3.22a and Figure 3.23a). Chlorophene (4-chloro-2-(phenylmethyl) phenol) is widely used as germicide in disinfectant solutions and soap formulations used in hospitals and household products (Werner et al., 1983). Structures for both identified markers are given in Figure 3.16. Clusters of chlorophene glucuronide $[M+M-H]^-$ was also detected as potential marker at *m/z* 787.1567 ($C_{38}H_{37}O_{14}Cl_2$) (data not shown); the formation of clusters in the source can be further justified by the fact that the molecular ion for chlorophene glucuronide in the mass spectrum showed a high abundance (i.e. due to a high concentration of ions in the source, they are likely to be combined as clusters).

Q-TOFMS fragmentation experiments led to the formation of signals at *m/z* 217.0428 and *m/z* 247.0537 which were identified as the corresponding aglycone [M-Glu-H] for chlorophene and methoxy chlorophene, respectively, along with the glucuronide moiety fragments (Figures 3.22b and 3.23b). One further fragment at *m/z* 232.0282 was obtained for methoxy chlorophene due to loss of a methyl group. In order to verify the identity of this marker and its metabolite, GC-MS analysis was performed on relevant fractions previously subjected to enzymatic hydrolysis. Fragmentation patterns of the compounds of interest were compared to those obtained from the commercial standard i.e. the TMS derivative of chlorophene standard. In Figure 3.22c, the EI mass spectrum of the chlorophene TMS derivative clearly shows signals at *m/z* 290 (molecular ion), *m/z* 275 (loss of CH_3 from trimethylsilyl group), *m/z* 255 (loss of Cl from molecular ion) together with ions at *m/z* 197, 165, 152, 135, 91 and *m/z* 75, which have not been fully characterized. On the other hand, methoxy metabolite of chlorophene TMS derivative revealed signals at *m/z* 320 (molecular ion), *m/z* 305 (loss of CH_3 either from TMS group or from methoxy group), *m/z* 290 (loss of OCH_3 from molecular ion), and *m/z* 255 which derived from the loss of CH_3Cl, as previously observed in the methoxy metabolite of chloroxylenol compound (Figure 3.23c).

Some other xenobiotics in the bile were also detected by chance during confirmation of the marker chemical structures during GC-MS analysis of the hydrolysed bile fractions. For instance, dichlorophene was not detected either as conjugated or as non-conjugated form in bile extracts by UPLC-TOFMS. However, GC-MS analysis of its TMS derivative (Figure 3.24) showed the characteristic ions, RT and isotopic pattern of the TMS derivatives of dichlorophene standard (i.e. the molecular ion of *m/z* 412 and a signal at *m/z* of 377 due to loss of a methyl from TMS group, in addition to the characteristic fragment at *m/z* 73 which corresponds to $((CH_3)_3Si)$. The base peak at *m/z* 73 corresponded to a typical fragment of TMS derivatives, indicating the occurrence of efficient derivatization of the two hydroxyl groups in the molecule.

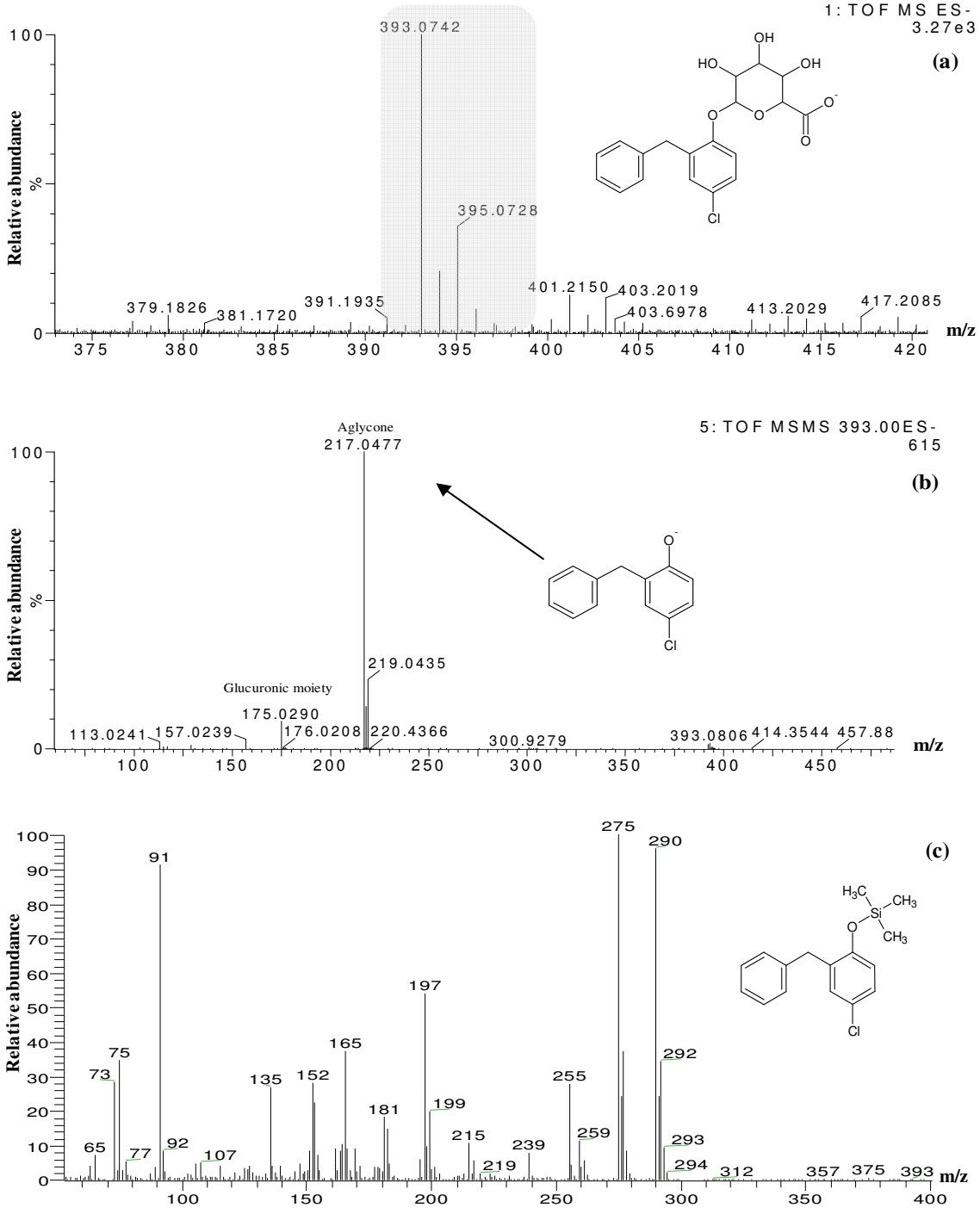


Figure 3.22:
(a) TOFMS mass spectrum of the glucuronide conjugate of chlorophene detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A and A+2.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of chlorophene obtained applying collision energy of 20 eV.
(c) EI mass spectra of chlorophene as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

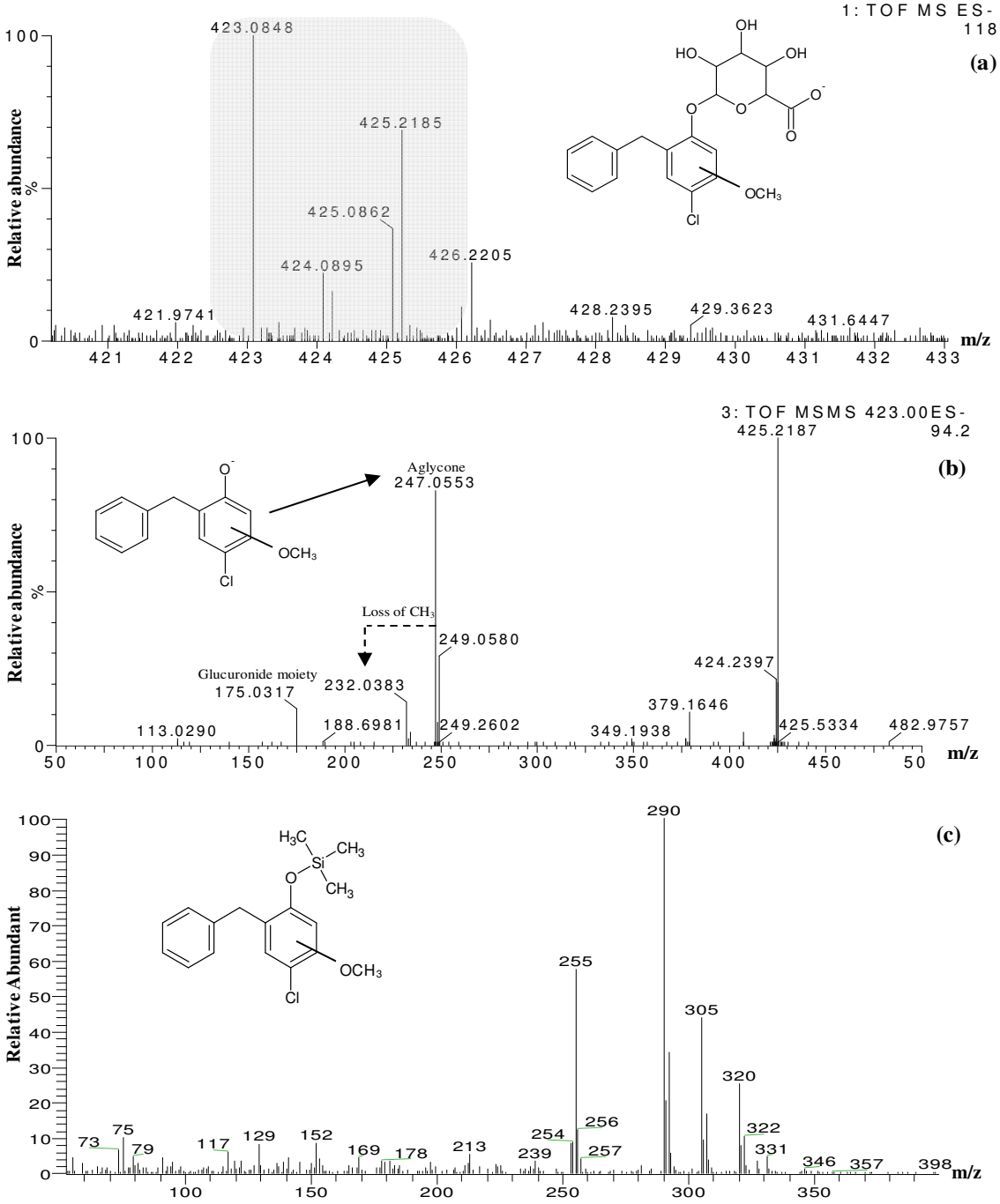


Figure 3.23:
(a) TOFMS mass spectrum of the glucuronide conjugate of methoxy chlorophene detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A and A+2.
(b) Q-TOFMS mass spectrum of the glucuronide conjugate of methoxy chlorophene obtained applying collision energy of 20 eV.
(c) EI mass spectra of methoxy chlorophene as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

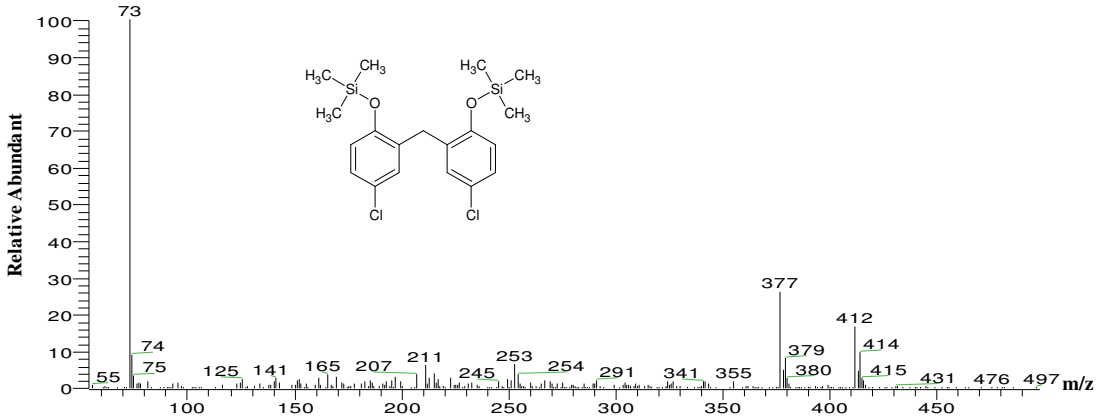


Figure 3.24: Electron impact (EI) mass spectrum of dichlorophene as the TMS derivative in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

3.3.4.8 Chlorinated phenoxyphenol and their metabolites

Chlorinated phenoxyphenol compounds (i.e. triclosan) were detected in bile from effluent-exposed fish. Triclosan is one of the most common broad spectrum antibacterial, antimicrobial and preservative agents. It is primarily used in personal care products and likewise incorporated in plastic kitchenware, footwear and clothing (Dann and Hontela, 2011).

Amongst the selected markers responsible for the class separation between control and exposed trout, two ions were observed at *m/z* 429.0146 (marker 10, Table 3.9) and *m/z* 462.9759 (marker 11, Table 3.9). These markers were only detected in – ESI and were then identified as glucuronide conjugates of diclosan and triclosan (5-chloro-2-(2,4-dichlorophenoxy) phenol), respectively, (Figure 3.16). In addition, two other RT-*m/z* signals co-eluting with triclosan glucuronide were extracted as potential markers from the S-plot. According to their elemental composition, they were confirmed as the aglycone ion $[M\text{-}Glu\text{-}H]^-$ and the cluster $[M+M\text{-}H]^-$ for triclosan glucuronide, showing a typical 3- and 6-chlorine isotopic pattern, respectively (data not shown). As observed in Figures 3.25a and 3.26a UPLC-TOFMS profiling (full scan mode; collision energy 10eV) shows the mass spectrum relative to the glucuronated compounds. Q-TOFMS fragmentation experiments gave rise mainly to the deprotonated aglycone ions, which showed signals at *m/z* 252.9823 and *m/z* 286.9433 for diclosan and triclosan, respectively, as well as other fragments characteristic of the glucuronide moiety (Figures 3.25b and 3.26b).

In the EI MS spectrum of the phenoxyphenol TMS derivatives (Figures 3.25c and 3.26c), signals at *m/z* 326 and *m/z* 360 were obtained, corresponding to the relative molecular ions $[M]^{\bullet+}$ of diclosan and triclosan, respectively. The $[M-15]^{+}$ fragment ion (*m/z* 311 and *m/z* 345) was generated for both markers after loss of methyl from the TMS group. Elimination of CH_3Cl from the two molecular ions was observed, giving rise to peaks at *m/z* 276 (diclosan) and *m/z* 310 (triclosan). Both chemical markers showed the same base peak at *m/z* 200 with one chlorine pattern. Sequential losses of CH_3 from this ion led to the ions observed at *m/z* 185 and *m/z* 170 with characteristic one Cl-pattern. RT time and fragmentation pattern were compared for confirm the identity of the mentioned compounds with those obtained from a commercial standard of triclosan (as the TMS derivative) showing good matching results.

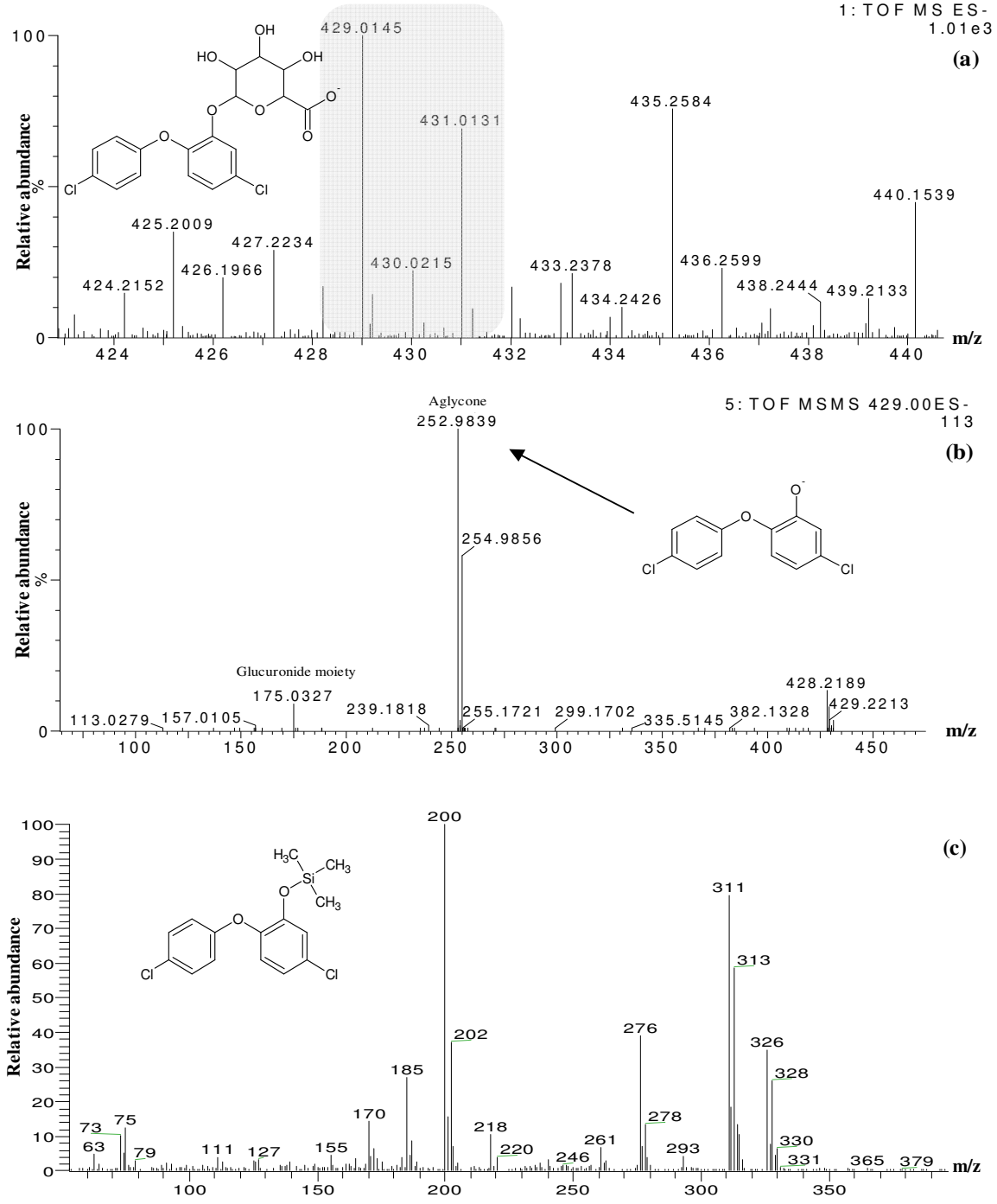


Figure 3.25:
(a) TOFMS mass spectrum of diclosan glucuronide detected in bile of effluent-exposed fish. The molecular ion shows the characteristic chlorine isotopic distribution at A, A+2 and A+4.
(b) Q-TOFMS mass spectrum of glucuronide conjugate of diclosan obtained applying collision energy of 20 eV.
(c) EI mass spectra of diclosan as the TMS derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

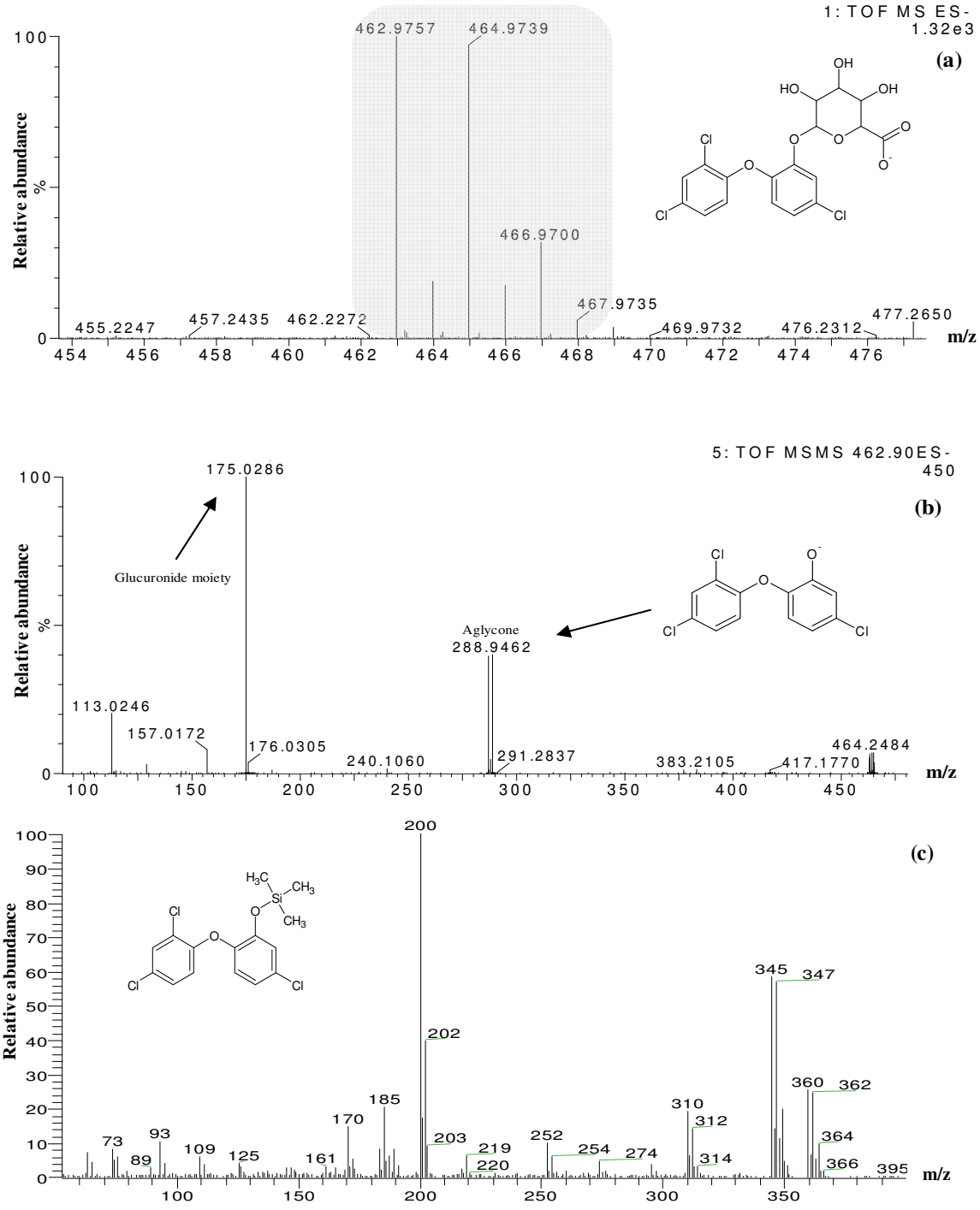


Figure 3.26:
(a) TOFMS mass spectrum of triclosan glucuronide identified in bile of effluent-exposed fish. The molecular ion shows the characteristic 3-chlorine isotopic distribution for A:A+2:A+4 signals, matching with the theoretical ratio of (3:3:1).
(b) Q-TOFMS mass spectrum of glucuronide conjugate of triclosan obtained applying a collision energy of 20 eV.
(c) EI mass spectrum of the TMS triclosan derivative detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

3.3.4.9 Chlorinated parabens

A water disinfection by-product (dichloromethylparaben) was found in exposed-effluent trout bile. Chlorine is known to generate disinfection by-products by reacting with organic components. Chlorinated parabens can be formed when parabens come in contact with chlorinated water, and is more likely for parabens containing phenolic groups. Parabens are commonly used as preservatives in personal care products and pharmaceuticals due to their antimicrobial activity (Canosa et al., 2006). However dichloromethylparaben was only identified by chance during GC-MS analysis of the whole bile sample and was not a marker detected from UPLC-TOFMS analysis of trout bile. Nevertheless GC-MS analyses showed that it was present in bile from effluent-exposed trout but not present in the same fraction from control trout. Since no commercial standard was available for this compound, in order to confirm its identity, the corresponding EI-MS spectrum of the derivatized compound was compared to the one obtained from the synthesised standard characterized in previous studies (Terasaki and Makino, 2008). As shown in Figure 3.27, there are several similarities between the two EI mass spectra: specifically, the presence of the molecular ion at *m/z* 292, the base peak corresponding to the loss of a methyl group at *m/z* 277, the ion at *m/z* 261 due to loss of the methoxy group from the ester side chain, and the characteristic ion at *m/z* 93 corresponding to the dimethylsilyl chloride fragment $[^{35}Cl\text{-}Si(CH_3)_2]^+$. However, some additional fragment ions can be observed in the bile sample: the ion at *m/z* 227 generated after loss of CH_3Cl from the base peak at *m/z* 277, and further ions at *m/z* 203 and at *m/z* 135. The mass spectrum of this metabolite presented a highly characteristic M+2 isotopic distribution patterns for 2-chlorine atoms.

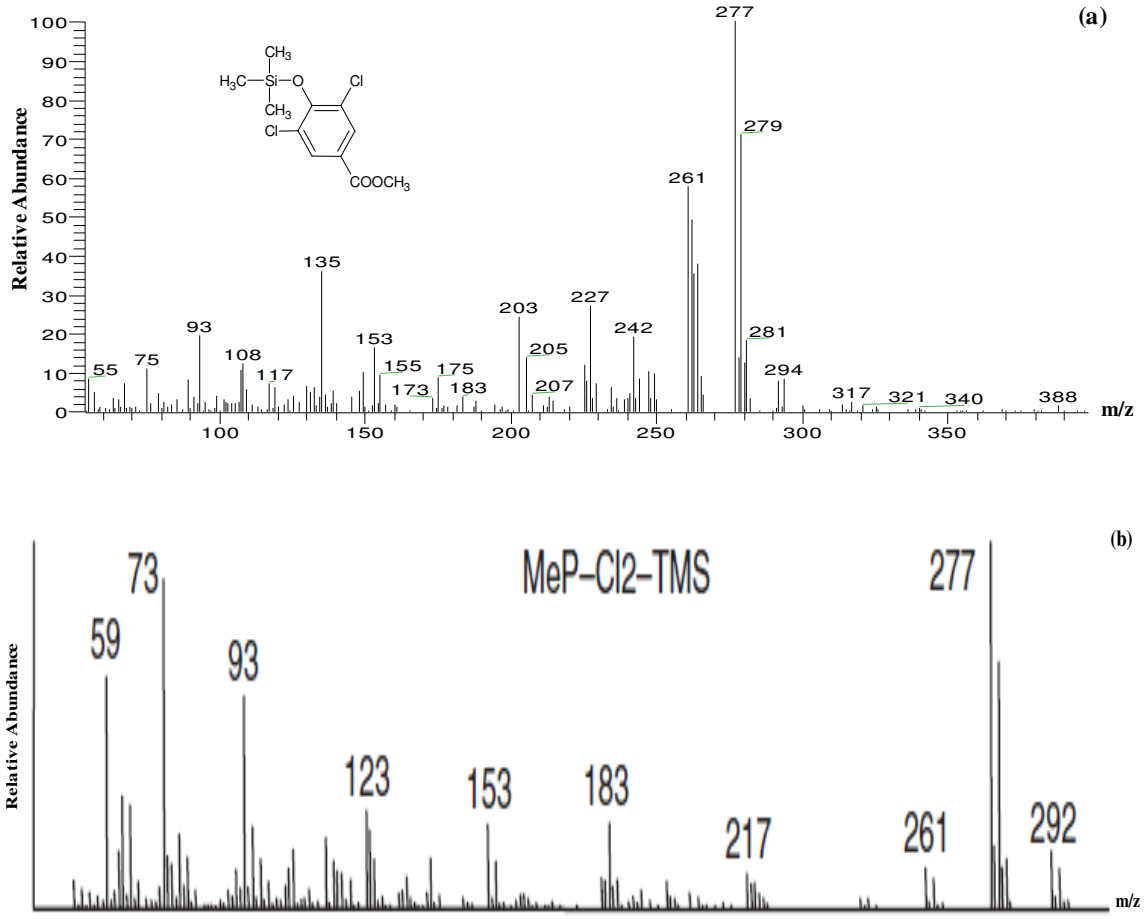
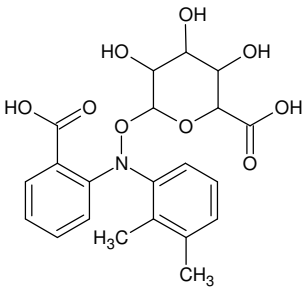


Figure 3.27: Comparison of mass spectra for methyldichloroparaben (a) in an UPLC fraction of bile subjected to enzymatic hydrolysis of fish exposed to effluent and (b) adopted from previous study (Terasaki and Makino, 2008).

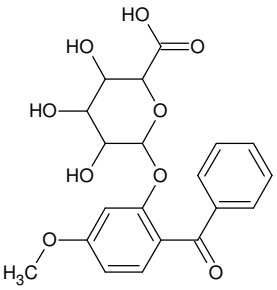
3.3.4.10 Metabolites of aromatic hydrocarbons

A group of aromatic hydrocarbons was also detected in this study. Within this class of contaminants, naphthol and 1-hydroxypyrene were found present in bile from effluent-exposed trout. Naphthol was detected as marker metabolite from UPLC-TOFMS profiling whilst the 1-hydroxypyrene was identified only by GC-MS analysis (Figure 3.28). Additional information obtained from GC-MS analysis confirmed the presence of naphthol in the bile as two different isomers, 1-naphthol and 2-naphthol. UPLC-TOFMS analysis in –ESI mode revealed the presence of a potential marker at *m/z* 319.0815 (marker 1, Table 3.10), which was attributed to the glucuronide conjugate of naphthol. Q-TOFMS experiments selecting a parent ion at *m/z* 319.0815 did not provide an informative fragmentation pattern; however, a low signal for a fragment ion at *m/z* 143.0497 corresponding to $[M\text{-}Glu\text{-}H]^-$ was observed (data not shown). High

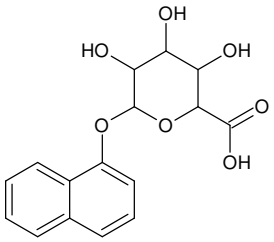
collision energy did not give any further fragmentation of the aglycone ion due to the high stability of the polycyclic aromatic molecule. The above mentioned marker was further characterized by GC-MS analysis of the relevant UPLC fraction after TMS derivatization of the hydroxyl group. Two peaks (at retention time 16.3 min and 16.7 min) were observed (Figure 3.29), showing the same characteristic ion: this indicates the presence of two isomers of the same metabolite. EI fragmentation pattern were compared to the one obtained from standards of 1-napthol and 2-naphthol and resulted comparable mass spectra. The MS spectrum showed a number of different informative ion peaks: the molecular ion $[M]^{\bullet+}$ (*m/z* 216), the ion $[M\text{-}CH_3]^+$ (*m/z* 201) due to the loss of methyl from the TMS group, and the ion $[M\text{-}31]^+$ (*m/z* 185) due to the cleavage of the trimethylsilyl ether moiety followed by cyclization involoving the silyl group to obtain a five member ring (Schummer et al., 2009). This mechanism is characteristic of compounds having another aromatic ring in the α-position to the hydroxyl group. The same loss $[M\text{-}31]^+$ (*m/z* 185) was also found in the fragmentation pattern of 2-naphtol, although it should not have been present if considered the mechanism mentioned above. However, in this specific case the ion *m/z* 185 showed lower intensity if compared to the spectrum of 1-naphthol (base peak), indicating lower ion stability. This result might be explained by a different mechanism leading to the same loss and the formation of a four member ring instead, which would result less stable due to the higher ring tension. Additional unidentified fragments were observed at *m/z* 141 $[M\text{-}75]^+$, *m/z* 115 $[M\text{-}101]^+$ for 1-naphthol and at *m/z* 145 and *m/z* 115 for 2-naphthol.



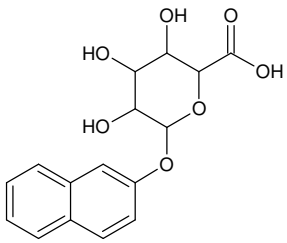
Glucuronide conjugate of mefenamic acid



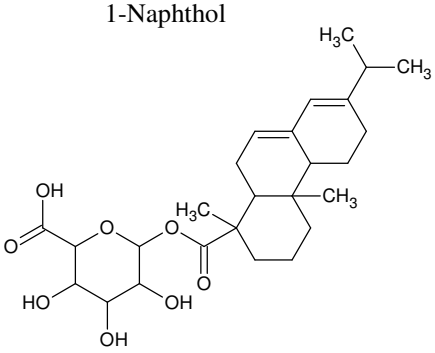
Oxybenzone glucuronide



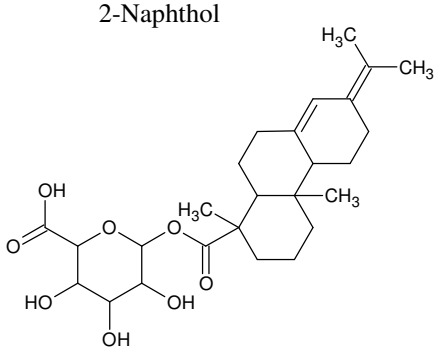
1-Naphthol



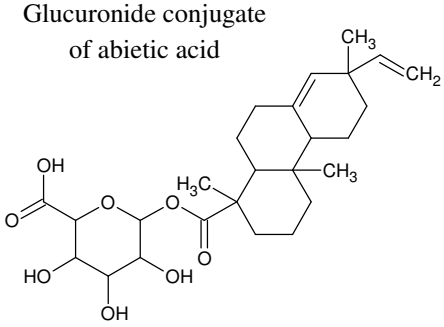
2-Naphthol



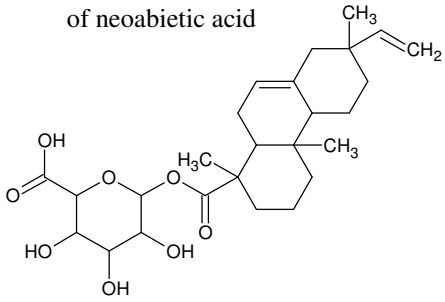
Glucuronide conjugate of abietic acid



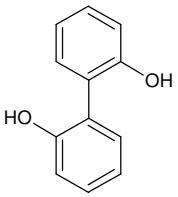
Glucuronide conjugate of neoabietic acid



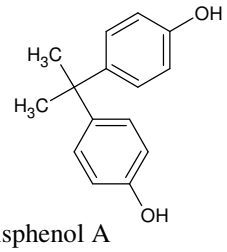
Glucuronide conjugate of pimaric acid



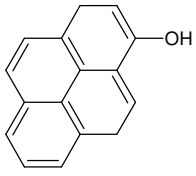
Glucuronide conjugate of isopimaric acid



2,2'-dihydroxybiphenyl



Bisphenol A



1-Hydroxypyrene

Figure 3.28: Chemical structures of non-chlorinated markers identified in trout bile.

Table 3.10: Markers identified in trout bile in both ionization modes (+/-ESI).

Marker No.	ESI mode	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments (plus fragments from GCMS analysis of the derivatized aglycone)[#]	Parent compound formula[§]	Ion form	Putative identity
1	-ESI	319.0815	6.19	$C_{16}H_{15}O_7$	319.0818	-0.9	1.0	175.0246, 113.0244, 143.0493 (216, 201,185,144,115,73)	$C_{10}H_8O$	$[M-H]^-$	Glucuronide conjugates of 1-naphthol and 2-naphthol isomers.
2	+ESI	229.0866 405.1195	7.06	$C_{14}H_{13}O_3$ $C_{20}H_{21}O_9$	229.0865 405.1186	0.4 2.3	1.7 2.2	151.0398, 105.0339 (299,283,225,73)	$C_{14}H_{12}O_3$	$[M+H]^+$	Glucuronide conjugate of oxybenzone (2-hydroxy-4-methoxybenzophenone)
3	+ESI	418.1505	12.39	$C_{21}H_{24}NO_8$	418.1502	0.7	0.1	400.1392, 242.1184, 224.1076, 209.0831	$C_{15}H_{15}NO_2$	$[M+H]^+$	Glucuronide conjugate of mefenamic acid
	-ESI	416.1345	12.39	$C_{21}H_{22}NO_8$	416.1345	0.0	1.3	240.1026, 196.1120,175.0244, 113.0244 (313,298,223,208, 180)	$C_{15}H_{15}NO_2$	$[M-H]^-$	
4	-ESI	477.2492	19.80	$C_{26}H_{37}O_8$	477.2488	0.8	0.5	301.2165, 175.0245 (374, 359,257,241,73)	$C_{20}H_{30}O_2$	$[M-H]^-$	Glucuronide conjugate of a resin acid mixture (abietic acid, neoabietic acid, pimaric acid, isopimaric acid)

+/-ESI: positive/negative electrospray ionization; m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm and additional fragments observed from GC-MS of the aglycone and after derivatization to TMS ether; [§]aglycone formula; $[M+H]^+$: protonated ion; $[M-H]^-$: deprotonated ion.

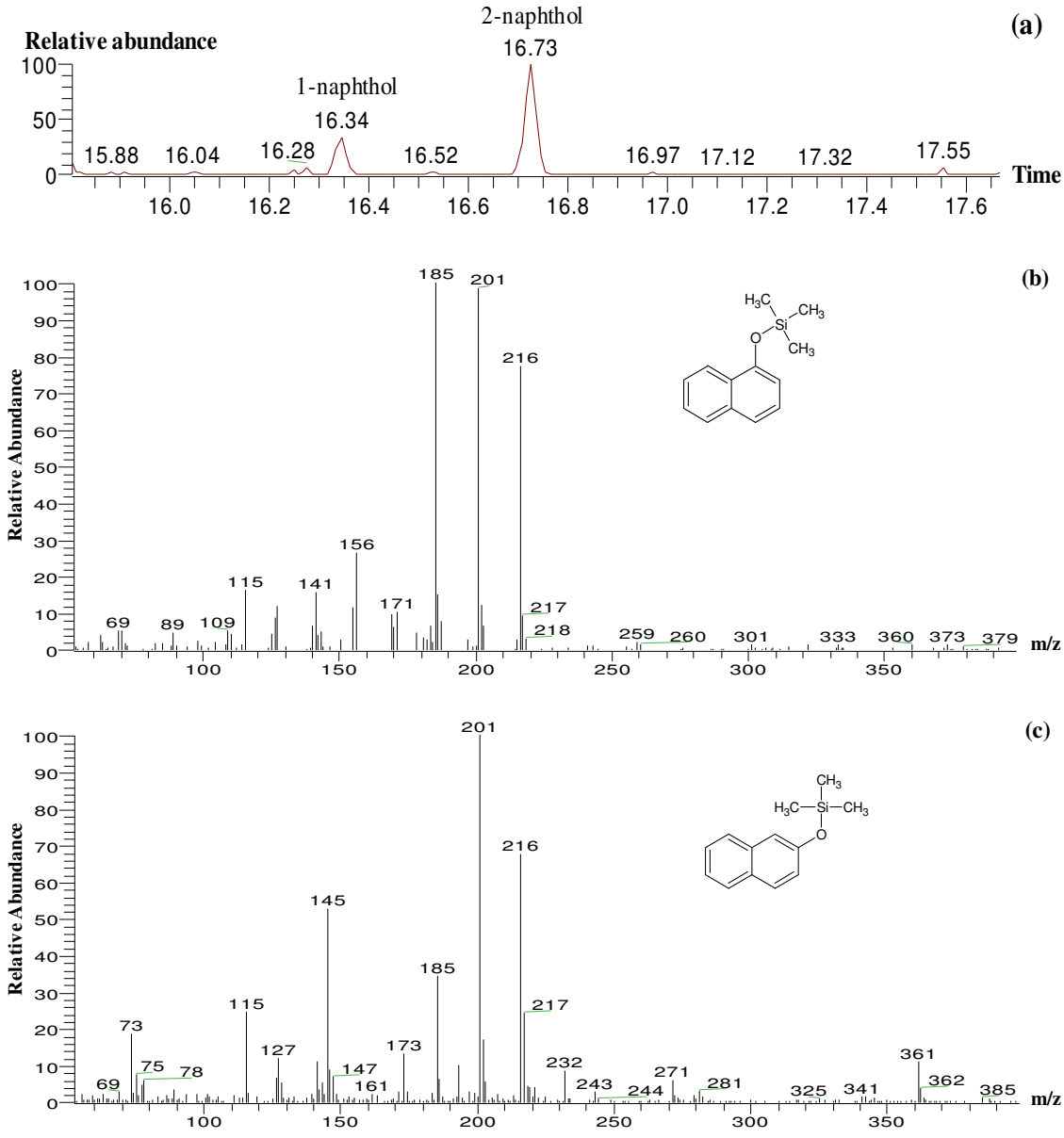


Figure 3.29: (a) GC-MS extracted ion chromatogram of derivatised naphthol isomers; mass spectra of (b) 1-naphthol and (c) 2-naphthol as TMS derivatives identified in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis.

In the GC-MS analyses (but not in the UPLC-TOFMS), another peak only present in bile from effluent-exposed fish and not in control fish was identified as the TMS derivative of 1-hydroxypyrene (Figure 3.28). To confirm the identity of this compound, its retention time and mass spectrum were compared to those of the TMS derivative of a pure standard. The diagnostic ions for the TMS derivative of this compound included the molecular ion at *m/z* 290 (base peak) and the ion at *m/z* 275 due to the loss of CH_3 from the trimethylsilyl group (Figure 3.30a). A further ion at *m/z* 259 could be associated to the cleavage of the trimethylsilyl group followed by cyclization to a five member ring as mentioned above. An abundant peak at *m/z* 189 $[M-101]^+$ was present in the spectrum but not identified.

Further examination of the whole bile sample subjected to enzymatic hydrolysis highlighted the presence of an another compound at retention time 18.9 min (GC-MS analysis); this peak corresponded to a derivatized phenolic compound, 2,2'-dihydroxybiphenyl (Figure 3.28). The compound identity was confirmed comparing the experimental mass spectrum to the mass spectrum of a pure standard. Good agreement was found for the main diagnostic ions: *m/z* 330 (molecular ion), *m/z* 315 (loss of CH_3 from TMS), and an abundant ion at *m/z* 73, indicating that both hydroxyl group were efficiently derivatized (Figure 3.30b).

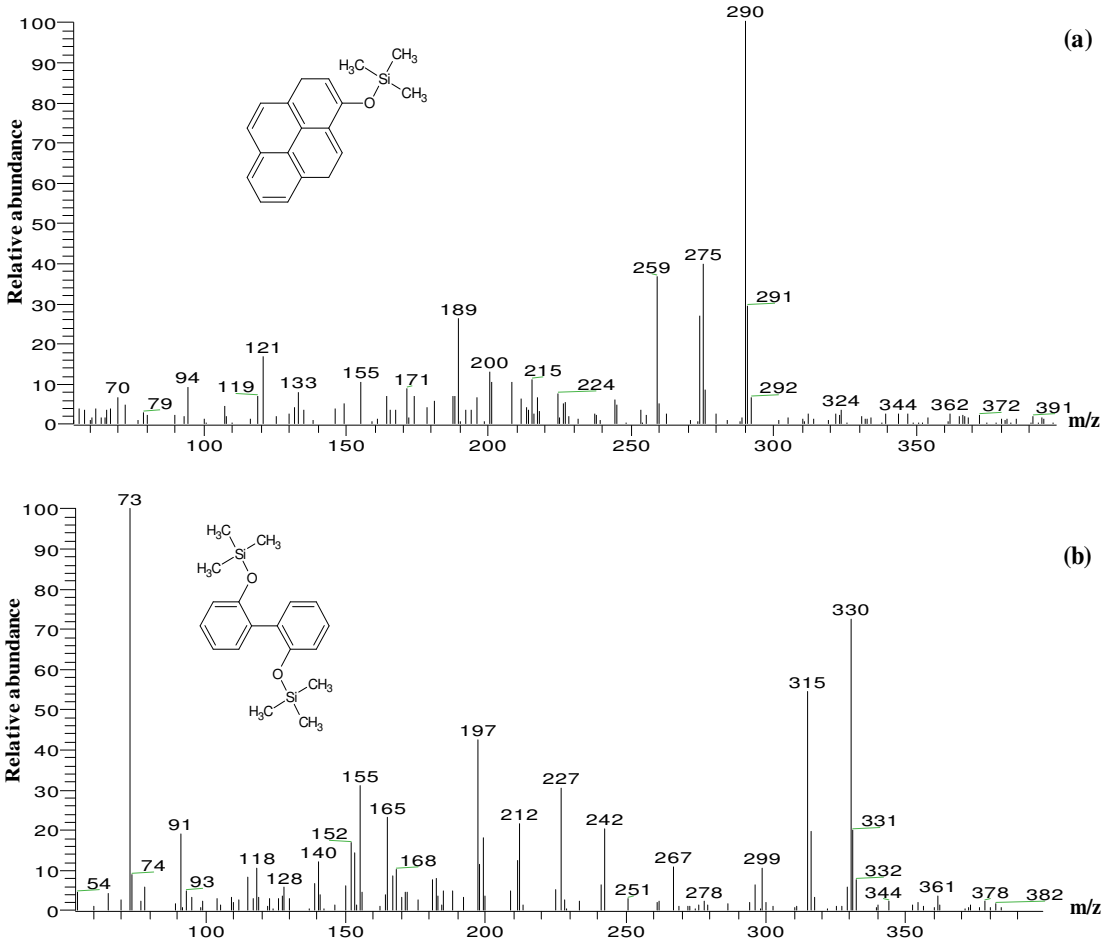


Figure 3.30: Electron impact (EI) mass spectra of (a) 1-hydroxypyrene and (b) 2,2'-dihydroxybiphenyl as TMS derivatives identified in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis.

3.3.4.11 Metabolite of a sunscreen product

Analysis of marker signals from the UPLC-TOFMS profiling revealed that a sunscreen product was another marker metabolite responsible for the class separation in the OPLS model (Figure 3.28). Oxybenzone was detected only in +ESI mode, giving the protonated molecule at *m/z* 229.0863 (aglycone). A careful examination of the obtained spectra led to the detection of a small peak at *m/z* 405.1195, corresponding to the glucuronide conjugate of oxybenzone (Table 3.10, marker 2, Figure 3.31a,b). This result indicates that the glycoside bond between oxybenzone and glucuronic acid is not strong enough to avoid the cleavage of the glycoside bond in the source at the applied collision energy (10 eV). In order to further investigate this molecular structure, Q-TOFMS experiments were performed on the aglycone ion (*m/z* 229.0865), giving fragment ions at *m/z* 151.0395 ($C_8H_7O_3$) and *m/z* 105.0340 (C_7H_5O). The ion at *m/z* 151.0395 was probably generated after cleavage of the phenyl group, whilst the ion at *m/z* 105.0340 could correspond to a bicyclic ion as shown in Figure 3.31c.

The EI MS fragmentation pattern of oxybenzone TMS derivative (Figure 3.32) presented two abundant ions at *m/z* 299 $[M]^{\bullet+}$ and *m/z* 73 $[(CH_3)_3Si]^+$ and two less intense ions at *m/z* 283 ($[M-16]^+$ possibly due to loss of CH_4) and *m/z* 225 ($[M-74]^+$ can be due to neutral loss of the TMS group. Due to high similarity of the RT and MS spectra between the sample and a pure standard, the marker *m/z* 229.0866 was therefore assigned to the structure of the oxybenzone.

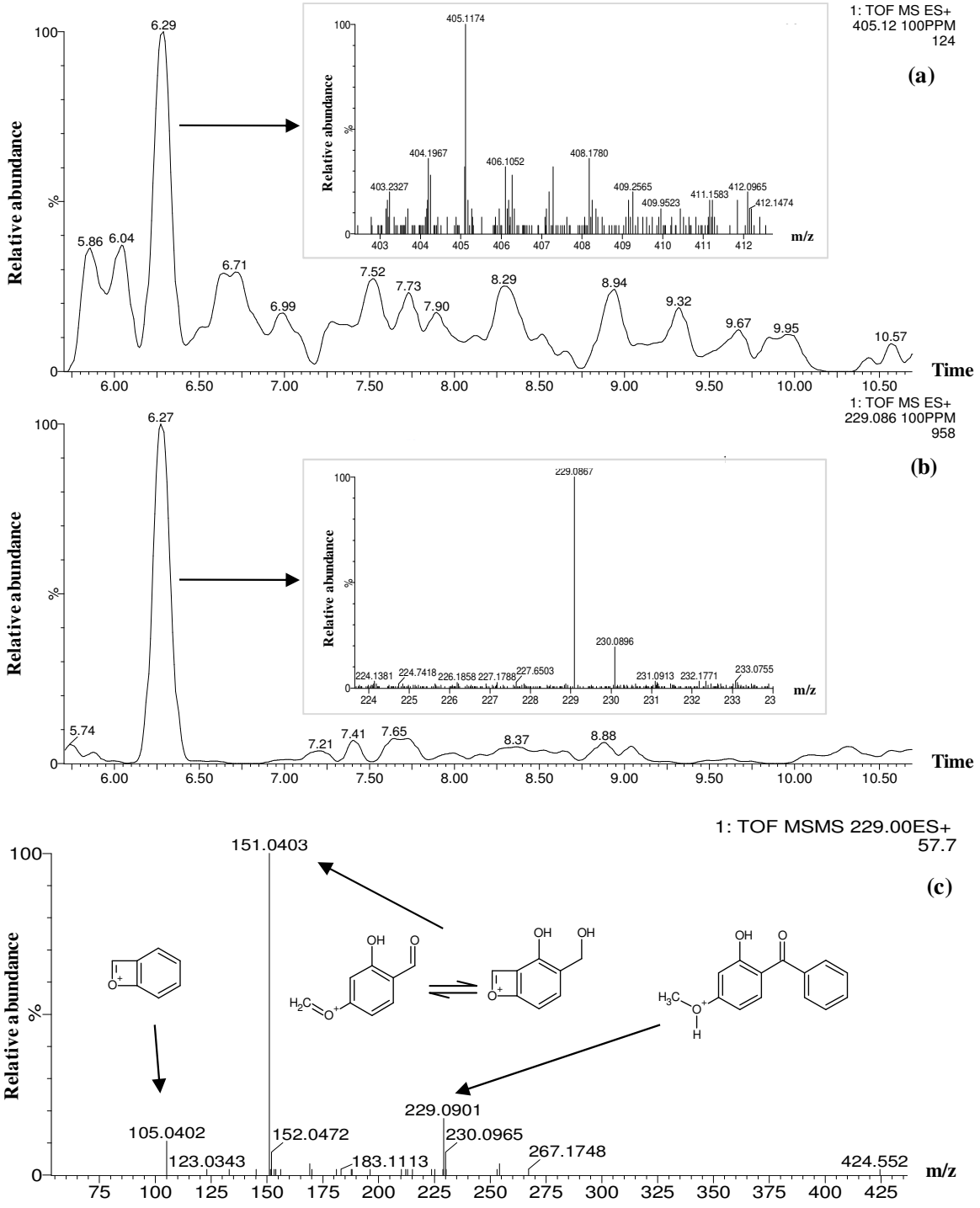


Figure 3.31:
Ion chromatograms and mass spectra of (a) oxybenzone glucuronide and (b) the aglycone of oxybenzone glucuronide identified in bile from effluent-exposed fish analyzed in +ESI mode.
(c) Q-TOFMS mass spectrum of the aglycone of oxybenzone glucuronide obtained with collision energy of 20 eV.

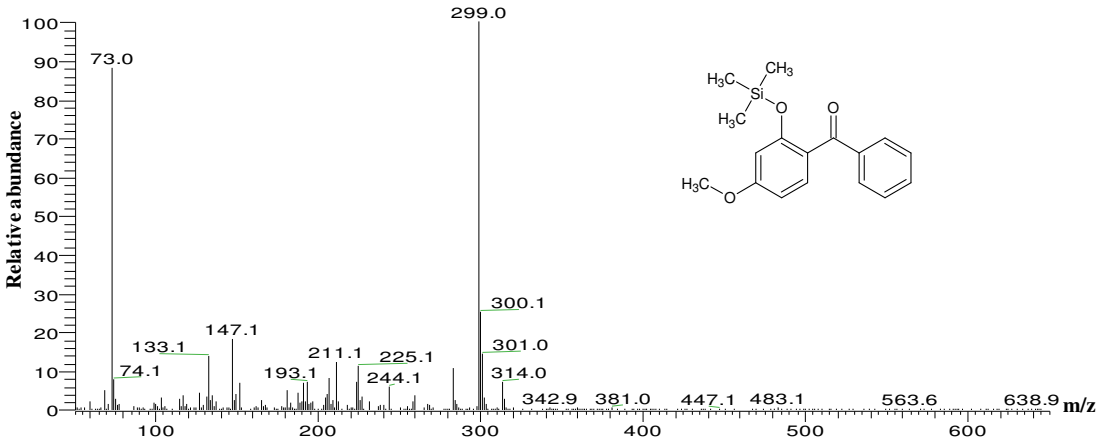


Figure 3.32: EI mass spectra of silylated oxybenzone as the TMS derivative identified in bile from an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

3.3.4.12 Metabolite of anti-inflammatory pharmaceutical

Glucuronide conjugate of mefenamic acid was indicated as potential marker by S-plot analysis in both +/-ESI modes (Figure 3.28). Mefenamic acid is non-steroidal anti-inflammatory drug which reduces inflammation and treats pain in primary care (e.g. back pain, migraine, fever or arthritis). This metabolite was detected as deprotonated ion $[M+Glu-H]^-$ at *m/z* 416.1345 in -ESI, whilst the protonated ion $[M+Glu+H]^+$ at *m/z* 418.1505 was predominant in +ESI (marker 3, Table 3.10).

Two additional RT-*m/z* signals (different mass but same RT) were also detected as potential markers at *m/z* 240.1027 (-ESI) and *m/z* 242.1181 (+ESI), respectively, corresponding to the relative aglycone ions in both +/-ESI modes. These ions were significantly more abundant than their conjugated form at the same experimental conditions (CE: 10 eV). As mentioned previously for the oxybenzone, the aglycone was generated after in-source fragmentation due to the weak linkage to the glucuronide moiety. TOFMS spectrum and high collision energy Q-TOFMS spectrum of mefenamic acid are shown in Figure 3.33a,b. Q-TOFMS fragmentation experiments assisted in the characterization of these two ions. In +ESI Q-TOFMS analysis fragments were observed at *m/z* 400.1392 (loss of H_2O), *m/z* 242.1184 (loss of glucuronide moiety), *m/z* 224.1076 (loss of H_2O from the aglycone), and *m/z* 209.0831 (loss of CH_3) (Figure 3.33b) whilst aglycone at *m/z* 240.1026 and further loss of CO_2 at *m/z* 196.1120 along with glucuronide moiety were obtained from -ESI Q-TOFMS experiment (data not shown).

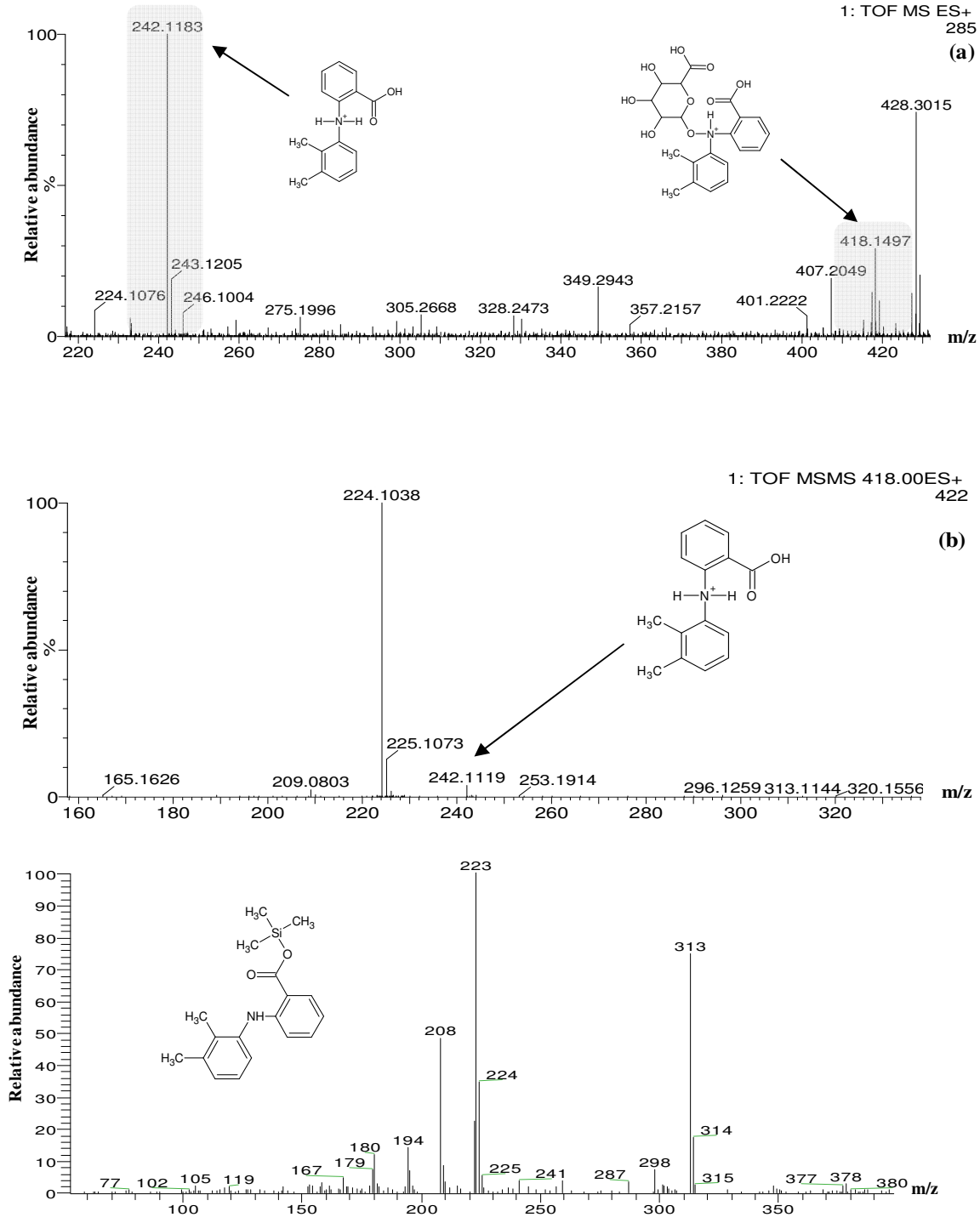


Figure 3.33:
Mass spectra in +ESI mode of mefenamic acid as (a) glucuronide conjugate and (b) aglycone identified in bile from effluent-exposed fish analyzed.
(c) Q-TOFMS mass spectrum of the glucuronide conjugate of mefenamic acid obtained using a collision energy of 20 eV.
(c) EI mass spectrum of mefenamic acid as the TMS derivative identified in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

The GC-MS hydrolysed EI mass spectrum of mefenamic acid obtained from a bile fraction that had been subjected to enzymatic hydrolysis (Figure 3.33c) exhibited ions at *m/z* 313 $[M]^{\bullet+}$ and *m/z* 298 $[M-15]^{+}$. Neutral loss of the trimethylsilylnol group $[(CH_3)_3SiO]$ from the TMS derivative led to the formation of a very stable ion (base peak) stabilized by resonance at *m/z* 223 $[M-90]^{+}$. This fragment showed sequential loss of a methyl group attached to the benzene ring and loss of CO, giving rise to the ions at *m/z* 208 and *m/z* 180, respectively. Mass fragments and retention times were compared to a pure standard and good agreement was obtained between the two set of data.

3.4.10 Metabolites of resin acids

A mixture of different resin acid isomers were found in effluent-exposed fish bile (Figure 3.28). Resin acids (RAs) are tricyclic diterpenes which arise naturally in the tree wood and bark. These acids can be released to wastewater effluent during pulping processes (Lindon et al., 2007). The glucuronide conjugates of different RAs isomers gave a number of non-resolved peaks on UPLC-TOFMS separation and they were only detected in –ESI as glucuronide conjugate (Table 3.10; marker 4, Figure 3.34a). The spectrum showed the molecular ion at *m/z* 477.2492 (its cluster ion at *m/z* 955.5056 $[M+M-H]^{-}$ was also observed). The aglycone ion at *m/z* 301.2165 was generated by Q-TOFMS fragmentation and the typical fragment ion corresponding to the glucuronide moiety was detected at *m/z* 175.0245 (Figure 3.34b).

Representative GC chromatograms and EI MS spectra of the RAs aglycone mixture are given in Figure 3.35. The extracted ion chromatogram showed a total of 8 isomeric RAs peaks; three of them were positively identified by comparison to relative standards as pimaric, isopimaric and abietic acid. The molecular ion $[M]^{\bullet+}$ was observed at *m/z* 374 and loss of CH_3 either from the trimethylsilyl group or from the molecular ion itself generated the fragment ion at *m/z* 359. The fragment ion at *m/z* 256 was probably due to loss of the trimethylsilyl group plus the carboxylic group, followed afterwards by further loss of CH_3 producing a signal at *m/z* 241. The presence of an ion at *m/z* 213 could be linked to the loss of the isopropyl group from the ion at *m/z* 256.

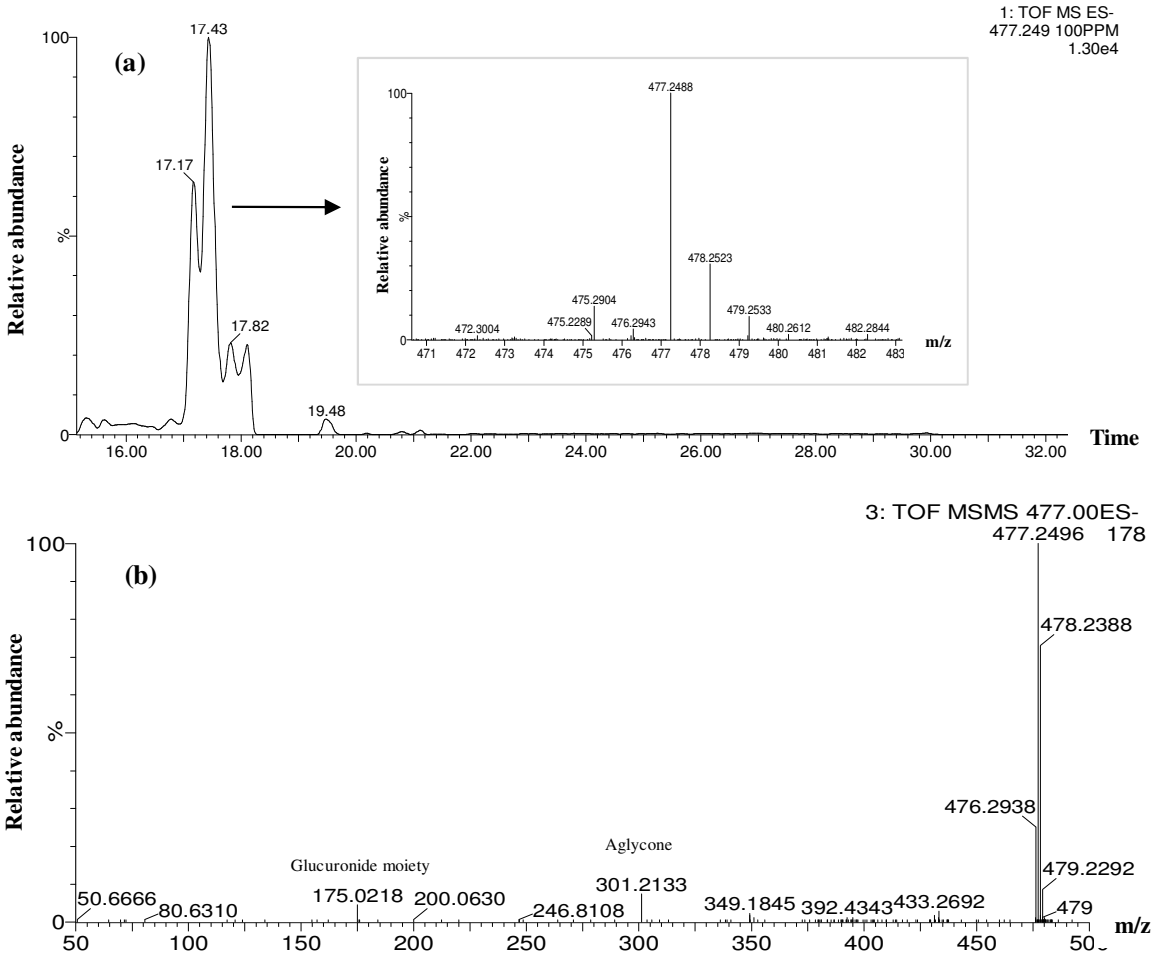


Figure 3.34:
(a) Ion chromatogram and mass spectrum of glucuronide conjugate of mixture of resin acid identified in bile from effluent-exposed fish and analyzed in -ESI mode.
(b) Q-TOFMS mass spectrum of glucuronide conjugate of resin acid obtained with collision energy of 10 eV.

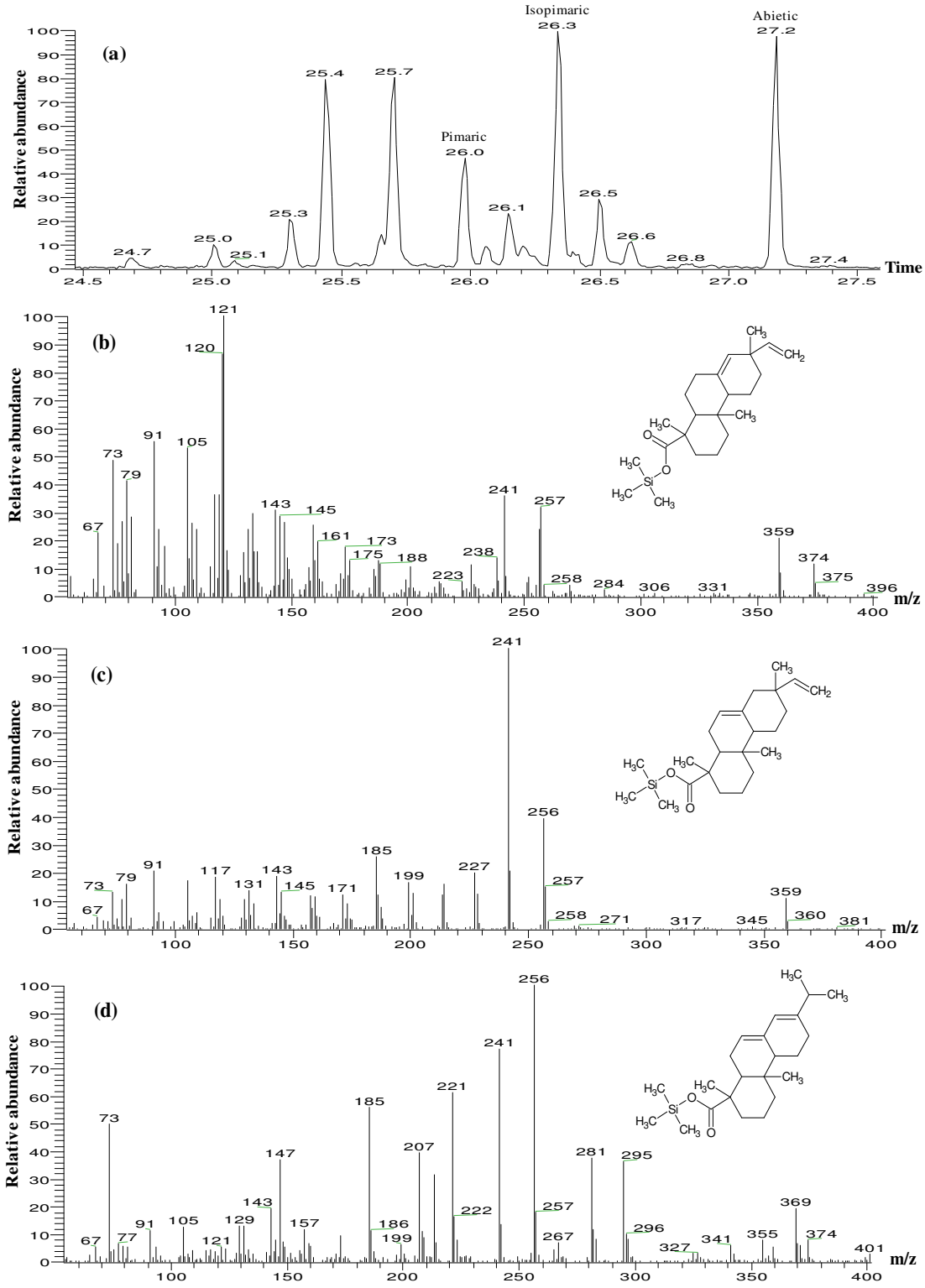


Figure 3.35: (a) GC-MS ion chromatogram of resin acids isomers detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis. Mass spectra of (b) pimaric acid, (c) isopimaric acid and (d) abietic acid as TMS derivatives.

3.3.4.14 Metabolite of bisphenol A

Bisphenol A (BPA) was detected in bile extract of effluent-exposed trout (Figure 3.28). BPAs are a broad-spectrum product which is used as antioxidants in plastics, coatings on cans, additives in thermal paper, powder paints and in dental fillings (Robertson, 2005). BPA was only identified by GC-MS analysis of the whole bile sample subjected to enzymatic hydrolysis and not by UPLC-TOFMS analysis. The di-TMS derivative of the bisphenol A was detected in the bile sample by GC-MS analysis (Figure 3.36). The EI MS spectrum showed the relative molecular ion at *m/z* 372, the base peak at *m/z* 357 due to loss of methyl from the trimethylsilyl group, and the ion at *m/z* 73 corresponding to the trimethylsilyl fragment ion $[(CH_3)_3Si]^+$; RT and fragmentation pattern were confirmed by GC-MS analysis of the relative commercial standard as TMS derivative.

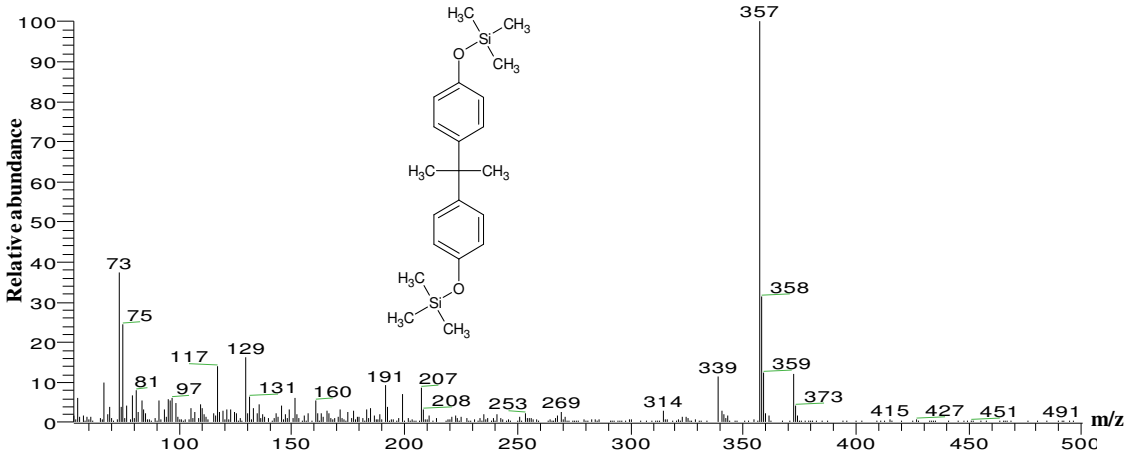


Figure 3.36: Mass spectra of bisphenol A detected in an UPLC fraction of bile from fish exposed to effluent. The bile had been subjected to enzymatic hydrolysis prior to GC-MS analysis.

Table 3.11 summarizes the identified markers detected in exposed fish bile. Appendix 3.12 summarizes the unidentified markers detected in the bile extracts in both +/-ESI modes. Only some of these markers could be putatively identified while the rest could not be characterised due to time limitations and poor spectrum abundance.

Table 3.11: Chemical markers present in bile of trout exposed to a WwTW effluent.

Class of chemical	Chemical identity	Fold change[#] C_{10} vs E_{10}	Relative abundance in E_{10} in either +ESI or –ESI datasets (mean±S.E)	% Decrease during depuration period	
				4 day	11 day
Metabolites of linear alkylbenzene sulfonic acid (LAS) anionic surfactants	C5-SPC	336.8±24.9	3.8±0.3	77.1	100
	C6-SPC	178.1±30.8[¶] (754.6)	8.5±0.2	88.8	98.8
	C7-SPC	57.1±11.7[¶] (1678.4)	19.0±1.7	91.3	99.1
	C8-SPC	717.7±79.6	8.1±0.9	100	100
	C9-SPC	1100.7±119.6	12.4±1.4	98.5	100
	C10-SPC	470.4±31.4	5.3±0.4	100	100
	Monohydroxylated C10-SPC	377.0±9.7	4.3±0.1	100	100
	Putative sulfophenyl dicarboxylate C11-SPdC	116.4±29.2[¶] (254.2)	2.9±0.1	100	100
	Dihydroxylated C10-LAS glucuronide	96.7±24.3[¶] (269.7)	3.0±0.1	97.7	100
Metabolites of nonylphenol polyethoxylate (NPEO) non-ionic surfactants	Nonylphenol glucuronide	20.9±1.6[¶] (4458.0)	50.4±1.1	81.6	96.4
	Nonylphenol-1EO glucuronide	15.4±1.9[¶] (3465.8)	39.2±1.5	67.7	100
	Nonylphenol-2EO glucuronide	1889.5±122.9	21.4±1.4	82.0	99.5
	Nonylphenol-3EO glucuronide	12.8±2.4[¶] (1158.7)	13.1±0.8	92.4	98.5
	Nonylphenol-4EO glucuronide	26.2±12.1[¶] (428.9)	4.8±0.3	100	100
	Nonylphenol-5EO glucuronide	35.9±12.3	0.4±0.1	100	100
	Nonylphenol-6EO glucuronide*	41.6±20.1	3.8±0.4	100	100
Metabolites of alcohol polyethoxylate (AEO) nonionic surfactants	Dodecanol glucuronide	195.1±5.6	2.2±0.1	80.2	100
	Dodecanol-1EO glucuronide	644.5±22.6	7.3±0.3	94.8	100
	Dodecanol-2EO glucuronide	1174.8±61.1	13.3±0.7	94.0	100
	Dodecanol-3EO glucuronide	590.8±35.1	6.7±0.4	100	100
	Dodecanol-4EO glucuronide	191.8±8.7	2.2±0.1	100	100
	Dodecanol-5EO glucuronide	107.0±5.3	1.2±0.6	100	100
	Tridecanol glucuronide	398.2±13.5	4.5±0.2	88.0	100
	Tridecanol-1EO glucuronide	141.3±49.4[¶] (2310.3)	26.1±0.4	73.3	100
	Tridecanol-2EO glucuronide	72.9±16.1[¶] (2734.1)	30.9±1.6	91.3	100
	Tridecanol-3EO glucuronide	2312.9±92.0	26.1±1.0	98.5	100
	Tridecanol-4EO glucuronide	947.6±61.2	10.7±0.7	99.5	100
	Tridecanol-5EO glucuronide	711.3±48.3	8.0±0.5	99.6	100
	Tridecanol-6EO glucuronide	176.0±13.4	2.0±0.2	100	100
	Tridecanol-7EO glucuronide	76.4±7.6	0.9±0.1	100	100
	Tetradecanol glucuronide	155.3±5.9	1.8±0.1	100	100
	Tetradecanol-1EO glucuronide	390.4±15.6	4.4±0.2	80.6	100
	Tetradecanol-2EO glucuronide	445.4±15.3	5.0±0.2	96.9	100
	Tetradecanol-3EO glucuronide	2466.0±186.4	27.9±2.1	91.0	100
	Tetradecanol-4EO glucuronide	95.8±32.1[¶] (2833.8)	32.0±1.8	95.9	98.8
	Tetradecanol-5EO glucuronide	2996.9±178.6	33.9±2.0	97.3	100
	Tetradecanol-6EO glucuronide	883.0±81.9	10.0±0.9	100	100
	Tetradecanol-7EO glucuronide	447.0±50.9	5.1±0.6	100	100
	Tetradecanol-8EO glucuronide*	311.4±26.9	2.0±0.2	100	99.7
	Pentadecanol glucuronide	117.7±15.2	1.3±0.2	100	100
	Pentadecanol-1EO glucuronide	288.4±11.9	3.3±0.1	74.1	100
	Pentadecanol-2EO glucuronide	242.5±52.4	2.7±0.6	97.2	100
	Pentadecanol-3EO glucuronide	198.6±8.0	2.2±0.1	100	100
	Pentadecanol-4EO glucuronide	2174.6±153.1	24.6±1.7	100	100
	Pentadecanol-5EO glucuronide	2623.5±173.0	29.6±2.0	100	100
	Pentadecanol-6EO glucuronide	996.9±91.7	11.3±1.0	100	100
	Pentadecanol-7EO glucuronide	607.2±63.8	6.9±0.7	99.7	100
	Pentadecanol-8EO glucuronide	181.0±18.1	2.0±0.2	100	100

C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; *p* values were based on one tank only; Bonferroni correction threshold used was 1.1×10^{-6} in +ESI and 9.3×10^{-7} in –ESI; *p* values of all markers were< 3.3×10^{-9} with exception of pentadecanol glucuronide was 3.2×10^{-6}.

[¶] Compounds detected in the control bile samples (fish exposed to river water) at levels > LOD, where LOD= 0.013 in –ESI; LOD=0.0064 in +ESI LOD units are relative to the intensities of the normalized total mass spectra signals in the mass chromatograms of the datasets (concentrations were detected in reference trout due to exposure to contaminated control river water).

All markers were from –ESI datasets except those labelled by * were from +ESI datasets.

[#] Fold change values are reported as mean ± standard error (mean±S.E); in brackets are fold change values assuming an LOD value in the control samples was detected.

Table 3.11: (continued) Chemical markers present in bile of trout exposed to a WwTW effluent.

Class of chemical	Chemical identity	Fold change[#] C_{10} vs E_{10}	Relative abundance in E_{10} in either +ESI or –ESI datasets (mean±S.E)	% Decrease during depuration period	
				4 day	11 day
Metabolites of alcohol polyethoxy carboxylates (AECs)	Dodecanol-5EC*	26.7±7.3[¶] (707.4)	4.5±0.1	59.2	94.8
	Tridecanol-4EC*	858.6±63.6	5.5±0.4	26.0	100
	Tridecanol-5EC*	2183.4±113.5	14.0±0.7	58.2	100
	Tridecanol-6EC*	1690.6±85.6	10.8±0.5	80.5	100
	Tridecanol-7EC*	1122.6±46.2	7.2±0.3	100	100
	Tetradecanol-4EC	111.0±6.8	1.3±0.1	11.4	100
	Tetradecanol-5EC	193.1±10.4	2.2±0.1	77.0	98.9
	Tetradecanol-6EC	152.9±83.4[¶] (356.6)	4.0±0.2	86.4	98.6
	Tetradecanol-7EC	390.4±24.3	4.4±0.3	96.3	100
	Tetradecanol-8EC	166.9±10.6	1.9±0.1	100	100
	Pentadecanol-4EC*	398.3±34.6	2.5±0.2	8.0	100
	Pentadecanol-5EC	78.5±8.3	0.9±0.1	100	100
	Pentadecanol-6EC	315.8±14.1	3.6±0.2	79.7	100
	Pentadecanol-7EC	513.1±30.6	5.8±0.3	95.6	100
	Pentadecanol-8EC	244.5±13.6	2.8±0.2	98.7	100
	Pentadecanol-9EC*	621.3±20.0	4.0±0.1	90.2	100
	Pentadecanol-10EC*	338.9±12.5	2.2±0.1	100	100
Metabolites of aromatic hydrocarbons	Glucuronide conjugates of 1-naphthol and 2-naphthol isomers.	189.0±6.6	2.1±0.1	75.3	100
Chlorinated phenols	Glucuronide conjugate of 2,4-dichlorophenol	237.4±10.4	2.7±0.1	88.8	100
	Glucuronide conjugate of an isomer trichlorophenol	232.6±16.8	2.6±0.2	100	100
Chlorinated xylenols	Chloroxylenol glucuronide (4-Chloro-3,5-dimethyl-phenol)	104.3±28.1[¶] (653.4)	7.4±0.3	92.3	100
	Glucuronide conjugate of a methoxy metabolite of chloroxylenol	293.2±12.7	3.3±0.1	93.0	100
	Dichloroxylenol glucuronide	353.4±21.5	4.0±0.2	95.3	100
	Glucuronide conjugate of putative a methoxy metabolite of a dichloroxylenol isomer	341.2±18.1	3.9±0.2	68.2	87.1
	Glucuronide conjugate of a methoxy metabolite of a dichloroxylenol isomer	52.3±2.0	0.6±0.02	71.9	97.7
Chlorophenes	Chlorophene glucuronide (2-Benzyl-4-chlorophenol)	1558.2±346.0	17.6±3.9	96.2	99.8
	Glucuronide conjugate of a methoxy metabolite of chlorophene	206.3±49.5	2.3±0.6	97.1	100
Chlorinated phenoxyphenols	Diclosan glucuronide	121.9±6.0	1.4±0.1	92.6	100
	Triclosan glucuronide	46.4±5.9[¶] (4447.4)	50.3±1.2	85.2	96.0
Pharmaceutical	Glucuronide conjugate of mefenamic acid	121.9±3.2	1.4±0.03	100	100
Sunscreen product	Oxybenzone (2-hydroxy-4-methoxybenzophenone) glucuronide*	325.3±13.5	2.1±0.1	100	100
Resin acids	Glucuronide conjugate of resin acid mixture	7.1±0.7[¶] (1848.6)	20.9±0.7	93.6	97.1

C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; *p* values were based on one tank only; Bonferroni correction threshold used was 1.1×10^{-6} in +ESI and 9.3×10^{-7} in –ESI; *p* values of all markers were< 3.3×10^{-9} with exception of pentadecanol glucuronide was 3.2×10^{-6}.

[¶] Compounds detected in the control bile samples (fish exposed to river water) at levels > LOD, where LOD= 0.013 in –ESI; LOD=0.0064 in +ESI LOD units are relative to the intensities of the normalized total mass spectra signals in the mass chromatograms of the datasets (concentrations were detected in reference trout due to exposure to contaminated control river water).

All markers were from –ESI datasets except those labelled by * were from +ESI datasets.

[#] Fold change values are reported as mean ± standard error (mean±S.E); in brackets are fold change values assuming an LOD value in the control samples was detected.

3.3.5 Elimination of markers of effluent exposure from fish bile after depuration

The elimination rate for each marker metabolite was calculated as the decrease of the relative metabolite concentration in the fish bile during the depuration period (see Table 3.7). The rates of elimination ranged from 8.0 % to 100% after 4 days and from 87% to 100% after 11 days of depuration in river water (Table 3.11).

The relative concentrations of all the LAS metabolites (i.e. SPC metabolites) in the analyzed bile were reduced by between 77-100% in the fish which were transferred to river water (non-contaminated water) for 4 days, and were all reduced by at least 98% after 11 days of depuration. This result indicates that longer periods of depuration allow the fish to excrete all of these metabolites efficiently from the organism. Nonylphenol polyethoxylates (NPEOs) were reduced by between 67-100% after 4 days and 96% after 11 days of depuration. The higher EO oligomers were eliminated more efficiently than NP, NP_1EO and NP_2EO glucuronides. Similarly alcohol polyethoxylates (AEOs) (i.e $C_{12}EOs$, $C_{13}EOs$, and $C_{14}EOs$) were eliminated by at least 73% and 98% after 4 and 11 days, respectively. Elimination of alcohol polyethoxy carboxylates (AECs) ranged from 8% to 100% after 4 days, and was at least 94% after 11 days. For all these surfactant molecules, there was a tendency for more efficient elimination of the longer chain oligomers compared with the short chains after 4 days depuration. However, after 11 days depuration, concentrations of all the remaining surfactant residues in the bile were negligible. For aromatic hydrocarbons (i.e naphthol isomers), > 75% of the original concentrations in the bile were reduced after 4 days and the compounds were not detected after 11 days. Clearance of the chlorinated compound (chlorophenols, chloroxylenols, chlorophens and chlorophenoxyphenols) ranged from 68% to 100% after 4 days of depuration and from 87% to 100% after 11 days of depuration. Mefenamic acid and oxybenzone glucuronide were not detected in bile samples of fish after depuration in river water for 4 days and this indicated that these compounds are unlikely to be persistent in the fish bile. After 4 days of depuration, over 93% of resin acid isomers were eliminated and more than 97% after 11 days in river water.

In summary, none of the identified xenobiotic markers of effluent exposure were highly persistent in the fish bile, and remaining concentrations after a 11 day depuration period were generally less than 2% of the original concentrations present in the effluent-exposed fish.

3.4 Discussion

3.4.1 The analytical approach

The metabolomics approach used in this study has allowed non-targeted chemical profiling of bile fluid in order to analyse the variety of xenobiotics in fish exposed to a wastewater effluent. Mass spectrometry-based approaches are one of the commonly used analytical techniques for profiling of (bio)chemicals present in organisms. In the literature many metabolomics studies have already been focused on the applicability of different hyphenated MS based techniques such as CE-MS (Ramautar et al., 2011) and GC-MS (Joachim, 2006) but LC-MS is usually the technique of choice for this kind of applications (Metz et al., 2008, Theodoridis et al., 2008, Bowen and Northen, 2010). However, the data obtained from the UPLC-TOFMS analysis are too complex to be interpreted by simple visualization of the chromatograms. In fact, metabolomics studies have to deal with a number of variables much higher than the number of available observations (samples) and this could lead to overfitting of any models used to analyse the data. However, if used with care, then bioinformatic and chemometeric tools are essential for the analysis of metabolomics data sets (Trygg et al., 2007, Shulaev, 2006). Principal component analysis (PCA) was initially conducted to overview general clustering, trends, and to spot outlier samples in the bile extracts (i.e. the observations). In order to further identify the diversely expressed metabolites responsible for the separation between the different classes (treatment groups), supervised techniques known as partial least square-discriminant analysis (PLS-DA) and orthogonal partial least squares analysis (OPLS) were applied to the datasets of the bile extracts.

The choice of fish bile as sample matrix was due to several reasons: generally a small sample volume is required, it allows an estimate of the uptake process and it gives a measure of the internal exposure of bioavailable compounds to the fish (Pettersson et al., 2006). Additionally, analysis of bile from fish exposed to contaminated water facilitates the structural identification of bioavailable contaminants present in the surrounding water, since xenobiotics can accumulate in the fish bile resulting in concentrations several orders of magnitude higher than in the water (Gibson et al., 2005a, Fenlon et al., 2010). A wide variety of xenobiotics can be excreted in fish bile including antibiotics, insecticides, dyes, herbicides, polyaromatic hydrocarbons, metals and natural products (Di Giulio and Hinton, 2008).

The main goal of the present study was the identification of chemical markers of effluent exposure. The use of high resolution mass spectrometry can provide high mass accuracy values which allows for the calculation of the elemental composition of the putative structure. However, the elemental composition in not enough to assure a reliable identification of the compound of interest because many molecular structures can account for the same molecular formula. Therefore tandem mass spectrometry (MS/MS) information is essential to obtain extremely useful structural information in order to support the identification process. In the present research, MS/MS fragmentation using Q-TOF as analyzer allowed accurate mass estimation of the obtained fragments; this piece of information can be advantageous to increase the confidence for a correct assignment of the candidate identity. Furthermore, this study has showed how mass spectrometry can significantly help in the analysis of classes having common structural elements (i.e. phase II metabolites; glucuronides).

3.4.2 Characterization of fish bile xenometabolome

Many parts of the UK are heavily populated and river and estuarine fish are exposed to effluent discharges on a daily basis, which implies that there is a chronic exposure to a complex mixture of chemicals. In this study, many different classes of widely used chemicals were detected in the bile of trout exposed to a final wastewater effluent. Six compound classes were fully characterised: surfactants (LAS, SPCs, NPEOs, AEOs, AECs), aromatic hydrocarbons, chlorinated compounds (phenols, xylenols, phenoxyphenols, chlorophenes, parabens), pharmaceuticals, sunscreen agents and resin acids. Most of these compounds were detected using chemical profiling by UPLC-TOFMS followed by multivariate modelling. However, some xenobiotics in the bile (e.g. dichlorophene) were also detected by chance during confirmation of the marker chemical structures using GC-MS analysis of bile fractions. This suggests that many other xenobiotics were present in the fish bile, that either were present in too low concentrations to be detected by UPLC-TOFMS or did not ionize readily and needed other approaches such as GC-MS to detect them.

3.4.2.1 Surfactants

Many different synthetic surfactants were detected in the bile of effluent-exposed fish. NPEO, LAS and AEO series were present together with their main degradation products (e.g. NP from NPEOs, SPCs from LAS and AECs from AEOs, respectively). These compounds are commonly used in household cleaning detergents,

personal care products, paints, textiles, and polymers and are therefore likely to be detected in contaminated water. Based on the data reported in 2007 by CESIO (European Committee of Organic Surfactants and their Intermediates), 1200 kilotons of anionic and 1400 kilotons of non-ionic surfactants were produced in Europe; this amount represents 87% of the total European production of synthetic surfactants (CESIO, 2007).

Nonylphenol ethoxylates (NPEOs) are widely used in industrial, institutional, commercial and household applications and thus are prevalent in influent and effluent of sewage treatment works due to their extensive use (Ying et al., 2002). They are manufactured with ethoxylate chain lengths between 2-25, and the alkylphenol moiety comprises a complex mixture of branched secondary and tertiary nonyl groups (Lee, 1999). Previous studies have showed that short chain NPEOs and NP result from the degradation of long chain polyethoxylates surfactants in WwTWs (Di Corcia et al., 2000). Biodegradation is an important process for the removal of surfactants in raw sewages in wastewater treatment plants reducing their impact on the environment. Degradation of surfactants in the environment occurs primarily through microbial activity. Surfactants can be either used as substrates by microorganisms for energy and nutrients or surfactants can be co-metabolized by microbial metabolic reactions. The degradation behaviour is surfactant class dependent. In the case of NPEOs, the biodegradation of APEOs is thought to start with a shortening of the ethoxylate chain, ultimately leading to short-chain APEOs with one or two ethoxylate units and nonylphenol (NP) itself (Maguire, 1999). In this study, the detection in fish bile of nonylphenol ethoxylates with ethoxymer numbers ranging from 1 to 6 (NP1EO to NP6EO) indicates the presence of poorly degraded residues in the WwTW effluents.

The physiochemical parameters of NP (log Kow=4.48 and water solubility=5.4mg/L) and short chain NPEOs (log Kow =4.17 and water solubility=3.0 mg/L for NP1EO and log Kow =4.20 and water solubility=3.4 mg/L for NP2EO) (Ahel and Giger, 1993) indicate that they are likely to bioaccumulate in aquatic organisms. This behaviour has already been documented in a bioconcentration study of 4-NP in marine animals: Ekelund *et al.* (1990) documented bioconcentration factor for shrimps *Crangon grangon L.* and mussels *Mytilus edulis L.* of 1300 and 3400, respectively. Other studies have revealed bioconcentration factors of NPEOs in bile of dosed roach using [^{14}C] radiolabelled NPEOs, averaging either 3 ethoxy units (NP_3avEO) or 7

ethoxy units (NP_7avEO), and these were 5457 and 3296, respectively (Smith and Hill, 2006). Amongst this class of surfactants, the NP degradation product from NPEOs (NP1EO and NP2EO) have raised particular scientific concern due to their estrogenic activity (Ying and Kookana, 2002). *In vitro* studies, NP is a 1000 times less active than 17-β-estradiol in affinity for the estrogen receptor however it is often present in the environment at 100 times higher concentration compared with steroidal estrogens (Jobling et al., 1996, Routledge and Sumpter, 1996). Alkylphenols can be rapidly metabolized by phase I and II transformations in fish. Arukwe *et al.* (2000) has found that in juvenile salmon 4-n-NP was mainly metabolized to its corresponding glucuronide conjugate, as in the present work, and to a minor extent to diverse hydroxylated and oxidised compounds. In the present study, NPEOs were only detected as the glucuronide conjugate in the bile indicating this was major metabolite in the trout. Although previous studies showed the presence of nonylphenol ethoxycarboxylates (NPEC) as NPEO metabolites, they were not detected in the characterized bile samples of effluent-exposed fish in this study.

LAS surfactants are employed in cleaning products and many studies have documented levels in marine and fresh water ranging from few µg/L to several hundred µg/L (Corada-Fernandez et al., 2011). They are manufactured as C_{10}-C_{13} carbon chain length. A previous study has demonstrated that LAS homologues were taken up by rainbow trout from water via the gills and then bioconcentrate in liver and other internal organs (Tolls et al., 2000). However, in the present study the majority of LAS metabolites were observed as short alkyl chain sulfophenyl carboxylates (SPCs) and only one LAS structure glucuronide conjugate of dihydroxylated C_{10}-LAS was detected. This metabolite could be formed from phase I biotransformation by β-oxidation of C_{10}-LAS and subsequently phase II conjugation. It would appear that in the wastewater treatment works, LAS are mainly transformed to short alkyl chain SPCs (C_5-C_{10}) as these were the predominant LAS metabolites detected in fish bile. Detecting short alkyl chain SPCs in the metabolome can indicate degradation of the long alkyl chain SPCs following ω-oxidation (the first step of phase I biotranformation) of LAS to form SPCs, followed by α- and/or β-oxidation, which leads to odd and/or even carbon chain SPC metabolites (Swisher, 1987, Alvarez-Munoz et al., 2010). Another explanation for detecting both odd and even carbon chain series of SPCs could be the degradation of the

long alkyl chain via β-oxidation starting from two different parent compounds (i.e. C_{10}-LAS & C_{11}-LAS).

Mono- and dicarboxylic C_{10} sulfophenyl acids were also observed in the bile of effluent-exposed trout and could be generated by ω-oxidation of the alkyl chain (terminal carbon) of LAS which followed by successive β-oxidation (Di Corcia et al., 1994, Yadav et al., 2001). A variety of SPCs have been reported with an alkyl chain length of 4 to 13 in sea water, interstitial water, sediments and fungi (González-Mazo et al., 1997, Yadav et al., 2001).

AEOs are manufactured with a carbon length (12-18) and EO (2-25) are used in industrial or household formulations. Studies have shown that the main source of input the environment for these compounds and their metabolites is by wastewater discharges and industrial activities, where they have been detected at concentrations between 1 and 30 μg/L in sewage plant influent and effluent samples (Matthijs et al., 1999, Dunphy et al., 2001, Eadsforth et al., 2006, Belanger et al., 2006). Eadsforth *et al*. (2006) reported similar results in a study monitoring AEOs in wastewater effluent in Europe and Canada, revealing overall mean concentration of 5.7 μg/L (range 1.0–22.7 μg/L). In a previous study, AEOs homologues with shorter alkyl chains (C_{12}) revealed higher relative percentages in water when compared with their longer chain homologues (C_{18}), which exhibited higher affinity for suspended solids and sediments. For AEO ethoxymers up to 9EOs were the most abundant in the water column, whilst the 11EOs was the most relevant in the sediment (Lara-Martín et al., 2008). The distribution of alkyl ethoxylates in the fish bile was very similar to that of some commercial mixture and comprised glucuronide conjugates of C12-C15 EO(1-10). The presence of up to 8 ethoxymers of AEOs indicates the presence of undegraded residues in the WwTW effluents. A number of different degradation pathway are possible for the AEOs but the aerobic biodegradation mechanism that is the most common, and comprises central cleavage of the molecule to form polyethylene glycols and free fatty alcohols (Ying, 2006). The three main routes of chain shortening for AEOs are shown in Figure 3.37.

In addition to AEOs, alkyl polyethoxy carboxylates (AEC*n* where A= alkyl groups and *n*= number of ethoxy units plus a terminal CH_2COOH moiety) were detected in trout bile. AECs could be formed from ω-oxidation of the ethoxylate group of the parent AEO during wastewater treatment (Di Corcia et al., 1998), however AECs are also used as anionic surfactants in the textile industry and in household and personal

care products and thus maybe an additional source of detergent-derived contaminants in WwTWs effluents.

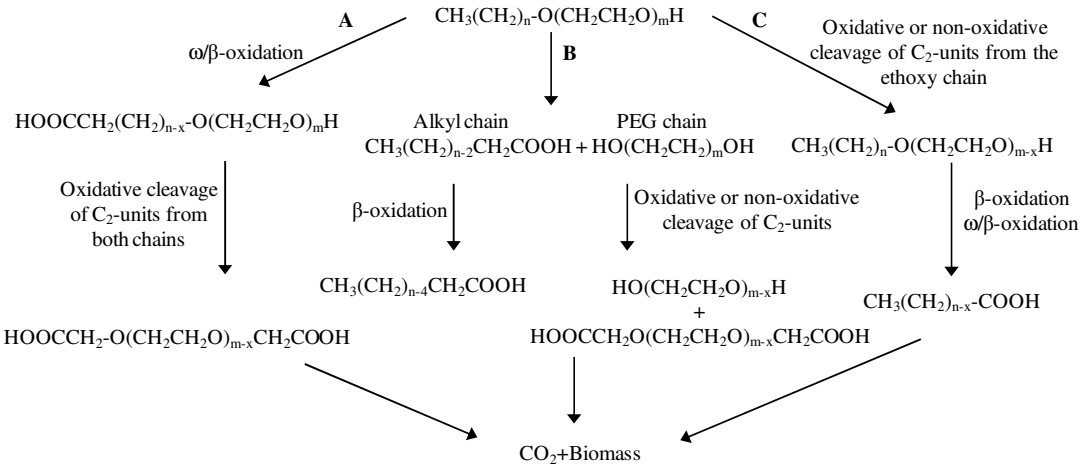


Figure 3.37: Three degradation routes have been described in the literature for AEOs: (a) degradation initiated on the alkyl chain with ω/β-oxidation followed by nonoxidative cleavage of C_2-units; (b) central fission into an alkyl chain and a polyethoxylate chain followed by individual degradation of the two moieties; (c) degradation is initiated at the polyethoxylate chain. Predominately the polyethoxylate is mineralised by oxidative cleavage of C_2-units, simultaneously the alkyl part is shortened by ω/β-oxidation. This latter route is most pronounced for branched secondary AEOs. Under anaerobic conditions only route (b) and (c) take place. Adapted from Krogh et al. (2003).

3.4.2.2 Aromatic compounds

Polycyclic aromatic hydrocarbons (PAH) have always raised wide scientific concern due to their mutagenic, carcinogenic and teratogenic characteristics. PAHs are the most wide-spread organic pollutants. In addition to their presence in fossil fuels, they are also generated by incomplete combustion of fuels (i.e. wood, coal, tobacco, or incense). In this study only two different PAHs metabolites were detected in the bile of effluent-exposed trout: 1-hydroxypyrene and the glucuronide conjugates of naphthols. Previous studies have already revealed the presence of 1-hydroxypyrene in bile of effluent-exposed fish e.g. (Mazéas and Budzinski, 2005). 1-Naphthol has been detected as glucuronide conjugate in rat, where most of the 1-naphthol glucuronide was excreted into the renal vein (Narukawa et al., 2004). Zang *et al.* (2010) demonstrated that naphthalene derivatives substituted at position 2 can be more toxic than those at position 1 while Viravaidya *et al.* (2004) reported that metabolites of naphthalene can decrease glutathione levels in lung cells. For this class of compounds, only the two isomers of

naphthol (1-naphthol and 2-naphthol) were detected by UPLC-TOFMS as glucuronide, whilst 1-hydroxypyrene was detected only by GC-MS analysis.

3.4.2.3 Chlorinated compounds

A number of chlorinated compounds were detected as markers of effluent exposure in this study. The characterized markers belonged to five main classes of compounds: chlorinated phenols, chlorinated xylenols, chlorophenes, chlorinated phenoxyphenols, chlorinated parabens.

Chlorinated phenols

In the analyzed bile samples, dichlorophenol and trichlorophenol were both detected both as glucuronide conjugates. The presence of chlorinated phenols in surface waters could be due to their formation from other contaminants. For instance studies have revealed that 2,4-dichlorophenol and 2,4,6-trichlorophenol could be present in the water as by-products of triclosan degradation in chlorinated water (K.L. Rule, 2005, P. Canosa, 2005). The presence of 2,4-dichlorophenol has recently been related to the fragmentation of triclosan after cleavage of the ether bond (Thomas and Kotchevar, 2010). However trichorophenol and dichorophenol are also used as bactericides, fungicides and preservatives. Concentrations of 2,4-dichlorophenol and 2,4,6-trichlorophenol have been reported to be ranged from <1.1 ng/L to 28650.0 ng/L in some Chinese surface waters (Gao et al., 2008). Chlorophenol compounds are toxic to aquatic species (Buikema Jr et al., 1979) and dichlorophenol is an uncoupler of mitochondrial respiration (Penttinen, 1995).

Chlorinated xylenols

Chloroxylenol is widely used in pharmaceuticals and cosmetic products as disinfectants and have been detected in bile of river fish in the Netherlands and in WwTW effluents (Houtman et al., 2004, Rostkowski et al., 2011). In this study, two chloroxylenols (i.e. chloroxylenol and dichloroxylenol) and their methoxy metabolites were detected as glucuronide conjugates in bile of effluent-exposed fish. The presence of methoxy metabolites can be explained as part of the detoxification process for lipophilic compounds. This process consists of an activation step (phase I) which involves the oxidation or hydroxylation of the toxic compound, followed by the phase II step (conjugation; methylation) via covalent binding to endogenous hydrophilic molecules.

Chlorophenes

Chlorophene is commonly used as germicide in disinfectant products and has been detected in surface water and WwTW effluents as reported in previous studies (Kasprzyk-Hordern et al., 2008, Martínez Bueno et al., 2007). Bile extracts of effluent-exposed fish showed the presence of two types of chlorophenes, chlorophene and dichlorophene. Chlorophene and its methoxy metabolites were detected as glucuronide conjugates by UPLC-TOFMS while the dichlorophene could be detected only by GC-MS in hydrolysed bile fractions. Previous studies have reported that dichlorophene can be accumulated in fish exposed to WwTW effluents and this compound was also detected in wastewater effluents too, a fact which indicates that its presence is mainly due to its widespread use as a bactericide and fungicide in a variety personal care product formulations (Hill et al., 2010, Rostkowski et al., 2011). Chlorophene is an androgen receptor antagonist (antiandrogenic) *in vitro* assays with a potency 13 fold higher than the standard antiandrogen flutamide. Alongside triclosan (see below) chorophene contributes 50% of the antiandrogen activity in bile of fish exposed to WwTW effluents (Rostkowski et al., 2011).

Chlorinated phenoxyphenols

Triclosan is widely used in cosmetic formulation as antibacterial and has been detected in wastewater treatment work effluents at concentrations between 50-100 ng/L (Rostkowski et al., 2011). Adolfsson-Erici *et al.* (2002) measured triclosan levels in rainbow trout (*Oncorhynchus mykiss*) exposed to WwTW in Sweden. Analysing fish bile revealed concentrations of 0.44–120 mg/kg in trout exposed to sewage water. Houtman *et al.* (2004) also identified triclosan in the bile of male bream (*Abramis brama*) in Dutch surface waters. Triclosan was detected at relatively high concentrations of 14 and 80 μg/ml of bile. In the present research the molecular ion $[M-H]^-$ associated with triclosan glucuronide was highly detected abundant in effluent-exposed fish bile indicating that it was a prominent marker of effluent exposure. The antiandrogenic activity of triclosan is 5 fold greater than flutamide indicating that it is a potent antiandrogen in *vitro (Rostkowski et al., 2011).*

A triclosan analogue (diclosan) was also detected in bile as glucuronide conjugate. Diclosan concentration has already been determined in treated effluent and activated sludge. Diclosan in the treated effluent was found 0.3, and in the activated sludge, <0.3.

Aquatic toxicity for this metabolite has been reported as well, showing LC50 value of 0.7 mg/L in fish (NICNAS, 2004).

Chloroxylenol, chlorophene and triclosan are broadly detected in surface waters and effluents. Previous study in the South Wales region of the UK has documented concentrations of 30-358ng/L for chloroxylenol, 3-16ng/L for chlorophene and 5-24ng/L for triclosan (Kasprzyk-Hordern et al., 2008). Previous metabolomics studies of chloroxylenol in dogs, and chlorophene and triclosan in rats have shown that the majority of these chlorinated antiseptics are excreted as glucuronide and sulphate conjugates in urine of higher vertebrates (rat) (Tulp et al., 1979, Kao and Birnbaum, 1986, Ridley et al., 1986, Dorantes and Stavchansky, 1992). This information indicate that these contaminants maybe hydrolysed to their non-conjuated forms in WwTW, however there are no studies to date to confirm this.

Chlorinated parabens

Dichloromethylparaben was also detected in effluent-exposed fish bile. Monochloro- and dichloro-methyl paraben (ClMeP and Cl_2MeP) have been documented in wastewaters at levels between 0.01 and 0.1 mg/L (González-Mariño et al., 2011). Parabens can react with free chlorine dissolved in the water after disinfection treatment giving rise to several halogenated by-products. Chlorinated parabens can be formed through chlorine substitution of the phenolic group by electrophiles such as HOCl and ClO^-, as a consequence of the high negative charge on the nucleophilic substrate (Ge et al., 2008, Terasaki and Makino, 2008).

The parabens chlorinated by-products have shown higher acute toxicity responses in the *Daphnia magna* (Terasaki et al., 2009) However, the occurrence of these derivatives in the environment has not been fully investigated yet and still little is known about the biodegradability of parabens and their by-products during wastewater treatments.

3.4.2.4 Pharmaceuticals

The majority of pharmaceutical compounds are introduced in the aquatic systems after ingestion and subsequent excretion by either humans or animals (veterinary pharmaceuticals). Mefenamic acid is a non-steroidal anti-inflammatory drug classified as “anthropogenic” pharmaceuticals (Werner et al., 2005). Glucuronide conjugate of mefenamic acid was identified and isolated from human urine (McGurk et al., 1996). In fish bile it was detected as the glucuronide form. Mefenamic acid has been

detected in wastewater influent, effluent and sewage sludge during wastewater treatment, revealing concentrations of 16 ng/L in influent, 5 ng/L in effluent and 0.4 ng/g in sludge, respectively (Jelic et al., 2011). This pharmaceutical has specifically been designed to inhibit prostaglandin biosynthesis; therefore it is tempting to speculate that this represents their primary mechanism of action also in the fish (Lemke et al., 2007).

3.4.2.5 Sunscreen agents

Oxybenzone is a commercial sunscreen product and is widely used in cosmetics and plastic applications. *In vitro* studies have shown that oxybenzone has estrogenic activity, as well as antiestrogenic and antiandrogenic (Kunz and Fent, 2006). Other studies have shown that oxybenzone can induce vitellogenin production in two fish species (i.e. juvenile rainbow trout and Japanese medaka) at similar concentrations and also significantly decreases fertilized eggs hatchability in Japanese medaka (Coronado et al., 2008). In trout bile sample this compound was detected as the glucuronide conjugate.

3.4.2.6 Resin acids

A mixture of resin acids (RAs) was detected as glucuronide conjugates. The group of isomers included abietic acid, pimaric acid, and isopimaric acid). RAs have already been detected in paper mill effluents at concentrations ranging from 20µg/L to 12000µg/L (Quinn et al., 2003). Although these isomers are mainly found in pulp and paper mill effluents they are also used as varnishes and can be therefore transferred to waste water. Different resins are used for specialty varnishes, for metal, paper, leather and to protect oil and watercolor paintings (Langenheim, 2003). The presence of resin acids in their glucuronide forms have been already reported in bile of fish exposed to pulp mill effluent (Meriläinen and Oikari, 2008). The RAs (abietic acid, pimaric acid, isopimaric acid, neoabietic acid) have been proved to have antiandrogenic activity with a relative potency when compared to the flutamide standard of 4.00 for the abietic acid, 2.73 for the pimaric acid and 5.00 for the isopimaric acid (Rostkowski et al., 2011).

3.4.2.7 Other compounds detected in effluent exposed fish bile by GC-MS analysis

Bisphenol A and 2,2'-dihydroxybiphenyl were only detected by GC-MS analysis in effluent-exposed fish bile. Bisphenol A is used in the manufacture of polycarbonate plastics, and has previously been detected in fish bile (Fenlon et al.,

2010). In UK concentration of bisphenol A in crude and treated effluents were 1.2µg/L and 0.046µg/L, respectively (Jiang et al., 2005, Ifelebuegu, 2011). Bisphenol A has shown endocrine disrupting activity in both *in vivo* and *in vitro* experiments (Welshons et al., 2003, Vom Saal and Hughes, 2005, Richter et al., 2007).

3.4.3 Toxicity implications

These analyses of the more highly abundant xenobiotics present in bile reveals that fish are exposed to a diverse group of chemical contaminants present in wastewaters. Most of the xenobiotics detected in the bile were present as glucuronide conjugates and these polar metabolites would be rapidly eliminated from the fish via the faeces. Glucuronide conjugates are mainly excreted via the bile of fish, and depuration studies revealed that with exception of some metabolite, >75% of the xenobiotic load was eliminated from the bile after 4 days in clean water and >99% after 11 days. Although some of this decrease is due to elimination via the bile there could also be renal excretion of the parent compound or other metabolites present in the blood. Gills and skin can also contribute to the elimination of the parent compound or its metabolite via diffusive elimination (Di Giulio and Hinton, 2008). Many of the glucuronide metabolites of xenobiotics are likely to be formed by conjugation of the parent compound within the fish itself as glucuronide conjugates formed by human metabolism (for instance the glucuronide conjugate of mefenamic acid) can be readily hydrolysed by bacteria during the WwTWs resulting in the detection of the parent compound in the final effluent (Gomes et al., 2009). Although the majority of xenobiotics were not persistent in the fish tissues, it is possible that exposure to such as variety of contaminants that are substrates for glucuronosyl transferase enzymes expressed in liver and other tissues of the fish could disrupt the metabolism and excretion of endogenous metabolites such as steroids and bile pigments that are also conjugated prior to excretion.

Although the complex mixtures of xenobiotics in the fish are not persistent, it is likely that in many UK catchments, some fish are semicontinuously exposed to these mixtures throughout their life cycle as contaminants from WwTW effluents can be detected many km downstream of the discharges (Chapter 1 Section 1.6). It is not clear what the overall health implications of repeated exposure to such as complex mixture of contaminants would be. The ethoxymer content of many of the surfactants was high enough to give detergent properties to the molecules, and so could disrupt the structure

and function of cell membranes in a variety of tissues or organs including the liver. Many of the chlorinated contaminants and the RAs have been shown to be androgen receptor antagonists (at least *in vitro*) and may disrupt sexual differentiation and gonad development in exposed fish. Likewise oxybenzone, bisphenol A and NP are estrogenic and could contribute the endocrine disrupting activity of the contaminant mixture in the fish. Mefenamic acid is a potent inhibitor of prostaglandin syntheses, and these metabolites play important roles in immune function, and reproduction. Therefore exposure to the complex mixture of xenobiotics present in WwTW effluent may result in disruption of function of a number of tissues and physiologies within the fish, and further studies are needed to determine the long term effects of exposure to these contaminants present in WwTW effluent on fish health. Other studies which could be conducted include an examination of fish exposure to a number of wastewater treatment effluents across the UK for varying lengths of time, to investigate how the xenometabolome and fish metabolite profile vary with effluent composition and to examine any assications between contaminant profiles and fish health.

CHAPTER 4: Analysis of Markers of Effluent Exposure in the Xenometabolome of Trout

4.1 Introduction

Biofluids such as urine and blood have been frequently used as biological matrices to investigate the systematic alteration of the metabolome (Fancy et al., 2006, Yu et al., 2007) and to evaluate the effects of xenobiotic exposure, as they are easily sampled and simple to analyze (Aresta et al., 2006, Lutz et al., 2008). For instance, Brown *et al.* (Brown et al., 2007) have used fish plasma as a biofluid model to assess the capability of pharmaceuticals to bioconcentrate into fish blood. The choice of the biofluid to be investigated influences the kind of contaminants which can be monitored in the selected matrix. Bile usually shows high concentration of glucuronic acid and other conjugated forms of lipophilic compounds whilst blood and urine usually are mainly characterized by the presence of either deconjugated forms or sulphate conjugates of the metabolites of interest. Therefore, in order to have a complete of the metabolome of a vertebrate and to encompass all the possible contaminants taken up by the organism, a combination of different biofluid matrices is required.

In chapter 3 the study was focused on the characterization of the metabolome in fish bile, identifying mainly lipophilic compounds in their glucuronide forms. In this chapter the attention will focus on the characterization of plasma as different matrix to check whether the same classes of compounds could have been detected and to investigate the presence of other possible chemical species which were not detected in the bile sample.

The blood was selected as biofluid to be investigated mainly because as reported in the literature, sulphate conjugates are more likely to be detected in this kind of matrix (Mulder and Scholtens, 1978). Sulphation of compounds predominates at low substrate concentrations and glucuronidation at high substrate concentrations when sulphation has been saturated (Pang et al., 1994). Thus, the glucuronidation of circulating xenobiotics is likely after saturation of the sulphotransferase enzymes. The analysis of blood for markers of effluent exposure is also advantageous as it is readily sampled without necessitating sacrifice of the animal. Therefore methodology to analyse markers of effluent exposure in small volumes of plasma samples could be useful for biomonitoring purposes.

The study in this chapter aimed to:

1) Investigate the presence of organic contaminants and their metabolites (the xenometabolome) in the plasma that are markers of effluent exposure in trout.
2) Assess the effect of depuration on persistence of markers of effluent exposure in the plasma.
3) Analyse the blood metabolome to investigate how the blood biochemistry is disrupted by exposure to wastewater effluent.
4) Determine the structures of the endogenous and exogenous metabolites and confirm the structures of contaminants using different mass spectrometry techniques.

Proteins in plasma samples were precipitated by methanol extraction method to a final ratio of 80:20 methanol/water (v/v). Plasma samples were profiled applying same methodology as for bile samples. Multivariate data analyses were performed using PCA, PLS-DA and OPLS-DA techniques. Accurate mass measurements were obtained in W mode for the selected discriminatory makers due to effluent exposure. Q-TOFMS experiments were then performed to obtain fragment information for structural identification.

4.2 Materials and Methods

4.2.1 Chemicals

Taurochenodeoxycholic acid (sodium salt), sphingosine, O-acetyl-L-carnitine hydrochloride, L-α-phosphatidylcholine from egg yolk (≥99%), L-α-phosphatidylethanolamine from egg yolk (~98%), and ammonium formate (99%) were purchased from Sigma Aldrich, UK.

4.2.2 Extraction of plasma samples

Blood samples were obtained by terminally anesthetizing the fish and sampling blood from the caudal vein at the end of the experimental exposure described in details in Chapter 3, section 3.2.2. Blood samples were transferred to Eppendorf tubes containing aprotinin and tubes were centrifuged at 10,000 g for 15 min at 4°C (Biofuge fresco, Heraeus, Germany). Plasma was then transferred to methanol rinsed glass vials and stored at -80°C for chemical profiling. In this study plasma extraction was performed adapting the protocol reported by Bruce *et al.* (Bruce et al., 2008). Frozen plasma samples (300µL) were thawed on ice. Proteins were precipitated by adding 1200µL of cold methanol in order to obtain a final ratio of 80:20 methanol/water (v/v)

(assuming that plasma is 100% aqueous). The mixture was vortexed for 1 min and then left for 1 h at -20°C. Afterwards, samples were centrifuged at 10,000 g for 5 min and the separated supernatant was kept at -20°C for further 30 min before further centrifugation at 10,000g for 5 min in order to remove any unprecipitated protein. The supernatant was evaporated to dryness using a SpeedVac concentrator (Savant Instruments, Inc., Holbrook, New York), and the residue was then reconstituted in 300µL methanol:water (1:1, v/v). Plasma extracts were filtered using a 96-well Strata Protein Precipitation Plates system (0.2µM, Phenomenex, Cheshire, UK) prior to UPLC-TOFMS analysis.

4.2.3 Chemical profiling of fish plasma samples

4.2.3.1 UPLC-TOFMS analysis

20µL aliquots of extracted plasma sample were analysed according to the experimental method described in Chapter 2 Section 2.3.2. Samples were injected in random order to avoid any instrument bias.

4.2.3.2 Data elaboration

MarkerLynx V 4.1 software package (Waters Corporation, Milford, MA, USA) was used to align and deconvolute spectral peaks. Required parameters for sample alignment were optimized as follows: minimum peak width 20s, minimum required intensity 50 counts, minimum signal to noise 10, maximum number of peaks 20, retention time tolerance 0.2 min, and mass window tolerance 0.03. Before exporting data to SIMCA-P software (Umetrics UK Ltd, Winkfield, Windsor Berkshire, UK) for subsequent multivariate analysis isotopic peaks were eliminated and datasets were normalised to the maximum spectral area observed in the sample set.

4.2.3.3 Multivariate data analysis

As described in Chapter 3 Section 3.2.3.3, data were exported into the multivariate analysis software SIMCA-P. Data were log transformed before analysis to reduce skewness and then pareto scaled. In order to obtain an overview of the whole datasets and to detect any outliers within the dataset, principle component analysis (PCA) was employed. Strong outliers were identified using Hotelling's T^2 which were then excluded for further analysis. The data were examined by partial least square-discriminate analysis (PLS-DA) followed by orthogonal projections to latent structures discriminant analysis (OPLS-DA). OPLS-DA approach gives an overview of the

existing relationships between variables (extracted metabolites) and observations (samples) as well as among the variables themselves and highlights deviating behaviour within the observations showing groups/trends. OPLS-DA is particularly useful in separating two classes of observations (such as samples from control and effluent exposed fish) along the first principal component and modelling other influences of the data on other orthogonal components (Wagner et al., 2007). The performance of the described models can be evaluated considering the R^2 and Q^2 values (high R^2 and Q^2 values are desirable). Cross-validation procedure was used to determine the number of significant principal components. Loadings plots of an OPLS-DA were analyzed to identify significant metabolites with the highest confidence and greatest contribution to the group separation.

4.2.3.4 Identification of metabolites

Metabolites identified as being altered in response (intensity) after exposure to effluent water were further analyzed using elemental composition tools of MassLynx 4.1 software in order to determine their elemental composition. Their structural identity was then verified performing Q-TOFMS approaches using different collision energies ranging from 12 eV to 50 eV. Collision energy was properly optimized for each considered marker. Ammonium formate was used as ion-pairing reagent to improve the fragmentation experiments replacing sodium with ammonium in the adduct molecule since ammonium adducts are usually less stable than the relative sodium adducts and therefore more likely to fragment. For this purpose, UPLC separation was then performed adding ammonium formate buffer (10 mM) adjusted to pH 3.0 by formic acid to solvent A and B. However, wherever possible deprotonated $[M-H]^-$ (negatively-charged compounds) or protonated $[M+H]^+$ (positively charged compounds) markers were mainly selected as product ions to obtain structural information after fragmentation. Commercial standards were used for the identification of the unconjugated markers by comparing their retention times, masses, and characteristic fragments with the target analytes. Sulfadimethoxine standard was added as a lock mass to compensate for the mass drift during the sample analysis. Metabolomic databases were also utilized in order to assist the metabolite identification as described in Chapter 3 Section 3.2.3.4.

4.3 Results

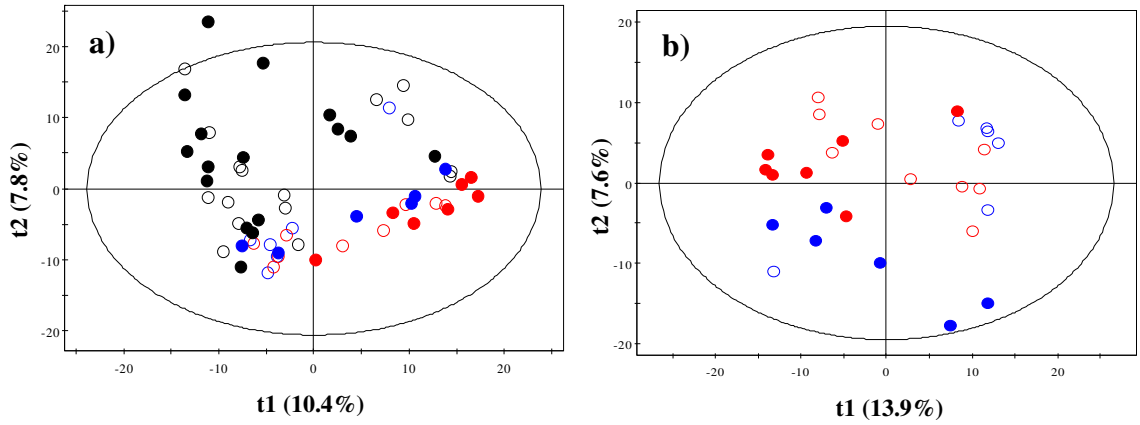
4.3.1 PCA overview of plasma sample in effluent-exposed fish

In –ESI mode, PCA analysis of the datasets showed clear separation between the control sample (C_{10}, control trout exposed to reference river water) and the treated sample (E_{10}, effluent exposed trout) after 10 day of exposure (Figure 4.1a); however, there was a significant overlap in model space between the controls and the different depuration groups (4 and 11 days of depuration in clean water after 10 days of effluent exposure) (Figure 4.1b).

The PCA overview of the same datasets did not reveal any groups of observations or trends in +ESI mode (Figure 4.1a,b). Therefore the comparison of the depurated groups did not provide clear class differentiation in both ESI modes. In some cases, auto-fitted PCA models showed no significant components; therefore, the models had to be forced to fit by calculating the first two components (Appendix 4.1). Poor values for the explained (R^2X) and predicted variation (Q^2) of all the datasets were obtained (Appendix 4.1). Two outliers (samples) were also detected using Hotelling's T^2 test (one outlier from class E_{10} in each ESI mode) (Figure 4.1a).

Univariate statistical analysis was utilized to confirm the results obtained from the multivariate analysis for the depurated groups (C_{14}, C_{21}, E_{14} and E_{21}). In both ESI modes, C_{14} versus C_{21} did not give a significant value of p. Differences between effluent groups (E_{14} versus E_{21}) were only significant in positive mode by the second component (PC2). This outcome is possibly due to high variation of endogenous metabolites present in plasma samples.

Positive mode



Negative mode

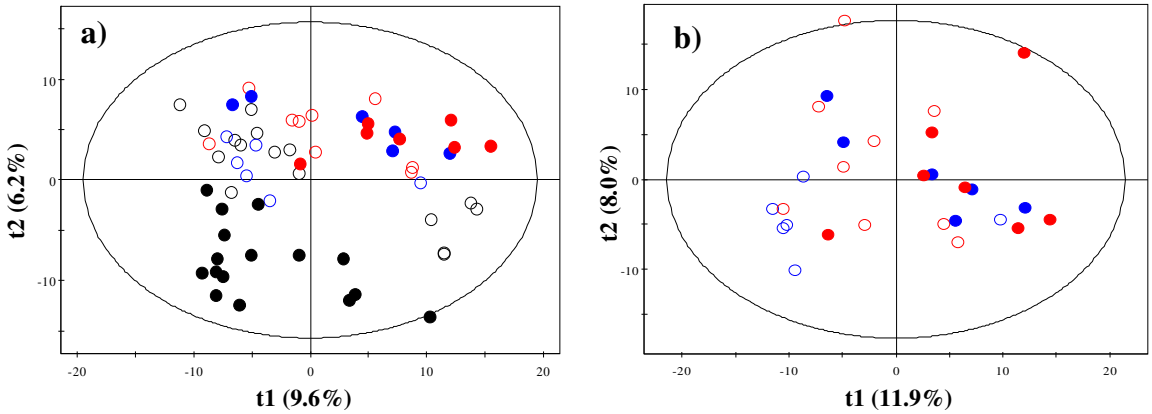


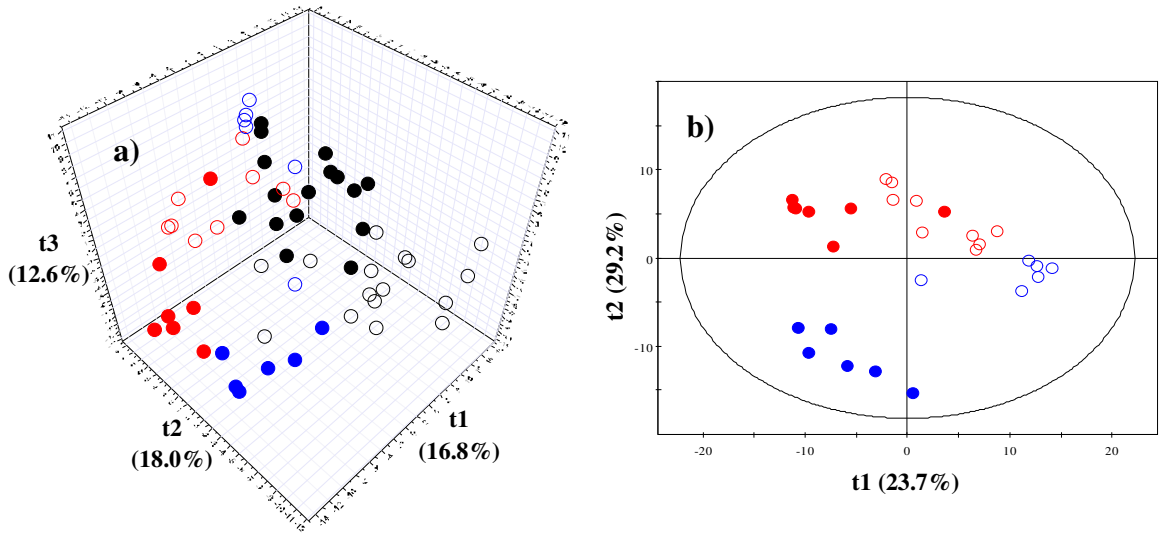
Figure 4.1: Principal component analyses (PCA) scores plots of the chemical profiles of plasma samples from trout exposed to either control river water or wastewater effluent in both ESI modes. On the left hand-side, (a) PCA of whole dataset [(○), (○) and (○) symbols represent C_{10}, C_{14} and C_{21} for control trout where n=16, 6, and 9 respectively. (●), (●) and (●) symbols represent E_{10}, E_{14} and E_{11} for effluent-exposed trout where n=16, 6 and 7 respectively] and on the right hand-side (b) PCA of dataset only for the depuration period. The percentages of explained variation (R^2X) for the first two components (t1 and t2) are displayed on the relative axes.

4.3.2 Overview of PLS-DA and OPLS-DA models for fish plasma

The PLS-DA approach was employed to further examine class separation. In both ESI models, PLS-DA models of all the classes revealed good separation between control and treated samples after 10 days of exposure and also between control and depurated samples after different depuration time points (4 and 11 days of depuration after 10 days of effluent exposure) (Figure 4.2). The obtained models showed good values of explained variation (R^2Y >0.98) and predicted variation (Q^2 >0.72) (Appendix 4.2). Reanalysis of the depurated groups (C_{14}, C_{21}, E_{14}, and E_{21}) omitting datasets for the 10 days of exposure, also resulted in a model with clear separation between the different classes and good predictive value (Q^2>0.65) (Figure 4.2 and Appendix 4.2). Subsequent OPLS-DA analyses between control and effluent exposed groups for each time point revealed good class separation between the two treatments in both ESI modes as described by their reasonable predictive value (Q^2>0.51) (Appendix 4.2 and Figure 4.3). However, class separation was poor for the control and treated groups at time point of 11 days of depuration after 10 days of effluent exposure (C_{21} versus E_{21}). This result highlights the high similarity between the E_{21} group and the control group C_{21} indicating that after 11 days of depuration most of the xenobiotics accumulated in the exposure period have been efficiently excreted and any other biochemical changes in the blood have been restored to normal physiological status.

In order to obtain statistically valid models and to ensure that high values of predictability were not occurring due to over-fitting on noise data, response permutation testing was performed. The models showed a classification achievement of >90% for almost of the datasets with the exception of the dataset of all treatments together (66% in +ESI and 78% in –ESI) and in the classification of the depurated datasets (C_{14}, C_{21}, E_{14}, and E_{21}), which was only 40% accuracy in both ESI modes (Appendix 4.2).

Positive mode



Negative mode

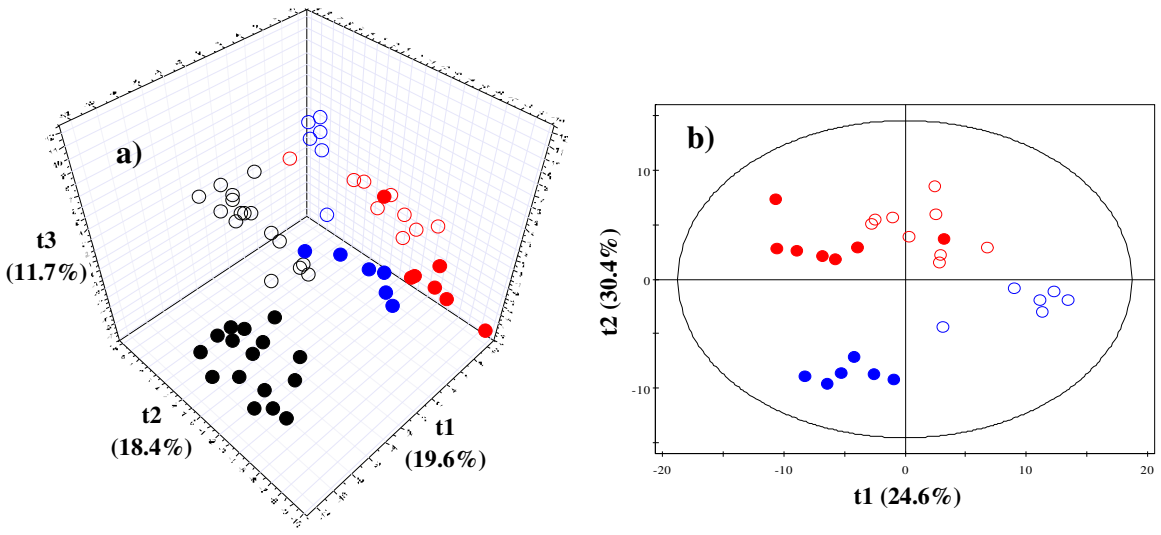


Figure 4.2: Partial least squares-discriminant analyses (PLS-DA) scores plots of the chemical profiles of plasma from trout exposed either to 100% wastewater effluent or control river water in both ESI modes. On the left hand-side, (a) PLS-DA of whole dataset [(○), (○) and (○) symbols represent C_{10}, C_{14} and C_{21} for control trout where n=16, 6, and 9 respectively. (●), (●) and (●) symbols represent E_{10}, E_{14} and E_{11} for effluent-exposed trout where n=16, 6 and 7 respectively] and on the right hand-side (b) PLS-DA of dataset only for the depuration period. The percentages of explained variation (R^2Y) modelled for the first two or three latent variables (t1, t2 and/or t3) are displayed on the related axes.

OPLS-DA analyses were conducted between each of the two groups (control and treated at different time points) to identify the loading variables influencing the classification. The OPLS-DA models of C_{10} versus E_{10} resulted in 1 predictive and 1 orthogonal significant component whilst the depurated groups showed no significance of their orthogonal components in both ESI modes (see Appendix 4.2). S-plots of the OPLS-DA models for the plasma datasets were used to extract the discriminatory chemicals responsible for group separation. Most of the detected markers were exogenous (xenobiotics and their metabolites) and already identified in fish bile. These variables were removed and the data remodelled to investigate changes in the levels of endogenous origin metabolites. A total of 77 variables (RT-*m/z*) were assigned as potential markers (exogenous and endogenous metabolites) in both ESI modes (Table 4.1a,b).

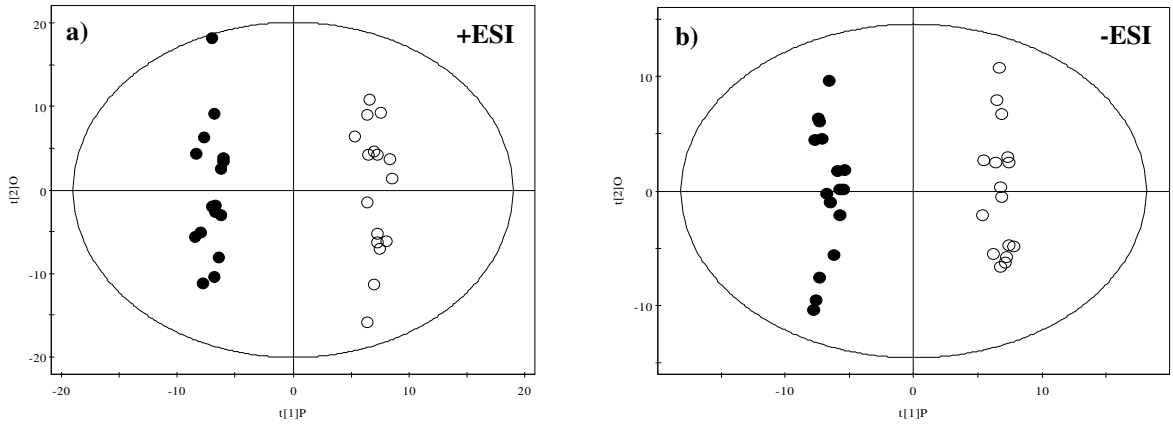


Figure 4.3: a) and b) OPLS-DA scores plots of the chemical profiles of plasma from trout exposed either to 100% wastewater effluent or control river water. The samples were profiled in both +/-ESI modes by UPLC-TOFMS; (○) and (●) symbols represent C_{10} for control and E_{10} for effluent-exposed trout for 10 days, respectively, analysed in both ESI modes.

4.3.3 Identified chemical marker classes

Biofluids (e.g. blood) are very complex matrices containing a huge number of potential biomarkers, i.e. endogenous metabolites which change in response to xenobiotic exposure, as well as xenobiotics or their metabolites. Compounds were detected that either increased or decreased in concentration as a result of effluent exposure. Accurate mass measurement and MS/MS fragmentation using Q-TOF as analysers (accurate mass for the fragments) were used to further characterise the candidate markers. Many of the markers were chemical contaminants already identified in fish bile: surfactants, chlorinated phenols, chlorinated xylenols, chlorophenes, and chlorinated phenoxyphenol (see Chapter 3). In addition to these, changes in concentrations of endogenously derived metabolites such as bile acids, and sphingolipids and acyl carnitines were also observed in plasma samples. Table 4.1a,b contains the full list of the detected markers in the plasma extracts *m/z* and relative *p*-values from the datasets studied: these were control versus effluent-exposed fish after 10 days of exposure (C_{10} vs E_{10}), 10 days effluent-exposure then 4 days depuration in river (C_{14} vs E_{14}) and 10 days effluent-exposure then 11 days depuration in river water (C_{21} vs E_{21}).

Table 4.1a: Markers of effluent exposure in rainbow trout (plasma extracts) in -ESI mode.

Marker No.	Observed ion (*m/z*)	RT	*P* value		
			C_{10} vs E_{10}	C_{14} vs E_{14}	C_{21} vs E_{21}
1	571.3837	4.39	6.77E-05	nd	nd
2	331.0587	7.39	3.33E-09	nd	nd
3	361.0691	7.83	3.33E-09	nd	nd
4	274.8743	7.85	2.48E-04	nd	nd
5	249.0218	8.04	4.84E-01	nd	nd
6	234.9831	8.06	5.09E-07	nd	nd
7	229.0561	8.07	3.33E-09	nd	nd
8	512.2705	8.34	9.98E-08	8.94E-01	4.12E-01
9	264.9940	8.67	3.33E-09	nd	nd
10	250.9783	8.77	5.09E-07	nd	nd
11	514.2841	8.78	3.04E-06	1.00E00	3.45E-01
12	393.0741	10.71	3.33E-09	nd	nd
13	498.2884	11.10	3.00E-03	5.15E-01	1.81E-01
14	531.2992	11.53	2.01E-04	6.97E-01	nd
15	327.1633	11.55	5.66E-08	nd	nd
16	296.9988	11.87	7.00E-03	nd	nd
17	462.9755	12.49	3.33E-09	nd	nd
18	482.2920	12.52	3.22E-06	nd	nd
19	332.9397	13.68	5.09E-07	nd	nd
20	409.9730	14.18	5.66E-08	nd	nd
21	366.9001	15.42	3.33E-09	2.00E-03	7.69E-01
22	347.9540	15.75	7.30E-06	2.00E-03	1.00E-03
23	395.2073	16.45	5.09E-07	nd	nd
24	397.9537	18.43	3.33E-09	2.00E-03	3.00E-03
25	497.9463	20.28	3.33E-09	2.00E-03	1.00E-03
26	262.8394	21.33	3.33E-09	nd	nd
27	450.9278	21.69	3.33E-09	1.00E00	4.29E-01
28	325.1838	22.17	4.63E-07	9.30E-02	1.00E00
29	569.9671	22.41	1.61E-05	2.00E-03	1.50E-02
30	321.2068	22.58	3.22E-06	nd	nd
31	339.1996	23.36	3.33E-09	8.18E-01	6.62E-01
32	403.3075	24.15	3.33E-09	nd	nd
33	329.2695	24.57	5.66E-08	nd	7.36E-01
34	373.2960	24.63	3.33E-09	nd	nd

ESI: electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control); E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; nd: not detected.

Table 4.1b: Markers of effluent exposure in rainbow trout (plasma extracts) in +ESI mode.

Marker No.	Observed ion (*m/z*)	RT	*P* value		
			C_{10} vs E_{10}	C_{14} vs E_{14}	C_{21} vs E_{21}
1	287.2014	4.30	6.77E-05	4.55E-01	nd
2	349.1902	4.30	3.33E-09	4.55E-01	6.59E-01
3	573.3970	4.31	3.22E-06	1.00E00	1.00E00
4	246.1707	4.48	3.04E-06	6.50E-02	4.34E-01
5	915.0385	4.10	4.73E-01	5.82E-01	5.43E-01
6	801.2529	4.27	3.28E-04	1.50E-02	6.62E-01
7	961.3014	4.27	1.00E-03	4.10E-02	7.55E-01
8	710.7023	4.69	2.00E-03	1.00E00	nd
9	711.0378	4.69	6.77E-05	1.00E00	nd
10	799.4448	5.18	6.40E-02	1.00E00	nd
11	849.7238	5.61	6.77E-05	1.00E00	6.64E-01
12	679.9779	5.63	1.00E-03	1.00E00	4.13E-01
13	566.8147	5.63	1.00E-03	8.48E-01	7.90E-01
14	445.3036	5.39	2.48E-04	nd	nd
15	498.2889	8.80	3.92E-06	7.06E-01	4.66E-01
16	398.3427	10.13	3.33E-09	2.00E-03	3.33E-04
17	400.3571	10.57	6.77E-05	2.00E-03	3.33E-04
18	464.2841	11.15	6.00E-03	nd	5.38E-01
19	300.2900	20.08	6.89E-06	9.30E-02	6.62E-01
20	327.0769	20.28	3.33E-09	nd	4.29E-01
21	330.3380	21.58	4.30E-02	nd	nd
22	454.2910	21.63	1.00E+00	1.00E+00	1.00E+00
23	284.3318	22.34	5.09E-07	6.10E-02	9.10E-02
24	355.2827	22.90	2.00E-03	nd	nd
25	344.3165	23.13	3.22E-06	nd	nd
26	251.0509	23.28	4.84E-01	nd	nd
27	427.3036	24.15	3.33E-09	nd	nd
28	471.3306	24.16	3.33E-09	nd	nd
29	515.3564	24.14	6.77E-05	nd	nd
30	559.3819	24.11	4.30E-02	nd	nd
31	369.2982	24.31	3.33E-09	nd	nd
32	413.3238	24.30	5.66E-08	nd	nd
33	457.3506	24.27	3.33E-09	nd	nd
34	501.3767	24.25	3.33E-09	6.97E-01	2.03E-01
35	545.4028	24.23	4.40E-01	2.10E-01	2.45E-01
36	441.3199	24.63	7.00E-03	nd	nd
37	485.3454	24.64	3.33E-09	nd	nd
38	529.3716	24.63	3.33E-09	nd	nd
39	573.3977	24.62	3.33E-09	1.00E00	1.00E00
40	427.3399	24.78	3.33E-09	nd	nd
41	471.3658	24.77	3.33E-09	nd	nd
42	515.3923	24.76	3.33E-09	nd	nd
43	559.4185	24.76	3.33E-09	nd	nd

ESI: electrospray ionization; *m/z*: mass to charge ratio; RT: retention time (min); C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control); E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; nd: not detected.

4.3.3.1 Chlorinated metabolites

UPLC-TOFMS profiling in –ESI mode showed the occurrence of many markers of effluent exposure relating to the presence of different classes of chlorinated compounds and their metabolites. Four main classes of chlorinated compounds were identified in plasma of effluent-exposed fish (Table 4.2): chlorinated phenols (a sulphate conjugate of trichlorophenol at *m/z* 274.8743), chlorinated xylenols (two peaks corresponding to the glucuronide conjugates of chloroxylenol and its methoxy metabolite at *m/z* 331.0587 and *m/z* 361.0691, respectively), and two peaks corresponding to the sulphate conjugate of chloroxylenol and its methoxy metabolite (at *m/z* 234.9831 and *m/z* 264.994, respectively), chlorophenes (two peaks corresponding to the glucuronide and sulphate conjugate of chlorophene at *m/z* 393.0741 and 296.9988, respectively) and chlorinated phenoxyphenols (one peak with a mass corresponding to the sulphate conjugate of diclosan at *m/z* 332.9397 and two peaks corresponding to the glucuronide and sulphate conjugate of triclosan at *m/z* 462.9755 and *m/z* 366.9001, respectively). The extracted ion chromatograms of the chlorinated metabolites detected in plasma samples from effluent-exposed trout plasma samples in –ESI mode are shown in Figure 4.4. None of these metabolites were detected in any control sample. Exact mass and fragmentation pattern were employed to fully characterize the structure of these metabolites. The retention time, exact mass, and Q-TOFMS fragmentation pattern of the glucuronide conjugates of chloroxylenol, a methoxy metabolite of chloroxylenol, chlorophene, and triclosan were the same as previously described to those detected in the bile samples (see Chapter 3). In –ESI mode, the Q-TOFMS spectra of the deprotonated glucuronide conjugates were dominated, as explained in Chapter 3, by the aglycone molecules together with the characteristic fragments of the glucuronide moiety (*m/z* 175.0243, *m/z* 157.0137, *m/z* 113.0239). In contrast, the sulphate esters of a trichlorophenol, chloroxylenol, a methoxy metabolite of chloroxylenol and triclosan showed the characteristic loss of 80 Da (loss of SO_3) leading to the fragment ion $[M\text{-}H\text{-}SO_3]^-$ for the detected chlorinated components. Sulphate conjugates can be formed when compounds with hydroxyl (especially phenolic groups) and amine groups come in contact with sulphuric acid after activation of 3′-phosphoadenosine-5′-phosphosulphate (PAPS) by the sulphotransferases (see Figure 4.5) (Levsen et al., 2005). The sulphotransferase enzyme catalyzes the transfer of the sulphonate group from PAPS to the hydroxyl group in a range of endogenous and exogenous substrates (Coughtrie, 2002). The identity of the sulphate conjugate of a trichlorophenol was based on the

presence of a small peak at *m/z* 194.9171, corresponding to the parent ion after loss of the SO_3 group (Figure 4.6 where the 3-Cl isotopic distribution for the molecular ion and the relative fragment are highlighted in yellow), whilst Q-TOFMS analysis was not possible for this compound due to its low response in full scan. This result indicated that the sulphate ester bond between the discussed chlorinated compounds and the sulphuric acid is not strong enough to avoid the cleavage of the ester bond in the source at the applied collision energy (10 eV). Parent ions at *m/z* 155.0264, *m/z* 185.0369 and 286.9433 were generated by Q-TOFMS fragmentation for sulphate conjugate of chloroxylenol, methoxychloroxylenol and triclosan, respectively, after loss of the SO_3 group (Figure 4.7, 4.8 and 4.9).

The sulphate conjugate of chlorophene and diclosan were putatively identified based on their exact mass obtained from UPLC-TOFMS analysis in –ESI mode (see Table 4.2). As observed in this study, the sulphate group when attached to phenylic groups can be easily cleaved from the bulk molecule (collision energy < 20 eV is required) whilst it is very hard to cleave when linked to a cyclic alkyl residual (e.g. cyprinol sulphate).

Table 4.2: Markers identified in trout plasma in -ESI mode.

Class of chemical	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula	Putative identity
Chlorinated phenols	274.8743	7.85	$C_6H_2O_4SCl_3$	274.8739	1.5	0.9	194.9165	$C_6H_3OCl_3$	Sulphate conjugate of a trichlorophenol
Chlorinated xylenols	331.0587	7.39	$C_{14}H_{16}O_7Cl$	331.0585	0.6	0.3	175.0247, 155.0267, 113.0243	C_8H_9OCl	Chloroxylenol glucuronide
	361.0691	7.83	$C_{15}H_{18}O_8Cl$	361.0690	0.3	0.8	185.0371, 175.0243, 113.0239	$C_9H_{11}O_2Cl$	Glucuronide conjugate of a methoxy metabolite of chloroxylenol
	234.9831	8.06	$C_8H_8O_4SCl$	234.9832	-0.4	1.5	155.0263	C_8H_9OCl	Chloroxylenol sulphate
	264.9940	8.67	$C_9H_{10}O_5SCl$	264.9937	1.1	0.3	185.0368	$C_9H_{11}O_2Cl$	Sulphate conjugate of a methoxy metabolite of chloroxylenol
Chlorophenes	393.0741	10.71	$C_{19}H_{18}O_7Cl$	393.0741	0.0	0.2	217.0418, 175.0242, 113.0243	$C_{13}H_{11}OCl$	Chlorophene glucuronide
	296.9988	11.87	$C_{13}H_{10}O_4SCl$	296.9988	0.0	2.3		$C_{13}H_{11}OCl$	Putative chlorophene sulphate
Chlorinated phenoxyphenols	332.9397	13.68	$C_{12}H_7O_5SCl_2$	332.9391	1.8	2.7		$C_{12}H_8O_2Cl_2$	Putative diclosan sulphate
	462.9755	12.49	$C_{18}H_{14}O_8Cl_3$	462.9754	0.2	0.0	286.9419, 175.0239, 113.0238	$C_{12}H_7O_2Cl_3$	Triclosan glucuronide
	366.9001	15.42	$C_{12}H_6O_5SCl_3$	366.9002	-0.3	0.8	286.9431	$C_{12}H_7O_2Cl_3$	Triclosan sulphate
Nonionic surfactants	395.2073	16.45	$C_{21}H_{31}O_7$	395.2070	0.8	0.4	219.1751, 175.0243	$C_{15}H_{24}O$	Nonylphenol glucuronide
Anionic surfactants	325.1839	22.17	$C_{18}H_{29}O_3S$	325.1837	0.6	1.7	183.0115, 170.0042	$C_{17}H_{30}O_3S$	C_{12}-LAS
	339.1996	23.36	$C_{19}H_{31}O_3S$	339.1994	0.6	0.8	197.0272, 183.0114, 170.0036	$C_{19}H_{32}O_3S$	C_{13}-LAS
	403.3070	24.15	$C_{22}H_{43}O_6$	403.3060	2.5	0.4	359.2791, 211.2053	$C_{22}H_{44}O_6$	Tetradecanol-E4C metabolite
Bile acids (biochemicals)	514.2841	8.78	$C_{26}H_{44}NO_7S$	514.2839	0.4	0.0	496.2724, 353.2473, 124.0066, 106.9807, 79.9564	$C_{26}H_{45}NO_7S$	Taurocholic acid
	498.2884	11.10	$C_{26}H_{44}NO_6S$	498.2889	-1.0	0.3	124.0068, 106.9807	$C_{26}H_{45}NO_6S$	Taurochenodeoxycholic acid
	531.2992	11.53	$C_{27}H_{47}O_8S$	531.2992	0.0	1.4	513.2864, 96.9592	$C_{27}H_{48}O_8S$	Cyprinol sulphate

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; LAS: linear alkylbenzene sulphonic acid.

Table 4.3: Markers identified in plasma of effluent-exposed trout in +ESI mode.

Class of Chemical	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Putative identity
Non-ionic surfactants	355.2827 *350.32*	22.90 *24.27*	$C_{19}H_{40}O_4Na$ *$C_{19}H_{44}NO_4$*	355.2824 *350.3270*	0.0	0.1	*151.0978*	$C_{19}H_{40}O_4$	Tridecanol-3EO; Na adduct *NH_4-adduct*
Non-ionic surfactants	369.2982 *364.34*	24.31 *24.52*	$C_{20}H_{42}O_4Na$ *$C_{20}H_{46}NO_4$*	369.2981 *364.3427*	0.3	0.7	*151.0977, 133.0872*	$C_{20}H_{42}O_4$	Tetradecanol-3EO; Na adduct *NH_4-adduct*
	413.3238 *408.36*	24.30 *24.49*	$C_{22}H_{46}O_5Na$ *$C_{22}H_{50}NO_5$*	413.3243 *408.3689*	-1.2	0.7	*195.1230, 177.1132, 133.0862*	$C_{22}H_{46}O_5$	Tetradecanol-4EO; Na adduct *NH_4-adduct*
	457.3506 *452.39*	24.27 *24.53*	$C_{24}H_{50}O_6Na$ *$C_{24}H_{54}NO_6$*	457.3505 *452.3951*	0.2	0.0	*239.1483, 221.1383, 177.1118, 133.0871*	$C_{24}H_{50}O_6$	Tetradecanol-5EO; Na adduct *NH_4-adduct*
	501.3767 *496.42*	24.25 *24.49*	$C_{26}H_{54}O_7Na$ *$C_{26}H_{58}NO_7$*	501.3767 *496.4213*	0.0	0.2	*283.1756, 265.1661, 221.1380, 177.1128, 133.0866*	$C_{26}H_{54}O_7$	Tetradecanol-6EO; Na adduct *NH_4-adduct*
	545.4028 *540.44*	24.23 *24.44*	$C_{28}H_{58}O_8Na$ *$C_{28}H_{62}NO_8$*	545.4029 *540.4475*	-0.2	2.7	*327.2025, 309.1905, 221.1394, 177.1131, 133.0864*	$C_{28}H_{58}O_8$	Tetradecanol-7EO; Na adduct *NH_4-adduct*
Non-ionic surfactants	427.3399 *422.38*	24.78 *25.07*	$C_{23}H_{48}O_5Na$ *$C_{23}H_{52}NO_5$*	427.3399 *422.3845*	0.0	0.9	*195.1229, 177.1128 133.0863*	$C_{23}H_{48}O_5$	Pentadecanol-4EO; Na adduct *NH_4-adduct*
	471.3658 *466.41*	24.77 *25.08*	$C_{25}H_{52}O_6Na$ *$C_{25}H_{56}NO_6$*	471.3662 *466.4108*	-0.8	0.2	*239.1501, 221.1392, 177.1131, 133.0867*	$C_{25}H_{52}O_6$	Pentadecanol-5EO; Na adduct *NH_4-adduct*
	515.3923 *510.43*	24.76 *25.05*	$C_{27}H_{56}O_7Na$ *$C_{27}H_{60}NO_7$*	515.3924 *510.4370*	-0.2	0.3	*283.1743, 265.1656, 221.1379, 177.1126, 133.0860*	$C_{27}H_{56}O_7$	Pentadecanol-6EO; Na adduct *NH_4-adduct*
	559.4185 *554.46*	24.76 *25.05*	$C_{29}H_{60}O_8Na$ *$C_{29}H_{64}NO_8$*	559.4186 *554.4632*	-0.2	0.8	*327.2017, 221.1400, 177.1125, 133.0869*	$C_{29}H_{60}O_8$	Pentadecanol-7EO; Na adduct *NH_4-adduct*

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; fragments obtained from ammonium adduct are labelled in italic font.

Table 4.3: (continued) Markers identified in plasma of effluent-exposed trout in +ESI mode.

Class of Chemical	Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Putative identity
Anionic surfactants	427.3036 *422.34*	24.15 *24.44*	$C_{22}H_{44}O_6Na$ *$C_{22}H_{48}NO_6$*	427.3036 *422.3482*	0.0	1.1	369.3009, 231.0855 *209.1021, 191.0921, 163.0970*	$C_{22}H_{44}O_6$	Tetradecanol-4EC metabolite; Na adduct *NH_4-adduct*
	471.3306 *466.37*	24.16 *24.38*	$C_{24}H_{48}O_7Na$ *$C_{24}H_{52}NO_7$*	471.3298 *466.3744*	1.7	0.1	413.3242, 275.1119 *253.1284, 235.1170, 207.1222*	$C_{24}H_{48}O_7$	Tetradecanol-5EC metabolite; Na adduct *NH_4-adduct*
	515.3564 *510.4005*	24.14 *24.38*	$C_{26}H_{52}O_8Na$ *$C_{26}H_{56}NO_8$*	515.3560 *510.4006*	0.8	2.3	457.3483, 319.1364 *297.1540, 279.1434, 251.1486*	$C_{26}H_{52}O_8$	Tetradecanol-6EC metabolite; Na adduct *NH_4-adduct*
	559.3819 *554.42*	24.11 *24.34*	$C_{28}H_{56}O_9Na$ *$C_{28}H_{60}NO_9$*	559.3822 *554.4268*	-0.5	0.9	 *341.1821, 295.1763*	$C_{28}H_{56}O_9$	Tetradecanol-7EC metabolite; Na adduct *NH_4-adduct*
Anionic Surfactants	441.3199 *436.36*	24.63 *24.95*	$C_{23}H_{46}O_6Na$ *$C_{23}H_{50}NO_6$*	441.3192 *436.3638*	1.6	0.9	 *209.1035, 191.0928, 163.0978*	$C_{23}H_{46}O_6$	Pentadecanol-4EC; Na adduct *NH_4-adduct*
	485.3454 *480.39*	24.64 *24.88*	$C_{25}H_{50}O_7Na$ *$C_{25}H_{54}NO_7$*	485.3454 *480.3900*	0.0	0.7	427.3390, 275.1094 *253.1284, 235.1182, 207.1227, 147.0651*	$C_{25}H_{50}O_7$	Pentadecanol-5EC metabolite; Na adduct *NH_4-adduct*
	529.3716 *524.41*	24.63 *24.87*	$C_{27}H_{54}O_8Na$ *$C_{27}H_{58}NO_8$*	529.3716 *524.4162*	0.0	0.4	471.3655, 319.1374 *297.1542, 279.1447, 251.1598, 191.0928*	$C_{27}H_{54}O_8$	Pentadecanol-6EC metabolite; Na adduct *NH_4-adduct*
	573.3977 *568.44*	24.62 *24.84*	$C_{29}H_{58}O_9Na$ *$C_{29}H_{62}NO_9$*	573.3979 *568.4425*	-0.3	0.6	515.3900, 363.1618 *341.1795, 323.1721, 295.1745, 235.1180*	$C_{29}H_{58}O_9$	Pentadecanol-7EC metabolite; Na adduct *NH_4-adduct*
Steroid alkaloid	398.3427	10.13	$C_{27}H_{44}NO$	398.3423	1.0	0.6	380.3514, 366.3158, 157.1012, 98.0962	$C_{27}H_{45}NO$	Solanidine
	400.3577	10.57	$C_{27}H_{46}NO$	400.3579	-0.5	0.6	382.3475, 161.1327, 98.0952	$C_{27}H_{47}NO$	Dihydrosolanidine
Bile acids	498.2889	8.80	$C_{26}H_{44}NO_6S$	498.2889	0.0	1.1	462.2689, 337.2533, 126.0221	$C_{26}H_{45}NO_7S$	Taurocholic acid (loss of H_2O)
	464.2841	11.15	$C_{26}H_{42}NO_4S$	464.2835	1.5	0.7	339.2672, 126.0221	$C_{26}H_{45}NO_6S$	Taurochenodeoxycholic acid (loss of $2H_2O$)
Acyl carnitines	246.1707	4.48	$C_{12}H_{24}NO_4$	246.1705	0.8	0.5	187.0975	$C_{12}H_{25}NO_4$	Acyl-L-carnitine (putative 2-methylbutyroylcarnitine)
Sphingolipids	300.2900	20.08	$C_{18}H_{38}NO_2$	300.2903	-1.0	0.0	282.2794, 264.2693, 252.2689	$C_{18}H_{37}NO_2$	Sphingosine

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; fragments obtained from ammonium adduct are labelled in italic font.

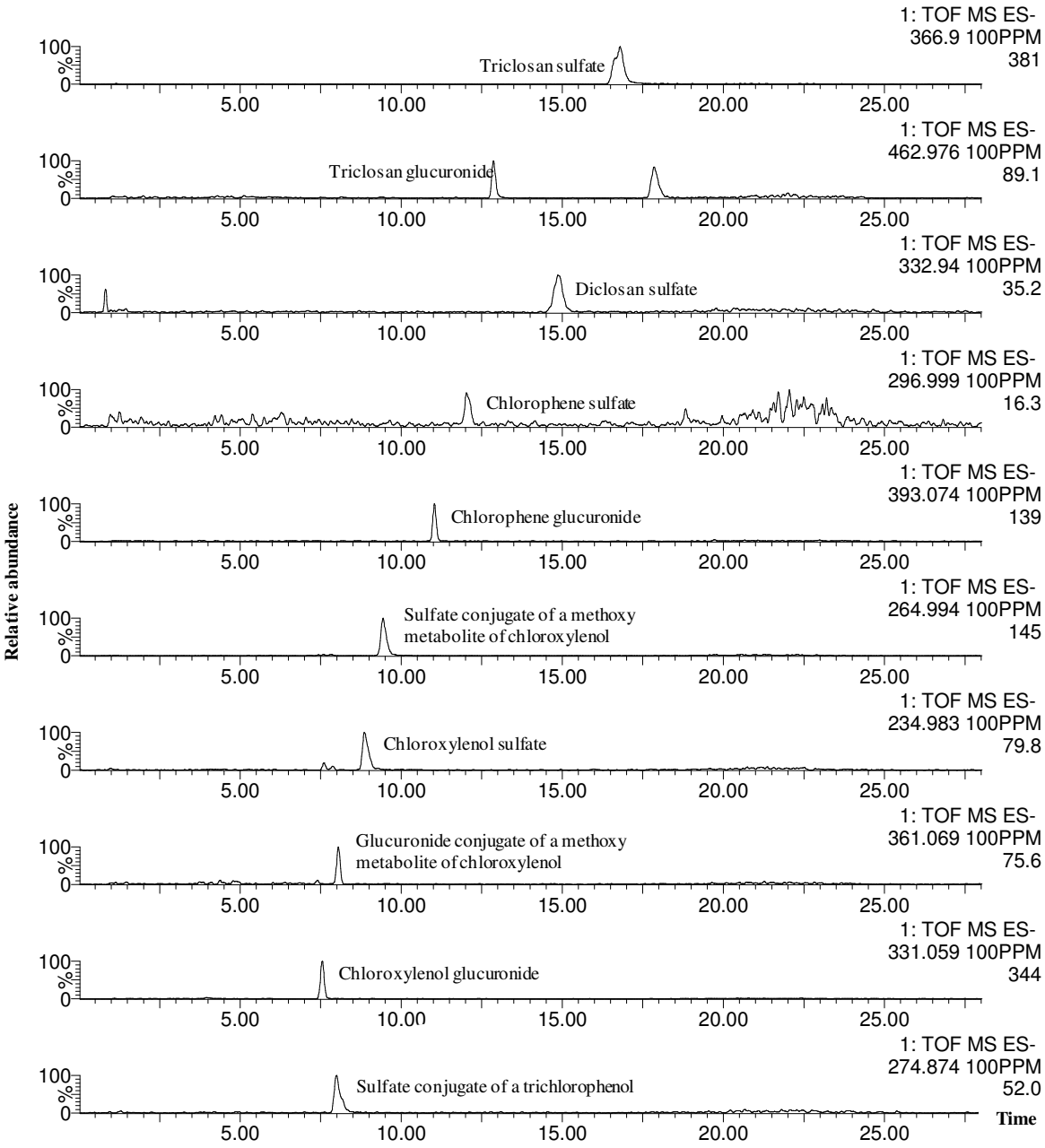


Figure 4.4: Extracted ion chromatograms of chlorinated metabolites detected in plasma from effluent-exposed trout in –ESI mode.

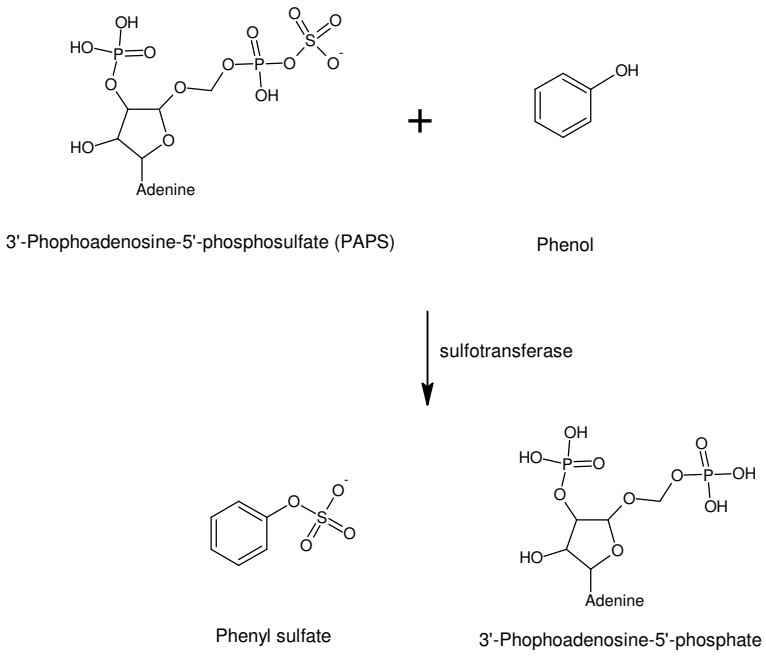


Figure 4.5: The enzymatic sulphation of phenol. Adapted from Levsen, Schiebel et al. (2005).

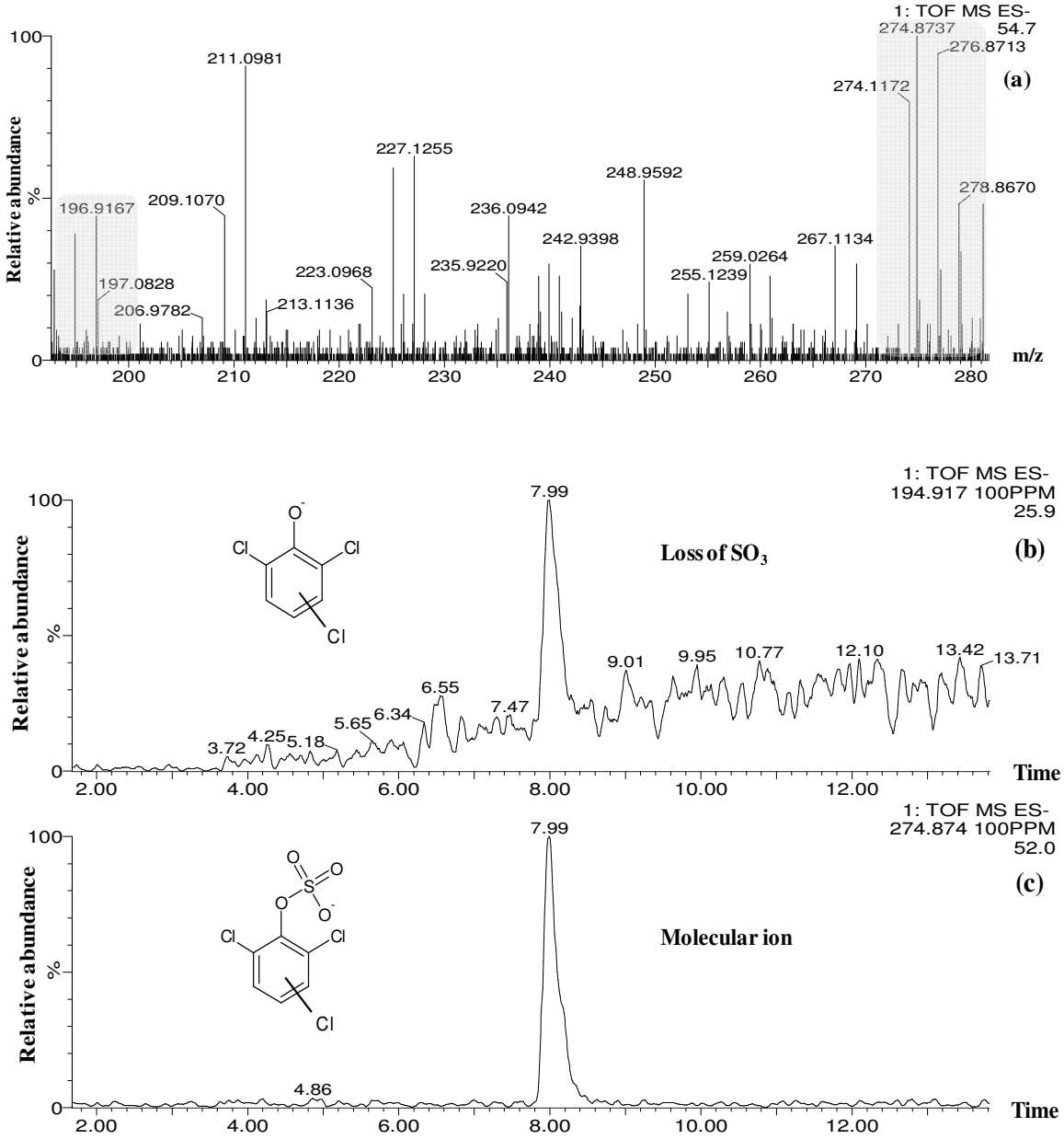


Figure 4.6: TOFMS spectra for isomer of trichlorophenol sulphate (a) and full scan extracted ion chromatograms for the molecular ion *m/z* 274.8740 (b) and in-source fragment *m/z* 194.917 (c) (collision energy=10 eV) detected in plasma from effluent-exposed trout in -ESI mode. 3-Cl isotopic pattern for the molecular ion and the fragment are highlighted in gray.

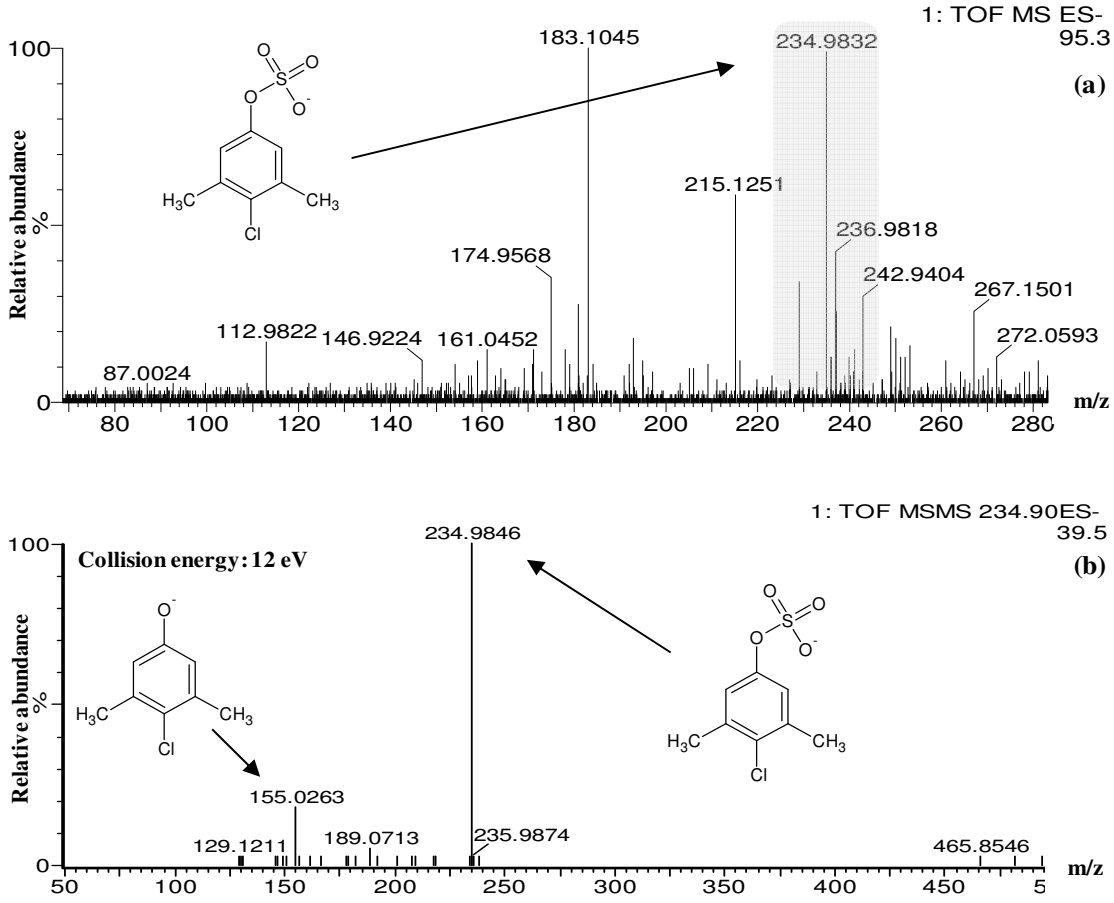


Figure 4.7: a) Mass spectrum of sulphate conjugate of chloroxylenol obtained with collision energy of 10 eV and b) its relative Q-TOFMS spectrum obtained with collision energy of 12 eV detected in plasma from effluent-exposed trout in –ESI mode.

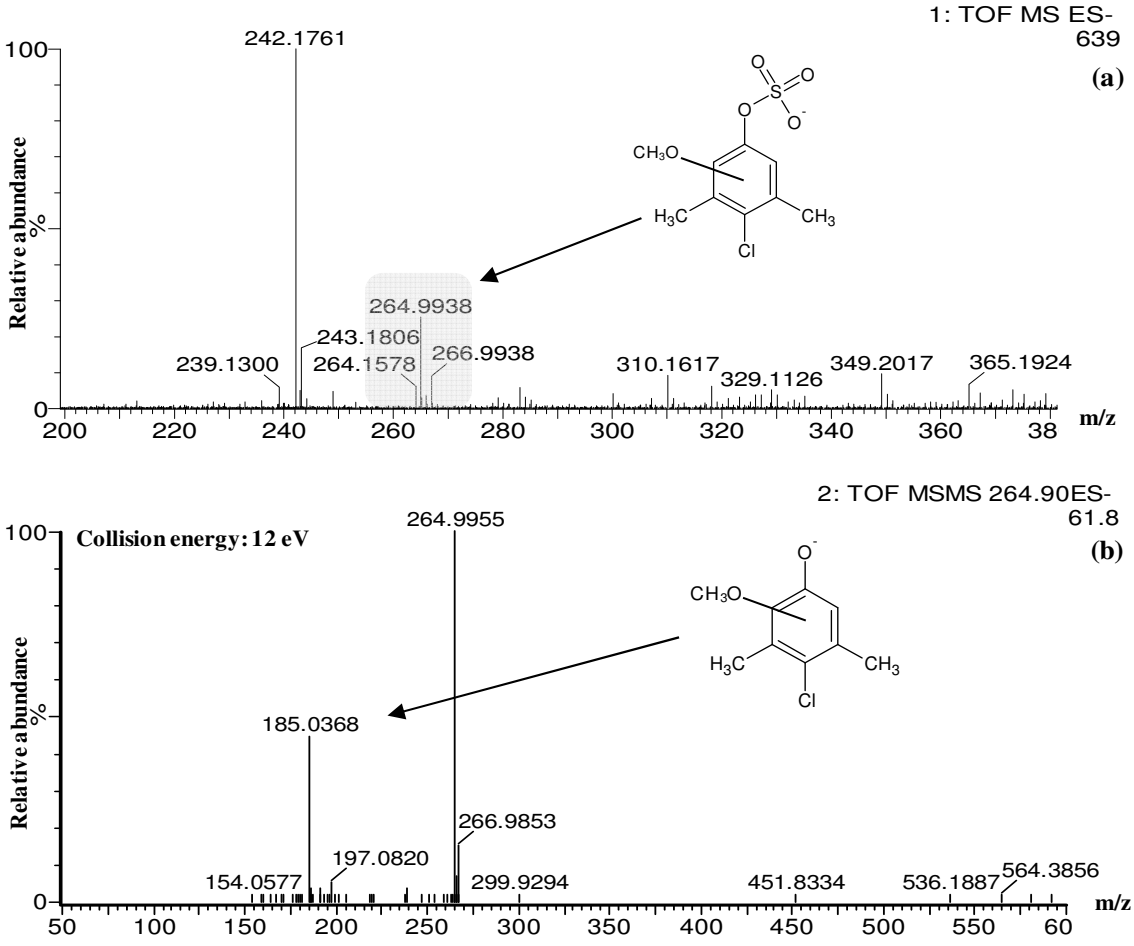


Figure 4.8: Mass spectrum of sulphate conjugate of a methoxy chloroxylenol obtained with collision energy of 10 eV (a) and its relative Q-TOFMS spectrum obtained with collision energy of 12 eV (b) found in plasma from effluent-exposed trout in –ESI mode.

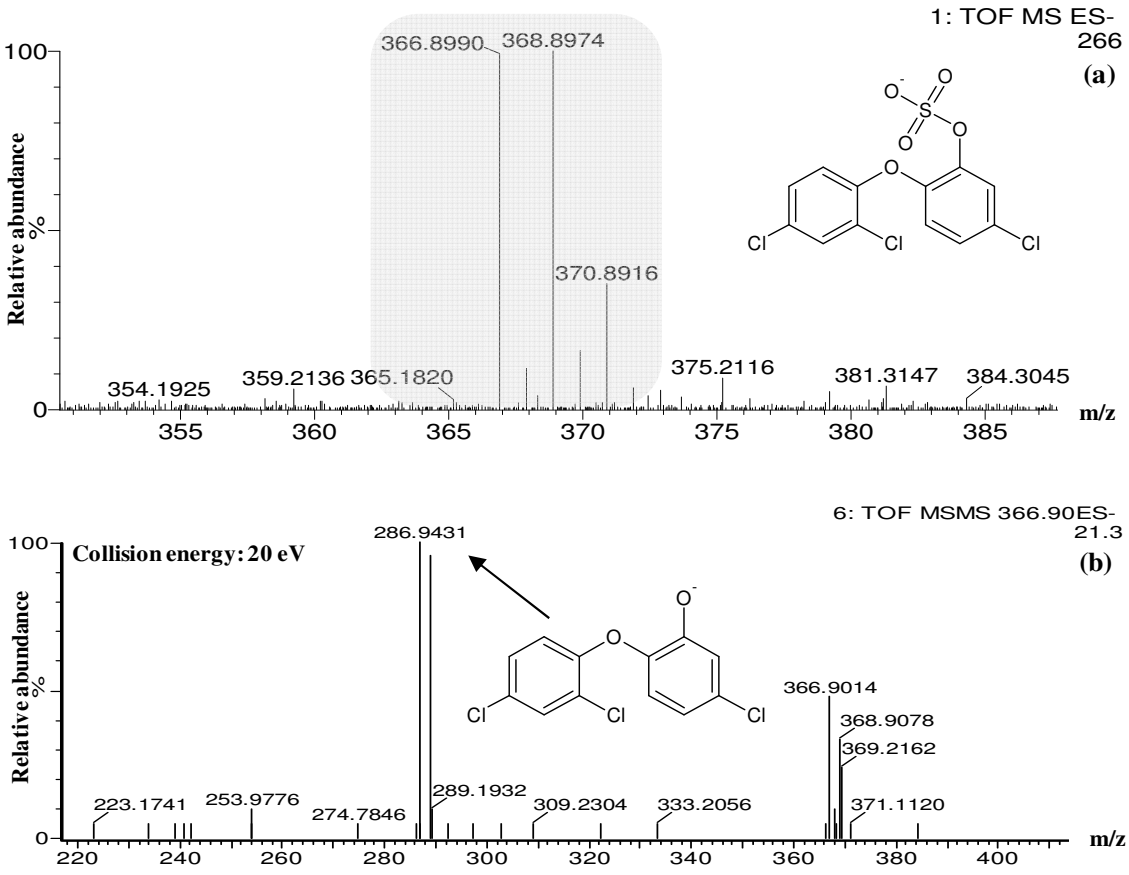


Figure 4.9: a) Mass spectrum of sulphate conjugate of triclosan obtained with collision energy of 10 eV and b) its relative Q-TOFMS spectrum obtained with collision energy of 20 eV found in plasma from effluent-exposed trout in –ESI mode.

4.3.3.2 Surfactant metabolites

Anionic and non-ionic surfactants were also detected as potential marker compounds in effluent-exposed fish plasma in both ESI modes. UPLC-TOFMS analysis showed two peaks of linear alkylbenzene sulfonic acid (LASs) which were identified as (C12-LAS and C13-LAS) in –ESI mode (Table 4.2). Q-TOFMS fragmentation spectra of these metabolites revealed the same characteristic ions, as described in Chapter 3, at *m/z* 183.0116 ($C_8H_7SO_3$), *m/z* 170.0038 ($C_7H_6SO_3$) and/or *m/z* 197.0272 ($C_9H_9SO_3$) (Figure 4.10).

Glucuronide nonylphenol (non-ionic surfactant) was detected in fish plasma in –ESI mode (Table 4.2) showing comparable RT, accurate mass and Q-TOFMS fragmentation pattern to the one detected in fish bile (see Chapter 3).

Alkyl ethoxylates (A*x*EO*n*; where *x*=13-15; n=3-7) were also detected in fish plasma but only in +ESI mode as sodium adducts $[M+Na]^+$ (Table 4.3). Unlike in bile samples, in the plasma samples these metabolites were not conjugated by the glucuronide moiety. As observed in Chapter 3, sodium adducts of AEOs series that were detected as markers of effluent exposure were too stable to give significant fragmentation, therefore, addition of ammonium formate was employed for a second batch of analysis (composite sample of control and effluent-exposed fish for 10 days of exposure, C_{10} and E_{10}) in order to replace the sodium ion with the ammonium adduct and therefore enhance the chance of fragmentation for the studied compounds. Selected ammonium adducts of AEOs resulted a good choice for further Q-TOFMS experiments as demonstrated in Chapter 3. Ammonium adducts gave rise to neutral loss of alkene (C_nH_{2n}; where n=13-15) plus NH_3 at *m/z* ranging from 151.0970 ($C_6H_{15}O_4$) to 327.2019 ($C_{14}H_{31}O_8$) (44 Da apart from each other). Furthermore, many fragments characteristic of the ethoxylated compounds were obtained at *m/z* ranging from 133.0865 ($C_6H_{13}O_3$) to 265.1651 ($C_{12}H_{25}O_6$) (each ion separated by 44 Da) (for details see Chapter 3). In Figure 4.11, an example is given for the fragmentation pattern of the ammonium adduct ions at *m/z* 452.3951 and *m/z* 466.4108, which were positively identified as tetradecanol-5EO and pentadecanol-5EO, respectively.

An additional series with *m/z* ranging from 427.3036 to 559.3819 and from 441.3199 to 573.3977 were also detected in +ESI mode as potential markers, differing by 44 Da from each other (EO unit) (Table 4.3). These peaks were identified as tetradecanol and pentadecanol polyethoxy carboxylates (C14EC*n* and C15EC*n*; where n=4-7), respectively. Characteristic ions were obtained from the fragmentation of the Na

and NH_4 adducts: ammonium adducts showed loss of alkene plus NH_3 followed by loss of H_2O and CO, whilst fragmentation of sodium adducts led to the loss of the acetate group (-58 Da) and alkene by central cleavage of the parent ion (-196 Da and -210 Da for the two series, respectively). These results were comparable to the ones obtained for the bile samples (see Chapter 3).

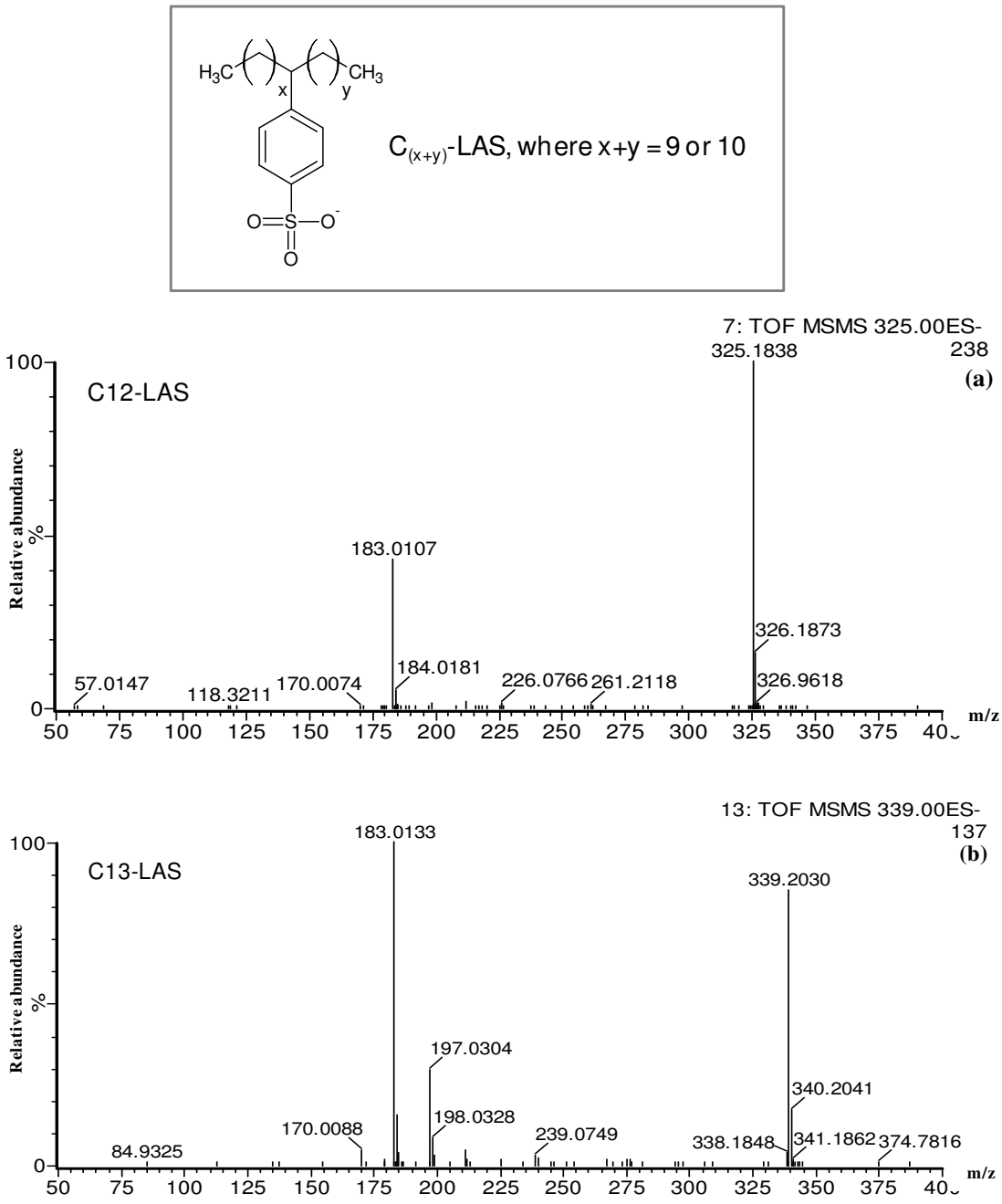


Figure 4.10: Molecular structure and relative Q-TOFMS spectra for linear alkylbenzene sulfonic acid (LASs): a) C12-LAS and b) C13-LAS detected in plasma from effluent-exposed trout in –ESI mode.

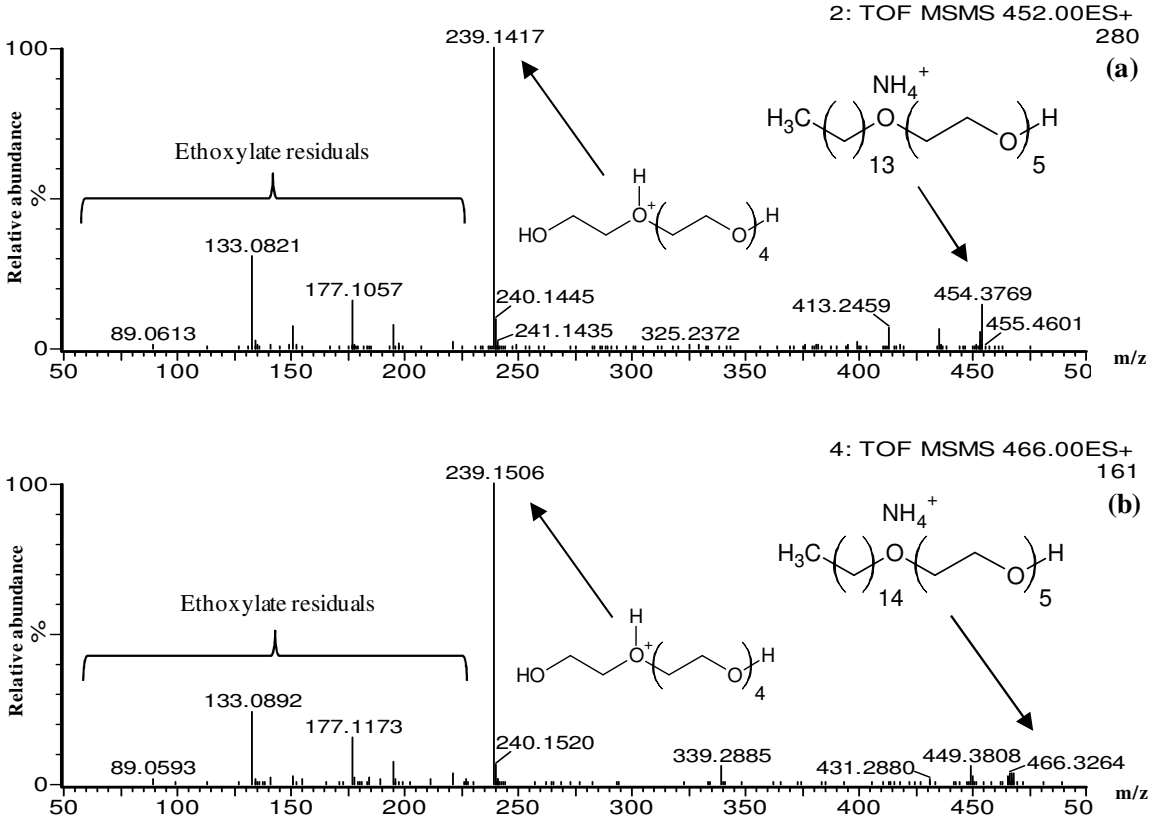


Figure 4.11: Q-TOFMS spectra of ammonium adducts for Alkyl polyethoxylates (AEOs): a) tetradecanol-5EO and b) pentadecanol-5EO detected in plasma from effluent-exposed trout in +ESI mode (collision energy=20 eV).

4.3.3.3 Steroid alkaloids

UPLC-TOFMS profiling of plasma extracts revealed the presence of two steroid alkaloids which could be detected only in +ESI mode giving rise to signals at *m/z* 398.3427 and *m/z* 400.3577, respectively (Table 4.3). These compounds were identified as solanidine and dihydrosolanidine, by means of exact mass data (*m/z* accuracy < 1ppm) and high energy collisional dissociation fragmentation patterns in full scan mode (CE: 50eV)(data not shown). An abundant in-source fragment ion at *m/z* 98.9869, corresponding to the heterocyclic ring ($C_6H_{12}N$), was obtained for both compounds as well as the signal at *m/z* 382.3474 generated by the loss of H_2O from the dihydrosolanidine molecule. Further mass spectrometry information will be given in Chapter 5 since these compounds were also present in roach plasma (detected as sulphate conjugates).

Analysis of the UPLC-TOFMS profiles for the plasma samples revealed the presence of highly saturated peak in all the treatments (C_{10}, C_{14}, C_{21}, E_{10}, E_{14}, E_{21}). This peak has been identified as the anaesthetic tricaine mesilate (MS222). In fact, tricaine mesilate, a methane sulfonate salt of 3-aminobenzoic acid ethyl ester, together with benzocaine (another popular anaesthetic and structural analogue), is one of the most widely used anaesthetics in marine and freshwater fish (Allen and Hunn, 1986, Meinertz et al., 1999) (Figure 4.12). MS222 was not included in the peak list because it was present in both control and treatment groups. This compound was detected only in +ESI mode, giving a signal corresponding to the protonated molecular ion at *m/z* 166.0868 $[C_9H_{12}NO_2]^+$. A careful examination of the obtained spectra led to the detection of another saturated peak at *m/z* 138.0555 $[C_7H_8NO_2]^+$, corresponding to the loss of C_2H_4 from the molecular ion. In addition, two signals (different *m/z* but same RT) due to the presence of MS222 were found at *m/z* 120.0449 $[C_7H_6NO]^+$ and *m/z* 94.0657 $[C_6H_8N]^+$ (Figure 4.13). This result indicates MS222 can be subjected to in source CID fragmentation under the experimental condition employed in this study. Furthermore, fragmentation of tricaine has been previously described by Scherpenisse and Bergwerff (2007).

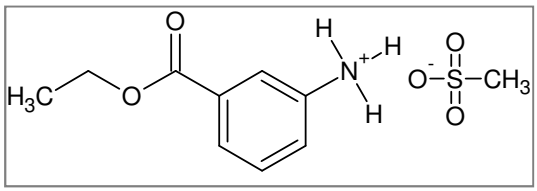


Figure 4.12: Chemical structure of tricaine mesilate (MS222).

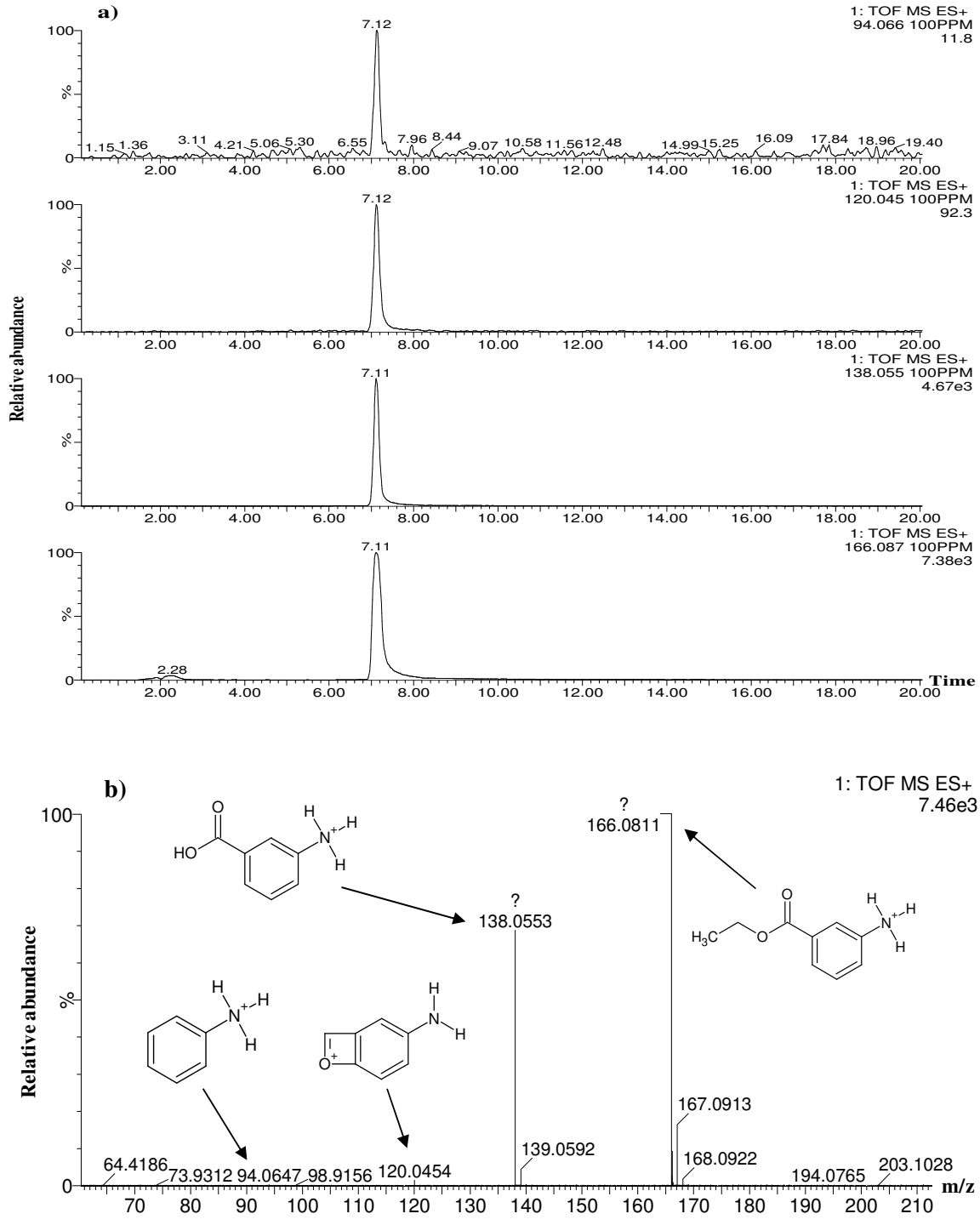


Figure 4.13: a) Full scan extracted ion chromatograms and b) spectrum of the in-source fragmented MS222 in plasma of trout in +ESI mode (collision energy = 10 eV).

4.3.3.4 Bile acids

Three different bile acids were detected in plasma from fish exposed to effluent wastewater: taurocholic acid (TC), taurochenodeoxycholic acid (TDC) and cyprinol sulphate (putative identity). TC and TDC were detected in both ESI modes (Table 4.2 and Table 4.3) whilst cyprinol sulphate could be detected only in –ESI mode (Table 4.2). Their intensities were increased in plasma samples after 10 days of effluent exposure (fold change and *p*-value for TC= 9.5;3.0×10^{-6}, for TDC 9 and 3.0×10^{-3} and for cyprinol sulphate 12.4 and 2.0×10^{-4}). Only the identity of TC and TDC was further confirmed by comparison to standards (retention time and mass spectra) since no standard for cyprinol sulphate was commercially available. Furthermore, TC and TDC standards were also analysed by Q-TOFMS experiments to obtain accurate mass for the fragments and the informative ions were used to confirm and characterize these bile acids in both ESI modes (Appendix 4.4).

Bile acids were effectively ionized in -ESI mode, producing a prominent $[M-H]^-$ ion at *m/z* 514.2841 (TC), *m/z* 498.2884 (TDC) and *m/z* 531.2992 (cyprinol sulphate), respectively.

Q-TOFMS spectra of the deprotonated ion for cyprinol sulphate (*m/z* 531.2992) showed a characteristic fragment at *m/z* 96.9596, which corresponds to the deprotonated sulphate group $[HSO_4]^-$, and an ion at *m/z* 513.2886 corresponding to loss of H_2O (Figure 4.14). High collision energy (50eV) was applied in order to obtain a detectable signal at *m/z* 96.9596 and no further fragmentation was observed. This poor fragmentation behaviour could be explained by the high stability of the molecular ion. Full characterization of this bile acid structures will be described in Chapter 5 because the signal for this compound in trout plasma was too poor to allow further detailed study.

The diagnostic fragmentation of taurine conjugates (TC and TDC) was obtained after selection of the corresponding deprotonated ion for Q-TOFMS experiments. The relative fragmentation pattern showed fragment ions at m/z 353.2481, *m/z* 124.0068, *m/z* 106.9803, and *m/z* 79.9568 (only for TC). These ions correspond to the loss of taurine moiety plus $2H_2O$ from the molecular ion $[C_{24}H_{34}O_2]^-$, taurine moiety $[C_2H_6NO_3S]^-$, ethenesulfonate $[C_2H_3O_3S]^-$, and sulphate $[SO_3]^-$, respectively. Neutral loss of H_2O from the molecular ion was also observed at *m/z* 496.2724 in the –ESI Q-TOFMS spectra (Figure 4.14).

In +ESI mode, three markers ions (different *m/z* but same RT) were assigned to TC (molecular ion $C_{26}H_{46}NO_7S$ at *m/z* 516.2991, loss of H_2O $C_{26}H_{44}NO_6S$ at *m/z* 498.2889 and loss of $2H_2O$ $C_{26}H_{42}NO_5S$ at *m/z* 480.2786, respectively) and one to TDC (loss of $2H_2O$ $C_{26}H_{44}NO_6S$ at *m/z* 464.2841). An ion at *m/z* 462.2667 ($C_{26}H_{40}NO_4S$) corresponding to the loss of 3 molecules of H_2O from the molecular ion was also present in the TC mass spectrum but was not indicated as marker by S-plot analysis. The most abundant ion amongst the three markers proposed for TC was the ion at *m/z* 498.2889 (loss of H_2O).

There was not enough plasma sample to undertake Q-TOFMS analysis so in order to confirm the fragmentation pattern of these bile acids in +ESI mode and potentially obtain additional structural information, high energy collisional dissociation fragmentation in full scan mode was performed on the same plasma samples previously profiled. A range of collision energies (15eV to 20eV) was used to obtain different fragmentation patterns for the molecular ions. As clearly shown in Figure 4.15 for TC molecule, the extracted ion chromatogram (XIC) for the in-source generated fragment ion at *m/z* 498.2889 (loss of H_2O) was aligned with the XICs of the fragment ions at *m/z* 480.2784, *m/z* 462.2689, *m/z* 337.2533, and *m/z* 126.0221 which correspond to loss of $2H_2O$, loss of $3H_2O$, loss of taurine moiety, and taurine moiety itself, respectively. TDC gave rise to the ion at *m/z* 337.2533 using a collision energy of 20eV, which corresponds to loss of taurine moiety from the fragment ion at *m/z* 464.2835 (loss of $2H_2O$) (Figure 4.16).

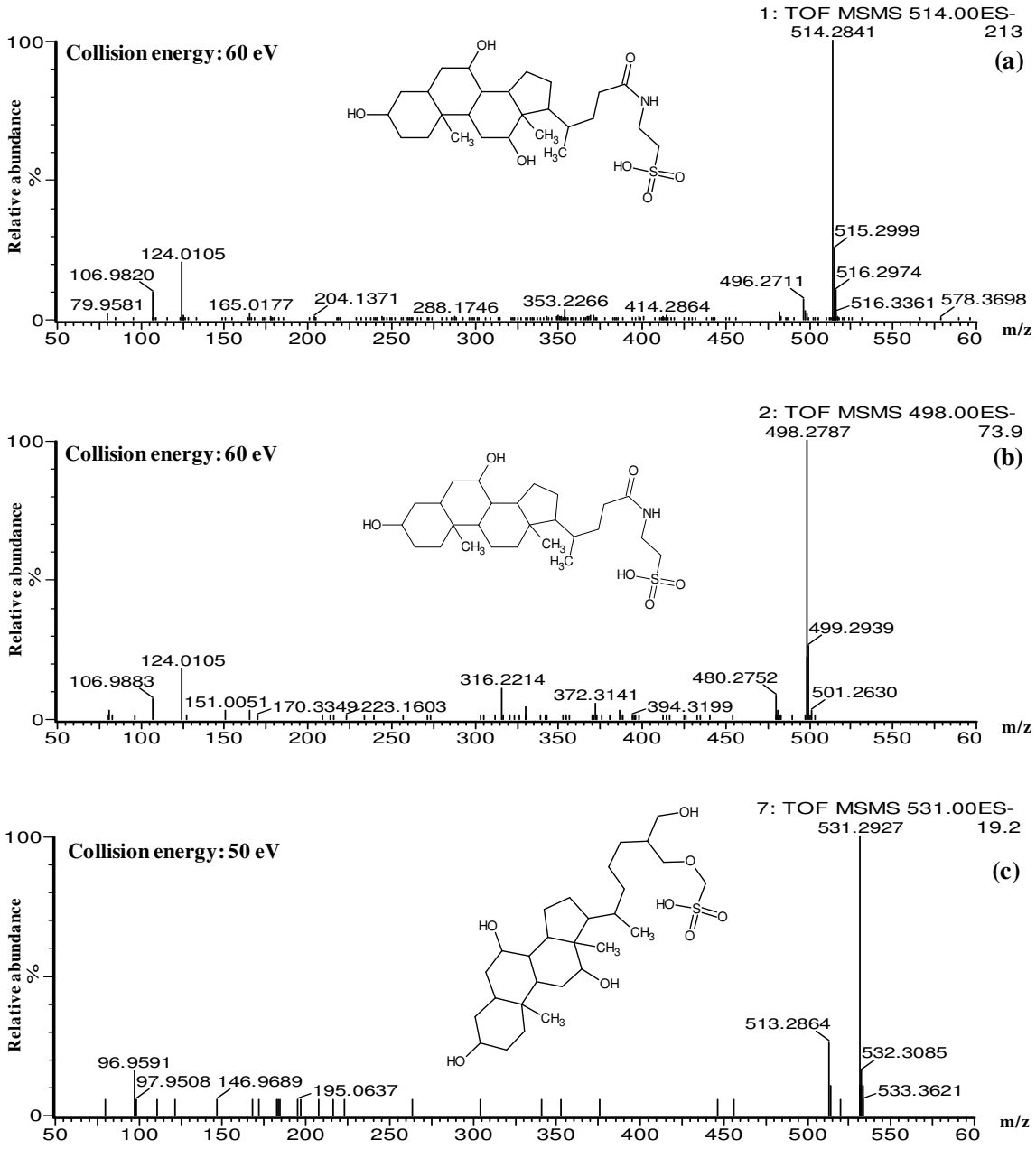


Figure 4.14: Q-TOFMS spectra for: a) taurocholic acid, b) taurochenodeoxycholic acid and c) cyprinol sulphate in plasma from effluent-exposed fish in -ESI mode.

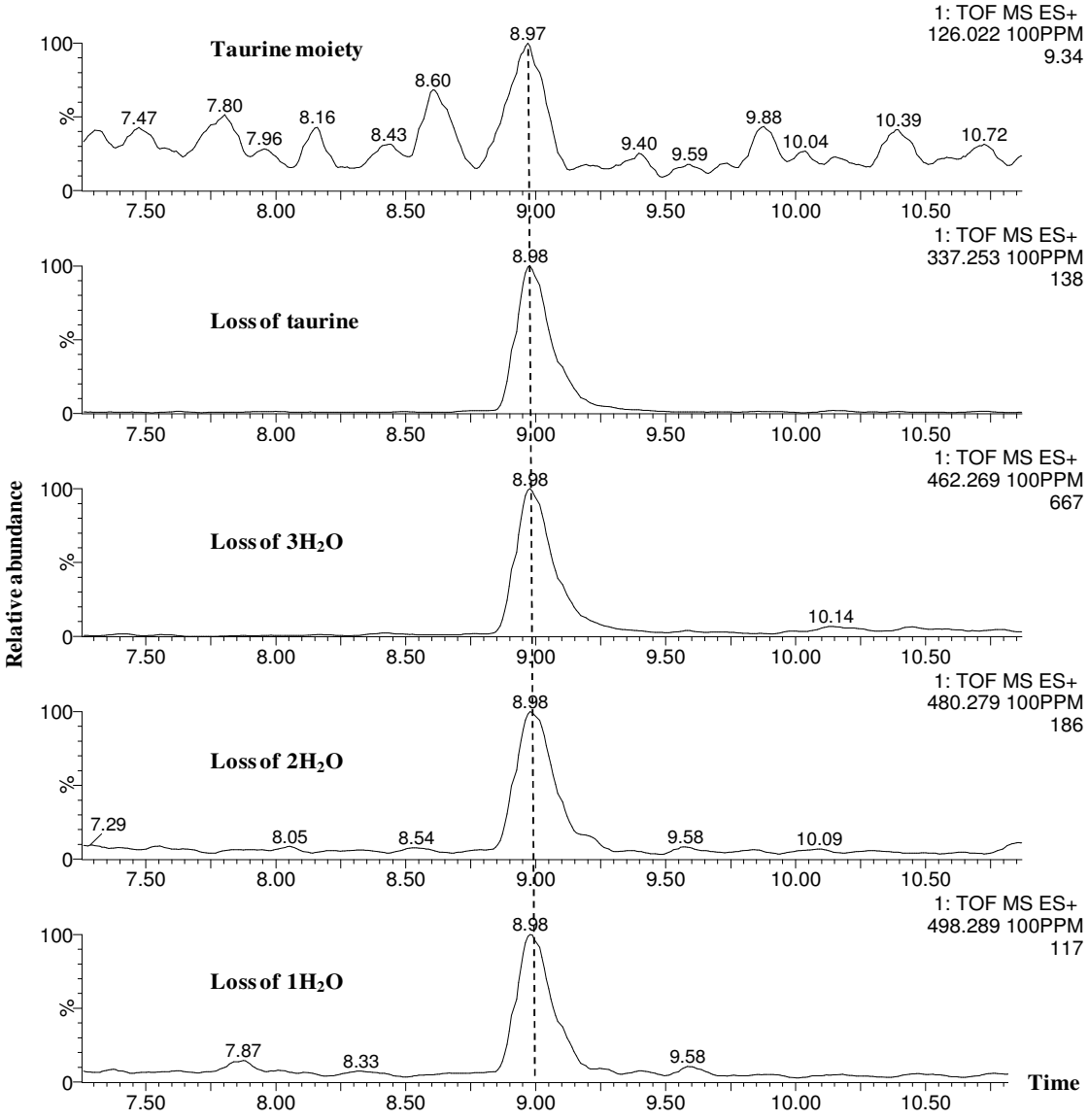


Figure 4.15: Full scan extracted ion chromatograms of the in-source fragmented taurocholic acid in plasma from effluent-exposed trout in +ESI mode (collision energy = 20 eV).

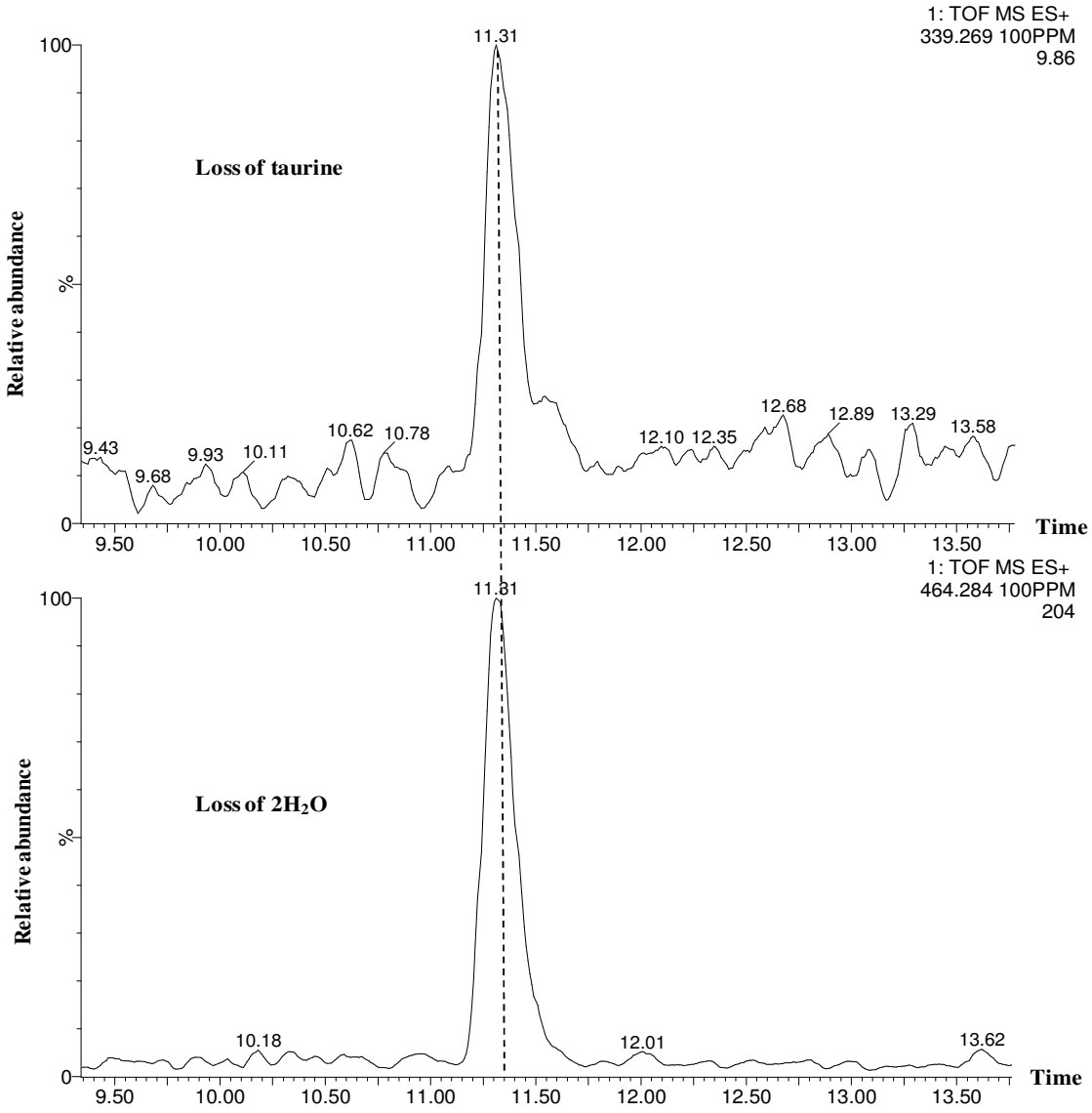


Figure 4.16: Full scan extracted ion chromatograms of the in-source fragmented Taurochenodeoxycholic acid in plasma from effluent-exposed trout in +ESI mode (collision energy = 20 eV).

4.3.3.5 Acylcarnitine

The detected acylcarnitine was putatively identified as methylbutryolcarnitine and it increased (3-fold; p= 3.04×10^{-6}) in plasma of effluent-exposed fish. Methylbutryolcarnitine was detected in +ESI mode as the protonated molecule [M+H] at *m/z* 246.1707 (Table 4.3). A careful examination of methylbutryolcarnitine spectrum obtained from the UPLC-TOFMS profiling led to the detection of a small peak at m/z 187.0970 at the same RT corresponding to $C_9H_{15}O_4$ (Figure 4.17).

Q-TOFMS experiments selecting the ion at *m/z* 246.1707 as parent ion also confirmed a small fragment ion at *m/z* 187.0975 (Figure 4.18). The neutral loss of 59 Da can be related to the loss of C_3H_9N (corresponding to the polar head of the molecule). Since the relative commercial standard was not available, in order to further investigate this molecular structure, Q-TOFMS experiments of a similar standard (O-acetyl-L-carnitine) were performed (Figure 4.18). Results showed the presence of both characteristic fragment ion at *m/z* 85.0279 and characteristic loss of 59 Da as also described for the fragmentation of carnitines by McClellan et. al. (2002). However, exact structure elucidation can only be confirmed by further MS^n experiment. The daughter ion at *m/z* 187.0996 could be fragmented resulting in another generation of daughter ions to assist the structure elucidation. However, this experiment could not be performed in the present study.

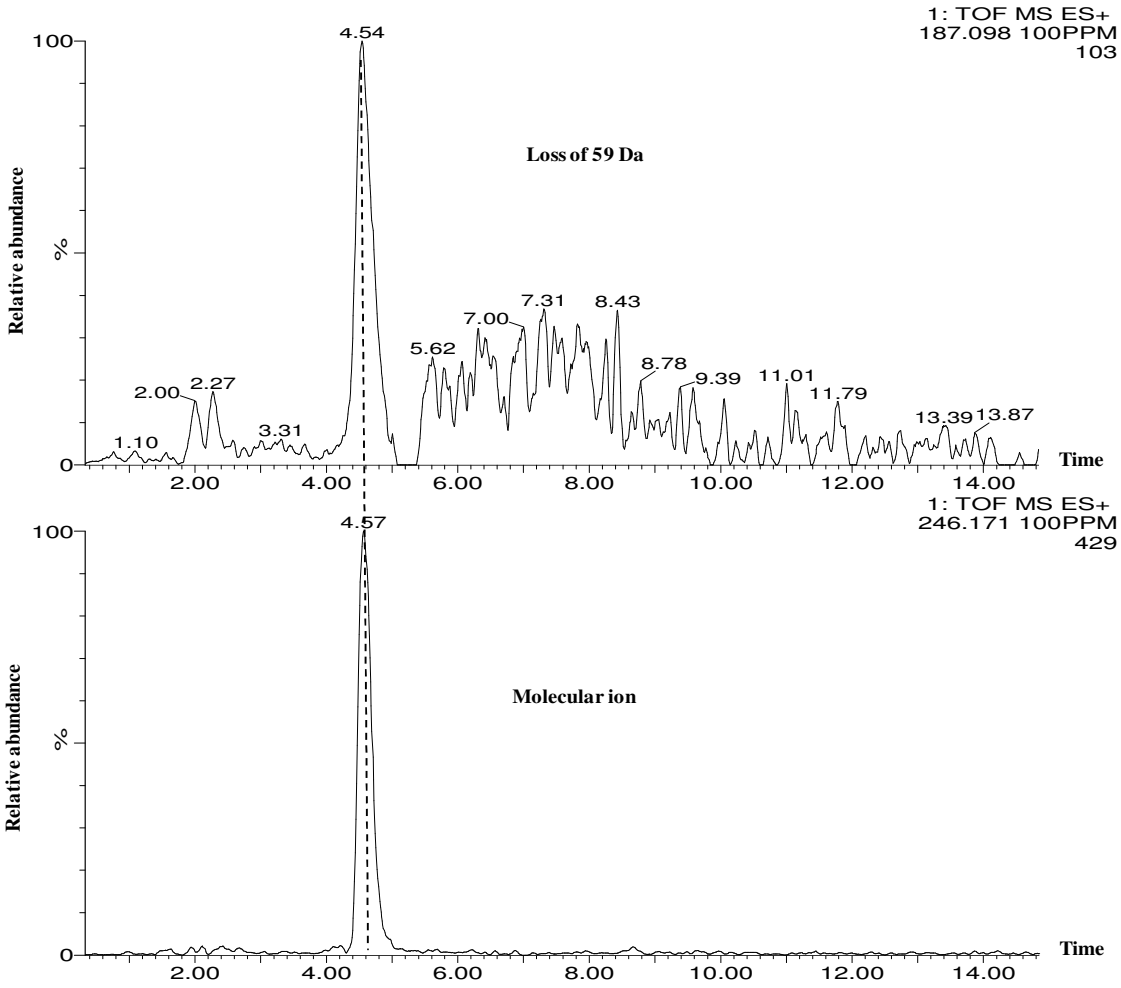


Figure 4.17: Full scan extracted ion chromatograms for putative methylbutryolcarnitine in plasma from effluent-exposed trout in +ESI mode as molecular ion *m/z* 246.1710 and in-source fragment *m/z* 197.0980 (collision energy = 10 eV).

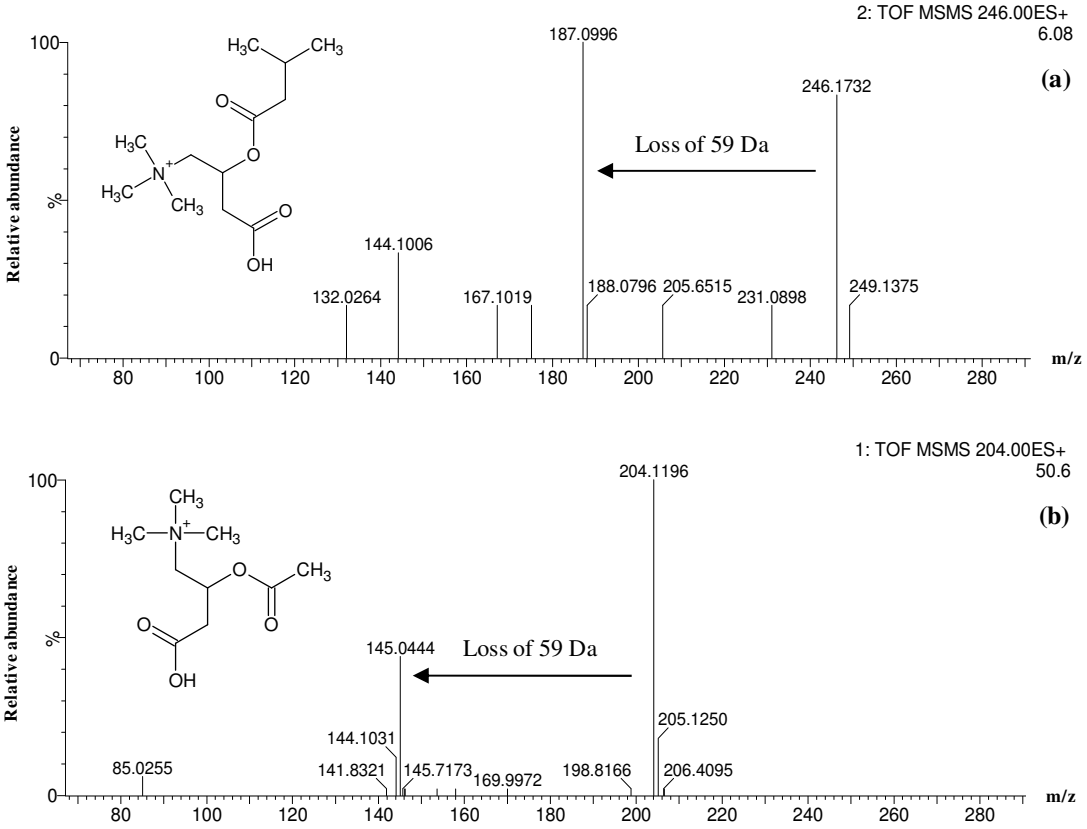


Figure 4.18: Comparison of Q-TOFMS spectra for acylcarnitine (methylbutryolcarnitine) in (a) plasma from effluent-exposed trout and (b) standard acetyl carnitine in +ESI mode (collision energy = 15 eV).

4.3.3.6 Sphingolipids

Sphingosine revealed a concentration decrease (2-fold; p= 6.89×10^{-06}) when compared to the control level. It was detected by S-plot as potential marker only in +ESI mode as molecular ion at *m/z* 300.2900 (Table 4.3). A very abundant in-source ion at *m/z* 282.2800 due to loss of H_2O from the molecular ion (CE 10 eV) was also proposed as marker at the same RT. Q-TOFMS experiment of the sphingosine molecule in +ESI mode showed a very abundant ion at *m/z* 282.2797 ($C_{18}H_{35}NO$), corresponding to the loss of H_2O. Other fragment ions at *m/z* 264.2691 ($C_{18}H_{33}N$) and *m/z* 252.2691 ($C_{17}H_{33}N$), corresponding to loss of H_2O and formaldehyde group respectively, were also present in the fragmentation pattern (Figure 4.19). METLIN database assisted in the confirmation of the obtained fragmentation pattern for this biomarker, and in order to fully characterize its structure, a comparison to sphingosine standard was then performed, showing same RT, exact mass and MS-MS fragmentation pattern as the sample.

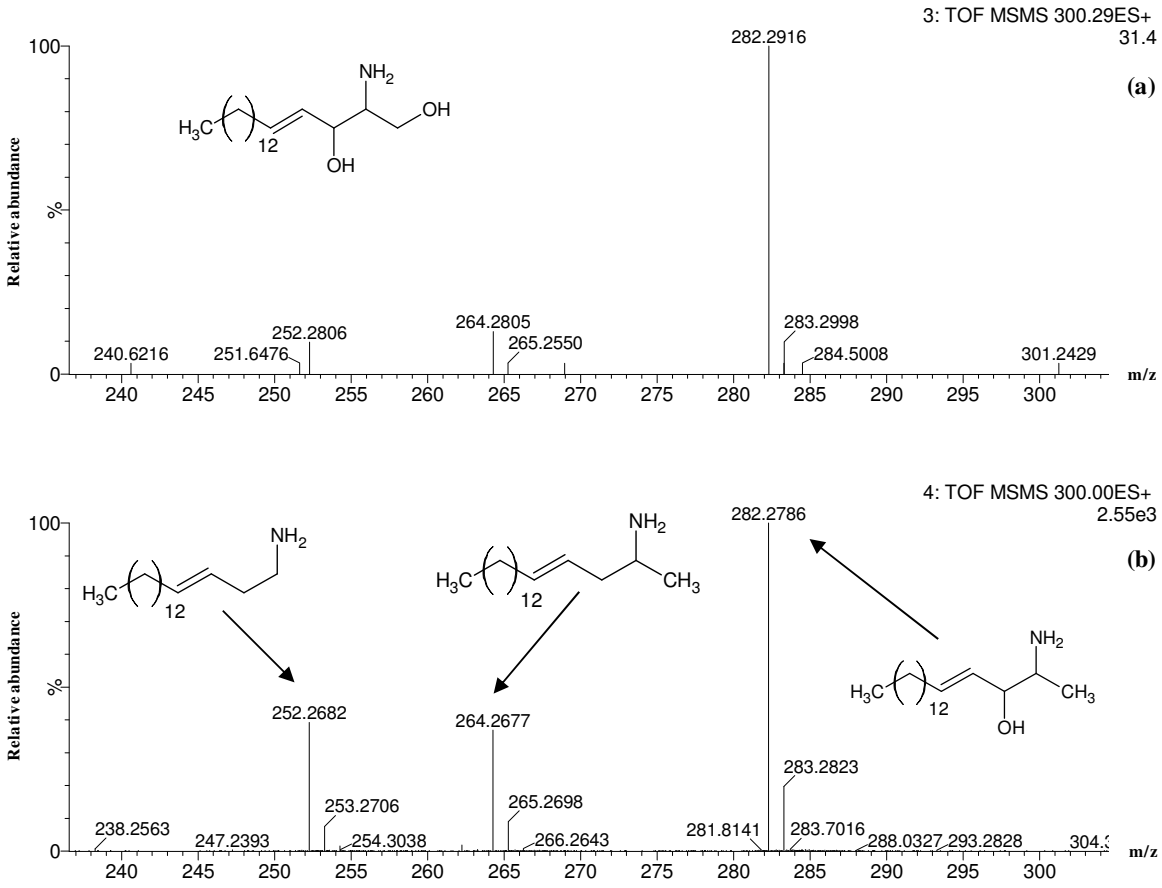


Figure 4.19: Comparison of Q-TOFMS spectra for sphingosine and structures of relative fragment ions in (a) plasma from effluent-exposed trout and (b) standard compound in +ESI mode at collision energy of 15 eV.

Table 4.4 summarizes the identified markers detected in exposed fish plasma. Other contaminants and possible biochemicals were also detected as potential markers but they could not be fully characterized either due to poor signal in the sample or limit of time. However, some of them appeared to belong to the classes of bile acids, ceramides and phospholipids. Further work is needed though in order to confirm the identity of these potential markers (see Appendix 4.3 for full list of unidentified markers).

Table 4.4a: Chemical markers present in plasma of trout exposed to a WwTW effluent.

Class of chemical	Chemical identity	Fold change[#] C_{10} vs E_{10}	*p*-value C_{10} vs E_{10}	Relative abundance in E_{10} in either +ESI or –ESI datasets (mean±S.E)	% Decrease during depuration period	
					4 day	11 day
Chlorinated phenol metabolites	Sulphate conjugate of a trichlorophenol	23.6±5.4	2.5×10^{-4}	0.4±0.1	100	100
Chlorinated xylenol metabolites	Chloroxylenol glucuronide	91.4±7.0	3.3×10^{-9}	1.5±0.1	100	100
	Glucuronide conjugate of a methoxy metabolite of chloroxylenol	29.2±1.7	3.3×10^{-9}	0.5±0.0	100	100
	Chloroxylenol sulphate	6.4±2.9	5.1×10^{-7}	0.1±0.0	100	100
	Sulphate conjugate of a methoxy metabolite of chloroxylenol	73.1±5.3	3.3×10^{-9}	1.2±0.1	100	100
Chlorophene metabolites	Chlorophene glucuronide	54.4±3.7	3.3×10^{-9}	0.9±0.1	100	100
	Putative chlorophene sulphate	12.6±3.8	7.0×10^{-3}	0.2±0.1	100	100
Chlorinated phenoxyphenols	Triclosan glucuronide	46.8±3.1	3.3×10^{-9}	0.8±0.1	100	100
	Putative diclosan sulphate	27.5±3.4	5.1×10^{-7}	0.5±0.1	100	100
	Triclosan sulphate	369.6±37.6	3.3×10^{-9}	6.2±0.6	88.8	92.3
Linear alkylbenzene sulphonic acid (LAS)	C12- LAS	2.6±0.4[¶] (627.5)	4.6×10^{-7}	10.6±0.8	49.5	67.5
	C13-LAS	7.6±1.4[¶] (3374.6)	3.3×10^{-9}	57.0±6.1	84.6	89.0
Metabolites of nonylphenol polyethoxylate (NPEO)	Nonylphenol glucuronide	86.7±14.1	5.1×10^{-7}	1.5±0.2	100	100
Alcohol polyethoxylate (AEO)	Tridecanol-3EO*	73.1±20.2	2.0×10^{-3}	0.2±0.1	100	100
	Tetradecanol-3EO*	65.1±20.1	3.3×10^{-9}	0.4±0.1	100	100
	Tetradecanol-4EO*	97.3±25.7	5.7×10^{-8}	0.5±0.1	100	100
	Tetradecanol-5EO*	187.6±34.4	3.3×10^{-9}	1.0±0.2	100	100
	Tetradecanol-6EO*	9.1±4.2[¶] (116.0)	3.3×10^{-9}	0.6±0.2	87.6	92.6
	Tetradecanol-7EO*	4.1±2.0[¶] (60.7)	1.1×10^{-6}	0.3±0.1	50.0	59.0
	Pentadecanol-4EO*	89.4±14.6	3.3×10^{-9}	0.5±0.1	100	100
	Pentadecanol-5EO*	121.4±15.2	3.3×10^{-9}	0.7±0.1	100	100
	Pentadecanol-6EO*	302.0±28.2	3.3×10^{-9}	0.7±0.1	100	100
	Pentadecanol-7EO*	115.0±10.5	3.3×10^{-9}	0.6±0.1	100	100
Alcohol polyethoxy carboxylates (AECs)	Tetradecanol-4EC*	89.3±26.5	3.3×10^{-9}	0.5±0.1	100	100
	Tetradecanol-5EC*	358.9±40.6	3.3×10^{-9}	0.3±0.1	100	100
	Tetradecanol-6EC*	111.9±21.3	6.8×10^{-5}	0.3±0.1	100	100
	Tetradecanol-7EC*	32.1±12.6	4.3×10^{-2}	0.9±0.0	100	100
	Pentadecanol-4EC*	50.4±16.3	7.0×10^{-3}	0.3±0.1	100	100
	Pentadecanol-5EC*	41.1±12.3	3.3×10^{-9}	0.2±0.1	100	100
	Pentadecanol-6EC*	89.5±5.4	3.3×10^{-9}	0.2±0.0	100	100
	Pentadecanol-7EC*	65.7±4.8	3.3×10^{-9}	0.1±0.0	100	100
Steroid alkaloid	Solanidine	84.7±8.8	3.3×10^{-9}	84.7±8.8	100	100
	Dihydrosolanidine	37.7±7.1	6.8×10^{-5}	37.7±7.1	100	100

C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; *p* values were based on one tank only; Bonferroni correction threshold used was 4.1×10^{-6} in –ESI and 3.7×10^{-6} in +ESI.

[¶] Compounds detected in the control plasma samples (fish exposed to river water) at levels > LOD, where LOD=0.0054 in –ESI; LOD=0.0169 in +ESI. LOD units are relative to the intensities of the normalized total mass spectra signals in the mass chromatograms of the datasets (concentration were detected in reference trout due to exposure to contaminated control river water).

* Markers detected in +ESI mode as $[M+Na]^+$; all other markers detected in –ESI mode as $[M-H]^-$.

[#] Fold change values are reported as mean±standard error (mean±S.E); in brackets are fold change values assuming an LOD value in the control samples was detected.

Table 4.4b: Biochemical markers present in plasma of trout exposed to a WwTW effluent.

Class of chemical	Chemical identity	Fold change[#] C_{10} vs E_{10}	*p*-value C_{10} vs E_{10}	Relative abundance in E_{10} in either +ESI or –ESI datasets (mean±S.E)	% Metabolite change	
					4 day	11 day
Bile acids	Taurocholic acid	9.5±5.1[¶]	3.0×10^{-6}	29.2±9.3	88.4↓	89.7↓
	Taurochenodeoxycholic acid	9.0±4.2[¶]	3.0×10^{-3}	8.9±3.7	91.5↓	87.0↓
	Cyprinol sulphate	12.4±9.2[¶]	2.0×10^{-4}	1.0±0.3	75.8↓	100↓
Acylcarnitines	2-Methylbutyroylcarnitine*	3.0±0.4[¶]	3.0×10^{-6}	1.0±0.1	64.5↓	84.2↓
Sphingolipids	Sphingosine*	0.5±0.1[¶]	6.9×10^{-6}	0.1±0.0	37.8↑	33.0↑

C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; *p* values were based on one tank only; Bonferroni correction threshold used was 4.1×10^{-6} in –ESI and 3.7×10^{-6} in +ESI.

[¶] Compounds detected in the control plasma samples (fish exposed to river water) at levels > LOD, where LOD=0.0054 in –ESI; LOD=0.0169 in +ESI. LOD units are relative to the intensities of the normalized total mass spectra signals in the mass chromatograms of the datasets.

* Markers detected in +ESI mode as $[M+H]^+$; all other markers detected in –ESI mode as $[M-H]^-$.

[#] Fold change values are reported as mean±standard error (mean±S.E)

↑Metabolites where depuration period caused a percentage increase; all other metabolites reported as percentage decrease ↓.

4.3.4 The rate of elimination of markers of effluent exposure from trout blood

The rate of elimination of the metabolites in the fish blood during the depuration periods were calculated as the percent decrease in the metabolite concentration in the fish blood relative to the amounts present after 10 days of effluent exposure (see Table 4.4). Many of the xenobiotic compounds could not be detected after 4 days depuration indicating that they were rapidly eliminated from the fish blood (equivalent to 100% decrease in concentration during the depuration period). However other compounds such as the LAS metabolites, and certain long chain tetradecanol-EO surfactants were detected in the plasma after 4 days depuration. These metabolites, e.g. C12-LAS and tetadecanol-7EO, were also present after 11 days depuration. A high percentage clearance were observed after 4 days of depuration for most compounds (ranging between 84% and 100%) with the exception of the non-ionic surfactant C14(EO)7 (46%) and C12-LAS (49%). After 11 days of depuration, 100% elimination was achieved for all metabolites at this sampling time with exception of C14(EO)6 (92%), C13-LAS (89%), C14(EO)7 (67%) and C14(EO)7 (59%). This suggests that these xenobiotics were rapidly metabolised and/or eliminated from the fish body. The relative concentrations in the analyzed plasma for the endogenously derived markers TC and TDC were reduced by at least 88% after 4 days of depuration in river water and by 87% after 11 days of depuration. This result indicates that 4 days of depuration is the time required by the organism to restore to near normal physiological status after changes due to the effluent exposure. In the case of cyprinol sulphate, more than 75% of the original amount found in the plasma of fish exposed to effluent for 10 days was reduced

after 4 days of depuration whilst the compound was back to normal levels after 11 days of depuration. This trend also suggests that the effluent exposure induced temporal increases in cyprinol sulphate blood levels in female trout under the exposure conditions used. Amounts of 2-methylbutyroylcarnitine were reduced by over 64% after 4 days of depuration and more than 82% after 11 days in river water. The relative concentration of sphingosine was increased by at least 37% and 33% after 4 days and 11 days, respectively.

4.4 Discussion

Many studies have already demonstrated that bile can efficiently concentrate a varaity of contaminants from the environment at levels which facilitate chemical analysis and quantification (Larsson et al., 1999, Gibson et al., 2005a). The levels of contaminants in the bile are often much higher (up to 1000 fold) when compared to the corresponding circulating blood levels (Smith and Hill, 2006). However the analysis of blood or plasma is necessary in order to determine which contaminants are actually present in the circulatory system and can be therefore bioavailable to the other tissues.

The xenometabolomes obtained from bile and plasma samples of effluent-exposed fish showed some similarities and some differences (see Tables 3.11 and 4.4). Glucuronide conjugates of some chlorinated compounds (chloroxylenol, chlorophene, triclosan) and nonylphenol were detected in both matrices. However other xenobiotics that were detected in bile (as glucuronide conjugates) such as mefenamic acid, resin acids, oxybenzone and naphthols were not detected as markers of effluent exposure in the plasma. In addition, not all the surfactants or their metabolites that were found in the bile were detected in the plasma. For instance SPCs, NPEOs, dodecanol EOs, short chain ethoxymers of the other AEOs, and tridecanol ECs were not detected in the plasma as either conjugates or the parent compound. This is despite that observation that tetradecanol 3-7EOs, tetradecanol and pentadecanol ECs were detected in both bile and plasma as either the free parent compound or as the glucuronide conjugate. The plasma samples had been concentrated 20 fold unlike the bile samples which were analysed diluted i.e. 1 µL of bile equivalent and 20 µL of plasma equivalents were profiled on column. Despite preconcentration of the plasma samples, the data suggest that a number of xenobiotics that could be present in the plasma were not detected, either as a result of too low concentrations in the blood, or due to ion suppression caused by the plasma matrix. This result highlights the need of further optimization for

the method used in this study in order to maximize the concentration and detection of the full range of factor for significant markers. Sulphate conjugates of some xenobiotics were detected only in plasma samples and these were the conjugates of chloroxylenol, triclosan, diclosan, chlorophene and trichlorophenol. Their absence in the bile could be due to their excretion via the urine, especially as they are small molecules.

It is well established that xenobiotic conjugates whose molecular weight are usually above the threshold for biliary elimination, are preferentially translocated into bile and only a small proportion of the xenobiotic mass that is conjugated as the glucuronide (as well as sulphate conjugates) can be transferred to the systemic circulation and to finally reach the kidneys to be excreted via urine (Hirom et al., 1972, Chipman and Walker, 1981). In this study, glucuronide conjugates of triclosan, chlorophene and chloroxylenol were present in the bile (corresponding to molecular weights of 462, 393 and 331) whereas both the glucuronide and sulphate conjugates (molecular weight of sulphates 336, 296 and 234) of these xenobiotics were predominant metabolites in the plasma samples.

As explained in details in section 4.4, the cofactor for the sulphonation reaction is 3′-phosphoadenosine-5′-phophosulphate (PAPS) (Figure 4.5). This compound is synthesized from adenosine triphosphate (ATP) and inorganic sulphate, and because two molecules of ATP are needed in this biotransformation process (a high energy content is required) (Falany, 1997), the cellular levels of this cofactor are usually low as a result of relatively slow rate of PAPS formation. The low level of this cofactor may explain why, in many studies, the relatively low effectiveness of sulphonation compared to other biotransformation pathways when the organism is exposed to an unexpected increase of potential substrates (Di Giulio and Hinton, 2008). Sulphotransferase activity in the channel catfish (Ictalurus punctatus) has already been investigated for phenolic and alcoholic xenobiotic substrates. Sulphonation of phenolic compounds revealed higher activity than alcoholic substrates (Tong and James, 2000). The same results were observed in this study where chlorinated phenolic compounds (e.g. chloroxylenol and chlorophene) were metabolized to their glucuronide and sulphate conjugates while alcoholic surfactants were just detected as parent molecules. In previous studies triclosan was proved to inhibit glucuronidation and sulphonation processes for phenolic xenobiotics (Wang et al., 2004). In this study, the presence of triclosan in the plasma

samples could have affected the metabolism of the other phenolic compounds present in the same matrix.

Although both sulphate and glucuronide conjugates can be directly excreted from the fish via urine or bile, they can also hydrolysed back to the parent xenobiotic and recirculated in the blood system prior to reconjugation and excretion. Sulphatase enzyme activity is generally more relevant in the liver, but sulphation process might become more important in the intestine either when substrate concentrations are very low or when some specific species are involved (e.g. catfish) (Tong and James, 2000, Di Giulio and Hinton, 2008). Glucuronide conjugates can be hydrolysed in the lower gut, and the parent xenobiotic can then pass through the epithelial wall and recirculate in the body prior to reconjugation in the liver (where it may undergo another round of biliary cycling); a process known as enterohepatic circulation (Jandacek and Tso, 2007). These processes can increase the persistence and toxicity of the parent compound within the organism.

The detection of a complex mixture of xenobiotics circulating in the blood of effluent exposed fish suggests that there could be a number of toxicological effects on fish health. For instance, as discussed in Chapter 3, many of the chlorinated compounds are antiandrogenic and have the capacity to disrupt sexual differentiation of male fish. Triclosan can also disrupt lipid biosynthesis in bacteria and may target related enzymes in higher organisms. Specifically, triclosan inhibits bacterial fatty acid synthesis binding at the enoyl-acyl carrier protein reductase (FabI) as primary site for action (Heath et al., 1999). Binding of triclosan to enoyl reductases has been reported in both gram-positive (*Staphylococcus aureus*) and gram-negative (*Escherichia coli*, *Pseudomonas aeruginosa*, and *Haemophilus influenzae*) bacteria (McDonnell, 2007). There is less information on the mode of action and toxicology of chlorophene and chloroxylenol in vertebrates.

Most of the surfactants detected in the plasma were detected as non-conjugated free compounds. The toxicity of non-ionic surfactants mainly depends on their hydrophobicity, which generally increases as the number of ethoxymer units decreases (Boeije et al., 2006, Roberts et al., 2007). Surfactants can interact with the biological membranes modifying the lipid and protein metabolism and therefore causing increase in permeability and disruption of trans-membrane solute transport. (Müller et al., 1999). The mechanisms of toxicity of AEOs are similar to other non-ionic surfactants acting as

general narcotics, and their toxicity depends on their hydrophobicity, generally increasing with the decreasing of EO content (Boeije et al., 2006, Roberts et al., 2007). Previous studies proved that the structure of mitochondria in gill tissue can be altered due to exposure to surfactants, leading to a decrease in the oxygen consumption (Cardellini and Ometto, 2001).

Anionic surfactants may behave differently to non-ionic surfactants, Verge *et al.* (2001) tested LAS acute toxicity to *D. magna* and toxicity increased either with the alkyl chain length or with the homologue molecular weight. This behaviour was explained by the higher interaction of heavier homologues with cell membranes.

The analysis of plasma extracts of effluent-exposed fish also revealed that some endogenous metabolites were potential markers of effluent exposure (Table 4.4). These included bile acids, a sphingolipids and an acyl carnitine metabolite. Bile acids (BAs) are the major components of bile and are synthesized in hepatocytes starting from the cholesterol molecule. BAs are usually conjugated with glycine or taurine, and secreted into the small intestine via bile (Chiang, 2003, Hofmann, 1999). Bile acids have important roles in the cholesterol homeostasis, lipid absorption, and intestinal signalling (Gu et al., 1992). Recently, BAs have been recognized as signaling molecules, which can activate several target genes, such as the farnesoid X receptor, the pregnane X receptor and the vitamin D receptor as well as selected G protein receptor and a number of other signalling pathways in liver cells (Chiang, 2002, Zollner et al., 2006). However, BAs can also be cytotoxic (Latta et al., 1993) and can cause hepatotoxicity (Hofmann, 1999). Taurocholic acid (TC) is the major component of bile acids in rainbow trout and to a lesser extent so is taurochenodeoxycholic acid (TDC) (Yeh and Hwang, 2001). Cyprinol sulphate is a bile acid specific to carp bile (detailed in chapter 5); however, this component was also found at detectable levels in juvenile female trout. This research revealed a significant increase of all three bile acids levels in plasma from effluent-exposed fish and this trend may indicate liver and gastrointestinal diseases in the animal, which might be partially related to the effluent exposure. As bile acids can be released from dysfunctional hepatocytes into the blood, cholestasis and other types of liver injury can also increase bile acids concentrations in plasma (Pauli-Magnus and Meier, 2006, Zollner and Trauner, 2008). It is possible that exposure to the xenobiotics such as the surfactants in the effluent have resulted in disruption of liver cell membranes and also disrupted transport of bile acids from the hepatocyte to the bile canaliculis.

Elevated serum or plasma bile acids have therefore long been used as prognostic and diagnostic markers for liver injury and diseases in mammals (Xiang et al., 2010) and may indicate similar toxicologies in the fish.

Acylcarnitine was also detected as a possible endogenous marker of effluent exposure in the plasma samples. A 3-fold increase of an acylcarnitine was observed in plasma from effluent-exposed trout (putatively identified as 2-methylbutyroylcarnitine). L-carnitine plays a key role in fatty acid oxidation; it facilitates the transport of fatty acids into mitochondria where oxidation takes place (Evans and Fornasini, 2003). It can be conjugated to a range of short-, medium- and long-chain fatty acids to form acylcarnitines. Acylcarnitines have become important biomarkers for diverse kinds of diseases, including inborn errors of metabolism (e.g. deficiency of 3-methylcrotonyl-CoA carboxylase, isobutyryl-CoA dehydrogenase, 2-methyl-3-hydroxybutyryl-CoA dehydrogenase, or 3-ketothiolase) (Pasquali et al., 2006, Jones et al., 2010) and diabetes mellitus type 2 (Möder et al., 2003). Patients with 2-3 times higher of total acylcarnitines (i.e. short-chain and medium chain acylcarnitines) are usually diagnosed with these types of metabolic disorders (Millington et al., 1989). Little is known about the effect of chemical exposure on the levels of acylcarnitines in vertebrates, however their levels are regulated via peroxisome proliferator-activated receptor (PPAR) signalling and the PPAR is activated by a number of different chemical contaminants such as di-(2-ethylhexyl)phthalate (DEPH) (Feige et al., 2009), diisobutyl phthalate (Boberg et al., 2008), trichloroethylene, di- and trichloroacetic acids (Christopher Corton, 2008), bisphenol A (Kwintkiewicz et al., 2009), butylparaben (Boberg et al., 2008), perfluoroalkyl acids (PFAAs) (Wolf et al., 2008), and organotins (Hiromori et al., 2009).

Sphingolipids are constituted of a sphingoid base (sphingosine, sphinganine or other species) which is derived from the combination of serine and a fatty acyl-coenzyme A (Figure 4.20). Sphingolipids can then be converted into more complex compounds (ceramides, phosphosphingolipids, glycosphingolipids and protein adducts) which are keys to the structures of cell membranes, lipoproteins, and the lamellar water barrier of the skin. Many sphingolipids, both complex and simple structures, are highly bioactive as extra- and intra-cellular regulators of growth, differentiation, migration, survival and numerous cellular responses to stress (Hirabayashi et al., 2006). In this study, the concentration of sphingosine was 2 times lower (2-fold decrease) in the

effluent-exposed fish samples than in the controls. This result indicated that exposure to contaminants in the effluent either resulted in a decrease in sphingosine synthesis or an increase in its metabolism to sphingolipids (Figure 4.20). However to date there is no information on whether these pathways are targeted by environmental contaminants.

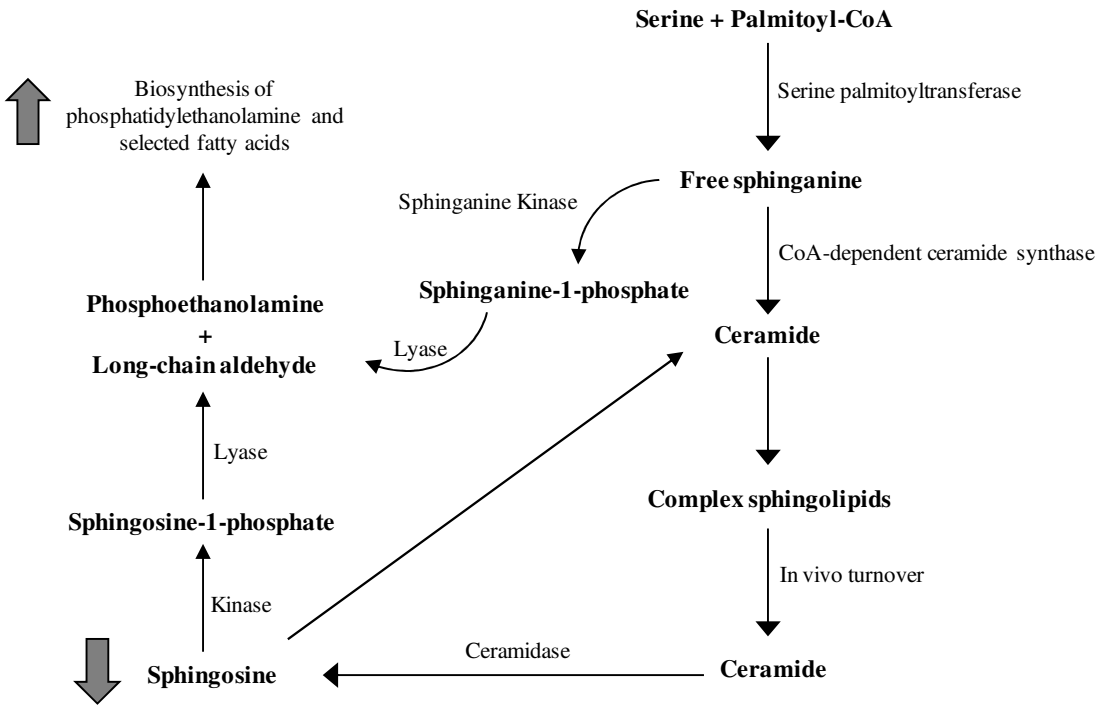


Figure 4.20: A simplified scheme of sphingolipids biosynthetic pathway and their metabolizing enzymes. Bold arrows reflect changes in plasma concentrations of metabolites of effluent-exposed fish as a result of effluent exposure.

4.4.3 Rate of elimination of endogenous compounds and markers of effluent exposure from trout blood

The elimination of many of the xenobiotics detected in the plasma suggested a very low systemic persistence of xenobiotics after exposure to effluent wastewater treatment; this can be due to efficient hepatic metabolism and biliary or urine excretion of their conjugated forms. Some xenobiotics (C12-LAS, tetradecanol-7EO) were present at detectable concentrations even after 11 days of depuration and were therefore not completely eliminated by the organism despite this long depuration period in river water; this could be possibly due to the high abundance of these chemicals present after 10 days of exposure, for which 11 days might have not been enough to fully excrete the contaminants from the organism body. More likely it was also due to the fact that the

river water used during the depuration period contained significant concentrations of these chemicals.

After 11 days depuration the bile acids decreased back to near normal levels indicating the return to the normal physiological status, however sphingosine levels were still markedly decreased in the plasma indicating possible longer term effect on this metabolite.

The selection of endogenous and xenobiotic markers of effluent exposure in the trout plasma was based on univariate statistics using the *p*-value. The *p* values were generally an order of magnitude higher for endogenous markers than for the xenobiotics and ranged from 3.0×10^{-6} – 3.0×10^{-4} for the endogenous markers but in many cases was 3.3×10^{-9} for the exogenous xenobiotic derived markers. However this statistical analysis was based on one tank replicate per treatment and therefore based on pseudoreplication. Therefore this work on the identification of markers of effluent exposure in fish plasma should be repeated with replicate tanks to confirm the findings.

CHAPTER 5: Analysis of the Xenometabolome and Identification of Markers of Chemical Contamination in Plasma from Roach Exposed to a Wastewater Effluent

5.1 Introduction

In the previous two chapters, a chemical profiling approach was developed to detect chemical contaminants and their metabolites bioaccumulated in biofluids of juvenile rainbow trout (*Oncorhynchus mykiss*) after exposure to a wastewater treatment work effluent. In addition in chapter 4 it was found that exposure to WwTW effluent altered the concentrations of endogenously derived metabolites (i.e. bile acids, carnitines and sphingolipids) in plasma of the trout after 10 days of effluent exposure. However, a full toxicological evaluation of the results could not be achieved due to lack of tank replicates and, in order to perform reliable investigation of the alteration of metabolic pathways, higher number of replicates is usually required to support statistical tests. These previous observations were followed up by investigating the impact of effluent exposure on both the xenometabolome and metabolome in plasma of sexually mature roach (*Rutilus rutilus*).

Roach are found in many lowland Europe rivers, and have already been shown to be affected by WwTW effluents in terms of endocrine disruption. Accordingly, Jobling *et al.* (1998) demonstrated a high incidence of intersexuality in populations of roach (*Rutilus rutilus*) in the United Kingdom.

Features of the roach that make it appropriate as a sentinel species to assess effluent exposure impact on fish include (Tyler et al., 2007):

1. It is a member of the carp family (Cyprinidae), one of the most ecologically important groups of freshwater fish, commonly used for ecotoxicological studies.
2. It is a gonochoristic species (developing as either male or female).
3. It is widespread and abundant in the UK and in Europe, where it lives naturally in lowland rivers in waters often contaminated by effluent sewage.
4. It shows tolerance to wide and varying chemical and physical conditions.
5. Its high abundance allows for destructive sampling without adverse effect on the fishery.
6. Good knowledge of it's normal reproductive development has been established.

One of the goals of toxicology is the study of adverse effects due to xenobiotics exposure on living organisms. The toxicity of a certain compound depends on its ability to alter some biological functions at a given level of biological organization (i.e. cell, tissue, or organ). It is usually associated with the amplitude and the duration of the exposure and also to the degree of uptake of the substance by the organism, its distribution, biotransformation and elimination or accumulation (Roux et al., 2011). A full understanding and describing the mechanism of a toxic event in vivo is an extremely challenging task and many *in vitro*, and cell and animal models are used to investigate toxicity mechanism but they cannot be easily extrapolated to real environment conditions. Biomarkers are a useful tool to predict toxic events and to monitor exposure to toxic agents (biomarkers of exposure) including the impacts of complex mixtures of contaminants with the same mechanism of action, e.g. the use of vitellogenin to monitor exposure to mixtures of oestrogenic pollutants.

The impact on fish due to effluent exposure will depend on the proportion of effluent present in the river at the time of the exposure. In the UK, a considerable proportion of the flow of rivers is made up of treated WwTW effluent; a river concentration of 10% WwTW effluent is a common level of contamination, but in some rivers it can reach more than 50% of the flow. In extreme cases in England, and generally in the summer months during periods of low rainfall, treated wastewater effluent can make up the entire flow of some rivers. For this reason, it is important to investigate how the effluent proportion can differently affect the fish health status. For this purpose, in this study two different effluent proportions (50% and 100%) were considered, in order to mimic extreme conditions in the environment.

The study reported in the following chapter aims to:

1. Identify the profile of xenometabolites present in plasma of roach exposed to a WwTW effluent.
2. Identify metabolic biomarkers indicating specific changes in the plasma metabolome caused by the effluent exposure.
3. Determine whether there are sex specific differences in xenometabolite or responses of endogenously derived metabolites of roach to effluent exposure.
4. Investigate the correlation between the levels of xenobiotics and the disruption of metabolite concentrations in plasma of roach exposed to WwTW effluent.

5.2 Materials and Methods

5.2.1 Chemicals

All materials used are reported in Chapter 3 Section 3.2.1 and Chapter 4 Section 4.2.1.

5.2.2 Fish exposure

The wastewater effluent used for this study was the same as described in Chapter 3 Section 3.2.2. A population of sexually mature roach (*Rutilus rutilus*) were obtained from same farm reported in Chapter 3 Section 3.2.2 with a mean±SEM length of 10.85±0.20 cm and weight 22.96±1.20 g. Roach were caged in duplicate tanks and exposed to either charcoal treated river water (n=16 females, n=21 males), 50% effluent from the WwTWs diluted with river water (n=12 females, n=21 males), or 100% effluent (n=17 females, n=23 males) for 28 days. At harvest, fish were anesthetized and blood was obtained from the caudal vein.

5.2.3 Sample treatment

Plasma samples (60µL) were obtained from roach exposed to river water, 50% effluent and 100% effluent. Final ratio of 80% of cold methanol (240 µL) were added to the samples in order to precipitate plasma protein; precipitation were performed in duplicate to ensure remove most of the interfering protein. Internal standards (1ng/10µL) of P-d_9 & E2-d_4-S were added to the samples. Samples were then processed and transferred to HPLC vials prior to chemical analysis as described in Chapter 4 Section 4.2.2.

5.2.4 UPLC-TOF-MS chemical profiling

Plasma samples (10 µL plasma equivalent) were separated using an Acquity UPLC BEH C18 column and the analytes were detected using a Micromass TOF-MS. Samples were analysed by UPLC-TOF-MS system in both ESI modes with acid in the mobile phases performing the same gradient program described in Chapter 2 Section 2.3.2.

5.2.5 Data analysis

Raw data obtained from UPLC-TOF-MS was preprocessed using MarkerLynx. The parameters used for the detection of the spectral signals were optimised similarly as in Chapter 4 Section 4.2.2 with exception of the tolerance of the retention time window which was set at 1.0 min only in the +ESI mode due to machine drift of the UPLC separation for these analyses. Data was then exported to SIMCA P-11 software for

multivariate analysis. Principal Component Analysis (PCA) was performed to provide an overview of the data followed by partial least squares projections to latent structures discriminant analysis (PLS-DA) to find class-separating differences in the data sets. To filter the information that was only due to class separation, the models were reanalyzed using orthogonal partial least squares to latent structures discriminant analysis (OPLS-DA). Shared and unique structure plots (SUS-plots) were also conducted to highlight the differences between two different treatments when compared to a common reference (e.g. control) and to locate the class-distinguished variables (metabolites) contributing to the class separation. SUS-plots were constructed in order to discover the difference between the female and male metabolome in their response to effluent exposure in comparison to the same reference control group. In the SUS-plot, the correlation from the predictive component of each model is plotted against each other using the combination of two predictive loadings obtained from two different OPLS-DA models. The SUS-plot visualizes both the shared and unique information; shared markers are the markers that are varying in the same direction in both models whilst unique markers vary in a unique direction. After the selection of potential markers from the S-plots and SUS-plots, univariate analysis (i.e. Mann-Whitney test) was employed for every variable (marker) to check for significant differences between treatments, and due to the likely number of false positives (Type 1 errors), a false discovery rate (FDR) correction was then applied to determine the significance of the marker. The Benjamini and Hochberg (BH) procedure was applied to control the false positives that pass a statistical test; the individual *p*-value of the variables was sorted from smallest to largest value for the FDR calculation.

The significance level was calculated for each marker as following:

$$= 0.05 \times i/m$$

, where i indicates the ranked *p*-value of the variable and m the total number of variables. This approach starts from the largest *p*-value and continues along all variables in rank order. The difference is considered significant only if it's *p*-value is equal or less than the corrected significance level (Benjamini et al., 2001).

In order to investigate correlations between selected variables, Spearman's correlation of pairwise comparisons among individual variables (metabolites) was employed. This correlation is based on ranked intensity of variables and used when the data have violated parametric assumptions such as non-normally distributed data.

The structural information of each chemical and biochemical marker detected in effluent exposed fish was based on their accurate mass measurements (W mode), in-source fragmentation by applying different collision energy (W mode), and wherever possible performing Q-TOF-MS on selected parent ion and then searching in metabolomic databases for the structural confirmation of unknown markers (detailed in Chapter 3 and 4).

5.3 Results

5.3.1 Multivariate analysis of plasma samples

In order to obtain an overview of the dataset (outliers and trends) from plasma of roach exposed to wastewater effluent, principal component analysis PCA was employed. Scores plots of PCA models for roach plasma datasets using two principal components (t1 and t2 in –ESI mode whilst t3 and t4 in +ESI mode due to instrumental drift) are shown in Appendix 5.1. In –ESI mode, the first principal component (t1) was responsible of the separation of the effluent-exposed fish samples (E_{100}: 100% effluent exposure) from the other two treatment groups (C: river water exposure as control and E_{50}: 50% effluent exposure). Discrimination between the C, E_{50} or E_{100} treatment groups was observed in both +/-ESI modes (Appendix 5.1). However, the percentages of the explained (R^2X <40%) and predicted variation (Q^2 < 0.3) were low in both +/-ESI modes (see Appendix 5.2). Outliers were observed in both +/-ESI mode datasets, but were not excluded for further analysis since their exclusion caused the generation of other outliers in these preliminary PCA studies (data not shown).

PLS-DA models were used for the exploration of the differences between the treatment groups. Clear discrimination between treatments (C, E_{50} or E_{100}) was observed in both +/-ESI modes (Figure 5.1). The first three components (t1, t2 and t3) were used to generate a three-dimensional scores plot of a PLS-DA model in –ESI mode whilst in +ESI mode t3, t4 and t5 were selected since t1 and t2 had to be excluded due to high variability in this dataset due to instrumental drift during UPLC separation.

Response permutation testing was used in order to confirm that the high values of predictability in PLS-DA models had not occurred because of over-fitting. This test

revealed more than 70% accuracy for the classification of treatment groups with the only exception of the dataset in +ESI mode for the model of all the three treatments together which gave a 38% accuracy (see Appendix 5.3).

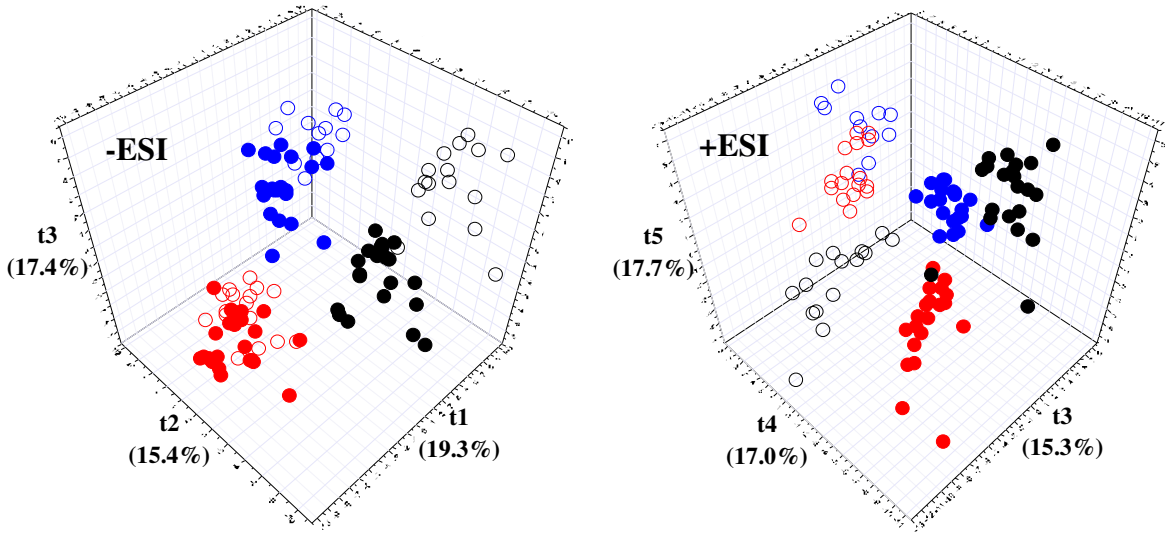


Figure 5.1: Partial least squares-discriminant analyses (PLS-DA) scores plots of the chemical profiles of plasma samples from roach exposed either to wastewater effluent or control river water. The samples were profiled in both +/-ESI modes by UPLC-TOFMS. Control (C)= river water exposure where n=16 females, n=21 males. E_{50}=exposure to 50% effluent where n=12 females, n=21 males. E_{100}= exposure to 100% effluent where n=17 females n=23 males. Female roach symbols: (○), (○) and (○) represent C, E_{50} and E_{100} respectively. Male roach symbols: (●), (●) and (●) symbols represent C, E_{50} and E_{100}, respectively. The percentages of explained variation (R^2Y) modelled for three latent variables (t1, t2 and t3) in –ESI and (t3, t4, and t5) in +ESI mode are displayed on the related axes.

Since duplicate tanks were used for each of the treatment considered, it was necessary to check the consistency between the two sample datasets. No difference was observed in the clustering of the samples between the two tanks in both +/-ESI modes for any of the treatment groups of the plasma from reference roach or those exposed to wastewater effluent (Figure 5.2).

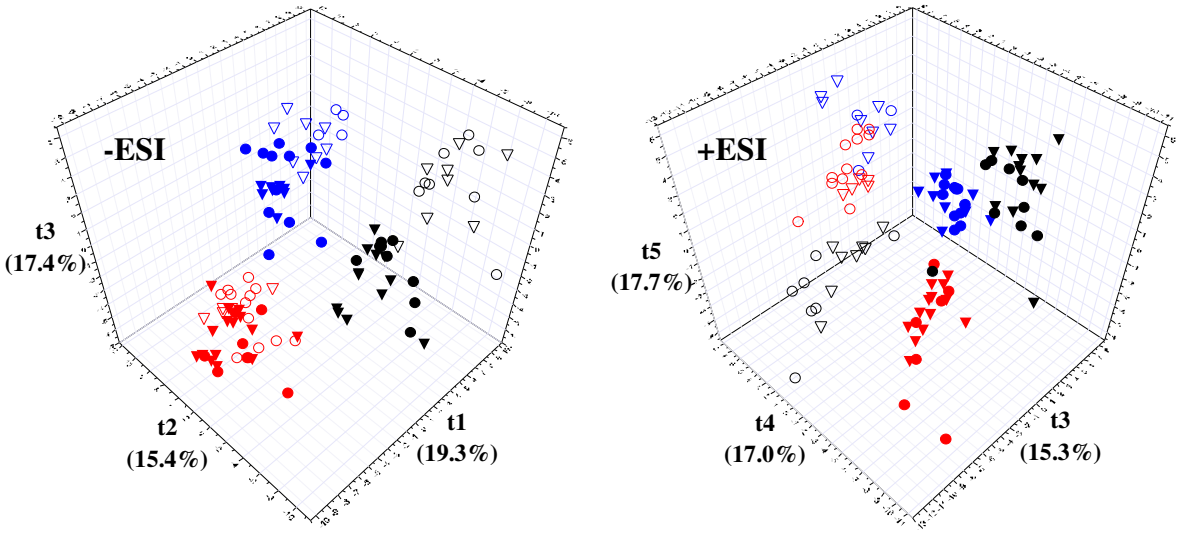


Figure 5.2: Partial least squares-discriminant analyses (PLS-DA) scores plots of the chemical profiles of plasma samples from roach exposed either to wastewater effluent or control river water. The samples were profiled in both +/-ESI modes by UPLC-TOFMS. Fish were kept in duplicate tanks (tank 1: circle, tank 2: triangle). Control (C): river water exposure where n=16 Females, n=21 Males. E_{50}: exposure to 50% effluent where n=12 Females, n=21 Males. $E_{100\%}$: exposure to 100% effluent exposures where n=17 Females, n=23 Males. Female roach symbols (○/▽), (○/▽) and (○/▽) represent C, E_{50} and E_{100}, respectively. Male roach symbols (●/▼), (●/▼) and (●/▼) represent C, E_{50} and E_{100}, respectively. The percentages of explained variation (R^2Y) modelled for three latent variables (t1, t2 and t3) in –ESI and (t3, t4, and t5) in +ESI mode are displayed on the related axes.

5.3.2 Metabolite differences due to WwTW exposure

The most discriminative variables (RT-*m/z*) were extracted from the S plots obtained from the OPLS-DA models for each two treatment groups: (FC versus FE_{50}), (FC versus FE_{100}), (MC versus ME_{50}) and (MC versus ME_{100}), where C is the 28 days river water exposure control, E_{50} the 28 days 50% effluent exposure, E_{100} the 28 days 100% effluent exposure, FC the female control, MC the male control, FE_{50} the female 50% effluent exposure, ME_{50} the male 50% effluent exposure, FE_{100} the female 100% effluent exposure and ME_{100} the male 100% effluent exposure. The most discriminative markers between C, E_{50} or E_{100} for both females and males are summarised in Tables (5.1, 5.2 and 5.3) and Appendix 5.6.

5.3.4 Metabolite differences due to sex

As shown in Figure 5.1, a distinct separation according to the sex (males versus females) was observed in control tanks in –ESI using the first three components (t1, t2 and t3). However, in –ESI mode a PLS-DA scores plot using t2, t3 and t4 components was also checked and it revealed a good separation as well between the two sexes in all the three treatments considered (C, E_{50} or E_{100}) (data not shown). The values of the

predicted and explained variation were improved by reanalysing the datasets using PLS-DA considering the two sexes separately (see Appendix 5.3). In order to determine whether there were sex specific differences in the roach plasma metabolome, OPLS-DA models were employed. The OPLS-DA model for female (n=16) roach versus male (n=21) from the control tank (C, river water) gave a high explained ($R^2Y>0.9$) and predicted ($Q^2>0.4$) variation in –ESI mode, whilst a poor model was obtained in +ESI (Appendix 5.3 and 5.4). The same approach was employed for the 50% and 100% effluent-exposed roach plasma samples to discover the discriminative variables between females and males (Appendix 5.3. and 5.4). These models revealed poor percentages of predicted and explained variation in +ESI mode but models improved in –ESI mode.

Discriminative variables (RT-*m/z*) were found at the extreme of the S-plot obtained from the OPLS-DA model for female control roach versus male controls and also for females and males at the two different effluent exposures. The discriminative variables are listed in Appendix 5.5 but their exact structures were not analysed due to time limitation.

An additional statistical approach "shared and unique structure 'SUS' plot, explained in details in Section 5.2.5, was used to compare two different treatments using the same reference group (Wiklund et al., 2008). SUS-plots were obtained by plotting two OPLS-DA models against each other. In order to test differences due to the sex of the fish, a reference group considered as the sum of female and male control samples was used. SUS plots were obtained by plotting the OPLS-DA model for the reference group and 100% effluent-exposed fish (female) versus OPLS-DA model for the reference group and 100% effluent exposed fish (male), respectively (Figure 5.3a). The same approach was used to test the 50% effluent-exposed samples (Figure 5.3b).

No unique variables from the two SUS plots were extracted and most of the variables associated with effluent exposure were shared i.e. common to both females and males. As the SUS plots did not give informative results, this indicated that there are no variables unique to one sex of the fish but rather (from the S plots obtained from the OPLS-DA models) there were differences in the concentration of metabolites between the two sexes.

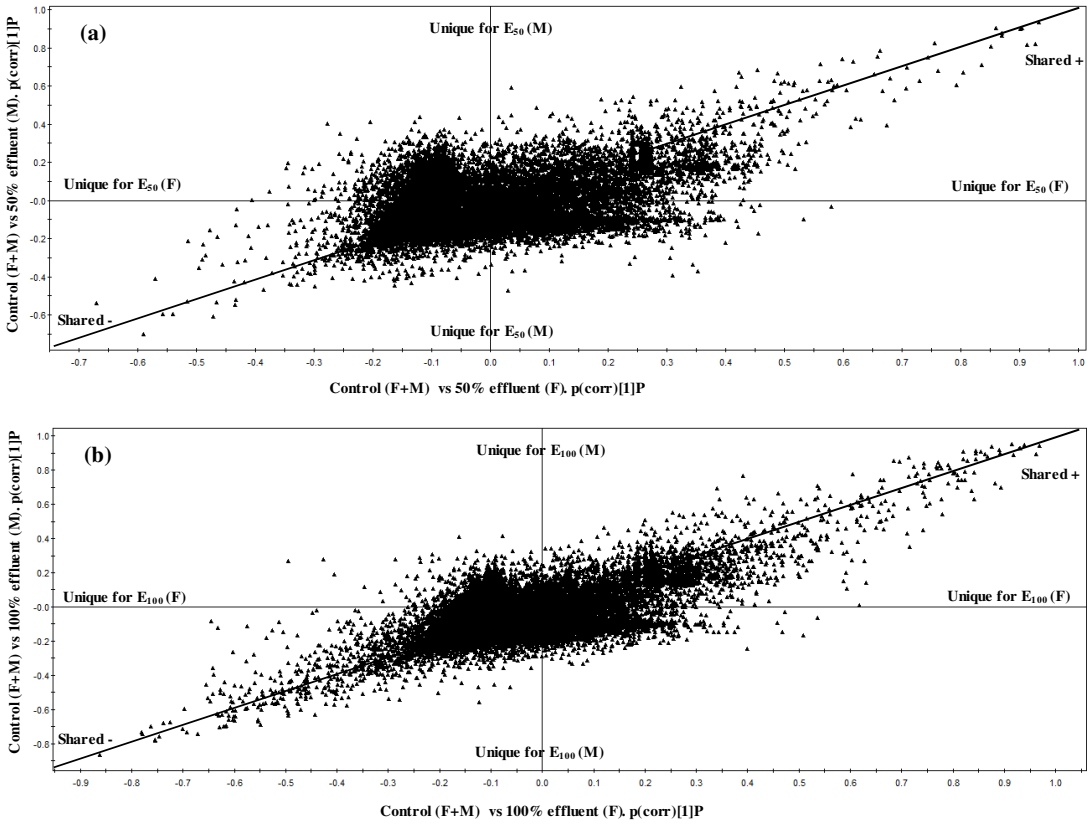


Figure 5.3: Shared and unique structure (SUS) plot analyses of the correlation (reliability) of loading variable from two OPLS models (a) 50% effluent female and 50% effluent male, both sharing a common reference (control female+male), (b) 100% effluent female and 100% effluent male, both sharing a common reference (control female+male) of the chemical profiles of plasma samples from roach exposed either to wastewater effluent or control river water. The samples were profiled in -ESI mode by UPLC-TOFMS. Variables located on x-axis (right and left direction) corresponded to signals unique for either (a) E_{50} treatment (b) E_{100} treatment (female roach). Variables located on y-axis (up and down direction) corresponded to signals unique for (a) E_{50} (b) E_{100} treatment (male roach).

5.3.5 Identification of metabolites due to wastewater effluent exposure

The markers identified in the roach plasma metabolome responsible for the discrimination between control and effluent-exposed fish could be classified in two main groups: exogenous compounds and their metabolites (steroid alkaloid, chlorinated compounds and surfactant) and endogenous compounds (bile acids, sphingolipids see Table 5.1 and phospholipids Table 5.3).

Markers which were only putatively identified were summarized in Table 5.2. Other potential markers whose signals were not good enough to allow full structural characterization were summarized in Appendix 5.6. The classes of chemicals which all these markers belonged to will be described in details in the following sections. In the present study, Q-TOF-MS analysis of individual parent ions was not possible for most of the detected markers due to lack of enough plasma sample. Therefore, additional

structural information for the markers due to effluent exposure was based on data obtained applying different collision energies in full scan mode (collision energy: CE= 25eV).

5.3.5.1 Exogenous metabolites

Identified exogenous markers in plasma from 50% and 100% effluent-exposed roach (females and males) encompassed some chlorinated compounds, surfactants and some other putatively identified compounds.

The two chlorinated markers were detected only in -ESI mode giving signals at *m/z* 366.9002 and *m/z* 264.9937; they were positively identified as sulphate conjugates of triclosan and a methoxy metabolite of chloroxylenol, respectively. Both markers were detected in 50% and 100% effluent-exposed fish samples (females and males). Higher signals for both markers were observed in fish exposed to 100% effluent; male fish showed higher fold changes than female for the triclosan sulphate (Table 5.1). Retention times and exact masses were compared to those found in plasma from effluent-exposed trout and resulted in comparable data. As shown in Figure 5.4, high collision energy mass spectrum (CE: 25eV) revealed signals for both conjugated (triclosan sulphate *m/z* 366.9002) and triclosan itself (*m/z* 286.9437), showing the typical 3-Cl isotopic distribution.

Two anionic surfactants were selected as potential markers in the OPLS-DA loading plot of plasma from effluent-exposed roach and were only detected in –ESI mode. These markers were identified as linear alkylbenzene sulphonic acids (C12-LAS and C13-LAS). For C12-LAS, signals were more abundant in 100% effluent-exposed fish (females > males) than 50% effluent-exposed fish. C13-LAS showed similar abundance in both females and males but was more abundant in 100% effluent exposed fish plasma. These two compounds showed same RT and exact masses as found in effluent-exposed trout.

Q-TOF-MS experiments were possible only for C13-LAS since C12-LAS signals were too low. Q-TOF-MS fragmentation gave rise to the same characteristic fragment ions (*m/z* 183.0116, *m/z* 170.0038 and/or *m/z* 197.0272) as fully explained in Chapter 4 Section 4.3.1.2.

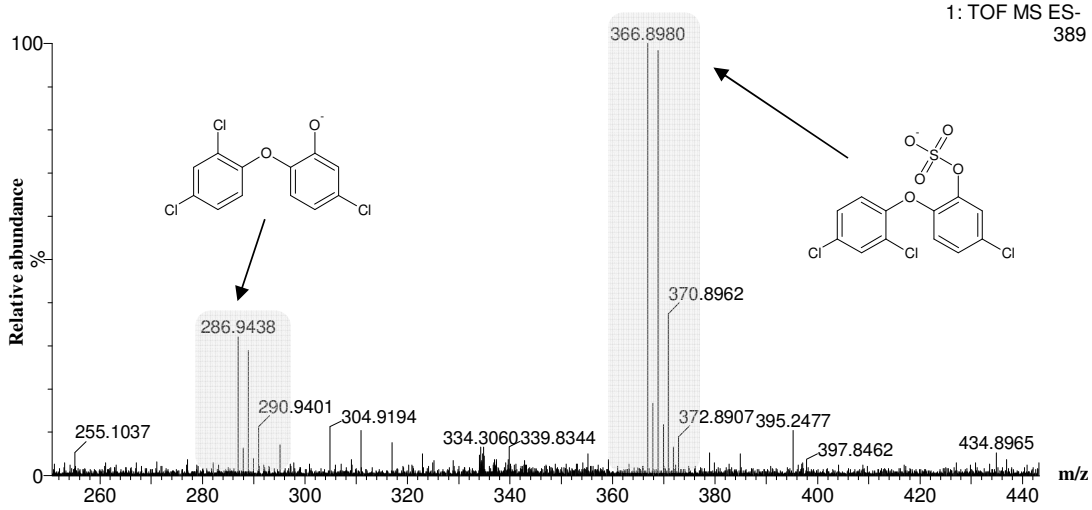


Figure 5.4: TOFMS spectra of triclosan sulphate and triclosan in full scan mode (collision energy=25 eV) detected in -ESI mode in plasma from effluent-exposed roach. Three Cl-isotopic pattern for the molecular ion and the fragment are highlighted in gray.

Table 5.1: Chemical and biochemical markers identified in plasma of effluent-exposed roach in both +/-ESI modes.

Class of chemical	Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOF-MS or high energy collisional dissociation fragments[#]	Parent compound formula	Gender	Control vs 50%effluent		Control vs 100%effluent		Putative identity
										Fold change	*p*-value	Fold change	*p*-value	
Chlorinated xylenol	264.9937	8.60	$C_9H_{10}O_5SCl$	264.9937	0.0	0.5		$C_9H_{11}O_2Cl$	F	7.1	2.40E-02	67.4	1.71E-09	Sulphate conjugate of a methoxy metabolite of chloroxylenol $[M-H]^-$
									M	7.6	2.10E-02	59.7	1.01E-06	
Chlorinated phenoxyphenol	366.9002	15.40	$C_{12}H_6O_5SCl_3$	366.9002	0.0	0.3	286.9437	$C_{12}H_7O_2Cl_3$	F	364.0	5.59E-07	855.0	1.71E-09	Triclosan sulphate $[M-H]^-$
									M	565.0	8.18E-11	1378.0	9.94E-13	
Steroid alkaloid	478.2991	8.2	$C_{27}H_{44}NO_4S$	478.2991	0.0	1.1	398.3452, 380.3315, 98.0967	$C_{27}H_{43}NO$	F	470.2	1.29E-07	987.7	3.54E-09	Sulphate conjugate of solanidine $[M+H]^+$
									M	349.3	3.72E-12	767.4	3.72E-12	
	480.3146	8.3	$C_{27}H_{46}NO_4S$	480.3148	-0.4	0.2	400.3589, 382.3478, 98.0968	$C_{27}H_{45}NO$	F	193.2	1.60E-05	421.0	1.17E-07	Sulphate conjugate of dihydrosolanidine $[M+H]^+$
									M	161.5	4.06E-06	417.4	2.80E-11	
Surfactant	325.1837	22.20	$C_{18}H_{29}O_3S$	325.1837	0.0	0.2		$C_{18}H_{30}O_3S$	F	2.0[¶]	4.00E-03	5.0[¶]	1.12E-08	C12-LAS $[M-H]^-$
									M	1.8[¶]	1.42E-05	2.6[¶]	3.48E-09	
	339.1994	23.38	$C_{19}H_{31}O_3S$	339.1994	0.0	0.5	197.0267, 183.0122, 170.0039	$C_{19}H_{32}O_3S$	F	7.3[¶]	4.31E-06	10.6[¶]	5.76E-06	C13-LAS $[M-H]^-$
									M	3.5[¶]	3.41E-08	10.4[¶]	1.69E-07	
Bile acid	516.2994	8.78	$C_{26}H_{46}NO_7S$	516.2995	-0.2	0.8	498.2887, 480.2780, 462.2678, 337.2532, 319.2426	$C_{26}H_{45}NO_7S$	F	2.9[¶]	2.80E-02	4.9[¶]	4.00E-03	Taurocholic acid $[M+H]^+$
									M	2.2[¶]	6.78E-01	3.3[¶]	1.70E-01	
	480.2780	8.78	$C_{26}H_{42}NO_5S$	480.2784	-0.8	0.3		$C_{26}H_{45}NO_7S$	F	3.0[¶]	8.60E-02	5.2[¶]	3.00E-03	$[M+H-2H_2O]^+$
									M	2.1[¶]	6.38E-01	3.5[¶]	1.94E-01	
	538.2814	8.78	$C_{26}H_{45}NO_7SNa$	538.2814	0.0	1.5		$C_{26}H_{45}NO_7S$	F	2.9[¶]	2.10E-02	4.9[¶]	2.00E-03	$[M+Na]^+$
									M	2.1[¶]	5.00E-01	2.8[¶]	2.93E-01	
	514.2841	8.74	$C_{26}H_{44}NO_7S$	514.2838	0.5	0.6	124.0062, 106.9805	$C_{26}H_{45}NO_7S$	F	3.0[¶]	6.30E-02	5.6[¶]	2.00E-03	$[M-H]^-$
									M	2.3[¶]	6.13E-01	3.8[¶]	1.57E-01	
	533.3146	11.5	$C_{27}H_{49}O_8S$	533.3148	-0.4	0.7		$C_{27}H_{48}O_8S$	F	2.1[¶]	1.00E-03	2.8[¶]	4.27E-05	Cyprinol sulphate $[M+H]^+$
									M	1.7[¶]	1.70E-02	2.3[¶]	3.00E-03	
	555.2969	11.5	$C_{27}H_{48}O_8SNa$	555.2968	0.2	0.6	475.3397	$C_{27}H_{48}O_8S$	F	2.0[¶]	1.00E-03	2.2[¶]	2.01E-04	$[M+Na]^+$
									M	1.4[¶]	3.70E-02	1.7[¶]	2.00E-03	
	550.3412	11.5	$C_{27}H_{52}NO_8S$	550.3414	-0.4	0.2	470.3850, 453.3577, 435.3469, 417.3370, 339.3261	$C_{27}H_{48}O_8S$	F	2.1	2.3E-04	2.6	4.9E-05	$[M+NH_4]^+$
									M	1.5	6.6E-02	1.9	3.0E-03	
	531 .2990	11.46	$C_{27}H_{47}O_8S$	531.2992	-0.4	0.2	513.2879, 96.9595	$C_{27}H_{48}O_8S$	F	2.3[¶]	1.32E-04	2.5[¶]	5.56E-05	$[M-H]^-$
									M	1.4[¶]	1.68E-01	1.7[¶]	8.00E-03	
Sphingolipids	282.2794	20.05	$C_{18}H_{36}NO$	282.2797	-1.1	0.7		$C_{18}H_{37}NO_2$	F	0.7[¶]	8.00E-02	0.5[¶]	2.23E-05	Sphingosine $[M+H-H_2O]^+$
									M	0.9[¶]	5.43E-01	0.5[¶]	1.71E-06	

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit. [#] fragments were obtained from Q-TOFMS or high collisional dissociation fragmentation in full scan mode within ±5ppm; F: female; M: male; $[M+H]^+$: protonated ion; $[M-H]^-$: deprotonated ion; $[M+Na]^+$: sodium adduct; $[M+H-H_2O]^+$: loss of water. Fold Mean fold change calculated from relative concentrations of the compound in plasma from roach exposed to either 50% or 100% effluent for 28 days compared with the levels in control roach held in river water (n=16-20 fish for each gender). [¶]Concentrations above the limit of detection were detected in plasma of the control roach. Non-normally distributed data (determined as such by the Kolmogorov-Smirnov test) were log transformed prior to statistical analysis; *p*-value was calculated from *t*-test between control and the effluent exposure, and values below the BH threshold are statistically significant with a false discovery rate of <5% (Benjamini and Hochberg (BH) threshold was 4.0×10^{-02}).

Two further potential markers exogenously derived were detected in plasma from roach exposed to wastewater effluent. The two markers were detected only in +ESI mode, giving signals as $[M+H]^+$ at *m/z* 478.2991 and *m/z* 480.3146 (Table 5.1). These compounds were identified as sulphate conjugates of solanidine and dihydrosolanidine, respectively. Both markers were higher in 100% effluent-exposed fish than 50% effluent-exposed fish. In order to confirm the identity of these two markers, high energy collisional dissociation fragmentation in full scan mode was employed on plasma samples to obtain their relative fragmentation patterns. In Figure 5.5a,b, alignment was observed between the extracted ion chromatograms for the molecular ions (*m/z* 478.2991 for solanidine and *m/z* 480.3146 for dihydrosolanidine) and for the corresponding in-source fragment ions (*m/z* 398.3452-*m/z* 380.3315) for solanidine and (*m/z* 400.3589 and *m/z* 382.3478) for dihydrosolanidine, respectively. The two different fragment ions for both solanidine and dihydrosolanidine were generated by loss of SO_3 and loss of H_2O, respectively. Furthermore, an additional fragment at *m/z* 98.0970 generated by the cleavage of the ring containing the nitrogen atom and corresponding to the formula $C_6H_{12}N$ was obtained by applying a collision energy of 50eV to both solanidine and dihydrosolanidine parent ions (Cahill et al., 2010). However, the identity of these two markers was not fully confirmed due to unavailability of relative commercial standards. These two markers where also detected in plasma of trout exposed to the same wastewater effluent, but were found as non-conjugated forms.

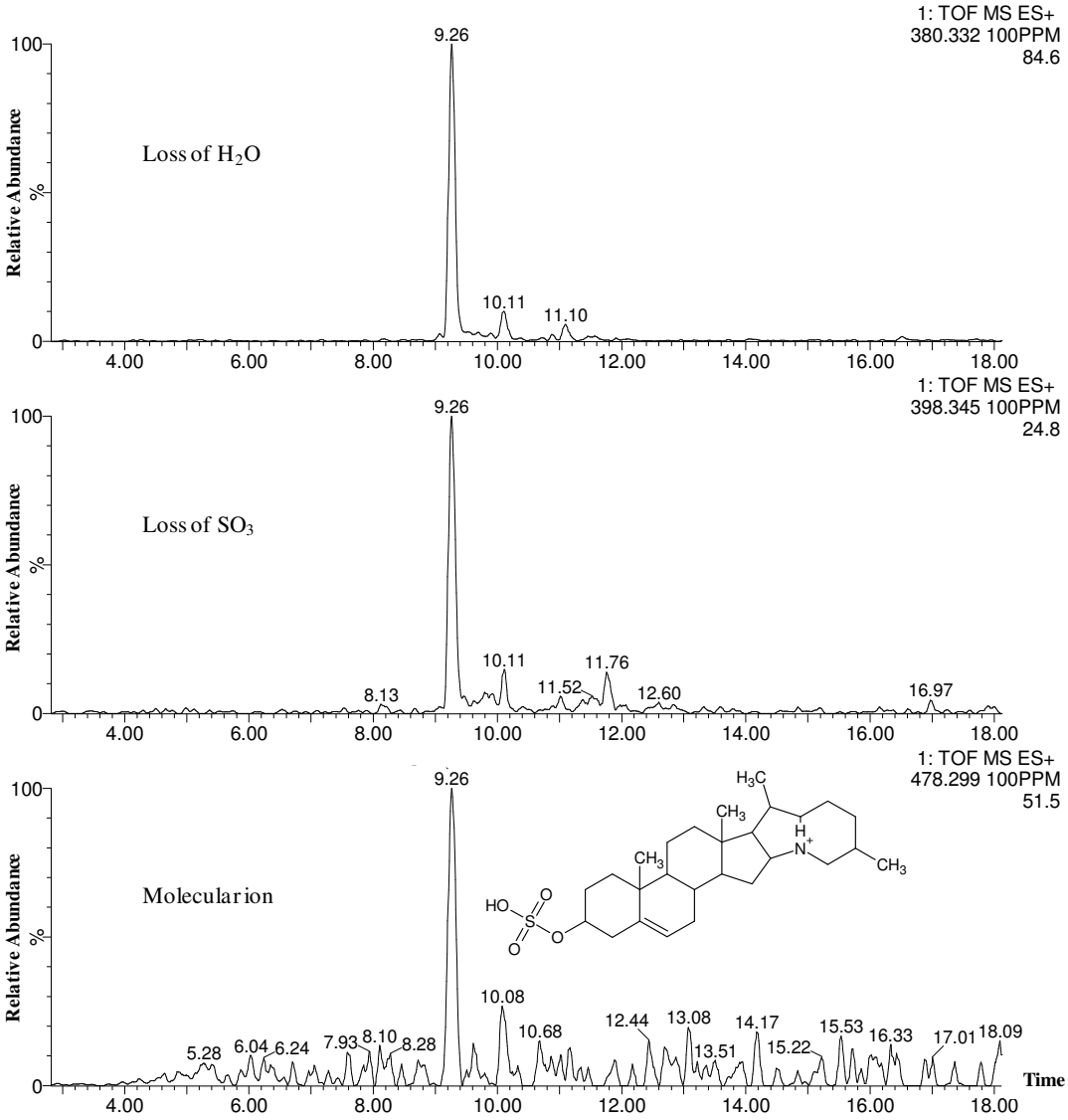


Figure 5.5a: Full scan extracted ion chromatograms of the in-source fragmented of the putative sulphate conjugate of solanidine in plasma from effluent-exposed roach in +ESI mode (collision energy = 25 eV).

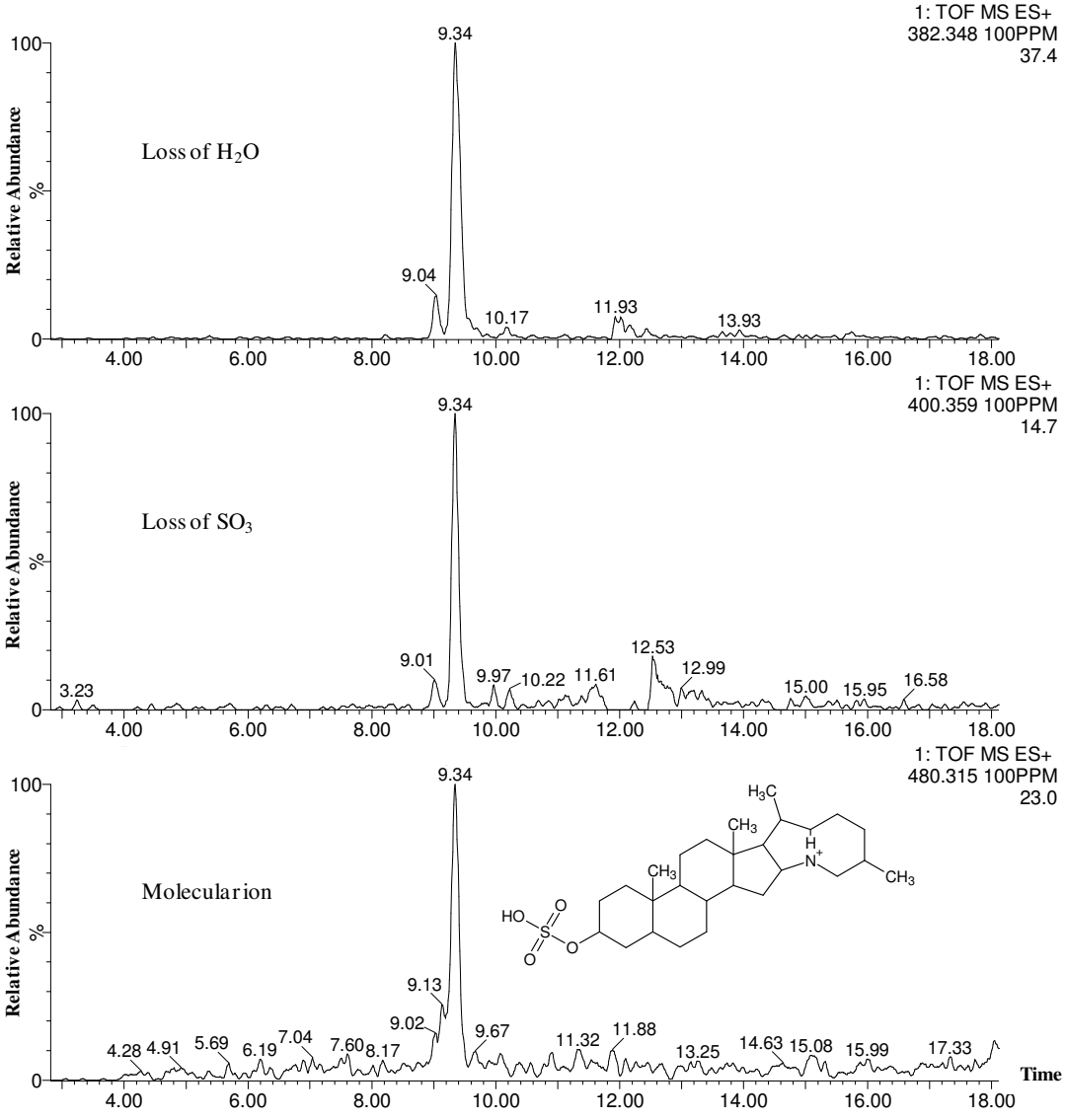


Figure 5.5b: Full scan extracted ion chromatograms of the in-source fragmented of the putative sulphate conjugate of dihydrosolanidine in plasma from effluent-exposed roach in +ESI mode (collision energy = 25 eV).

Other signals (RT-*m/z*) were extracted in the S-plot which theoretically corresponded to other potential xenobiotic compounds since they were not detected in any of the control samples, showed very low *p*-values and increased in effluent-exposed fish to >100 fold compared with controls (see Appendix 5.6). The elemental composition tool assisted in the calculation of the theoretical chemical formula of the selected protonated ions. Ions of some markers at *m/z* 366.3007, *m/z* 302.3061, *m/z* 328.3215 and *m/z* 356.3530 corresponded to elemental composition $C_{22}H_{40}NO_3$, $C_{18}H_{40}NO_2$, $C_{20}H_{42}NO_2$ and $C_{22}H_{46}NO_2$, respectively. Q-TOF-MS analysis (CE: 20eV) of the marker at *m/z* 366.3007 gave rise to fragment ions at *m/z* 348.2901 (loss of H_2O) and *m/z* 274.2509 (loss of 92 Da from the molecular ion) (Appendix 5.7a). The other two molecular ions at *m/z* 302.3061 and *m/z* 356.3530 gave only fragments due to loss of H_2O from the molecular ion (Appendix 5.7b,c). There was not enough mass spectrometry information to ascertain the structures of these compounds but it is possible that $C_{18}H_{40}NO_2$, $C_{20}H_{42}NO_2$ and $C_{22}H_{46}NO_2$ could be fatty acid ethanolamide surfactants corresponding to hexadecoylethanolamide, stearoylethanolamide and arachidoylethanolamides, respectively. These are common detergents used in shampoos and soaps.

It is likely that many other xenobiotics were also present in the plasma samples from effluent-exposed fish but they could not either be detected or identified due to lack of sufficient plasma sample and due to the limit of sensitivity using ESI TOFMS (Appendix 5.6).

Analysis of the UPLC-TOFMS profiles for the plasma samples revealed the presence of highly saturated peak in all the treatments (C, E_{50}, E_{100}). This peak has been corresponeded to the anaesthetic tricaine mesilate (MS222) (detailed in Chapter 4).

Table 5.2: Chemical and biochemical markers putatively identified in plasma from effluent-exposed roach in both +/-ESI modes.

Class of chemical	Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOF-MS or high energy collisional dissociation fragments[#]	Parent compound formula	Gender	Control vs 50% effluent		Control vs 100% effluent		Putative identity
										Fold change	*p*-value	Fold change	*p*-value	
Steroid	541.2642	5.6	$C_{27}H_{41}O_{11}$	541.2649	-1.3	2.0	523.2530, 505.2445, 347.2216, 329.2123, 311.2017, 293.1905, 281.1904, 269.1887	$C_{21}H_{32}O_5$	F	1.1[¶]	6.02E-01	0.5[¶]	6.00E-03	Glucuronide conjugate of putative tetrahydrocortisone $[M+H]^+$
									M	1.4[¶]	4.80E-02	0.6[¶]	8.00E-03	
	539.2490	5.5	$C_{27}H_{39}O_{11}$	539.2492	-0.4	0.1			F	1.9[¶]	4.81E-04	0.4[¶]	1.86E-04	$[M-H]^-$
									M	1.1[¶]	7.48E-01	0.5[¶]	1.00E-03	
Bile acid	500.3414	10.1	$C_{27}H_{50}NO_5S$	500.3410	0.8	0.8	465.3036, 402.3738, 367.3368, 257.2276	$C_{27}H_{49}NO_5S$	F	2.1[¶]	2.70E-02	3.5[¶]	2.00E-03	Bile acid sulphate like $[M+H]^+$
									M	0.8[¶]	1.82E-01	1.5[¶]	1.20E-02	
	498.3253	10.0	$C_{27}H_{48}NO_5S$	498.3253	0.0	0.8	96.9598	$C_{27}H_{49}NO_5S$	F	1.8[¶]	1.27E-01	2.8[¶]	1.00E-03	$[M-H]^-$
									M	0.7[¶]	2.60E-02	1.4[¶]	2.80E-02	

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit. [#] fragments were obtained from Q-TOFMS or high collisional dissociation fragmentation in full scan mode within ±5ppm;
F: female; M: male; $[M+H]^+$: protonated ion; $[M-H]^-$: deprotonated ion.
Fold Mean fold change calculated from relative concentrations of the compound in plasma from roach exposed to either 50% or 100% effluent for 28 days compared with the levels in control roach held in river water (n=16-20 fish for each gender). [¶]Concentrations above the limit of detection were detected in plasma of the control roach.
Non-normally distributed data (determined as such by the Kolmogorov-Smirnov test) were log transformed prior to statistical analysis; *p*-value was calculated from *t*-test between control and the effluent exposure, and values below the BH threshold are statistically significant with a false discovery rate of <5% (Benjamini and Hochberg (BH) threshold was 4.0×10^{-02}).

5.3.5.2 Endogenous metabolites

From the S-plot analyses, a number of biochemical markers were detected in plasma from roach exposed to wastewater effluent. Endogenous compounds revealed in this research mainly belonged to four different classes: steroids, bile acids, sphingolipids and phospholipids.

A putatively identified glucuronide conjugate of tetrahydrocortisone was identified in both +/-ESI modes as a marker upregulated by effluent exposure. Elemental compositions were calculated for both protonated (*m/z* 541.2642) and deprotonated ion (*m/z* 539.2490) as listed in Table 5.2. In +ESI TOFMS analysis (CE: 10eV) fragments were observed at *m/z* 523.2530 (M-H_2O), *m/z* 505.2445 (M-$2H_2O$) and *m/z* 347.2216 (M-Glu-H_2O) (Appendix 5.8). An increase of collision energy to a value of 25eV led to additional fragment ions at *m/z* 329.2123, *m/z* 311.2017, *m/z* 293.1905, *m/z* 281.1904 and *m/z* 269.1887. The fragmentation pattern obtained for the aglycone (CE: 25eV) from the roach plasma sample resulted very similar to the pattern of the tetrahydrocortisone standard available in METLIN database, therefore this marker was putatively assigned to the structure of the glucuronide conjugate of tetrahydrocortisone.

Two bile acids (cyprinol sulphate and taurocholic acid) were detected in plasma of effluent-exposed roach (upregulated markers). These two bile acids were detected in a number of forms in both +/-ESI modes (Table 5.1): deprotonated ions $[M-H]^-$, protonated ions $[M+H]^+$, Na/NH_4 adducts $[M+Na]^+$/$[M+NH_4]^+$ and in-source fragments such as $[M+H-H_2O]^+$ and $[M+H-2H_2O]^+$. TOFMS analysis of cyprinol sulphate in +ESI mode at collision energy of 10eV showed a very abundant ion at *m/z* 555.2969 (Na-adduct) and less so at *m/z* 550.3412, which corresponds to the NH_4-adduct of cyprinol sulphate. The protonated ion at *m/z* 533.3146 was also detected as potential marker in +ESI mode. The Q-TOFMS analysis of the Na-adduct gave only the fragment at *m/z* 475.3397 ($C_{27}H_{48}O_5Na$) corresponding to neutral loss of the sulphate moiety SO_3 (M-80) (Figure 5.6a). On the other hand, NH_4-adduct gave rise to a number of different fragment ions (Figure 5.6b): *m/z* 470.3850 ($C_{27}H_{53}NO_5$) due to loss of SO_3 from the molecular ion, *m/z* 453.3577 ($C_{27}H_{49}O_5$) due to loss of NH_3, *m/z* 435.3469 ($C_{27}H_{47}O_4$) due to loss of $1H_2O$, *m/z* 417.3370 ($C_{27}H_{45}O_3$) due to loss of $2H_2O$ and *m/z* 399.3261 ($C_{27}H_{43}O_2$) due to loss of $3H_2O$ (Figure 5.6b). Q-TOFMS analysis of cyprinol sulphate in –ESI mode (CE: 50eV) revealed only two fragments at *m/z* 513.2879 corresponding to $[M-H_2O-H]^-$ and at *m/z* 96.9595 corresponding to $[HSO_4]^-$ (Figure 5.6c).

Other *m/z*-RT signals were detected in +ESI mode as upregulated markers at *m/z* 538.2814, *m/z* 516.2994 and *m/z* 480.2780, which corresponded to $[M+Na]^+$, $[M+H]^+$ and $[M+H-2H_2O]^+$ ions, respectively of taurocholic acid. The Na-adduct was more abundant than other forms at the experimental conditions used (CE: 10 eV). Retention time, exact mass and TOFMS fragmentation pattern (full scan mode) were compared to those found in plasma samples of trout and also a commercial available standard and resulted in the same data (see Chapter 4 Section 4.3.1.3).

Amongst the markers responsible for the class separation between control and effluent-exposed samples, two ions were present at *m/z* 500.3414 (+ESI mode) and *m/z* 498.3253 (-ESI mode). These metabolite markers were putatively assigned to a sulphate conjugate of a bile acid (Table 5.2). In order to further investigate the molecular structure, TOFMS experiments in full scan mode at collision energy of 25eV were performed on the roach plasma sample. In the +ESI TOFMS analysis fragment ions at *m/z* 465.3036 ($C_{27}H_{45}O_4S$), *m/z* 402.3738 ($C_{27}H_{48}NO$), *m/z* 367.3368 ($C_{27}H_{43}$) and *m/z* 257.2276 ($C_{19}H_{29}$) were detected (Appendix 5.9). The ion at *m/z* 465.3036 could have been generated by sequential losses of H_2O plus NH_3 (loss of 35 Da) from the molecular ion. The ion at *m/z* 402.3738 could be due to the neutral loss of H_2SO_4 from the molecular ion (*m/z* 500.3414). In –ESI Q-TOFMS analysis, the product ion at *m/z* 96.9598, corresponding to the deprotonated sulphate group $[HSO_4]^-$, was the only fragment detected at CE 50eV. The exact structure of this bile acid is currently unknown and it could not be found in the available databases either, but it is likely to be a sulphated C_{27} sterol.

Sphingosine was also detected in roach plasma in +ESI mode as a marker downregulated by effluent exposure (Table 5.1). It was identified from it's RT which was comparable to that obtained in trout plasma and also from a pure standard (see Chapter 4 Section 4.3.1.5) and from the in-source fragment ion at *m/z* 282.2800 due to loss of H_2O from the molecular ion.

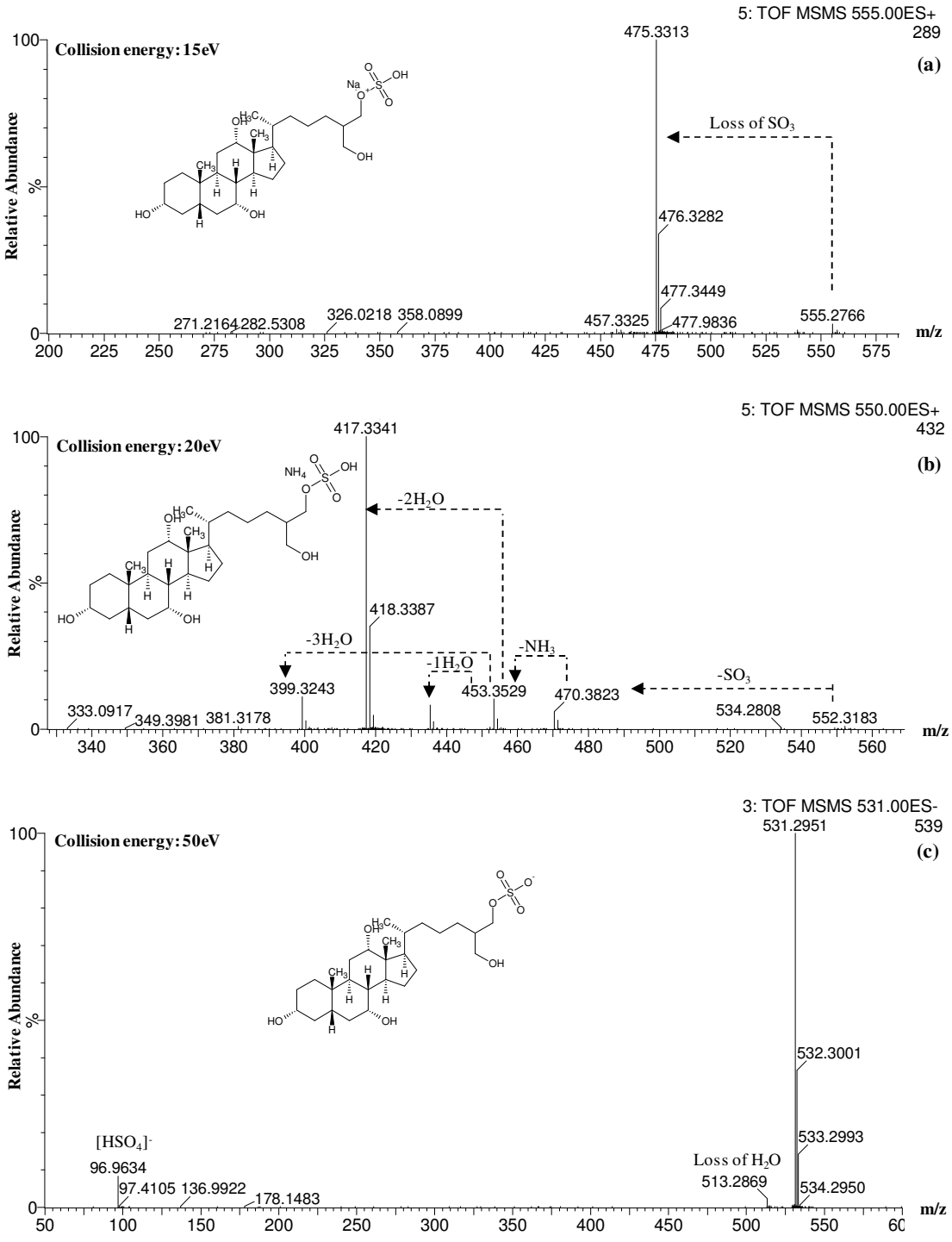


Figure 5.6: Q-TOFMS mass spectra of cyprinol sulphate detected as Na-adduct (a), NH_4-adduct (b) in +ESI and as $[M-H]^-$ (c) in –ESI mode in plasma of effluent-exposed roach.

A number of lysophospholipids (LP) were detected in both ESI modes in roach plasma and which were increased as a result of effluent exposure. Lysophosphatidylethanolamine (LPE) was detected as positively charged $[M+H]^+$ in +ESI and as negatively charged $[M-H]^-$ in –ESI (Table 5.3). Lysophosphatidylcholine (LPC) co-eluted with LPE and was detected in both ESI modes as $[M+H]^+$, $[M-H]^-$ and formate adduct $[M+FA-H]^-$ where FA=formic acid. Significant increase in LPE concentrations was detected, with fold-change values ranging between 1.4-2.3 for female and 1.3-1.7 for male of roach exposed to 100% effluent compared with control samples. LPC concentrations increased as well, giving fold-change values ranging between 1.2-2.0 for female and 1.2-1.7 for male of roach exposed to 100% effluent compared with control samples (Table 5.3). In most cases LPE and LPC signals in 100% effluent-exposed fish were higher than in 50% effluent-exposed fish and the fold increase in females was generally higher than males for both LPCs and LPEs (Table 5.3).

Wherever possible, the precise structure of the LPs was determined from fragments formed from Q-TOFMS. Phospholipids can be usually fragmented at the glycerol-phosphate bond, leading to the diglyceride ion and the ion corresponding to the polar head group (Fang and Barcelona, 1998). The presence of the polar head group leads to a typical neutral loss which results in an abundant fragment ion which gives extremely valuable information for the characterization of these compounds in Q-TOFMS experiments. For instance, loss of 141 Da $[M-141]^+$ constitutes very strong evidence to support the identification of the compound as glycerophosphoethanolamine lipid (Fang and Barcelona, 1998, Murphy, 2002). An example of fragmentation pattern for different LPE structures is shown in Figure 5.7a,b. This figure shows the Q-TOFMS spectra for the protonated ions at *m/z* 454.2937 and *m/z* 480.3092 of PE (16:0/0:0) and PE (18:1/0:0), respectively. Q-TOFMS of the parent ion at *m/z* 454.2937 showed a signal at *m/z* 436.2816 ($C_{21}H_{43}NO_6P$) due to loss of H_2O and a base peak at *m/z* 313.2733 $[M-141]^+$ due to loss of the ethanolaminephosphate head group (Figure 5.7a). Elimination of ethenamine (C_2H_5N) was also observed from the fragment ion $[M+H-H_2O]^+$ at *m/z* 436.2827 generating the ion at *m/z* 393.2407. This ion fragmentation pathway is characteristic of phophatidylethanolamines, and can be used together with the information deriving from other fragments to fully characterize the structure of the phospholipids-PE. PE (18:1/0:0) showed a fragmentation pattern similar to the one

explained for PE (16:0/0:0): the protonated parent ion at *m/z* 480.3092 gave fragment ions at *m/z* 462.2930, *m/z* 419.2543, and *m/z* 339.2878 (Figure 5.7b). A fragmentation pattern example for a selected PE from Lipidmap database [PE(18:1(9Z)/0:0)] is shown in Figure 5.7c, where loss of H_2O at *m/z* 462.32, loss of ethenamine at *m/z* 419.28, and loss of head group at *m/z* 339.12 were observed. Q-TOFMS analysis in –ESI mode of the PE (16:0/0:0) deprotonated ion *m/z* 452.2776 (Figure 5.8a) revealed a base peak at *m/z* 255.2330 ($C_{16}H_{31}O_2$) corresponding to the carboxylate anion (fatty acyl; 16:0) and a peak at *m/z* 196.0381 ($C_5H_{11}NO_5P$) corresponding to the diagnostic polar head group for the phosphatidylethanolamines. The same fatty acyl (16:0) and polar head anion were observed as well in the spectrum obtained from Lipidmap database for [PE(16:0/0:0)] (Figure 5.8b).

Due to sample size limitations, LPC fragmentation patterns were mainly investigated in –ESI mode selecting the formate adduct as parent ion. Q-TOFMS spectra of the deprotonated ions $[M+FA-H]^-$ for [PC (16:1/0:0)] and [PC (18:2/0:0)] showed four peaks: a peak due to loss of formate plus CH_3, a peak corresponding to the fatty acyl group and a peak characteristic of the head group (Figure 5.9a,b). The loss methyl formate from the deprotonated parent ion at *m/z* 538.3145 gave rise to the fragment at *m/z* 478.2931 ($C_{23}H_{45}NO_7P$) (Myers et al., 2011). Fatty acyl and the polar head group ions can be recognized at *m/z* 253.2166 ($C_{16}H_{29}O_2$) and *m/z* 224.0692 ($C_7H_{15}NO_5P$), respectively (Figure 5.9a). The Q-TOFMS mass spectrum for a PC structure from Lipidmap database [PC(18:1(9Z)/0:0)] revealed similar fragmentation pattern with the exception of the loss of methyl chloride instead of methyl formate (Figure 5.9c). This difference can be explained by the fact that spectra reported in the Lipidmap database are relative to chloride adducts while in this study the same compound was detected as formate adduct due to the addition of formic acid as modifier in the mobile phase.

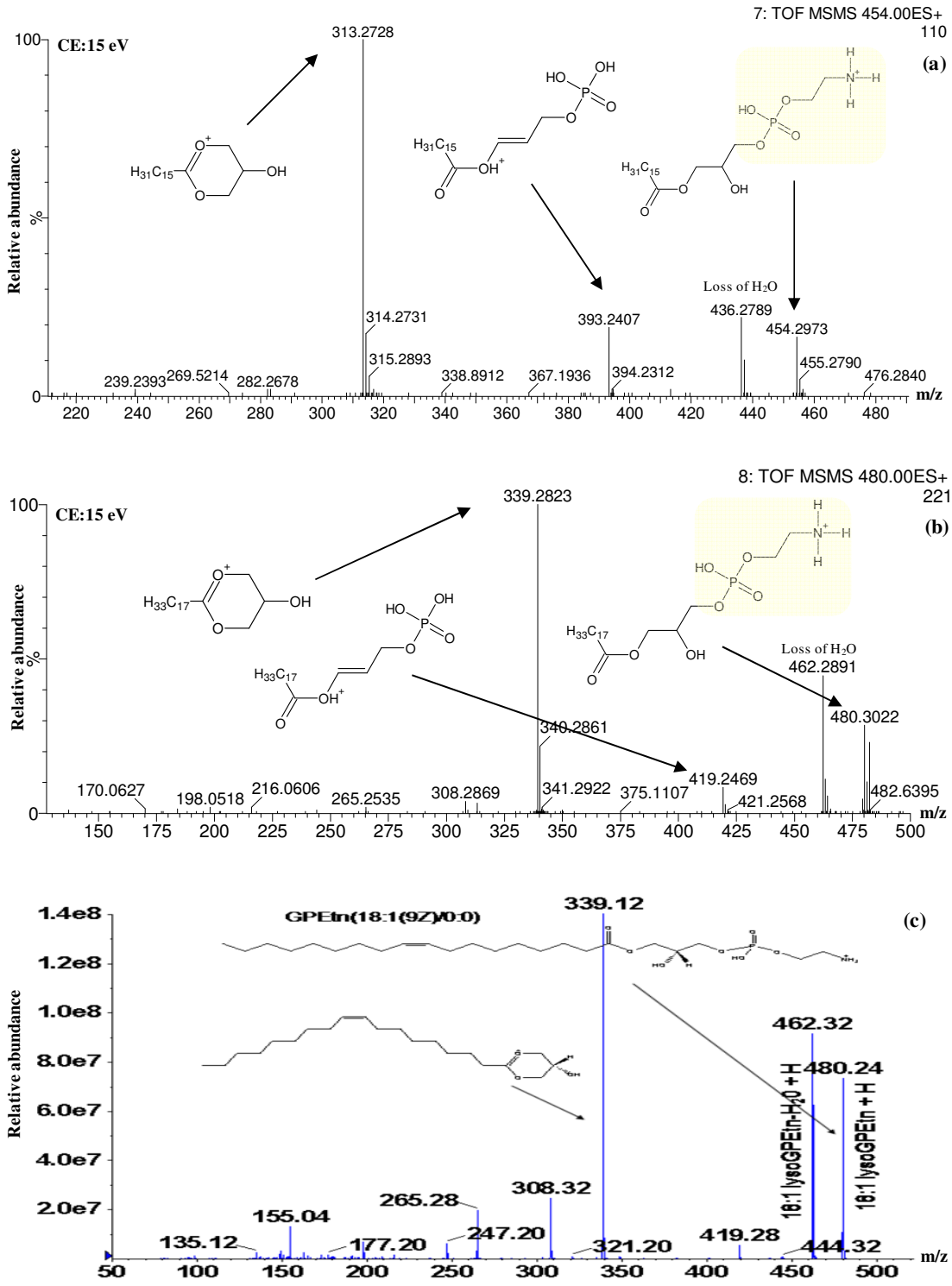


Figure 5.7: Comparison between +ESI mode Q-TOFMS spectra (CE=15eV) for Lysophosphatidylethanolamine (a) [PE (16:0/0:0)] or (b) [PE (18:1/0:0)] in plasma from effluent-exposed roach and (c) 1-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine [PE(18:1(9Z)/0:0)] from Lipidmap database (EPI mode, CE=30eV).

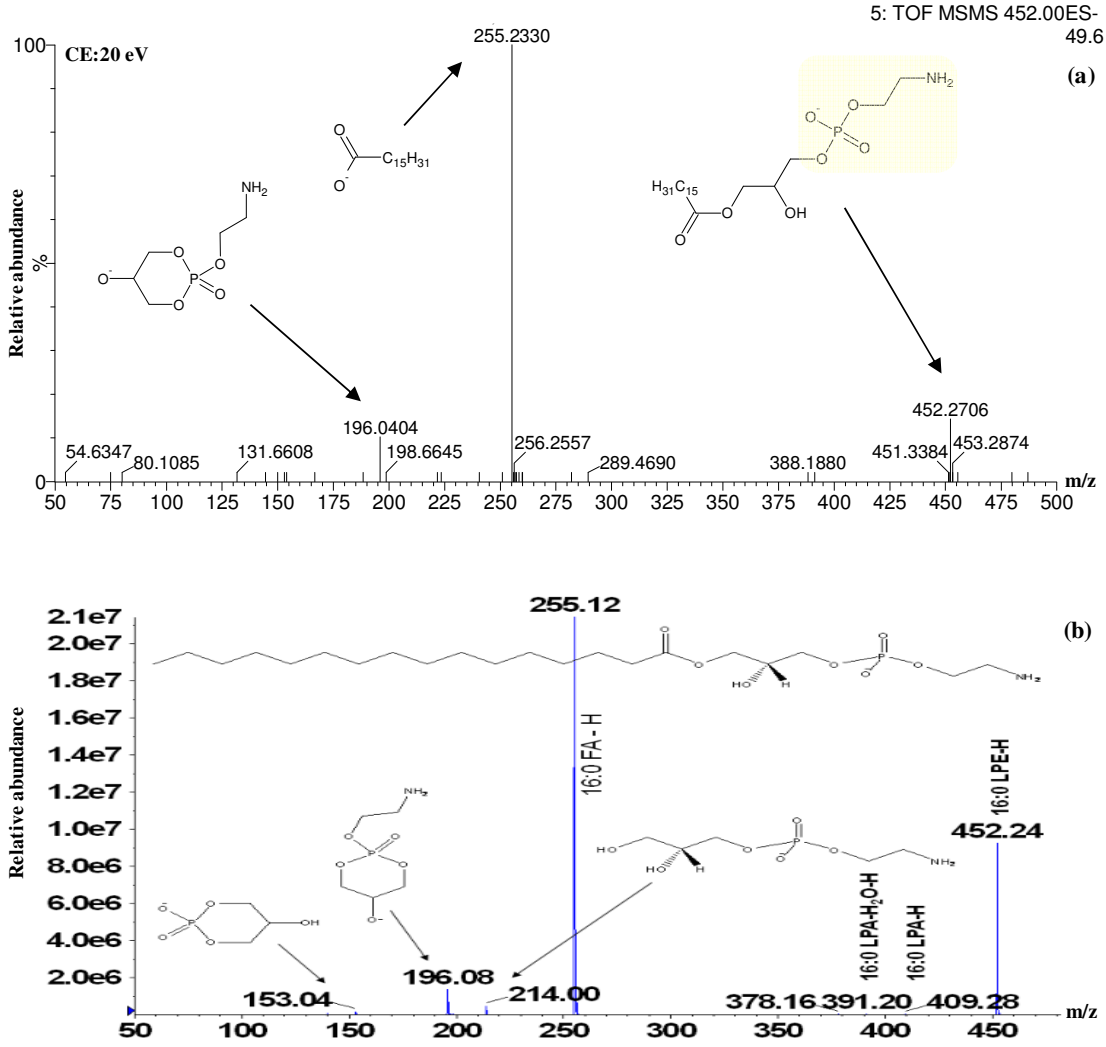


Figure 5.8: Comparison between -ESI mode Q-TOFMS spectra (CE=20eV) for (a) Lysophosphatidylethanolamine [PE (16:0/0:0)] in plasma from effluent-exposed roach and (b) 1-hexadecanoyl-sn-glycero-3-phosphoethanolamine [PE(16:0/0:0)] from Lipidmap database (EPI mode, CE=30 eV).

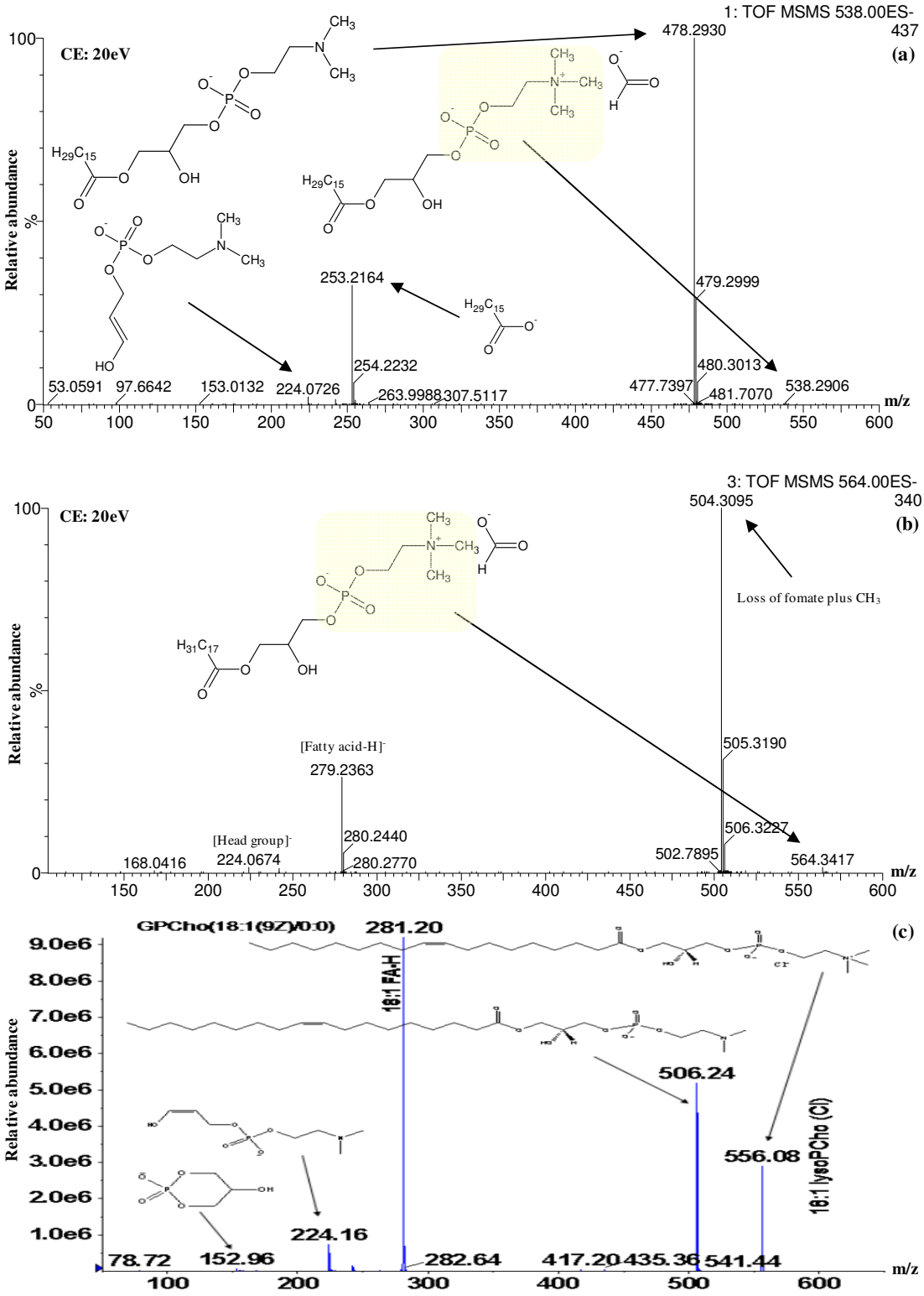


Figure 5.9: Comparison between -ESI Q-TOFMS spectra (CE=20eV) for Lysophosphatidylcholine (a) [PC (16:1/0:0)] or (b) [PC (18:2/0:0)] in plasma from effluent-exposed roach and (c) 1-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine [PC(18:1(9Z)/0:0)] from Lipidmap database (EPI mode, CE=40eV).

Table 5.3: Phospholipid markers detected in roach plasma in both +/-ESI modes.

Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS or high energy collisional dissociation fragments[#]	Parent compound formula	Gender	Control vs 50%effluent Fold change	Control vs 50%effluent *p*-value	Control vs 100%effluent Fold change	Control vs 100%effluent *p*-value	Putative identity
452.2777	20.64	$C_{21}H_{43}NO_7P$	452.2777	0.0	1.9		$C_{21}H_{42}NO_7P$	F	1.3[¶]	8.70E-02	1.9[¶]	6.06E-05	Lyso PE 16:1 $[M+H]^+$
								M	1.0[¶]	5.97E-01	1.3[¶]	2.10E-02	
494.3252	20.80	$C_{24}H_{49}NO_7P$	494.3247	1.0	1.5		$C_{24}H_{48}NO_7P$	F	1.4[¶]	8.70E-02	2.0[¶]	4.27E-06	Lyso PC 16:1 $[M+H]^+$
								M	1.2[¶]	1.28E-01	1.7[¶]	5.55E-05	
450.2621	20.60	$C_{21}H_{41}NO_7P$	450.2621	0.0	0.2		$C_{21}H_{42}NO_7P$	F	1.0[¶]	3.47E-01	1.4[¶]	3.00E-03	Lyso PE 16:1 $[M-H]^-$
								M	0.9[¶]	3.43E-01	1.0[¶]	4.01E-01	
538.3145	20.77	$C_{25}H_{49}NO_9P$	538.3145	0.0	0.2	478.2931, 253.2166, 224.0692	$C_{24}H_{48}NO_7P$	F	1.1[¶]	1.33E-01	1.3[¶]	2.41E-05	Lyso PC 16:1 $[M+FA-H]^-$
								M	1.0[¶]	8.03E-01	1.3[¶]	1.62E-04	
478.2934	21.32	$C_{23}H_{45}NO_7P$	478.2934	0.0	1.0		$C_{23}H_{44}NO_7P$	F	1.3[¶]	2.70E-02	2.1[¶]	1.80E-06	Lyso PE 18:2 $[M+H]^+$
								M	1.2[¶]	9.30E-02	1.4[¶]	2.80E-02	
520.3402	21.47	$C_{26}H_{51}NO_7P$	520.3403	-0.2	0.2		$C_{26}H_{50}NO_7P$	F	1.2[¶]	3.30E-01	1.8[¶]	4.14E-04	Lyso PC 18:2 $[M+H]^+$
								M	1.1[¶]	3.58E-01	1.1[¶]	7.00E-02	
476.2776	21.30	$C_{23}H_{43}NO_7P$	476.2777	-0.2	0.0		$C_{23}H_{44}NO_7P$	F	1.1[¶]	3.71E-01	1.6[¶]	3.00E-05	Lyso PE 18:2 $[M-H]^-$
								M	1.2[¶]	2.1E-02	1.1[¶]	8.60E-02	
564.3300	21.44	$C_{27}H_{51}NO_9P$	564.3301	-0.2	0.7	504.3095, 279. 2321, 224.0674	$C_{26}H_{50}NO_7P$	F	1.1[¶]	2.23E-01	1.4[¶]	3.69E-05	Lyso PC 18:2 $[M+FA-H]^-$
								M	1.0[¶]	7.84E-01	1.1[¶]	5.30E-01	
526.2939	21.22	$C_{27}H_{45}NO_7P$	526.2934	1.0	0.4		$C_{27}H_{44}NO_7P$	F	1.7[¶]	2.00E-03	1.6[¶]	2.00E-03	Lyso PE 22:6 $[M+H]^+$
								M	1.4[¶]	1.20E-02	1.3[¶]	6.60E-02	
524.2778	21.21	$C_{27}H_{43}NO_7P$	524.2777	0.2	0.1		$C_{27}H_{44}NO_7P$	F	1.3[¶]	8.00E-03	1.2[¶]	5.80E-02	Lyso PE 22:6 $[M-H]^-$
								M	1.3[¶]	2.00E-03	1.1[¶]	4.56E-01	
454.2937	21.52	$C_{21}H_{45}NO_7P$	454.2934	0.7	0.7	436.2816, 393.2407, 313.2733	$C_{21}H_{44}NO_7P$	F	1.2[¶]	9.90E-01	1.7[¶]	2.16E-04	LysoPE 16:0 $[M+H]^+$
								M	1.1[¶]	6.56E-01	1.1[¶]	4.07E-01	
452.2776	21.50	$C_{21}H_{43}NO_7P$	452.2777	-0.2	0.4	255.2330, 196.0381	$C_{21}H_{44}NO_7P$	F	1.1[¶]	3.24E-01	1.4[¶]	1.00E-03	LysoPE 16:0 $[M-H]^-$
								M	1.0[¶]	9.40E-01	1.0[¶]	7.27E-01	
528.3090	20.88	$C_{27}H_{47}NO_7P$	528.3090	0.0	12.0		$C_{27}H_{46}NO_7P$	F	1.8[¶]	1.00E-03	2.3[¶]	7.32E-06	Lyso PE 22:5 $[M+H]^+$
								M	1.5[¶]	1.00E-03	1.5[¶]	5.00E-03	
570.3560	20.96	$C_{30}H_{53}NO_7P$	570.3560	0.0	0.6		$C_{30}H_{52}NO_7P$	F	1.6[¶]	3.20E-02	1.7[¶]	1.00E-03	Lyso PC 22:5 $[M+H]^+$
								M	1.3[¶]	1.70E-02	1.4[¶]	1.00E-02	

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit. [#] fragments were obtained from Q-TOFMS or high collisional dissociation fragmentation in full scan mode within ±5ppm. Fold Mean fold change calculated from relative concentrations of the compound in plasma from roach exposed to either 50% or 100% effluent for 28 days compared with the levels in control roach held in river water (n=16-20 fish for each gender). [¶]Concentrations above the limit of detection were detected in plasma of the control roach. Lyso PC: lysophosphatidylcholine; Lyso PE: lysophsophatidylethanolamine; F: female; M: male; $[M+H]^+$: protonated ion; $[M-H]^-$: deprotonated ion; $[M+H-H_2O]^+$: loss of water molecule; $[M+FA-H]^-$: formate adduct.
Non-normally distributed data (determined as such by the Kolmogorov-Smirnov test) were log transformed prior to statistical analysis; *p*-value was calculated from *t*-test between control and the effluent exposure, and values below the BH threshold are statistically significant with a false discovery rate of <5% (Benjamini and Hochberg (BH) threshold was 4.0×10^{-02}).

Table 5.3: (continued) Phospholipid markers detected in roach plasma in both +/-ESI modes.

Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS or high energy collisional dissociation fragments[#]	Parent compound formula	Gender	Control vs 50%effluent		Control vs 100%effluent		Putative identity
									Fold change	*p*-value	Fold change	*p*-value	
526.2935	20.85	$C_{27}H_{45}NO_7P$	526.2934	0.2	1.2		$C_{27}H_{46}NO_7P$	F	1.5[¶]	3.00E-03	1.7[¶]	2.41E-05	Lyso PE 22:5 $[M-H]^-$
								M	1.4[¶]	1.00E-03	1.3[¶]	1.80E-02	
614.3458	20.86	$C_{31}H_{53}NO_9P$	614.3458	0.0	0.7		$C_{30}H_{52}NO_7P$	F	1.4[¶]	1.00E-01	1.5[¶]	1.74E-04	Lyso PC 22:5 $[M+FA-H]^-$
								M	1.3[¶]	4.00E-03	1.2[¶]	2.30E-02	
480.3092	20.78	$C_{23}H_{47}NO_7P$	480.3090	0.4	0.7	462.2930, 419.2543, 339.2878	$C_{23}H_{46}NO_7P$	F	1.4[¶]	8.00E-03	1.7[¶]	1.12E-04	LysoPE 18:1 $[M+H]^+$
								M	1.2[¶]	1.90E-01	1.4[¶]	1.60E-02	
478.2934	20.77	$C_{23}H_{45}NO_7P$	478.2934	0.0	0.2		$C_{23}H_{46}NO_7P$	F	1.3[¶]	3.7E-02	1.5[¶]	5.56E-05	Lyso PE 18:1 $[M-H]^-$
								M	1.0[¶]	4.70E-01	1.1[¶]	1.75E-01	
506.3327	21.30	$C_{25}H_{49}NO_7P$	506.3247	0.0	0.1		$C_{25}H_{48}NO_7P$	F	1.6[¶]	1.00E-03	1.9[¶]	3.23E-6	LysoPE 20:2 $[M+H]^+$
								M	1.3[¶]	1.80E-02	1.5[¶]	5.00E-03	
548.3719	21.39	$C_{28}H_{55}NO_7P$	548.3716	0.5	1.7		$C_{28}H_{54}NO_7P$	F	1.2[¶]	1.48E-01	1.6[¶]	1.00E-03	Lyso PC 20:2 $[M+H]^+$
								M	1.1[¶]	6.70E-01	1.4[¶]	1.60E-02	
504.3090	21.28	$C_{25}H_{47}NO_7P$	504.3090	0.0	9.0		$C_{25}H_{48}NO_7P$	F	1.2[¶]	8.20E-02	1.5[¶]	1.00E-03	Lyso PE 20:2 $[M-H]^-$
								M	1.2[¶]	5.90E-02	1.1[¶]	3.63E-01	
592.3615	21.28	$C_{29}H_{55}NO_9P$	592.3614	0.2	0.6	532.3391, 307.2644, 224.0701	$C_{28}H_{54}NO_7P$	F	1.1[¶]	2.23E-01	1.4[¶]	4.66E-07	Lyso PC 20:2 $[M+FA-H]^-$
								M	1.1[¶]	5.17E-01	1.1[¶]	1.40E-01	
464.3145	21.70	$C_{23}H_{47}NO_6P$	464.3141	0.9	0.3		$C_{23}H_{48}NO_7P$	F	1.4[¶]	1.00E-03	1.7[¶]	1.71E-05	Lyso PE 18:0 $[M+H-H_2O]^+$
								M	1.2[¶]	3.80E-02	1.5[¶]	1.00E-03	
482.3250	21.70	$C_{23}H_{49}NO_7P$	482.3247	0.6	1.5		$C_{23}H_{48}NO_7P$	F	1.7[¶]	1.00E-03	2.0[¶]	1.32E-6	Lyso PE 18:0 $[M+H]^+$
								M	1.3[¶]	5.50E-02	1.7[¶]	1.00E-03	
480.3089	21.67	$C_{23}H_{47}NO_7P$	480.3090	-0.2	0.0		$C_{23}H_{48}NO_7P$	F	1.3[¶]	1.94E-01	1.7[¶]	1.00E-02	Lyso PE 18:0 $[M-H]^-$
								M	1.3[¶]	8.20E-02	1.6[¶]	7.00E-03	
508.3407	22.01	$C_{25}H_{51}NO_7P$	508.3403	0.8	62.0		$C_{25}H_{50}NO_7P$	F	1.7[¶]	5.00E-03	2.0[¶]	9.12E-05	Lyso PE 20:1 $[M+H]^+$
								M	1.2[¶]	4.73E-01	1.2[¶]	3.16E-01	
550.3869	22.08	$C_{28}H_{57}NO_7P$	550.3873	-0.7	3.0		$C_{28}H_{56}NO_7P$	F	1.3[¶]	1.09E-01	1.5[¶]	3.00E-03	Lyso PC 20:1 $[M+H]^+$
								M	0.9[¶]	5.61E-01	1.2[¶]	2.18E-01	
506.3246	22.02	$C_{25}H_{49}NO_7P$	506.3247	-0.2	0.6		$C_{25}H_{50}NO_7P$	F	1.3[¶]	7.51E-01	2.0[¶]	9.00E-02	Lyso PE 20:1 $[M-H]^-$
								M	1.2[¶]	7.29E-01	0.8[¶]	4.20E-01	
594.3770	22.09	$C_{29}H_{57}NO_9P$	594.3771	-0.2	0.5		$C_{28}H_{56}NO_7P$	F	1.3[¶]	7.40E-02	1.3[¶]	2.00E-03	Lyso PC 20:1 $[M+FA-H]^-$
								M	1.0[¶]	7.09E-01	1.0[¶]	9.81E-01	

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit. [#] fragments were obtained from Q-TOFMS or high collisional dissociation fragmentation in full scan mode within ±5ppm. Fold Mean fold change calculated from relative concentrations of the compound in plasma from roach exposed to either 50% or 100% effluent for 28 days compared with the levels in control roach held in river water (n=16-20 fish for each gender). [¶]Concentrations above the limit of detection were detected in plasma of the control roach. Lyso PC: lysophosphatidylcholine; Lyso PE: lysophsophatidylethanolamine; F: female; M: male; $[M+H]^+$: protonated ion; $[M-H]^-$: deprotonated ion; $[M+H-H_2O]^+$: loss of water molecule; $[M+FA-H]^-$: formate adduct. Non-normally distributed data (determined as such by the Kolmogorov-Smirnov test) were log transformed prior to statistical analysis; *p*-value was calculated from *t*-test between control and the effluent exposure samples, and values below the BH threshold are statistically significant with a false discovery rate of <5% (Benjamini and Hochberg (BH) threshold was 4.0×10^{-02}).

5.3.6 Correlation analysis

Since Spearman (rank) correlation does not depend on linearity and is more robust towards outliers (e.g. abnormality in concentration levels between replicates of biological samples), this pair-wise statistic was used to check correlation between datasets. Data were not normally distributed and sample size was >10, thus Spearman's analysis requirements were fully fulfilled (Camacho et al., 2005). In this study the sample size was >30 for each exposure (control, E_{50} and E_{100}): Spearman's analyses revealed positive and negative significant correlations in plasma from female roach exposed to river water (control), 50% or 100% effluent, ranging from very strong (1.0-0.9) to strong (0.9-0.7) and moderate (0.7-0.5) (see Table 5.4 as matrix correlation for selected markers).

Firstly, the correlation between xenobiotics was investigated. The Spearman's correlation coefficient values ranged from 0.9 to 0.5 for both female and male groups, with just two exceptions, the correlation between C12-LAS versus methoxy chloroxylenol sulphate and the correlation between C13-LAS versus methoxy chloroxylenol sulphate, whose Spearman's correlation coefficient values were 0.3 and 0.4, respectively. High correlation was observed between xenobiotics having similar molecular structures. For instance, Spearman's correlation coefficient >0.9 was obtained from the correlation between the sulphate conjugates of solanidine and dihydrosolanidine (Figure 5.10). Triclosan sulphate and solanidine sulphate highlighted a significant positive correlation in both females and males (female: r=0.939, $p=1.3\times10^{-20}$; male: r=0.885, $p=3.2\times10^{-21}$; Tables 5.4a,b). This result suggests that uptake and/or metabolism pathway is clearly correlated between different xenobiotics. The correlation between the endogenously derived metabolites was examined as follows:

Bile acids: In female roach, Spearman's coefficient of >0.5 was observed for the correlation between cyprinol sulphate and the listed xenobiotics in Table 5.4a,b while taurocholic acid showed poor correlation with the same array of xenobiotics, with the exception of C13-LAS and methoxy chloroxylenol sulphate, which gave Spearman's coefficient values of >0.5. In roach males, correlation coefficients varied between 0.1 to 0.3 for the correlation of cyprinol sulphate versus the selected xenobiotics, whereas taurocholic acid revealed no correlation at all. Positive correlation was observed between cyprinol sulphate and taurocholic acid revealed a strong correlation for females

and moderate correlation for males as revealed by their Spearman's and *p*-values (female: r=0.801, $p=1.1\times10^{-10}$; male: r=0.627, $p=6.3\times10^{-8}$).

Sphingosine: The decrease in concentration of sphingosine was also tested by Spearman's analysis. A negative correlation was observed between sphingosine and the listed xenobiotics with Spearman's coefficient values ranging from -0.7 to -0.4 for female roach and from -0.5 to -0.2 for male roach, respectively. The Spearman's coefficient for the correlation of sphingosine versus bile acids (cyprinol sulphate and taurocholic acid) resulted less than -0.5 for female roach and -0.2 for male roach.

Phospholipids: Spearman's correlation coefficients were computed to investigate the existing correlation between lysophospholipids (LPs) levels in the fish and xenobiotics or other upregulated endogenous compounds already detected in this study in both females and males of effluent-exposed roach. The results varied depending on the LPs type (LPEs or LPCs). Correlation between LPs and xenobiotics gave coefficients ranging between 0.7 and 0.1 for roach females, while in males correlation coefficient resulted less than 0.4. LPs versus bile acids showed poor correlation for both females and males (Spearman's coefficient range 0.1-0.4 for females and >0.3 for males). No correlation was observed between LPs and sphingosine in both females and males. LPCs versus LPEs resulted in correlation coefficients ranging between 0.8 and 0.4 for female roach and between 0.7 and 0.2 for male roach.

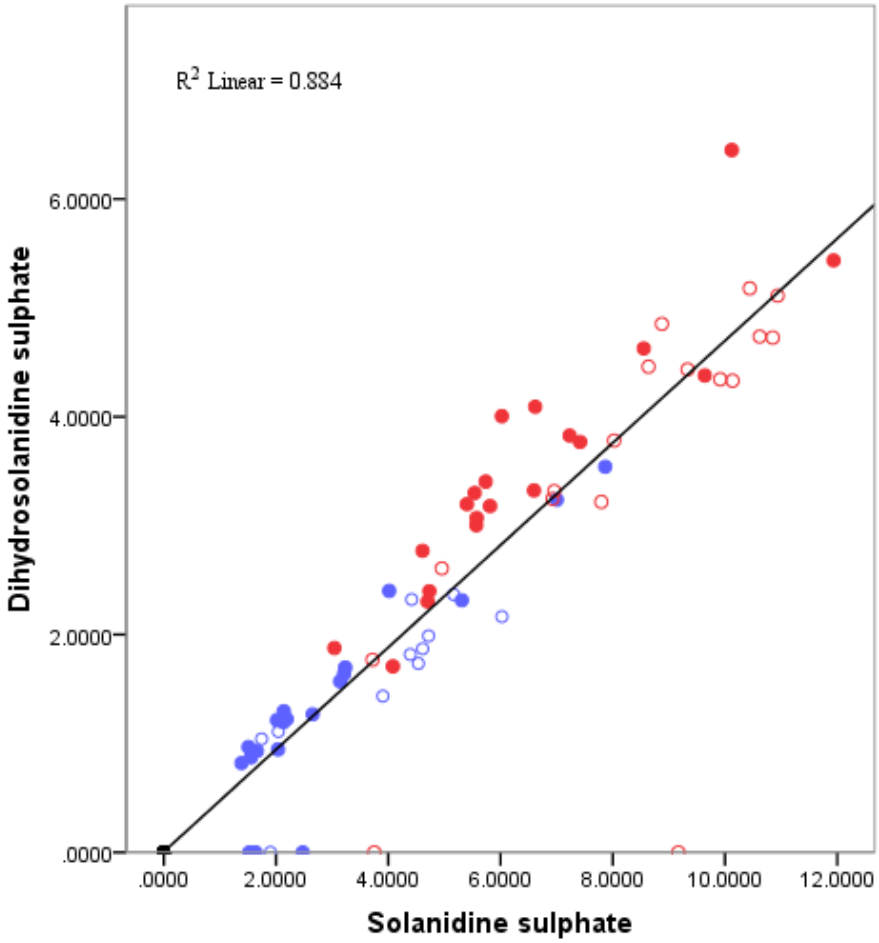


Figure 5.10: Correlation between the concentrations of the sulphate conjugates of solanidine and dihydrosolanidine in plasma from roach exposed to river water, 50% or 100% effluent. The samples were profiled by UPLC-TOFMS in +ESI mode. Female roach symbols: (○), (○) and (○) represent C, E_{50} and E_{100} exposure, respectively. Male roach symbols: (●), (●) and (●) represent C, E_{50} and E_{100} exposure, respectively. The Spearman's coefficient for females was 0.91 (n=43, *p* (2-tailed) = 1.3×10^{-17}) and for males was 0.97 (n=61, *p* (2-tailed) = 2.9×10^{-23}).

Table 5.4a: Correlation matrix of concentrations of xenobiotics and metabolites in plasma from female roach exposed to river water, 50% or 100% effluent.

		$SD+SO_3$	$DSD+SO_3$	Surf	$Xyl+SO_3$	$TCS+SO_3$	C12-LAS	C13-LAS	$Cyp+SO_3$	TC	$BA+SO_3$	SP	LPE18:2	LPC18:2	LPE22:5	LPC22:5	LPE20:1	LPC20:1
Spearman's Coeff.	(+ESI)Solanidine sulphate ($SD+SO_3$)	1	0.914	0.916	0.815	0.939	0.762	0.584	0.600	0.395	0.552	-0.549	0.689	0.557	0.652	0.606	0.591	0.404
p-value		X	1.3×10^{-17}	6.6×10^{-18}	3.0×10^{-11}	1.3×10^{-20}	3.0×10^{-9}	3.9×10^{-5}	2.1×10^{-5}	9.0×10^{-3}	1.3×10^{-4}	1.4×10^{-4}	3.2×10^{-7}	1.0×10^{-4}	2.2×10^{-6}	1.7×10^{-5}	3.0×10^{-5}	7.0×10^{-3}
Spearman's Coeff.	(+ESI)Dihydrosolanidine sulphate ($DSD+SO_3$)	0.914	1	0.901	0.689	0.847	0.665	0.514	0.526	0.323	0.500	-0.426	0.741	0.645	0.645	0.628	0.578	0.489
p-value		1.3×10^{-17}	X	2.0×10^{-16}	3.2×10^{-7}	8.3×10^{-13}	1.1×10^{-6}	4.2×10^{-4}	2.9×10^{-4}	3.4×10^{-2}	6.5×10^{-4}	4.0×10^{-3}	1.3×10^{-8}	3.0×10^{-6}	3.1×10^{-6}	6.4×10^{-6}	4.9×10^{-5}	1.0×10^{-3}
Spearman's Coeff.	(+ESI)366.3007 surfactant (Surf)	0.916	0.901	1	0.824	0.925	0.756	0.595	0.598	0.402	0.543	-0.583	0.724	0.597	0.611	0.547	0.615	0.526
p-value		6.6×10^{-18}	2.0×10^{-16}	X	1.1×10^{-11}	7.5×10^{-19}	4.7×10^{-9}	2.6×10^{-5}	2.3×10^{-5}	8.0×10^{-3}	1.7×10^{-4}	4.0×10^{-5}	4.2×10^{-8}	2.4×10^{-5}	1.4×10^{-5}	1.5×10^{-4}	1.1×10^{-5}	2.9×10^{-4}
Spearman's Coeff.	(-ESI)methoxy xylenol sulphate ($Xyl+SO_3$)	0.815	0.689	0.824	1	0.873	0.731	0.588	0.663	0.614	0.553	-0.704	0.622	0.431	0.527	0.422	0.447	0.291
p-value		3.0×10^{-11}	3.2×10^{-7}	1.1×10^{-11}	X	2.4×10^{-14}	2.6×10^{-8}	3.3×10^{-5}	1.3×10^{-6}	1.2×10^{-5}	1.2×10^{-4}	1.4×10^{-7}	8.5×10^{-6}	4.0×10^{-3}	2.8×10^{-4}	5.0×10^{-3}	3.0×10^{-3}	5.8×10^{-2}
Spearman's Coeff.	(-ESI)Triclosan sulphate ($TCS+SO_3$)	0.939	0.847	0.925	0.873	1	0.762	0.644	0.646	0.436	0.556	-0.658	0.633	0.506	0.610	0.532	0.550	0.396
p-value		1.3×10^{-20}	8.3×10^{-13}	7.5×10^{-19}	2.4×10^{-14}	X	2.9×10^{-9}	3.1×10^{-6}	2.8×10^{-6}	3.0×10^{-3}	1.1×10^{-4}	1.7×10^{-6}	5.2×10^{-6}	1.0×10^{-3}	1.4×10^{-5}	2.4×10^{-4}	1.3×10^{-4}	9.0×10^{-3}
Spearman's Coeff.	(-ESI)C12-LAS	0.762	0.665	0.756	0.731	0.762	1	0.604	0.549	0.313	0.604	-0.668	0.462	0.277	0.259	0.314	0.366	0.116
p-value		3.0×10^{-9}	1.1×10^{-6}	4.7×10^{-9}	2.6×10^{-8}	2.9×10^{-9}	X	1.8×10^{-5}	1.4×10^{-4}	4.1×10^{-2}	1.8×10^{-5}	1.0×10^{-7}	2.0×10^{-3}	7.2×10^{-2}	9.4×10^{-2}	4.0×10^{-2}	1.6×10^{-2}	4.6×10^{-1}
Spearman's Coeff.	(-ESI)C13-LAS	0.584	0.514	0.595	0.588	0.644	0.604	1	0.584	0.508	0.439	-0.440	0.435	0.185	0.554	0.403	0.601	0.318
p-value		3.9×10^{-5}	4.2×10^{-4}	2.6×10^{-5}	3.3×10^{-5}	3.1×10^{-6}	1.8×10^{-5}	X	3.9×10^{-5}	1.0×10^{-3}	3.0×10^{-3}	3.0×10^{-3}	4.0×10^{-3}	2.3×10^{-1}	1.2×10^{-4}	7.0×10^{-3}	2.1×10^{-5}	3.7×10^{-2}
Spearman's Coeff.	(-ESI)Cyprinol sulphate ($Cyp+SO_3$)	0.600	0.526	0.598	0.663	0.646	0.549	0.584	1	0.801	0.409	-0.449	0.468	0.227	0.491	0.263	0.400	0.192
p-value		2.1×10^{-5}	2.9×10^{-4}	2.3×10^{-5}	1.3×10^{-6}	2.8×10^{-6}	1.4×10^{-4}	3.9×10^{-5}	X	1.1×10^{-10}	7.0×10^{-3}	3.0×10^{-3}	2.0×10^{-3}	1.4×10^{-1}	1.0×10^{-3}	8.9×10^{-2}	8.0×10^{-3}	2.2×10^{-1}
Spearman's Coeff.	(-ESI)Taurocholic acid (TC)	0.395	0.323	0.402	0.614	0.436	0.313	0.508	0.801	1	0.306	-0.397	0.441	0.162	0.491	0.283	0.316	0.200
p-value		9.0×10^{-3}	3.4×10^{-2}	8.0×10^{-3}	1.2×10^{-5}	3.0×10^{-3}	4.1×10^{-2}	1.0×10^{-3}	1.1×10^{-10}	X	4.6×10^{-2}	8.0×10^{-3}	3.0×10^{-3}	3.0×10^{-1}	1.0×10^{-3}	6.6×10^{-2}	3.9×10^{-2}	2.0×10^{-1}
Spearman's Coeff.	(-ESI)Bile acid sulphate ($BA+SO_3$)	0.552	0.500	0.543	0.553	0.556	0.604	0.439	0.409	0.306	1	-0.380	0.451	0.406	0.333	0.285	0.203	0.310
p-value		1.3×10^{-4}	6.5×10^{-4}	1.7×10^{-4}	1.2×10^{-4}	1.1×10^{-4}	1.8×10^{-5}	3.0×10^{-3}	7.0×10^{-3}	4.6×10^{-2}	X	1.2×10^{-2}	2.0×10^{-3}	7.0×10^{-3}	2.9×10^{-2}	6.4×10^{-2}	1.9×10^{-1}	4.3×10^{-2}
Spearman's Coeff.	(+ESI)Sphingosine (SP)	-0.549	-0.426	-0.583	-0.704	-0.658	-0.668	-0.440	-0.449	-0.397	-0.380	1	-0.237	-0.146	-0.074	-0.009	-0.159	-0.138
p-value		1.4×10^{-4}	4.0×10^{-3}	4.0×10^{-5}	1.4×10^{-7}	1.7×10^{-6}	1.0×10^{-7}	3.0×10^{-3}	3.0×10^{-3}	8.0×10^{-3}	1.2×10^{-2}	X	1.3×10^{-1}	3.5×10^{-1}	6.4×10^{-1}	9.6×10^{-1}	3.0×10^{-1}	3.8×10^{-1}
Spearman's Coeff.	(+ESI)LPE18:2	0.689	0.741	0.724	0.622	0.633	0.462	0.435	0.468	0.441	0.451	-0.237	1	0.699	0.681	0.605	0.615	0.500
p-value		3.2×10^{-7}	1.3×10^{-8}	4.2×10^{-8}	8.5×10^{-6}	5.2×10^{-6}	2.0×10^{-3}	4.0×10^{-3}	2.0×10^{-3}	3.0×10^{-3}	2.0×10^{-3}	1.3×10^{-1}	X	1.9×10^{-7}	5.1×10^{-7}	1.7×10^{-5}	1.2×10^{-5}	1.0×10^{-3}
Spearman's Coeff.	(+ESI)LPC18:2	0.557	0.645	0.597	0.431	0.506	0.277	0.185	0.227	0.162	0.162	-0.146	0.699	1	0.561	0.750	0.489	0.663
p-value		1.0×10^{-4}	3.0×10^{-6}	2.4×10^{-5}	4.0×10^{-3}	1.0×10^{-3}	7.2×10^{-2}	2.3×10^{-1}	1.4×10^{-1}	3.0×10^{-1}	3.0×10^{-1}	3.5×10^{-1}	1.9×10^{-7}	X	9.0×10^{-5}	7.1×10^{-9}	1.0×10^{-3}	1.3×10^{-6}
Spearman's Coeff.	(+ESI)LPE22:5	0.652	0.645	0.611	0.527	0.610	0.259	0.554	0.491	0.491	0.333	-0.074	0.681	0.561	1	0.804	0.700	0.514
p-value		2.2×10^{-6}	3.1×10^{-6}	1.4×10^{-5}	2.8×10^{-4}	1.4×10^{-5}	9.4×10^{-2}	1.2×10^{-4}	1.0×10^{-3}	1.0×10^{-3}	2.9×10^{-2}	6.4×10^{-1}	5.1×10^{-7}	9.0×10^{-5}	X	8.5×10^{-11}	1.7×10^{-7}	4.2×10^{-4}
Spearman's Coeff.	(+ESI)LPC22:5	0.606	0.628	0.547	0.422	0.532	0.314	0.403	0.263	0.283	0.285	-0.009	0.605	0.750	0.804	1	0.530	0.498
p-value		1.7×10^{-5}	6.4×10^{-6}	1.5×10^{-4}	5.0×10^{-3}	2.4×10^{-4}	4.0×10^{-2}	7.0×10^{-3}	8.9×10^{-2}	6.6×10^{-2}	6.4×10^{-2}	9.6×10^{-1}	1.7×10^{-5}	7.1×10^{-9}	8.5×10^{-11}	X	2.6×10^{-4}	1.0×10^{-3}
Spearman's Coeff.	(+ESI)LPE20:1	0.591	0.578	0.615	0.447	0.550	0.366	0.601	0.400	0.316	0.203	-0.159	0.615	0.489	0.700	0.530	1	0.652
p-value		3.0×10^{-5}	4.9×10^{-5}	1.1×10^{-5}	3.0×10^{-3}	1.3×10^{-4}	1.6×10^{-2}	2.1×10^{-5}	8.0×10^{-3}	3.9×10^{-2}	1.9×10^{-1}	3.0×10^{-1}	1.2×10^{-5}	1.0×10^{-3}	1.7×10^{-7}	2.6×10^{-4}	X	2.2×10^{-6}
Spearman's Coeff.	(+ESI)LPC20:1	0.404	0.489	0.526	0.291	0.396	0.116	0.318	0.192	0.200	0.310	-0.138	0.500	0.663	0.514	0.498	0.652	1
p-value		7.0×10^{-3}	1.0×10^{-3}	2.9×10^{-4}	5.8×10^{-2}	9.0×10^{-3}	4.6×10^{-1}	3.7×10^{-2}	2.2×10^{-1}	2.0×10^{-1}	4.3×10^{-2}	3.8×10^{-1}	1.0×10^{-3}	1.3×10^{-6}	4.2×10^{-4}	1.0×10^{-3}	2.2×10^{-6}	X

+/-ESI: electrospray ionization; LAS: linear alkylbenzene sulphonic acid; LPE: lysophosphatidylethanolamine; LPC: lysophosphatidylcholine; xenobiotics are labelled in red; bile acids in green; sphingolipid in blue and phospholipids in purple.

Table 5.4b: Correlation matrix of concentrations of xenobiotics and metabolites in plasma from male roach exposed to river water, 50% or 100% effluent.

		SD	DSD	Surf	Xyl+SO_3	TCS+SO_3	C12-LAS	C13-LAS	Cyp+SO_3	TC	BA+SO_3	SP	LPE18:2	LPC18:2	LPE22:5	LPC22:5	LPE20:1	LPC20:1
Spearman's Coeff.	(+ESI)Solanidine sulphate (SD)	1	0.971	0.916	0.604	0.885	0.682	0.716	0.356	0.230	0.183	-0.484	0.296	0.253	0.367	0.354	0.200	0.204
p-value		X	2.9×10^{-38}	4.2×10^{-25}	2.5×10^{-7}	3.2×10^{-21}	1.4×10^{-9}	8.8×10^{-11}	5.0×10^{-3}	7.4×10^{-2}	1.6×10^{-1}	7.9×10^{-5}	2.1×10^{-2}	5.0×10^{-2}	4.0×10^{-3}	5.0×10^{-3}	1.2×10^{-1}	1.1×10^{-1}
Spearman's Coeff.	(+ESI)Dihydrosolanidine sulphate (DSD)	0.971	1	0.905	0.548	0.879	0.678	0.694	0.344	0.227	0.201	-0.486	0.305	0.272	0.385	0.372	0.226	0.243
p-value		2.9×10^{-38}	X	1.4×10^{-23}	4.9×10^{-6}	1.3×10^{-20}	1.9×10^{-9}	5.8×10^{-10}	7.0×10^{-3}	7.8×10^{-2}	1.2×10^{-1}	7.1×10^{-5}	1.7×10^{-2}	3.4×10^{-2}	2.0×10^{-3}	3.0×10^{-3}	8.0×10^{-2}	6.0×10^{-2}
Spearman's Coeff.	(+ESI)366.3007 surfactant (Surf)	0.916	0.905	1	0.528	0.949	0.746	0.759	0.239	0.067	0.124	-0.565	0.213	0.177	0.303	0.287	0.110	0.141
p-value		4.2×10^{-25}	1.4×10^{-23}	X	1.2×10^{-5}	3.2×10^{-31}	5.1×10^{-12}	1.4×10^{-12}	6.3×10^{-2}	6.1×10^{-1}	3.4×10^{-1}	2.1×10^{-6}	9.9×10^{-2}	1.7×10^{-1}	1.8×10^{-2}	2.5×10^{-2}	4.0×10^{-1}	2.8×10^{-1}
Spearman's Coeff.	(-ESI)methoxy xylenol sulphate (Xyl+SO_3)	0.604	0.548	0.528	1	0.570	0.333	0.441	0.356	0.462	0.322	-0.462	0.182	0.141	0.127	0.167	0.051	0.161
p-value		2.5×10^{-7}	4.9×10^{-6}	1.2×10^{-5}	X	1.7×10^{-6}	9.0×10^{-3}	3.7×10^{-4}	5.0×10^{-3}	1.8×10^{-4}	1.1×10^{-2}	1.8×10^{-4}	1.6×10^{-1}	2.8×10^{-1}	3.3×10^{-1}	2.0×10^{-1}	7.0×10^{-1}	2.2×10^{-1}
Spearman's Coeff.	(-ESI)Triclosan sulphate (TCS+SO_3)	0.885	0.879	0.949	0.570	1	0.697	0.670	0.240	0.075	0.158	-0.563	0.258	0.229	0.341	0.367	0.117	0.163
p-value		3.2×10^{-21}	1.3×10^{-20}	3.2×10^{-31}	1.7×10^{-6}	X	4.5×10^{-10}	3.6×10^{-9}	6.2×10^{-2}	5.7×10^{-1}	2.2×10^{-1}	2.3×10^{-6}	4.5×10^{-2}	7.5×10^{-2}	7.0×10^{-3}	4.0×10^{-3}	3.7×10^{-1}	2.1×10^{-1}
Spearman's Coeff.	(-ESI)C12-LAS	0.682	0.678	0.746	0.333	0.697	1	0.779	0.143	-0.055	0.097	-0.243	0.135	-0.013	0.245	0.168	0.202	0.027
p-value		1.4×10^{-9}	1.9×10^{-9}	5.1×10^{-12}	9.0×10^{-3}	4.5×10^{-10}	X	1.4×10^{-13}	2.7×10^{-1}	6.8×10^{-1}	4.6×10^{-1}	5.9×10^{-2}	3.0×10^{-1}	9.2×10^{-1}	5.7×10^{-2}	2.0×10^{-1}	1.2×10^{-1}	8.3×10^{-1}
Spearman's Coeff.	(-ESI)C13-LAS	0.716	0.694	0.759	0.441	0.670	0.779	1	0.165	0.043	0.069	-0.413	0.134	-0.076	0.279	0.086	0.195	-0.055
p-value		8.8×10^{-11}	5.8×10^{-10}	1.4×10^{-12}	3.7×10^{-4}	3.6×10^{-9}	1.4×10^{-13}	X	2.1×10^{-1}	7.5×10^{-1}	6.0×10^{-1}	1.0×10^{-3}	3.0×10^{-1}	5.6×10^{-1}	2.9×10^{-2}	5.1×10^{-1}	1.3×10^{-1}	6.7×10^{-1}
Spearman's Coeff.	(-ESI)Cyprinol sulphate (Cyp+SO_3)	0.356	0.344	0.239	0.356	0.240	0.143	0.165	1	0.627	0.362	-0.194	0.255	0.219	0.167	0.094	-0.034	0.252
p-value		5.0×10^{-3}	7.0×10^{-3}	6.3×10^{-2}	5.0×10^{-3}	6.2×10^{-2}	2.7×10^{-1}	2.1×10^{-1}	X	6.3×10^{-8}	4.0×10^{-3}	1.4×10^{-1}	4.7×10^{-2}	9.0×10^{-2}	2.0×10^{-1}	4.7×10^{-1}	8.0×10^{-1}	5.0×10^{-2}
Spearman's Coeff.	(-ESI)Taurocholic acid (TC)	0.230	0.227	0.067	0.462	0.075	-0.055	0.043	0.627	1	0.321	-0.124	0.285	0.229	0.024	0.081	0.110	0.320
p-value		7.4×10^{-2}	7.8×10^{-2}	6.1×10^{-1}	1.8×10^{-4}	5.7×10^{-1}	6.8×10^{-1}	7.5×10^{-1}	6.3×10^{-8}	X	1.2×10^{-2}	3.4×10^{-1}	2.6×10^{-2}	7.6×10^{-2}	8.5×10^{-1}	5.4×10^{-1}	4.0×10^{-1}	1.2×10^{-2}
Spearman's Coeff.	(-ESI)Bile acid sulphate (BA+SO_3)	0.183	0.201	0.124	0.322	0.158	0.097	0.069	0.362	0.321	1	-0.168	0.185	0.306	-0.090	0.157	0.150	0.368
p-value		1.6×10^{-1}	1.2×10^{-1}	3.4×10^{-1}	1.1×10^{-2}	2.2×10^{-1}	4.6×10^{-1}	6.0×10^{-1}	4.0×10^{-3}	1.2×10^{-2}	X	2.0×10^{-1}	1.5×10^{-1}	1.6×10^{-2}	4.9×10^{-1}	2.3×10^{-1}	2.5×10^{-1}	4.0×10^{-3}
Spearman's Coeff.	(+ESI)Sphingosine (SP)	-0.484	-0.486	-0.565	-0.462	-0.563	-0.243	-0.413	-0.194	-0.124	-0.168	1	0.178	-0.071	0.171	0.091	0.275	-0.082
p-value		7.9×10^{-5}	7.1×10^{-5}	2.1×10^{-6}	1.8×10^{-4}	2.3×10^{-6}	5.9×10^{-2}	1.0×10^{-3}	1.4×10^{-1}	3.4×10^{-1}	2.0×10^{-1}	X	1.7×10^{-1}	5.9×10^{-1}	1.9×10^{-1}	4.9×10^{-1}	3.2×10^{-2}	5.3×10^{-1}
Spearman's Coeff.	(+ESI)LPE18:2	0.296	0.305	0.213	0.182	0.258	0.135	0.134	0.255	0.285	0.185	0.178	1	0.546	0.630	0.552	0.620	0.426
p-value		2.1×10^{-2}	1.7×10^{-2}	9.9×10^{-2}	1.6×10^{-1}	4.5×10^{-2}	3.0×10^{-1}	3.0×10^{-1}	4.7×10^{-2}	2.6×10^{-2}	1.5×10^{-1}	1.7×10^{-1}	X	5.3×10^{-6}	5.4×10^{-8}	3.9×10^{-6}	9.9×10^{-8}	1.0×10^{-3}
Spearman's Coeff.	(+ESI)LPC18:2	0.253	0.272	0.177	0.141	0.229	-0.013	-0.076	0.219	0.229	0.306	-0.071	0.546	1	0.355	0.741	0.311	0.676
p-value		5.0×10^{-2}	3.4×10^{-2}	1.7×10^{-1}	2.8×10^{-1}	7.5×10^{-2}	9.2×10^{-1}	5.6×10^{-1}	9.0×10^{-2}	7.6×10^{-2}	1.6×10^{-2}	5.9×10^{-1}	5.3×10^{-6}	X	5.0×10^{-3}	9.9×10^{-12}	1.5×10^{-2}	2.2×10^{-9}
Spearman's Coeff.	(+ESI)LPE22:5	0.367	0.385	0.303	0.127	0.341	0.245	0.279	0.167	0.024	-0.090	0.171	0.630	0.355	1	0.699	0.524	0.208
p-value		4.0×10^{-3}	2.0×10^{-3}	1.8×10^{-2}	3.3×10^{-1}	7.0×10^{-3}	5.7×10^{-2}	2.9×10^{-2}	2.0×10^{-1}	8.5×10^{-1}	4.9×10^{-1}	1.9×10^{-1}	5.4×10^{-8}	5.0×10^{-3}	X	3.7×10^{-10}	1.5×10^{-5}	1.1×10^{-1}
Spearman's Coeff.	(+ESI)LPC22:5	0.354	0.372	0.287	0.167	0.367	0.168	0.086	0.094	0.081	0.157	0.091	0.552	0.741	0.699	1	0.461	0.528
p-value		5.0×10^{-3}	3.0×10^{-3}	2.5×10^{-2}	2.0×10^{-1}	4.0×10^{-3}	2.0×10^{-1}	5.1×10^{-1}	4.7×10^{-1}	5.4×10^{-1}	2.3×10^{-1}	4.9×10^{-1}	3.9×10^{-6}	8.9×10^{-12}	3.7×10^{-10}	X	1.8×10^{-4}	1.2×10^{-5}
Spearman's Coeff.	(+ESI)LPE20:1	0.200	0.226	0.110	0.051	0.117	0.202	0.195	-0.034	0.110	0.150	0.275	0.620	0.311	0.524	0.461	1	0.560
p-value		1.2×10^{-1}	8.0×10^{-2}	4.0×10^{-1}	7.0×10^{-1}	3.7×10^{-1}	1.2×10^{-1}	1.3×10^{-1}	8.0×10^{-1}	4.0×10^{-1}	2.5×10^{-1}	3.2×10^{-2}	9.9×10^{-8}	1.5×10^{-2}	1.5×10^{-5}	1.8×10^{-4}	X	2.7×10^{-6}
Spearman's Coeff.	(+ESI)LPC20:1	0.204	0.243	0.141	0.161	0.163	0.027	-0.055	0.252	0.320	0.368	-0.082	0.426	0.676	0.208	0.528	0.560	1
p-value		1.1×10^{-1}	6.0×10^{-2}	2.8×10^{-1}	2.2×10^{-1}	2.1×10^{-1}	8.3×10^{-1}	6.7×10^{-1}	5.0×10^{-2}	1.2×10^{-2}	4.0×10^{-3}	5.3×10^{-1}	1.0×10^{-3}	2.2×10^{-9}	1.1×10^{-1}	1.2×10^{-5}	2.7×10^{-6}	X

+/-ESI: electrospray ionization; LAS: linear alkylbenzene sulphonic acid; LPE: lysophosphatidylethanolamine; LPC: lysophosphatidylcholine; xenobiotics are labelled in red; bile acids in green; sphingolipid in blue and phospholipids in purple.

5.4 Discussion

Chemical profiling of trout biofluids performed in Chapter 3 and 4 gave interesting and informative results about the possible effects in fish due to exposure to wastewater effluents. Nevertheless, the short term experiment and the availability of only one tank replicate would not have allowed the adequate determination of toxicity pathways and health effects. In this chapter the results of a further experiment are reported, where replication issues have been taken into account. The main aim of this study was to detect changes in the metabolome obtained from plasma samples of a different fish species (roach) exposed to the same wastewater effluent but for longer time (28 days) and employing statistically relevant replicates (2 tanks per each treatment). No depuration step was employed in this particular case since the preliminary trout study showed little persistence of the bioconcentrated xenobiotics. Different proportions of effluent in the exposure tanks were investigated (50% of effluent and 100% of effluent) in order to understand how different dilution conditions of effluents in the field can affect fish health status. The effluent contribution to watercourses can change drastically throughout the year, depending on the rainfall and seasonal conditions. Usually in the UK, in winter a 1 in 2 dilution of the effluent can be reached whilst during the summer, in some extreme cases, due to the much lower level of dilution, rivers can be made of 100% effluents. These seasonal changes in effluent concentration need to be taken into account when assessing the possible impacts of effluent discharges in wild fish (Harries et al., 1999). In this study the exposure of roach to either 50% or 100% effluent resulted in the detection of a number discriminatory metabolites of endogenous and exogenous origin in the plasma of the effluent-exposed roach. Many of these metabolites were similar to those reported as markers of effluent exposure in trout (Chapter 4).

The sulphate conjugate of triclosan and an isomer of the methoxy metabolite of chloroxylenol were detected in both fish exposures. However, other xenobiotics which were detected in trout plasma as glucuronide conjugates (e.g. triclosan, chlorophene and chloroxylenol), were not detected as markers of effluent exposure in roach plasma. Linear alkylbenzene sulphonic acids C12-LAS and C13-LAS were found in both fish species, but all other surfactants and their metabolites previously found in the trout plasma could not be detected in roach plasma. For instance, NP, AEOs, and ECs were not detected in roach plasma either as conjugates or as parent compounds. It is possible

that these surfactants had been taken up by the roach but they were present in the samples in concentrations below the detection limit. For instance, the amount of plasma extract injected on column was 2-fold lower for the roach samples compared with trout, due to the lack of available sample.

In addition to chlorinated organics and surfactant contaminants, the steroidal alkaloids putatively identified as solanidine and dihydrosolanidine were determined as sulphate conjugates in roach plasma whilst they were detected as parent compounds in the trout plasma. The presence of these metabolites as sulphate conjugates in roach plasma could be due to the higher efficiency of the sulphate conjugation of xenobiotics in cyprinids (Kobayashi et al., 1984). These steroidal metabolites are formed from the glycoalkaloids α-solanine and α-chaconine, which are natural toxins produced by potato plants. α-Solanine and α-chaconine share the same aglycone, solanidine, but have different carbohydrate component. These two glycoalkaloids, together with their aglycone solanidine, have been detected in soil extracts due to leakage of these compounds from potato tubers that are left on the field after harvesting; losses to soil could potentially impact shallow groundwater and freshwater ecosystems (Jensen et al., 2008). Both glycoalkaloids have been detected in human blood serum samples collected from volunteers after a meal of potatoes (Hellenäs et al., 1992). The elimination of α-chaconine and α-solanine is similar with the major metabolite being the aglycone, solanidine (Donald G, 2008). As reported previously, the human stomach hydrolyzes glycoalkaloid to the relative shared less toxic aglycone, solanidine and elimination occurs rapidly in the faeces and to a lesser extent in urine (Donald G, 2008). α-solanine and α-chaconine have biological half-lives in serum of about 11 hours and 19 hours, respectively suggesting they are rapidly eliminated from the body (Hellenäs et al., 1992). Unaltered α-chaconine and the major metabolite solanidine have been detected in the urine and faeces of golden hamsters orally administered labelled α-chaconine (Alozie et al., 1979).

As revealed by Hellenas and Donald studies, both glycoalkaloids can be present in human waste, together with their shared aglycone solanidine. Therefore, the presence of solanidine in wastewater could be related both to discharge of human waste already containing it and to microbial degradation of the original glycoalkaloids directly in the environment.

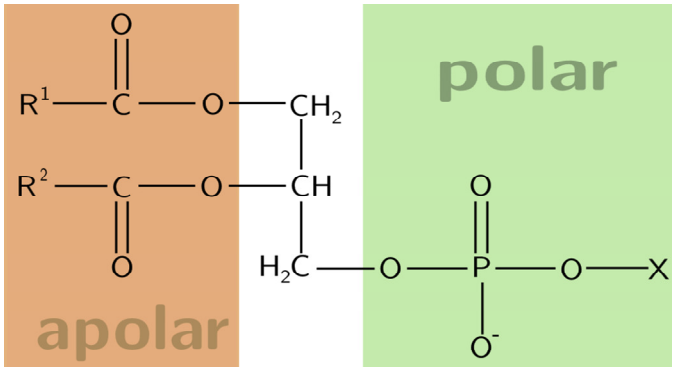
As already mentioned above, triclosan was found in both fish species: in trout samples it was present as both conjugated forms (glucuronide and sulphate) whilst in the roach samples the sulphate conjugate was the predominant form. In aquatic ecosystems, triclosan exists in the ionized form (Orvos et al., 2002) and it is mainly the neutral form which is responsible for most of its toxic effects. The half-life of triclosan in surface water is about 41 min; most of the parent compound is converted to 2,4-dichlorophenol, although degradation rates can vary considerably depending on the aquatic conditions (Reiss et al., 2002, Lyndall et al., 2010). As mentioned in chapter 3, 2,4-dichlorophenol was detected in trout bile in this research and so its presence could be hypothetically related to triclosan degradation. Numerous studies have already assessed the toxicity of triclosan in a variety of organisms, (e.g. algae, invertebrates, amphibians, fish, birds and mammals). The structural similarity of triclosan to known estrogenic and androgenic contaminants (e.g. polychlorinated biphenyls, polybrominated diphenyl ethers, and bisphenol A)) could suggest endocrine disruption behaviour, according to the structure–activity relationships (Veldhoen et al., 2006, Allmyr et al., 2008). Several studies have proved that triclosan has the ability to disrupt endocrine function in many species (reviewed by Hontela (2011)). This result raises a considerable public and scientific concern, since large amounts of triclosan are used on a regular basis by the population. Besides the aquatic environment (Chalew and Halden, 2009), triclosan has also been detected in human plasma (Hovander et al., 2002), breast milk (Adolfsson-Erici et al., 2002), and urine (Calafat et al., 2008).

Besides xenometabolites or metabolites of exogenous origin (i.e triclosan sulphate), endogenously derived metabolites were also detected in trout and roach samples. Bile acids (e.g. taurocholic acid and cyprinol sulphate) were found to be up-regulated in both trout and roach blood after effluent exposure. The bile acid cyprinol sulphate was predominant in roach, taurocholic acid was more abundant in trout while taurochenodeoxycholic acid was detected in trout but absent in roach (both in control and effluent-exposed samples). Another C_{27} bile acid like was detected only in roach plasma. The occurrence of C_{27} bile acid sulphates is common in the wild: they are present as the main biliary surfactants in cartilaginous fish (sharks, rays, and skates), herbivorous bony fish (carp, arapima, and angelfish), and amphibians as well (salamanders and frogs) (Goto et al., 2003). Cyprinol sulphate has been reported to be the major bile acid for grass carp and common carp (family Cyprinidae) (Yeh and

Hwang, 2001). C_{27} bile acids are more cytotoxic than the C_{24} bile acids found in more evolved animals and can be potent inhibitors of oxidative phosphorylation enhancing mitochondrial reactive oxygen species (ROS) production by inhibiting the respiratory chain (Ferdinandusse et al., 2009). 5α-Cyprinol sulphate was found to have highly enhance mucosal membrane permeability to water-soluble compounds (Murakami et al., 2000). 5α-Cyprinol and its sulphate can be toxic and can cause renal and hepatic failure after ingestion of goldfish or carp gallbladders (Yeh et al., 2002, Xuan et al., 2003). Goto et al., (2003) have found that 5α-cyprinol can inhibit taurocholate uptake, since hepatocellular bile salt uptake is mediated predominantly by the Na(+)-taurocholate cotransport proteins. 5α-Cyprinol sulphate inhibited taurocholate uptake in COS-7 cells transfected with rat *asbt*, the apical bile salt transporter of the ileal enterocyte.

A veriaty of phospholipids were up-regulated due to effluent exposure only in roach plasma. One of the possible reason of the lack of detection of these metabolites in trout plasma could be the use of different statistical approaches. In the roach study, *p*-test using the non-parametric test Mann-Whitney was used to screen for significant changes in the MS signals. The obtained *p*-values were adjusted to check for false discovery rates (FDR) by employing Benjamini and Hochberg (BH) procedure, whilst Bonferroni correction has been used in chapter 3 and chapter 4 data (trout samples). Reasons for using BH correction are its simple set up in Microsoft Excel and has a high efficiency in controlling FDR whereas the Bonferroni correction may over correct for the FDR (Robertson et al., 2006). In the Bonferroni correction case some true metabolites (as lysophospholipids in the roach samples) may not pass the Bonferroni corrected threshold, leading to an increase (inflation) of type II errors (Broadhurst and Kell, 2006).

Phospholipids consist of a glycerol back-bone esterified on the sn-1 and sn-2 positions with two fatty acids and a phosphate ester on the sn-3 position (Figure 5.9); the moiety esterified on the sn-3 position is referred to as polar head group, while the alcohol esterified on the phosphate group is referred as R.



Head group (X):	Phospholipid class:
$CH_2CH_2NH_3^+$	Phosphatidylethanolamine (PE)
$CH_2CH_2N(CH_3)_3^+$	Phosphatidylcholine (PC)

Figure 5.11: General structure of phospholipids.

Phospholipids (PLs) are essential constituents of cell membranes, but their importance in living organisms is not limited to this function (Vance, 2002; D'Arrigo and Servi, 2010). Among the different existing classes of lipids, lysophospholipids (LPs) represent only 0.5 to 6% of the total PL content of cell membranes (Birgbauer and Chun, 2006, Torkhovskaya et al., 2007) and their structural role is considered minor. However, they act as second messenger molecules in cellular signal transduction processes (Torkhovskaya et al., 2007, D'Arrigo and Servi, 2010) and have a wide range of biological activities. For this reason, they are among the most interesting lipids for pharmaceutical formulations (Moolenaar, 2000, Birgbauer and Chun, 2006). LP receptors are required for the correct development and functioning of cardiovascular, nervous, immune, respiratory, and reproductive systems in mammals (Rivera and Chun, 2008). They also seem to have a critical role in pathological operations such as inflammatory reactions, autoimmune diseases, cancer, and atherosclerosis (Mills and Moolenaar, 2003, Gardell et al., 2006). LPEs and LPCs can be generated from phosphatidylethanolamine (PE) and phosphatidylcholine (PC), respectively, and are cleavage product generated by increased activity of phospholipase A-type (PLA). Other studies have shown that excessive levels of LPC have been implicated in hepatocellular apoptosis (Han et al., 2008).

Previous studies which have provided strong evidence that halogenated pollutants can alter lipid profiles, altering the lipid-mediated storage and transport of

lipophilic contaminants. In this study, many different halogenated xenobiotics have been detected in both fish species (e.g. triclosan, chloroxylenol and chlorophene); therefore disruption of LP concentations maybe related to their presence in the fish. Capuzzo and Leavitt (1988) observed increases in lipid content and lipid:protein ratios in the tissues of PAH/PCB contaminated mussels and hypothesised that these contaminants mediated a decrease in the conversion of storage lipids to membrane lipids. Their analysis of lipid class composition revealed differences in response to contaminant gradients and reflected alterations in mobilization of triacylglycerols to phospholipid pools, sterol turnover and reductions in phospholipid content. In another study, exposure to the chlorinated insecticide toxaphene increased sterols and decreased triacylglycerol levels in hepatocytes of yellowtail flounder both *in vivo* (Scott et al., 2002) and *in vitro* (Fåhræus-Van Ree and Spurrell, 2000). Capuzzo *et al.* (1984) showed that exposure to hydrocarbons disrupted triacylglycerol synthesis in larval lobsters, while PCB exposure dramatically reduced whole body triacylglycerol levels in oysters (Mommsen and Moon, 2005).

Acylcarnitine was increased in concentration after exposure in trout plasma while no alteration in this metabolite was detected in roach plasma. Decreases in sphingosine levels were confirmed in both species (trout and roach plasma) after effluent-exposure (see Chapter 4 for further details).

Correlation between metabolites were investigated using Spearman's correlation test, conducting a correlation matrix in order to find out any kind of correlation between xenobiotics and endogenous metabolites (i.e. xenobiotics versus xenobiotics, endogenous versus endogenous and xenobiotics versus endogenous). Positive correlation was observed between the two bile acids cyprinol sulphate and taurocholic acid. These two bile acids are involved in primary bile acid biosynthesis, deriving both from cholesterol. Sphingosine revealed a negative correlation to some xenobiotics (e.g. triclosan sulphate). Sphingosine is generated by lysis of ceramide in the lysosome by ceramidase enzymes, leading to an increase in cellular sphingosine levels. However, in this specific study, the concentration of sphingonsine was decreased after effluent exposure; this trend could indicate a possible inhibition of the ceramidase enzymes, leading therefore to a depletion of the sphingosine levels in the cell. Further studies would be needed to determine which xenobiotic(s) accumulating in the fish cause the observed disruption in plasma concentrations of bile acids and lipid metabolites.

CHAPTER 6: General Discussion

Many parts of river catchments in the UK are densely populated and rivers and estuaries are contaminated by effluent discharges on a daily basis, with the implication that aquatic wildlife (e.g. fish) is exposed to complex mixtures of contaminants often throughout their life cycle. Furthermore, an increasing demand on water resources due to an expanding population, and a probable reduced river flow due to climate change are likely to result in even lower dilution of effluents in many UK catchments, increasing the pressure on the sustainability of wildlife (Johnson et al., 2009). In some UK river sites, especially during low rainfall in the summer, WwTW effluents discharges already comprise between 50% and 90% of the river flow. In contrast, in catchments that receive increased rainfall, the greater demands on the sewerage treatment system will inevitably lead to discharge of poorly treated sewage into receiving waters via storm overflow systems. In order to aid understanding of the impacts of effluent discharges into water bodies, this research aimed to characterise the array of organic contaminants and their metabolites present in biofluids (bile and plasma) of two species of fish (trout and roach) after exposure to a selected WwTW effluent. In this work, juvenile rainbow trout, commonly used as a model organism for ecotoxicology work, were initially used to investigate the (xeno)metabolome of fish exposed to a final wastewater effluent. As described in details in Chapter 2, preliminary work was focussed on the analytical approach optimization in order to provide the best methodology for correct interpretation of the information obtained from the metabolomics approach. A fit for purpose UPLC-TOFMS was developed and the biofluid matrices were fully characterized. Bioconcentration and depuration mechanisms for the xenobiotics detected in the bile of the effluent exposed trout were described in Chapter 3. This first exposure study investigated the effect of a 4 and 11 day depuration period on the trout xenometabolome. To do this, parallel control groups of trout were held in reference river water. Furthermore, in Chapter 4 profiles of the xenometabolome in blood samples from the same rainbow trout exposure were also investigated in order to evaluate changes in profiles of endogenous metabolites as a result of effluent exposure. It is worth mentioning that in this study, there was a clear separation between the different control treatment groups (C_{10}, C_{14} and C_{21}) in the PLSDA models of both the bile and blood datasets of trout exposed to the reference river water, proving the need to use a separate control group for each treatment. In these studies it was clear that the river

water abstracted upstream of the effluent contained low levels of contaminants as some of these were detected in the bile of the reference trout (see Chapter 3). It was possible that the contaminant of the river water changed during the exposure period. In addition, some of the differences between the control groups may have also been due to changes in the metabolic status of the trout at the different time points. However for this first study, samples from only one tank replicate were available and a follow up investigation was needed with duplicate exposures to confirm some of the results relating to the disruption of the endogenous metabolites comprising the metabolome. This point has been fully examined in Chapter 5, where the main aim was to detect changes in the plasma metabolome from different fish species exposed to the same wastewater effluent but for longer time and employing statistically relevant replicates. In this second study, roach was used as the test organism as it is a species that was the first to be discovered with a high incidence of intersex condition due to exposure to oestrogenic effluents (Purdom et al., 1994). Roach are therefore considered to be sensitive to the effects of contaminants in wastewater effluents, and they are also representative of cyprinid fish that occur in European rivers. In addition, sexually mature roach were available which meant that gender specific responses to effluent exposure could be evaluated. Exposures in duplicate tanks were undertaken for 28 days, and at two different concentrations of effluent which enabled responses to be determined for wastewater concentrations that represented low flow conditions in many UK river sites. This second approach confirmed most of the data observed in the first piece of work on the trout species. Additionally, it allowed the investigation of some of the pathways disrupted by effluent exposure and the determination of gender specific changes in the metabolome. PLS-DA models of the metabolomic profiles of roach plasma revealed that there was little variability between samples from the duplicate tanks for any one treatment (Chapter 5). This suggested that the differences in the endogenously derived metabolome of trout plasma (changes in bile acids, carnitine and sphingosine concentrations) were likely due to effluent exposure rather than variability in trout metabolite status between individual tanks.

6.1 Main implications of using a chemical profiling approach to identify bioavailable contaminants

In this study, a chemical profiling approach was used to identify not only changes in the metabolome but also uptake of xenobiotics into the fish as a result of exposure to a wastewater effluent. The approach was based on analytical (LC-MS) and multivariate methods used for metabolomics, and thus allowed analysis of xenometabolites as well as metabolites themselves. The concept of the xenometabolome has only been recently recognised and has been applied to the study of metabolites of a single drug as well as to the profile of drugs, pollutants, and dietary or other products that an organism is exposed to or ingested (Holmes et al., 2007, Holmes et al., 2008). A step further is to consider the exposome, defined as the cumulative effects of diverse environmental exposures over a life time and the organism's response to them (Wild, 2012). The exposome complements the genome in understanding the association between the health status of an organism in relation to specific environmental exposures (Wild, 2012). In this context, this study of the profile of xenometabolites bioconcentrating in bile or plasma of fish represents a snapshot of the fish exposome arising from exposure to a wastewater effluent. The fish bioconcentrated not only contaminants such as surfactants and chlorinated organics, but also dietary metabolites such as solanidine from the effluent. Further studies comprising exposure to other wastewater effluents, as well as to river sites in catchments will aid understanding of the exposome of fish in rivers receiving point and diffuse sources of pollutants and other inputs.

It should be recognized that a chemical profiling approach using comparative (xeno)metabolite profiling can be much more easier with wildlife than with human populations, since in the former case it is possible to obtain an uncontaminated reference population to compare xenometabolite profiles. With wildlife, reference populations can be obtained from organisms kept in clean water in the laboratory, or present at sites with little or low contamination. This is often not the case with the human population, where even the most geographically extreme sites, such as the Arctic Circle, contain contaminated indigenous populations (Van Oostdam et al., 2005). In addition, the use of a population that can be transferred from contaminated to clean waters allows the use of depuration experiments in order to evaluate persistence and

clearance rates of contaminants and the recovery of the organism health to normal status.

6.1.1 The analytical methodology

It was clear from this study that the chemical profiling approach used in this study encompassed a variety of bioavailable organic contaminants present in the effluent, however there were many contaminants and other xenobiotics that were not detected using UPLC-ESI TOFMS analysis of the bile and plasma. This could be due to a number of possible reasons: some compounds may have been excreted in urine rather than bile, or present at too low concentrations in the plasma to be detected without preconcentration. Some compounds may not have been ionised by ESI e.g. alkanes or chlorinated hydrocarbons such as polychlorinated biphenyls. Other compounds such as phenolics or steroidal oestrogens would not have been detected due to the use of an acidic mobile phase, and would only have been detected –ESI when they were ionised more efficiently in the presence of a basic modifier (see Chapter 2). Apart from steroidal oestrogens, it is likely that many pharmaceuticals that are known to be in wastewater effluents e.g. non-steroidal anti-inflammatory drugs such as diclofenac, antiepileptic drugs such as carbamazepine and selective serotonin reuptake inhibitors such as citalopram should have been detected in the fish (Zhang et al., 2008, Vasskog et al., 2008). It was also apparent in the current work that the range of contaminants which can be detected greatly depended on the choice of the biofluid to be investigated. Fish bile was chosen as sample matrix because it can bioconcentrate contaminants from ambient water by a factor of many 1000 fold, increasing chances for a successful identification of unknowns. However analysis of plasma may give a more representative picture of the range of xenobiotics circulating in the fish. However, it was clear that analysis of the plasma in this study, detected only a proportion of the compounds that were present in the bile, despite the fact that the amount injected for plasma on column was 20 times higher than bile. For instance, mefenamic acid, resin acids, oxybenzone and naphthols were detected (as glucuronide conjugates) in bile of effluent-exposed trout but were not detected in the plasma samples from the same fish (see chapters 3 and 4).

It is clear that the profiling method needs to be further optimized to detect a more complete range of contaminants and other xenobiotics in the fish plasma. This could be focused on reducing matrix effects in LC-ESI-MS analysis by improving

sample preparation techniques and/or increasing the resolution by using nanoflow LC. The use of nanospray technology instead of electrospray ionisation could also increase the sensitivity of analysis and lower the limit of detection by up to 100 fold. To detect contaminants that do not ionise using nanospray, then samples could be additionally analysed by GC-MS profiling after suitable derivatization. Together these measures could allow a more complete profiling of xenobiotics present in microlitre volumes of fish plasma.

6.1.2 Detection of xenobiotics in effluent-exposed fish

Many different classes of widely used chemicals were detected in the bile of trout exposed to a final wastewater effluent. Six compound classes were characterised as: surfactants (LAS, SPCs, NPEOs, AEOs, AECs), aromatic hydrocarbons, chlorinated compounds (phenols, xylenols, phenoxyphenols, chlorophenes, parabens), pharmaceuticals, sunscreen agents and resin acids. Most of these compounds were detected using UPLC-TOFMS; however, some xenobiotics in the bile (e.g. dichlorophene) were identified by chance during GC-MS analysis of bile fractions. Surfactants are widely detected in the environment due to their broad usage (e.g. in the formulation of household products). Series of nonionic (NPEOs and AEOs) and anionic surfactants (LAS) and were found alongside their potential degradation products (e.g. NP from NPEOs, SPCs from LAS and AECs from AEOs) in trout bile. The detection of SPCs as metabolites of LAS could have been due to either the fish metabolism or to degradation processes occurred in the effluent by microbial activity (Alvarez-Munoz et al., 2010). The presence of AECs in fish, which were found as non-conjugated forms in the bile samples, could be explained either by ω-oxidation of AEOs in effluent (Di Corcia et al., 1998) or by input into the effluent due to their usage as anionic detergents. In addition to SPC metabolites, mono- and dicarboxylic C_{10}-sulfophenyl acids were observed by analysis of bile from fish exposed to wastewater effluent. These metabolites were possibly generated by ω-oxidation of alkyl chain followed by β-oxidation. The detection of these surfactants as the parent compounds (i.e. LAS) in the plasma samples and as oxidised SPCs in the bile samples could be a proof of the occurrence of the metabolism of these contaminants within the fish. Many of the NPEO AEO and AECs were detected with ethoxymers of 6-10 EOs, which indicated that the effluent from the WwTW was poorly treated and failed to degrade these surfactants to shorter ethoxylate chains.

Analysis of bile and plasma matrices of effluent-exposed fish showed the presence of some chlorinated compounds. These belonged to five main classes of compounds: chlorinated phenols (i.e. 2,4-dichlorophenol and an isomer of trichlorophenol), chlorinated xylenols (chloroxylenol, dichloroxylenol and their methoxy metabolites), chlorophenes (chlorophene, methoxychlorophene and dichlorophene), chlorinated phenoxyphenols (diclosan and triclosan), chlorinated parabens (dichloromethylparaben). With the exception of chlorinated parabens, all chlorinated compounds were detected as glucuronide or sulphate conjugated metabolites. All these compounds are generally used as anti-microbial in personal care and household products and have been proved to be acutely toxic to aquatic species in various studies (for details see Section 3.4 Chapter 3). Amongst these xenobiotics, triclosan and chlorophene especially raised scientific concern due to their apparent androgen receptor antagonist activity (antiandrogen) in *in vitro* assays (Rostkowski et al., 2011). However, there is little information on the health effects of exposure of sublethal concentrations of these compounds to fish or whether they possess anti-androgen activity *in vivo*. The presence of chlorinated parabens can be explained as by-products of the reaction of parabens in the effluent with free chlorine dissolved in the water. Sulphate conjugates of some of these xenobiotics (chloroxylenol, triclosan, diclosan, chlorophene and trichlorophenol) were detected only in blood samples, whereas glucuronide conjugates of triclosan, chlorophene and chloroxylenol were present in the bile and in trout blood. This result could be explained either by the relatively low effectiveness of sulphation when compared to other biotransformation or by the excretion of sulphate conjugates via urine due to their relatively small molecular size.

Two PAHs metabolites, 1-hydroxypyrene and the glucuronide conjugates of naphthols, were found in the bile of effluent-exposed trout but not in the blood of both species (trout and roach). PAHs have widely been recognised as mutagenic, carcinogenic and teratogenic properties and may have been present in wastewater effluents thought road runoff of car exhaust emissions, and other sources of carbon combustion (see Section 3.4 for details).

The non-steroidal anti-inflammatory drug mefenamic acid was detected in fish bile as the glucuronide form. This drug has been designed to inhibit prostaglandin biosynthesis; therefore this might represent its primary mechanism of action also in the

fish (Lemke et al., 2007). Oxybenzone, a commercial sunscreen well known for its estrogenic activity, as well as antiestrogenic and antiandrogenic *in vitro* (Kunz and Fent, 2006) has been identified as well in the fish bile as a glucuronide conjugate. Furthermore, a mixture of resin acids (RAs, including abietic acid, pimaric acid, and isopimaric acid) were detected as glucuronide conjugates in the bile. RAs are usually detected in pulp mill effluents and these RA isomers have recently been proved to have relevant antiandrogenic activity (Rostkowski et al., 2011). Lower detection levels would be needed to detect mefenamic acid, oxybenzone and the mixture of resin acids in plasma as these compounds were not detected in blood plasma samples from either the trout or roach exposures.

Bisphenol A, dichlorophene and 2,2'-dihydroxybiphenyl were compounds which could only be detected in the bile samples by means of the GC-MS analysis and their relevance as xenobiotics is due to the fact that they show endocrine disrupting activity in *in vitro* experiments and (for bisphenol A) in *in vivo* studies too (Welshons et al., 2003, Vom Saal and Hughes, 2005, Richter et al., 2007, Rostkowski et al., 2011). Their identification by GC-MS analysis supports the conclusion that using only one analytical approach may not detect the whole array of contaminants present in the samples. For this reason the use of other techniques such as GC-TOFMS could be useful in order to cover the majority of xenobiotics taken up by the fish.

Not all the xenobiotics detected in the effluent exposed fish were chemical contaminants arising from household waste. Solanidine and dihydrosolanidine were likely metabolites of potato consumption in the human population. Interestingly, these were present as the parent compounds in trout plasma but were detected as sulphate conjugates in roach plasma. This observation may reflect differences in the metabolism of xenobiotics between salmonid and cyprinid fish and has been observed previously for metabolism of phenolic substrates (Ferreira-Leach and Hill, 2001, Pedersen and Hill, 2002).

As far as concerns the endogenously derived compounds, the bile acids (i.e. taurocholic acid and cyprinol sulphate), were found to be up-regulated in both trout and roach plasma after effluent exposure. Cyprinol sulphate was the most abundant bile acid in roach, whereas taurocholic acid was the predominant bile acid in trout and taurochenodeoxycholic acid was detected in trout but absent in roach. These results supported previous findings that cyprinol sulphate is the major bile acid in cyprinid fish

(Yeh and Hwang, 2001). Exposure to 100% effluent resulted in a 9-12 fold increase for bile acids in trout plasma compared with a 2-6 fold increase in roach. This result could indicate species differences in response to effluent exposure. However as the trout and roach exposures were carried out at different times, further work would be needed to determine whether there were differences between salmonid and cyprinid fish in bile acid biomarker responses to effluent exposure. Lysophospholipids were up-regulated predominantly in roach blood plasma after effluent exposure, whereas an acylcarnitine was observed as an up-regulated marker of effluent exposure exclusively in the trout blood sample. Again, further work is needed to determine whether these differences are due to fish species or to the length of time of the exposure period (trout 10 days, roach 28 days), or due to differences in effluent composition. Interestingly, with both fish species, exposure to 100% effluent resulted in a two-fold decrease in plasma sphingosine levels indicating that this could be a consistent response to this wastewater effluent.

6.2 Evaluation of the depuration process in trout exposed to wastewater effluent

Many of the xenobiotics detected in the trout plasma were completely eliminated after 4 days depuration and this trend suggests a very low systemic persistence of xenobiotics after exposure to effluent, possibly due to efficient hepatic metabolism and biliary or urine excretion of their conjugated forms. Many of these xenobiotics were present as glucuronide conjugates in bile and these polar metabolites can be rapidly eliminated from the fish via the faeces. Only some xenobiotics, such as some ethoxymers of AECs and NPEOs, were still present in the bile after a 4 day depuration period and this could be possibly related either to their very high concentrations after the initial effluent exposure, or due to contamination of the river water used for the depuration study. The use of a depuration treatment allowed the persistence of the changes in the levels of some endogenously derived metabolites in the plasma to be investigated. Bile acids that were increased by effluent exposure, were then decreased back to almost normal levels after a 4 day depuration period, indicating a rapid return to the normal physiological status of the fish. On the other hand, sphingosine levels were decreased two fold by effluent exposure and only slightly increased (by 30%) back to normal levels after either 4 and 11 days of depuration, indicating possible longer term effect of effluent exposure on the concentrations of this metabolite. This result could be explained either by a prolonged downregulation of enzymes involved in sphingosine

synthesis or by an upregulation of enzymes involved in the metabolism to sphingolipids (see chapter 4).

6.3 Toxicity implications of effluent exposure for fish health

The presence of the above mentioned bioavailable xenobiotics in the wastewater effluent may result in a number of toxic endpoints in fish:

6.3.1 Membrane damage

The composition of many of the surfactant structures detected in the fish, which contained ethoxymers of up to 10 units, could have caused disruption of the structure and function of cell membranes in different organs including the liver. A high concentration of bioavailable surfactant structures circulating in the plasma, and being metabolised in the liver, could have disrupted liver cell membranes and therefore affected the transport of bile acids from the hepatocyte to the bile canaliculi. This could result in liver toxicity and inflammation, and this is borne out by the fact increased bile acids levels were detected in fish blood plasma after 10 days of effluent exposure in trout and after 28 days exposure in roach. The liver damage however, may be reversible, since bile acids reached nearly normal levels back to the normal physiological status (>80%) after depuration. Some bile acids can act as ligands for a variety of receptors (the farnesoid X receptor, the pregnane X receptor and the vitamin D receptor, selected G protein receptor) and interact in a variety of signalling pathways in the liver cells (Chiang, 2002, Zollner et al., 2006). They may in turn affect lipid metabolism, and this maybe reflected by changes in lysophospholipids (i.e. LPC and LPE), which were increased in effluent-exposed roach when compared to the control (river water exposure). Additionally, increasing levels of acylcarnitine and decreasing levels of sphingosine could have been caused by disruption of lipid metabolism and mediated by an increase in plasma bile acid concentrations. However, there is strong evidence that exposure to halogenated pollutants can alter lipid profiles, and can alter the storage lipid metabolism in different species. Exposure to halogenated organic compounds can alter profiles of fatty acids, triacylglycerols, phospholipids and plasma lipids in rats (Mommsen and Moon, 2005). An increase in sterols and a decrease in triacylglycerol levels in hepatocytes of yellowtail flounder have been related to exposure to the chlorinated insecticide toxaphene in *in vivo* (Scott et al., 2002) and *in vitro* studies (Fåhræus-Van Ree and Spurrell, 2000). In this study in particular, a

number of different halogenated xenobiotics have been identified in both the fish exposures (e.g. triclosan, chloroxylenol and chlorophene). The observed alterations in plasma lysophospholipid concentrations might be related to the presence of halogenated contaminants in the fish. In the roach study some disruption of LPE concentrations were apparent even at exposure to 50% effluent; however, further work is needed in order to investigate effects with lower effluent levels at prolonged exposures. In addition, improved lipid profiling techniques should be used to investigate disruption of lipid metabolism in more detail.

This study also identified potential gender differences in response of roach to effluent exposure. Generally, females seemed to be more responsive to effluent exposure compared with males, and the increase of plasma bile acid and phospholipid levels were higher in females than males, whilst sphingosine levels were almost equally affected in both females and males. A correlation approach was used to further investigate the relationships between the concentrations of the xenobiotic and endogenous metabolites detected in roach plasma. Significant positive correlations were observed particularly between related metabolites (e.g. between the two bile acids cyprinol sulphate and taurocholic acid, or between solanidine and dihydroloanidine). However, although significant correlations were also observed between concentrations of many of the xenobiotics and endogenously derived metabolites, no one xenobiotic stood out as being strongly associated with disruption of a specific metabolite. It was possible that some of the disruption in bile acid and lipid levels were due to exposure to mixtures of the contaminants in the effluent. A particular problem in the interpretation of these types of profiling studies is the determination of the particular mixtures of contaminants that are responsible for the disruption of organism health and metabolite homeostasis.

6.3.2 Reproductive axis

Many of the chlorinated contaminants and RAs detected in this study have been shown to be androgen receptor antagonists (at least *in vitro*) and may disrupt sexual differentiation and gonad development in exposed fish. Mefenamic acid (detected in trout bile) belongs to the NSAIDs (non steroidal anti-inflammatory drugs), and in mammals is a potent inhibitor of prostaglandin synthesis (see Chapter 3 for details) and these latter metabolites play important roles in immune function and reproduction. Hence the reproductive axis of male fish may be affected by the presence of variety of

anti-androgenic contaminants: chlorophene, triclosan, resins acids, chloroxylenol and mefenamic acid). In addition, oxybenzone (Kunz and Fent, 2006), bisphenol A (Welshons et al., 2003, Vom Saal and Hughes, 2005, Richter et al., 2007) and NP (Ying and Kookana, 2002) are well-known estrogenic compounds which were detected in the trout exposures and could significantly contribute to the endocrine disrupting activity of the xenobiotic mixture present in the fish. Furthermore, there is the possibility that these xenobiotics may act in a combined manner to demasculinize male fish and to reduce male fertility. *In vitro* studies have proved that exposure to contaminant cocktail mixtures at concentrations below their individual no observed effect level (NOEC) can generate significant biological effects (Silva et al., 2002). Evidence has been given about the additive effects for the same class of EDCs assessing the combined effects of estrogen receptor agonists on induction of biomarkers in fish *in vivo* (Brian et al., 2007). However, little is known about combination effects for EDCs with different modes of action as occurs in the wild. There is the risk that fish reproductive health arising from the exposure to estrogens in surface waters may have been significantly underestimated because of the ability of anti-androgens to potentiate the estrogenic effects in the fish (Katsiadaki et al., 2006, Kiparissis et al., 2003). Molecular approaches studying changes in gene expression have revealed that the feminizing effects of estrogens and anti-androgens in fish share both common and distinct gene pathways (Filby et al., 2007a, Filby et al., 2007b). Modelling studies at sites below WwTWs effluents in UK rivers has shown a strong correlation between the incidence of feminized fish to levels of both anti-androgens and estrogens compared with exposure to estrogens alone. Therefore it is highly likely that exposure to this very complex mixture of endocrine disrupting xenobiotics present in WwTW effluents identified in these profiling studies may result in disruption of reproductive function of male fish. However, further work to develop more sensitive profiling approaches are needed to determine whether effluent exposure disrupted endocrine levels and in particular biosynthesis of sex steroids in the fish.

6.4 Potential for future work

It is clear from this work that using chemical profiling approaches could result in the identification of xenobiotic markers of exposure to WwTW effluents in fish. In addition, effluent exposure also results in disruption of the plasma metabolome and investigation of these associated pathways could also result in the identification of biomarkers of effluent exposure. However many metabolites isolated from either the

bile or plasma, and that were increased by effluent exposure, remain to be identified and were not present in metabolite or chemical databases. The establishment of a database incorporating xenobiotic structures as well as metabolites associated with fish metabolism would aid these types of profiling studies.

Most aquatic monitoring programmes based on collecting grab samples of water at a given time. Some pollutants are present at trace levels and in such a case large volumes of water need to be collected. Laboratory analysis of these samples can only provide a snapshot of the levels of pollutants at the time of sampling. The drawbacks of this approach are that the contaminant concentrations in the environment can vary over time. Passive sampling techniques have been used as tools for measuring aqueous concentrations of a wide range of priority pollutants (Vrana et al., 2005). Moreover, fish is one of the species of wildlife that can be used as indicators of the health of the environment, for instance, effects of disease, pollutants, and other stressors (e.g. climate change) can be assessed by monitoring fish species. For instance, National Marine Monitoring program (UK) uses several fish species to assess a verity of stressors on fish in the aquatic environment. Previous study has domenstrated that bile fluid from stationary fish can be effectively used in monitoring the exposure of fish to aquatic pollutants based on the high bioconcentration factor values obtained using a labelled phenolic compound (Wachtmeister et al., 1991). Therefore, based on the use of fish for biomonitoring of aquatic pollution, this high throughput approach used in this study supports the use of fish as a tool in environmental monitoring.

At present, one of the problems in using chemical profiling approaches is that there are not organised databases for metabolomics and xenobiotic metabolites and the general approach in the scientific community involves the construction of personal databases including data on detected masses, retention times although mass spectra and fragmentation information. This study however has led to valuable information regarding:

- xenometabolome/metabolome
- biological matrixes (bile and blood)
- fish species (roach, trout)
- analytical techniques (LC-MS, GC-MS)

and these data could be extremely useful and informative when compared to results obtained by different research groups. Most of the contaminant metabolites in bile were

identified but in the plasma samples, only approximately 60% of metabolites could be fully characterized. The identity of the remaining 40% could not be established. This is a very common problem in metabolomic analyses, particularly with wildlife species where many metabolites are different to those which are well characterised in humans and higher mammals. The non-identified and/or putatively identified metabolites could also be useful and informative if compared with similar structures discovered other research studies.

New techniques for metabolomic profiling are continually being developed complicated and one LC run cannot cover the diverse range of analyte polarities present in most biological samples (Theodoridis et al., 2008). Hence, both reversed phase and normal phase or hydrophilic interaction chromatography should be tested to provide a deeper insight into the metabolome. Furthermore, as analytes might ionise preferably in one ionization mode (either positive or negative depending on the functional moieties present in the molecule), analysis in both modes is recommended to ensure a wide detection of different metabolites in the sample. Unlikely electrospray ionization, atmospheric pressure chemical ionisation has shown potential in revealing apolar metabolites (Sana et al., 2008). Additionally, the employment of nano ESI would allow detection of metabolites at much lower concentrations due to the significant increase in sensitivity. For instance, a reduction in flow rate from electrospray (100μL/min) to nanospray (1μL/min) would increase the sensitivity of analyte detection by 100 fold. However use of many of these approaches entails multiple analyses of each sample. For large scale studies, even with relatively short analysis times this is clearly quite impractical, therefore alternative methods are needed to expand metabolome coverage by LC–MS. An alternative approach to improve the quality of the metabolomics data would be an optimization of the sample preparation. The use of micro-SPE could lead to a reduction of the matrix effect in the sample increasing the selectivity and sensitivity of the analysis. There is a clear need to develop more sensitive and comprehensive methods of analysis of (xeno)metabolites present in a few microliters of fish plasma. As discussed previously, this may utilise SPE or chemical extraction and preconcentration of low volumes of plasma, and the use of nanoflow linked nanospray TOFMS to increase sensitivity. The use of GC-TOFMS would also allow analysis of compounds that are poorly ionised using nanospray technology. The analysis of low volumes of plasma would also aid non destructive and repeated testing of wild or caged fish to

evaluate the effects of contaminant or effluent exposure. The development of more sensitive analyses would allow a more complete examination of the effects of lipid homeostasis to be examined, and also to determine effects on other endpoints such as the reproductive axis.

Further studies are needed to identify markers of effluent exposure using a variety of WwTW effluents with different industrial inputs and treatment processes. The effects of prolonged exposures to different concentrations of final effluent could be investigated. This study identified potential gender and species differences in response to effluent exposure and these could be examined in more detail using replicated treatments. Implementation of further exposure studies, in combination with improved analytical methods is likely to results in identification of contaminant and metabolite markers of recent exposure that could be used in monitoring fish health in river and estuarine habitats.

6.5 Conclusions

This study of the profiles of xenobiotics in fish biofluids revealed that fish are exposed to a diverse variety of contaminants via uptake from river water contaminated by wastewater effluent. In many UK catchments fish might be exposed to these complex mixtures of xenobiotics semi-continuously throughout their life cycle as wastewater contaminants can be detected many km downstream of the effluent discharge e.g. estrogens (Jobling et al., 2005). The xenobiotics detected in bile and plasma were often present as glucuronide and sulphate conjugates and were rapidly eliminated by the fish organism via faeces or urine. Many of the glucuronide metabolites of the detected xenobiotics were likely to be formed by conjugation of the parent compound within the fish itself, since glucuronide conjugates formed from human metabolism can be readily hydrolysed by bacteria during the wastewater treatment process. This study also revealed that effluent exposure resulted in perturbation of plasma lipid profiles and disruption of bile acid concentrations which could be associated with hepatotoxicity. Chemical profiling techniques can be used to screen for uptake of complex mixtures of contaminants into fish and also for the detection of natural metabolite pathways in the organism which are disrupted by exposure to multiple xenobiotics.

REFERENCE

Accurate Mass: Metabolomics Fiehn Lab. http://fiehnlab.ucdavis.edu/projects/Seven_Golden_Rules/Accurate_Mass/.

Adolfsson-Erici, M., Pettersson, M., Parkkonen, J. & Sturve, J. (2002) Triclosan, a commonly used bactericide found in human milk and in the aquatic environment in Sweden. *Chemosphere,* 46**,** 1485-1489.

Agency, E. (2002) The Water Framework Directive: Guiding Principles on the Technical Requirements. *Environment Agency.*

Ahel, M. & Giger, W. (1993) Aqueous solubility of alkylphenols and alkylphenol polyethoxylates. *Chemosphere,* 26**,** 1461-1470.

Albertsson, E., Kling, P., Gunnarsson, L., Larsson, D. G. & Förlin, L. (2007) Proteomic analyses indicate induction of hepatic carbonyl reductase/20beta-hydroxysteroid dehydrogenase B in rainbow trout exposed to sewage effluent. *Ecotoxicology and Environmental Safety,* 68**,** 33-9.

Allen, J. L. & Hunn, J. B. (1986) Fate and distribution studies of some drugs used in aquaculture. *Veterinary and Human Toxicology***,** 21-24.

Allmyr, M., Harden, F., Toms, L.-M. L., Mueller, J. F., McLachlan, M. S., Adolfsson-Erici, M. & Sandborgh-Englund, G. (2008) The influence of age and gender on triclosan concentrations in Australian human blood serum. *Science of the Total Environment,* 393**,** 162-167.

Alozie, S. O., Sharma, R. P. & Salunkhe, D. K. (1979) Excretion of α-chaconine-^{3}H, a steroidal glycoalkaloid from solanum tuberosum L. and its metabolites in hamsters. *Pharmacological Research Communications,* 11**,** 483-490.

Alvarez-Munoz, D., Gomez-Parra, A. & Gonzalez-Mazo, E. (2010) Influence of the molecular structure and exposure concentration on the uptake and elimination kinetics, bioconcentration, and biotransformation of anioic and nonionic surfactants. *Environmental Toxicology and Chemistry,* 29**,** 1727-1734.

Amster, I. J. (1996) Fourier transform mass spectrometry. *Journal of Mass Spectrometry,* 31**,** 1325–1337.

Ankley, G. T., Jensen, K. M., Kahl, M. D., Korte, J. J. & Makynen, E. A. (2001) Description and evaluation of a short-term reproduction test with the fathead minnow (*Pimephales promelas*). *Environmental Toxicology and Chemistry,* 20**,** 1276-1290.

Aresta, A., Carbonara, T., Palmisano, F. & Zambonin, C. G. (2006) Profiling urinary metabolites of naproxen by liquid chromatography–electrospray mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis,* 41**,** 1312-1316.

Arukwe, A., Goksøyr, A., Thibaut, R. & Cravedi, J. P. (2000) Metabolism and organ distribution of nonylphenol in Atlantic salmon (*Salmo salar*). *Marine Environmental Research,* 50**,** 141-145.

Astle, J., Ferguson, J. T., German, J. B., Harrigan, G. G., Kelleher, N. L., Kodadek, T., Parks, B. A., Roth, M. J., Singletary, K. W., Wenger, C. D. & Mahady, G. B. (2007) Characterization of proteomic and metabolomic responses to dietary factors and supplements. *The Journal of Nutrition,* 137**,** 2787-2793.

Balmer, M. E., Poiger, T., Droz, C., Romanin, K., Bergqvist, P.-A., Müller, M. D. & Buser, H.-R. (2003) Occurrence of methyl triclosan, a transformation product of the bactericide triclosan, in fish from various lakes in Switzerland. *Environmental Science & Technology,* 38**,** 390-395.

Balogh, M. P. (2004) Debating resolution and mass accuracy. *LC•GC Europe,* 17**,** 152–159.

Barnes, K., Kolpin, D., Furlong, E., Zaugg, S., Meyer, M. & Barber, L. (2008) A national reconnaissance of pharmaceuticals and other organic wastewater contaminants in the United States — I) groundwater. *Science of the Total Environment,* 402**,** 192–200.

Belanger, S. E., Dorn, P. B., Toy, R., Boeije, G., Marshall, S. J., Wind, T., Van Compernolle, R. & Zeller, D. (2006) Aquatic risk assessment of alcohol ethoxylates in North America and Europe. *Ecotoxicology and Environmental Safety,* 64**,** 85-99.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. (2001) Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research,* 125**,** 279-284.

Bernhard Lendl & Karlberg, B. (2005) Advancing from unsupervised, single variable-based to supervised, multivariate-based methods: A challenge for qualitative analysis. *Trends in Analytical Chemistry,* 24.

Bernhard, M., Müller, J. & Knepper, T. P. (2006) Biodegradation of persistent polar pollutants in wastewater: Comparison of an optimised lab-scale membrane bioreactor and activated sludge treatment. *Water Research,* 40**,** 3419-3428.

Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B. J., Mendes, P., Roessner-Tunali, U., Beale, M. H., Trethewey, R. N., Lange, B. M., Wurtele, E. S. & Sumner, L. W. (2004) Potential of metabolomics as a functional genomics tool. *Trends in Plant Science,* 9**,** 418-425.

Birgbauer, E. & Chun, J. (2006) New developments in the biological functions of lysophospholipids. *Cellular and Molecular Life Sciences,* 63**,** 2695-2701.

Bjergager, M.-B. A., Hanson, M. L., Lissemore, L., Henriquez, N., Solomon, K. R. & Cedergreen, N. (2011) Synergy in microcosms with environmentally realistic concentrations of prochloraz and esfenvalerate. *Aquatic Toxicology,* 101**,** 412-422.

Blanksby, S. J. & Mitchell, T. W. (2010) Advances in mass spectrometry for lipidomics. *Annual Review Analytical Chemistry,* 3**,** 433-65.

Boberg, J., Metzdorff, S., Wortziger, R., Axelstad, M., Brokken, L., Vinggaard, A. M., Dalgaard, M. & Nellemann, C. (2008) Impact of diisobutyl phthalate and other PPAR agonists on steroidogenesis and plasma insulin and leptin levels in fetal rats. *Toxicology,* 250**,** 75-81.

Boeije, G. M., Cano, M. L., Marshall, S. J., Belanger, S. E., Van Compernolle, R., Dorn, P. B., Gümbel, H., Toy, R. & Wind, T. (2006) Ecotoxicity quantitative structure–activity relationships for alcohol ethoxylate mixtures based on substance-specific toxicity predictions. *Ecotoxicology and Environmental Safety,* 64**,** 75-84.

Bowen, B. & Northen, T. (2010) Dealing with the unknown: Metabolomics and metabolite atlases. *Journal of The American Society for Mass Spectrometry,* 21**,** 1471-1476.

Breitling, R., Pitt, A. R. & Barrett, M. P. (2006) Precision mapping of the metabolome. *Trends in Biotechnology,* 24**,** 543-548.

Brereton, R. G. (2003) *Chemometrics Data Analysis for the Laboratory and Chemical Plant,* USA, John Wiley & Sons, Ltd.

Brian, J. V., Harris, C. A., Scholze, M., Kortenkamp, A., Booy, P., Lamoree, M., Pojana, G., Jonkers, N., Marcomini, A. & Sumpter, J. P. (2007) Evidence of estrogenic mixture effects on the reproductive performance of fish. *Environmental Science & Technology,* 41**,** 337-344.

Broadhurst, D. & Kell, D. (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics,* 2**,** 171-196.

Brown, J. N., Paxéus, N., Förlin, L. & Larsson, D. G. J. (2007) Variations in bioconcentration of human pharmaceuticals from sewage effluents into fish blood plasma. *Environmental Toxicology and Pharmacology,* 24**,** 267-274.

Bruce, S. J., Jonsson, P., Antti, H., Cloarec, O., Trygg, J., Marklund, S. L. & Moritz, T. (2008) Evaluation of a protocol for metabolic profiling studies on human blood plasma by combined ultra-performance liquid chromatography/mass spectrometry: From extraction to data analysis. *Analytical Biochemistry,* 372**,** 237-249.

Bruch, M. K. (1996) Chloroxylenol: An old-new antimicrobial. *Handbook of Disinfectants and Antiseptics.***,** 265-294. New York: Marcel Dekker, Inc.

Buerge, I. J., Buser, H.-R., Poiger, T. & Müller, M. D. (2006) Occurrence and fate of the cytostatic drugs cyclophosphamide and ifosfamide in wastewater and surface waters. *Environmental Science & Technology,* 40**,** 7242-7250.

Buikema Jr, A. L., McGinniss, M. J. & Cairns Jr, J. (1979) Phenolics in aquatic ecosystems: A selected review of recent literature. *Marine Environmental Research,* 2**,** 87-181.

Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E. & Trygg, J. (2006) OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics,* 20**,** 341-351.

Cahill, M. G., Caprioli, G., Vittori, S. & James, K. J. (2010) Elucidation of the mass fragmentation pathways of potato glycoalkaloids and aglycons using Orbitrap mass spectrometry. *Journal of Mass Spectrometry,* 45**,** 1019-1025.

Calafat, A. M., Ye, X., Wong, L.-Y., Reidy, J. A. & Needham, L. L. (2008) Urinary concentrations of triclosan in the U.S. population: 2003–2004. *Environmental Health Perspectives,* 116.

Camacho, D., de la Fuente, A. & Mendes, P. (2005) The origin of correlations in metabolomics data. *Metabolomics,* 1**,** 53-63.

Canosa, P., Rodríguez, I., Rubí, E., Negreira, N. & Cela, R. (2006) Formation of halogenated by-products of parabens in chlorinated water. *Analytica Chimica Acta,* 575**,** 106-113.

Cao, B., Aa, J., Wang, G., Wu, X., Liu, L., Li, M., Shi, J., Wang, X., Zhao, C., Zheng, T., Guo, S. & Duan, J. (2011) GC–TOFMS analysis of metabolites in adherent MDCK cells and a novel strategy for identifying intracellular metabolic markers for use as cell amount indicators in data normalization. *Analytical and Bioanalytical Chemistry,* 400**,** 2983-2993.

Capuzzo, J. M., Lancaster, B. A. & Sasaki, G. C. (1984) The effects of petroleum hydrocarbons on lipid metabolism and energetics of larval development and metamorphosis in the american lobster (*Homarus americanus* Milne Edwards). *Marine Environmental Research,* 14**,** 201-228.

Capuzzo, J. M. & Leavitt, D. F. (1988) Lipid composition of the digestive glands of *Mytilus edulis* and *Carcinus maenas* in response to pollutant gradients. *Marine Ecology Progress Series,* 46**,** 139-145.

Cardellini, P. & Ometto, L. (2001) Teratogenic and toxic effects of alcohol ethoxylate and alcohol ethoxy sulfate surfactants on *Xenopus laevis* embryos and Tadpoles. *Ecotoxicology and Environmental Safety,* 48**,** 170-177.

CESIO (2007) European Committee of Organic Surfactants and their Intermediates surfactants statistics for Western Europe.

Chace, D. (2001) Mass spectrometry in the clinical laboratory. *Chemical Reviews,* 101**,** 445-477.

Chalew, T. E. & Halden, R. U. (2009) Environmental exposure of aquatic and terrestrial biota to triclosan and triclocarban. *Journal - American Water Works Association,* 45**,** 4-13.

Chan, C. C., Lam, H., Lee, Y. C. & Zhang, X.-M. (Eds.) (2004) *Analytical Method Validation and Instrument Performance Verification,* Hoboken, New Jersey., John Wiley & Sons, Inc.

Chen, X., Carroll, J. A. & Beavis, R. C. (1998) Near-ultraviolet-induced matrix-assisted laser desorption/ionization as a function of wavelength. *Journal of The American Society for Mass Spectrometry,* 9**,** 885-891.

Chiang, J. L. (2003) Bile acid regulation of hepatic physiology: III. Bile acids and nuclear receptors. *American Journal Physiology. Gastrointestinal and Liver Physiology,* 284**,** 349-356.

Chiang, J. Y. L. (2002) Bile acid regulation of gene expression: Roles of nuclear hormone receptors. *Endocrine Reviews,* 23**,** 443-463.

Chipman, J. K. & Walker, C. H. (1981) Metabolism and route of excretion of the chloro-cyclodiene HCE in the Pigeon. *Archives of Environmental Contamination and Toxicology,* 10**,** 755-764.

Choi, S. M., Yoo, S. D. & Lee, B. M. (2004) Toxicological characteristics of endocrine-disrupting chemicals: Developmental toxicity, carcinogenicity, and mutagenicity. *Journal of Toxicology and Environmental Health, Part B,* 7**,** 1-23.

Christiansen, L. B., Pedersen, K. L., Pedersen, S. N., Korsgaard, B. & Bjerregaard, P. (2000) *In vivo* comparison of xenoestrogens using rainbow trout vitellogenin induction as a screening system. *Environmental Toxicology and Chemistry,* 19**,** 1867-1874.

Christopher Corton, J. (2008) Evaluation of the role of peroxisome proliferator-activated receptor α (PPARα in mouse liver tumor induction by trichloroethylene and metabolites. *Critical Reviews in Toxicology,* 38**,** 857-875.

Clarke, D. J., George, S. G. & Burchell, B. (1991) Glucuronidation in fish. *Aquatic Toxicology,* 20**,** 35-56.

Cleuvers, M. (2003) Aquatic ecotoxicity of pharmaceuticals including the assessment of combination effects. *Toxicology Letters,* 142**,** 185-194.

Cloarec, O., Dumas, M. E., Trygg, J., Craig, A., Barton, R. H., Lindon, J. C., Nicholson, J. K. & Holmes, E. (2005) Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H NMR spectroscopic metabonomic studies. *Analytical Chemistry,* 77**,** 517-526.

Coen, M., Holmes, E., Lindon, J. C. & Nicholson, J. K. (2008) NMR-based metabolic profiling and metabonomic approaches to problems in molecular toxicology. *Chemical Research in Toxicology,* 21**,** 9-27.

Cohen, A., Klint, K., Bowadt, S., Persson, P. & Jonsson, J. A. (2001) Routine analysis of alcohol and nonylphenol polyethoxylates in wastewater and sludge using

liquid chromatography-electrospray mass spectrometry. *Journal of Chromatography A,* 927**,** 103-110.

Comisarow, M. B. & Marshall, A. G. (1974a) Fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters,* 25**,** 282-283.

Comisarow, M. B. & Marshall, A. G. (1974b) Frequency-sweep fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters,* 26**,** 489-490.

Commission (2000) Commission E. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Commission of the European Communities.

Corada-Fernandez, C., Lara-Martin, P. A., Candela, L. & Gonzalez-Mazo, E. (2011) Tracking sewage derived contamination in riverine settings by analysis of synthetic surfactants. *Journal of Environmental Monitoring,* 13**,** 2010-2017.

Coronado, M., De Haro, H., Deng, X., Rempel, M. A., Lavado, R. & Schlenk, D. (2008) Estrogenic activity and reproductive effects of the UV-filter oxybenzone (2-hydroxy-4-methoxyphenyl-methanone) in fish. *Aquatic Toxicology,* 90**,** 182-187.

Coughtrie, M. W. H. (2002) Sulfation through the looking glass-recent advances in sulfotransferase research for the curious. *Pharmacogenomics Journal,* 2**,** 297-308.

Cowan, D. A., Kicman, A. T., Kubli-Garfias, C. & Welchman, H. J. (2008) Ion trap MS/MS of intact testosterone and epitestosterone conjugates-Adducts, fragile ions and the advantages of derivatisation. *Steroids,* 73**,** 621-628.

D'Arrigo, P. & Servi, S. (2010) Synthesis of lysophospholipids. *Molecules,* 15**,** 1354-1377.

Dann, A. B. & Hontela, A. (2011) Triclosan: Environmental exposure, toxicity and mechanisms of action. *Journal of Applied Toxicology,* 31**,** 285-311.

Daughton, C. G. & Ternes, T. A. (1999) Pharmaceuticals and personal care products in the environment: Agents of subtle change? *Environmental Health Perspectives,* 107**,** 907–942.

De Hoffmann, E. & Stroobant, V. (Eds.) (2007) *Mass Spectrometry: Principles and Applications,* England, John Wiley & Sons.

DeLorenzo, M. E. & Fleming, J. (2008) Individual and mixture effects of selected pharmaceuticals and personal care products on the marine phytoplankton species *Dunaliella tertiolecta*. *Archives of Environmental Contamination and Toxicology,* 54**,** 203-210.

Desbrow, C., Routledge, E. J., Brighty, G. C., Sumpter, J. P. & Waldock, M. (1998) Identification of estrogenic chemicals in STW effluent. 1. Chemical fractionation and *in vitro* biological screening. *Environmental Science & Technology,* 32**,** 1549-1558.

Dettmer, K., Aronov, P. A. & Hammock, B. D. (2007) Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews,* 26**,** 51-78.

Di Corcia, A., Cavallo, R., Crescenzi, C. & Nazzari, M. (2000) Occurrence and abundance of dicarboxylated metabolites of nonylphenol polyethoxylate surfactants in treated sewages. *Environmental Science & Technology,* 34**,** 3914-3919.

Di Corcia, A., Costantino, A., Crescenzi, C., Marinoni, E. & Samperi, R. (1998) Characterization of recalcitrant intermediates from biotransformation of the branched alkyl side chain of nonylphenol ethoxylate surfactants. *Environmental Science & Technology,* 32**,** 2401-2409.

Di Corcia, A., Samperi, R. & Marcomini, A. (1994) Monitoring aromatic surfactants and their biodegradation intermediates in raw and treated sewages by solid-phase extraction and liquid chromatography. *Environmental Science & Technology,* 28**,** 850-858.

Di Giulio, R. T. & Hinton, D. E. (Eds.) (2008) *The Toxicology of Fishes,* US, CRC Press, Taylor & Francis Group.

Ding, W.-H., Fujita, Y. & Reinhard, M. (1994) Chemical ionization mass spectra of linear alcohol polyethoxy carboxylates and polyethylene glycol dicarboxylates. *Rapid Communications in Mass Spectrometry,* 8**,** 1016-1020.

Donald G, B. (2008) Medical toxicology of natural substances: Foods, fungi, medicinal herbs, plants, and venomous animals. *John Wiley & Sons, Canada*.

Dorantes, A. & Stavchansky, S. (1992) Pharmacokinetic and metabolic disposition of *p*-chloro-*m*-xylenol (PCMX) in dogs. *Pharmaceutical Research,* 9**,** 677-682.

Douglas B, K. (2006) Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discovery Today,* 11**,** 1085-1092.

Dunn, W., Broadhurst, D., Deepak, S., Buch, M., McDowell, G., Spasic, I., Ellis, D., Brooks, N., Kell, D. & Neyses, L. (2007) Serum metabolomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. *Metabolomics,* 3**,** 413-426.

Dunn, W. B. & Ellis, D. I. (2005) Metabolomics: Current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry,* 24**,** 285-294.

Dunphy, J. C., Pessler, D. G., Morrall, S. W., Evans, K. A., Robaugh, D. A., Fujimoto, G. & Negahban, A. (2001) Derivatization LC/MS for the simultaneous determination of fatty alcohol and alcohol ethoxylate surfactants in water and wastewater samples. *Environmental Science & Technology,* 35**,** 1223-1230.

Eadsforth, C. V., Sherren, A. J., Selby, M. A., Toy, R., Eckhoff, W. S., McAvoy, D. C. & Matthijs, E. (2006) Monitoring of environmental fingerprints of alcohol ethoxylates in Europe and Canada. *Ecotoxicology and Environmental Safety,* 64**,** 14-29.

Ekelund, R., Bergman, Å., Granmo, Å. & Berggren, M. (1990) Bioaccumulation of 4-nonylphenol in marine animals— A re-evaluation. *Environmental Pollution,* 64**,** 107-120.

Ekman, D., Teng, Q., Villeneuve, D., Kahl, M., Jensen, K., Durhan, E., Ankley, G. & Collette, T. (2009a) Profiling lipid metabolites yields unique information on sex- and time-dependent responses of fathead minnows *Pimephales promelas* exposed to 17α-ethynylestradiol. *Metabolomics,* 5**,** 22-32.

Ekman, D. R., Teng, Q., Villeneuve, D. L., Kahl, M. D., Jensen, K. M., Durhan, E. J., Ankley, G. T. & Collette, T. W. (2008) Investigating compensation and recovery of fathead minnow (*Pimephales promelas*) exposed to 17α-ethynylestradiol with metabolite profiling. *Environmental Science & Technology,* 42**,** 4188-4194.

Ekman, R., Silberring, J. S., Westman-Brinkmalm, A. M. & Kraj, A. (Eds.) (2009b) *Mass Spectrometry: Instrumentation, Interpretation, and Applications,* Canada, John Wiley & Sons, Inc.

EPA Washington, D. (2004) Primer for municipal waste water treatment systems.

Eriksson, L., Johanson, E., Kettaneh-Wold, N., Trygg, J., Wikstrom, C. & Wold, S. (Eds.) (2006) *Multi- and megavariate data analysis. Part I: Basic principles and applications,* Sweden, Umetrics AB.

Eriksson, L., Johansson, E., Kettaneh-Wold, N. & Wold, S. (2001) *Multi- and Megavariate Data Analysis: Principles and Applications*, Umetrics AB, Sweden.

Evans, A. & Fornasini, G. (2003) Pharmacokinetics of L-carnitine. *Clinical Pharmacokinetics,* 42**,** 941-967.

The exposome: A powerful approach for evaluating environmental exposures and their influences on human disease. (2010) *Emerging Science for Environmental Health Decisions.*

Fåhræus-Van Ree, G. E. & Spurrell, D. R. (2000) Effect of toxaphene on isolated hepatocytes of the yellowtail flounder, *Pleuronectes ferrugineus* Storer. *Ecotoxicology and Environmental Safety,* 46**,** 289-297.

Falany, C. N. (1997) Enzymology of human cytosolic sulfotransferases. *The FASEB Journal,* 11**,** 206-216.

Fancy, S.-A., Beckonert, O., Darbon, G., Yabsley, W., Walley, R., Baker, D., Perkins, G. L., Pullen, F. S. & Rumpel, K. (2006) Gas chromatography/flame ionisation detection mass spectrometry for the detection of endogenous urine metabolites for metabonomic studies and its use as a complementary tool to nuclear magnetic resonance spectroscopy. *Rapid Communications in Mass Spectrometry,* 20**,** 2271-2280.

Fang, J. & Barcelona, M. J. (1998) Structural determination and quantitative analysis of bacterial phospholipids using liquid chromatography/electrospray ionization/mass spectrometry. *Journal of Microbiological Methods,* 33**,** 23-35.

Feige, J. N., Gerber, A., Casals-Casas, C., Yang, Q., Winkler, C., Bedu, E., Bueno, M., Gelman, L., Auwerx, J., Gonzalez, F. J. & Desvergne, B. (2009) The pollutant diethylhexyl phthalate regulates hepatic energy metabolism via species-specific PPARα-dependent mechanisms. *Environmental Health Perspectives,* 118.

Fenlon, K. A., Johnson, A. C., Tyler, C. R. & Hill, E. M. (2010) Gas-liquid chromatography-tandem mass spectrometry methodology for the quantitation of estrogenic contaminants in bile of fish exposed to wastewater treatment works effluents and from wild populations. *Journal of Chromatography A,* 1217**,** 112-118.

Fenn, J., Mann, M., Meng, C., Wong, S. & Whitehouse, C. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science,* 246**,** 64-71.

Fent, K., Weston, A. A. & Caminada, D. (2006) Ecotoxicology of human pharmaceuticals. *Aquatic Toxicology,* 76**,** 122-159.

Ferdinandusse, S., Denis, S., Dacremont, G. & Wanders, R. J. A. (2009) Toxicity of peroxisomal C27-bile acid intermediates. *Molecular Genetics and Metabolism,* 96**,** 121-128.

Ferreira-Leach, A. M. R. & Hill, E. M. (2001) Bioconcentration and distribution of 4-tert-octylphenol residues in tissues of the rainbow trout (*Oncorhynchus mykiss*). *Marine Environmental Research,* 51**,** 75-89.

Ferrer, I., Fernandez-Alba, A., Zweigenbaum, J. A. & Thurman, E. M. (2006) Exact-mass library for pesticides using a molecular-feature database. *Rapid Communications in Mass Spectrometry,* 20**,** 3659-3668.

Ferrer, I., Garcia-Reyes, J. F., Mezcua, M., Thurman, E. M. & Fernandez-Alba, A. R. (2005) Multi-residue pesticide analysis in fruits and vegetables by liquid chromatography-time-of-flight mass spectrometry. *Journal of Chromatography A,* 1082**,** 81-90.

Fiehn, O. (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics,* 2**,** 155-68.

Fiehn, O. (2002) Metabolomics – The link between genotypes and phenotypes. *Plant Molecular Biology,* 48**,** 155-171.

Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N. & Willmitzer, L. (2000) Metabolite profiling for plant functional genomics. *Nature Biotechnology,* 18**,** 1157-1161.

Fil., G. G. G. P. S. (2001) Introduction to Drug Metabolism. *Nelson Thornes*.

Filby, A. L., Santos, E. M., Thorpe, K. L., Maack, G. & Tyler, C. R. (2007a) Gene expression profiling for understanding chemical causation of biological effects for complex mixtures: A case study on estrogens. *Environmental Science & Technology,* 41**,** 8187-8194.

Filby, A. L., Thorpe, K. L., Maack, G. & Tyler, C. R. (2007b) Gene expression profiles revealing the mechanisms of anti-androgen-and estrogen-induced feminization in fish. *Aquatic Toxicology,* 81**,** 219-231.

Finne, E. F., Cooper, G. A., Koop, B. F., Hylland, K. & Tollefsen, K. E. (2007) Toxicogenomic responses in rainbow trout (*Oncorhynchus mykiss*) hepatocytes exposed to model chemicals and a synthetic mixture. *Aquatic Toxicology,* 81**,** 293-303.

Flores-Valverde, A. & Hill, E. (2008) Methodology for profiling the steroid metabolome in animal tissues using ultraperformance liquid chromatography-electrospray-time-of-flight mass spectrometry. *Analytical Chemistry,* 80**,** 8771-8779.

Flores-Valverde, A. M., Horwood, J. & Hill, E. M. (2010) Disruption of the steroid metabolome in fish caused by exposure to the environmental estrogen 17α-ethinylestradiol. *Environmental Science & Technology,* 44**,** 3552-3558.

Flores, A. (2008) The effects of the environmental endocrine disrupter, ethynyloestradiol, on steroid metabolism in the roach (*Rutilus rutilus*). University of Sussex.

Folmar, L. C., Denslow, N. D., Kroll, K., Orlando, E. F., Enblom, J., Marcino, J., Metcalfe, C. D. & Guillette, L. J. (2001) Altered serum sex steroids and vitellogenin induction in walleye (*Stizostedion vitreum*) collected near a metropolitan sewage treatment plant. *Archives Environmental Contamination Toxicology,* 40**,** 392-398.

Folmar, L. C., Denslow, N. D., Rao, V., Chow, M., Crain, D. A., Enblom, J., Marcino, J. & Guillette, L. J., Jr. (1996) Vitellogenin induction and reduced serum testosterone concentrations in feral male carp (*Cyprinus carpio*) captured near a major metropolitan sewage treatment plant. *Environmental Health Perspectives,* 104**,** 1096–1101.

Galloway, T. & Handy, R. (2003) Immunotoxicity of organophosphorous pesticides. *Ecotoxicology,* 12**,** 345-363.

Gao, J., Liu, L., Liu, X., Zhou, H., Huang, S. & Wang, Z. (2008) Levels and spatial distribution of chlorophenols – 2,4-Dichlorophenol, 2,4,6-trichlorophenol, and pentachlorophenol in surface water of China. *Chemosphere,* 71**,** 1181-1187.

Garcia-Reyero, N., Adelman, I., Liu, L. & Denslow, N. (2008) Gene expression profiles of fathead minnows exposed to surface waters above and below a sewage treatment plant in Minnesota. *Marine Environmental Research,* 66**,** 134-136.

García, M. C. (2005) The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography–electrospray mass spectrometry. *Journal of Chromatography B,* 825**,** 111-123.

Gardell, S. E., Dubin, A. E. & Chun, J. (2006) Emerging medicinal roles for lysophospholipid signaling. *Trends in Molecular Medicine,* 12**,** 65-75.

Ge, F., Zhu, L. & Wang, J. (2008) Distribution of chlorination products of phenols under various pHs in water disinfection. *Desalination,* 225**,** 156-166.

Geoff, B. (2001) The fate of xenobiotic organic compounds in wastewater treatment plants. *Water Research,* 35**,** 2523-2533.

George, S. G., Leaver, M. J. & Wright, J. (1998) Structural Studies of a UDP-glucuronosyltransferase gene from the plaice (*Pleuronectes platessa*). *Marine Environmental Research,* 46**,** 33-35.

Gibson, R., Smith, M. D., Spary, C. J., Tyler, C. R. & Hill, E. M. (2005a) Mixtures of estrogenic contaminants in bile of fish exposed to wastewater treatment works effluents. *Environmental Science and Technology,* 39**,** 2461-2471.

Gibson, R., Tyler, C. R. & Hill, E. M. (2005b) Analytical methodology for the identification of estrogenic contaminants in fish bile. *Journal of Chromatography A,* 1066**,** 33-40.

Gimeno, S., Komen, H., Gerritsen, A. G. M. & Bowmer, T. (1998) Feminisation of young males of the common carp, *Cyprinus carpio*, exposed to 4-tert-pentylphenol during sexual differentiation. *Aquatic Toxicology,* 43**,** 77-92.

Gimeno, S., Komen, H., Venderbosch, P. W. M. & Bowmer, T. (1997) Disruption of sexual differentiation in genetic male common carp (*Cyprinus carpio*) exposed to an alkylphenol during different life stages. *Environmental Science & Technology,* 31**,** 2884-2890.

Glinski, M. & Weckwerth, W. (2006) The role of mass spectrometry in plant systems biology. *Mass Spectrometry Reviews,* 25**,** 173-214.

Gomes, R. L., Scrimshaw, M. D. & Lester, J. N. (2009) Fate of conjugated natural and synthetic steroid estrogens in crude sewage and activated sludge batch studies. *Environmental Science & Technology,* 43**,** 3612-3618.

González-Mariño, I., Quintana, J. B., Rodríguez, I. & Cela, R. (2011) Evaluation of the occurrence and biodegradation of parabens and halogenated by-products in wastewater by accurate-mass liquid chromatography-quadrupole-time-of-flight-mass spectrometry (LC-QTOF-MS). *Water Research,* 45**,** 6770-6780.

González-Mazo, E., Honing, M., Barceló, D. & Gómez-Parra, A. (1997) Monitoring long-chain intermediate products from the degradation of linear alkylbenzene sulfonates in the marine environment by solid-phase extraction followed by liquid chromatography/ionspray mass spectrometry. *Environmental Science & Technology,* 31**,** 504-510.

Goodacre, R. (2005) Making sense of the metabolome using evolutionary computation: Seeing the wood with the trees. *Journal of Experimental Botany,* 56**,** 245-254.

Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G. & Kell, D. B. (2004) Metabolomics by numbers: Acquiring and understanding global metabolite data. *Trends in Biotechnology,* 22**,** 245-252.

Goto, T., Holzinger, F., Hagey, L. R., Cerrè, C., Ton-Nu, H.-T., Schteingart, C. D., Steinbach, J. H., Shneider, B. L. & Hofmann, A. F. (2003) Physicochemical and physiological properties of 5α-cyprinol sulfate, the toxic bile salt of cyprinid fish. *Journal of Lipid Research,* 44**,** 1643-1651.

Gross, J. H. (Ed.) (2004) *Mass Spectrometry: A Textbook,* Berlin Heidelberg, Springer-Verlag.

Gu, J. J., Hofmann, A. F., Ton-Nu, H. T., Schteingart, C. D. & Mysels, K. J. (1992) Solubility of calcium salts of unconjugated and conjugated natural bile acids. *Journal of Lipid Research,* 33**,** 635-46.

Guillarme, D., Nguyen, D. T. T., Rudaz, S. & Veuthey, J.-L. (2007) Recent developments in liquid chromatography—Impact on qualitative and quantitative performance. *Journal of Chromatography A,* 1149**,** 20-29.

Hall, R. (2006) Plant metabolomics: From holistic hope, to hype, to hot topic. *New Phytologist,* 169**,** 453-468.

Halling-Sorensen, B., Nors Nielsen, S., Lanzky, P. F., Ingerslev, F., Holten Lutzhoft, H. C. & Jorgensen, S. E. (1998) Occurrence, fate and effects of pharmaceutical substances in the environment- A review. *Chemosphere,* 36**,** 357-393.

Han, M. S., Park, S. Y., Shinzawa, K., Kim, S., Chung, K. W., Lee, J.-H., Kwon, C. H., Lee, K.-W., Lee, J.-H., Park, C. K., Chung, W. J., Hwang, J. S., Yan, J.-J., Song, D.-K., Tsujimoto, Y. & Lee, M.-S. (2008) Lysophosphatidylcholine as a death effector in the lipoapoptosis of hepatocytes. *Journal of Lipid Research,* 49**,** 84-97.

Harries, J. E., Janbakhsh, A., Jobling, S., Matthiessen, P., Sumpter, J. P. & Tyler, C. R. (1999) Estrogenic potency of effluent from two sewage treatment works in the United Kingdom. *Environmental Toxicology and Chemistry,* 18**,** 932-937.

Harvey Motulsky, M. D. (Ed.) (1995) *Intuitive Biostatistics,* New York, Oxford University Press.

Heath, R. J., Rubin, J. R., Holland, D. R., Zhang, E., Snow, M. E. & Rock, C. O. (1999) Mechanism of triclosan inhibition of bacterial fatty acid synthesis. *Journal of Biological Chemistry,* 274**,** 11110-11114.

Heberer, T. (2002) Occurrence, fate, and assessment of polycyclic musk residues in the aquatic environment of urban areas — A review. *Acta hydrochimica et hydrobiologica,* 30**,** 227-243.

Heberer, T. & Stan, H.-J. (1997) Detection of more than 50 substituted phenols as their t-butyldimethylsilyl derivatives using gas chromatography-mass spectrometry. *Analytica Chimica Acta,* 341**,** 21-34.

Hecker, M., Tyler, C. R., Hoffmann, M., Maddix, S. & Karbe, L. (2002) Plasma biomarkers in fish provide evidence for endocrine modulation in the Elbe River, Germany. *Environmental Science & Technology,* 36**,** 2311-2321.

Hellenäs, K. E., Nyman, A., Slanina, P., Lööf, L. & Gabrielsson, J. (1992) Determination of potato glycoalkaloids and their aglycone in blood serum by high-performance liquid chromatography: Application to pharmacokinetic studies in humans. *Journal of Chromatography B: Biomedical Sciences and Applications,* 573**,** 69-78.

Hemming, J. M., Waller, W. T., Chow, M. C., Denslow, N. D. & Venables, B. (2001) Assessment of the estrogenicity and toxicity of a domestic wastewater effluent flowing through a constructed wetland system using biomarkers in male fathead minnows (*Pimephales promelas* rafinesque, 1820). *Environmental Toxicology and Chemistry,* 20**,** 2268-2275.

HERA (2009a) Human & environmental risk assessment on ingredients of european household cleaning products, linear alkylbenzene sulphonate. http://www.heraproject.com/RiskAssessment.cfm.

HERA (2009b) Human and environmental risk assessment on ingredients of european household cleaning products, alcohol ethoxylates (AEOs). http://www.heraproject.com/RiskAssessment.cfm.

Hill, E. M., Evans, K. L., Horwood, J., Rostkowski, P., Oladapo, F. O., Gibson, R., Shears, J. A. & Tyler, C. R. (2010) Profiles and some initial identifications of (anti)androgenic compounds in fish exposed to wastewater treatment works effluents. *Environmental Science & Technology,* 44**,** 1137-1143.

Hines, A., Oladiran, G. S., Bignell, J. P., Stentiford, G. D. & Viant, M. R. (2007) Direct sampling of organisms from the field and knowledge of their phenotype: Key recommendations for environmental metabolomics. *Environmental Science & Technology,* 41**,** 3375-3381.

Hines, A., Staff, F. J., Widdows, J., Compton, R. M., Falciani, F. & Viant, M. R. (2010) Discovery of metabolic signatures for predicting whole organism toxicology. *Toxicological Sciences,* 115**,** 369-378.

Hirabayashi, Y., Igarashi, Y., Merrill, A. H. & Jr. (Eds.) (2006) *Sphingolipid Biology,* Tokyo, Springer-Verlag.

Hirom, P. C., Millburn, P., Smith, R. L. & Williams, R. T. (1972) Species variations in the threshold molecular-weight factor for the biliary excretion of organic anions. *Biochemical Journal,* 129**,** 1071-1077.

Hiromori, Y., Nishikawa, J.-i., Yoshida, I., Nagase, H. & Nakanishi, T. (2009) Structure-dependent activation of peroxisome proliferator-activated receptor

(PPAR) γ by organotin compounds. *Chemico-Biological Interactions,* 180**,** 238-244.

Hochachka, P. W. & Mommsen, T. P. (Eds.) (1995) *Biochemistry and Molecular Biology of Fishes, Metabolic Biochemistry,* Amsterdam, The Netherlands, Elsevier Science B.V.

Hofmann, A. F. (1999) The continuing importance of bile acids in liver and intestinal disease. *Archives of Internal Medicine,* 159**,** 2647-2658.

Hollywood, K., Brison, D. R. & Goodacre, R. (2006) Metabolomics: Current technologies and future trends. *Proteomics,* 6**,** 4716-23.

Holmes, E., Loo, R. L., Cloarec, O., Coen, M., Tang, H., Maibaum, E., Bruce, S., Chan, Q., Elliott, P., Stamler, J., Wilson, I. D., Lindon, J. C. & Nicholson, J. K. (2007) Detection of urinary drug metabolite (xenometabolome) signatures in molecular epidemiology studies via statistical total correlation (NMR) spectroscopy. *Analytical Chemistry,* 79**,** 2629-2640.

Holmes, E., Wilson, I. D. & Nicholson, J. K. (2008) Metabolic Phenotyping in Health and Disease. *Cell,* 134**,** 714-717.

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

Houde, F., DeBlois, C. & Berryman, D. (2002) Liquid chromatographic-tandem mass spectrometric determination of nonylphenol polyethoxylates and nonylphenol carboxylic acids in surface water. *Journal of Chromatography A,* 961**,** 245-256.

Houtman, C. J., Van Oostveen, A. M., Brouwer, A., Lamoree, M. H. & Legler, J. (2004) Identification of estrogenic compounds in fish bile using bioassay-directed fractionation. *Environmental Science and Technology,* 38**,** 6415-6423.

Hovander, L., Malmberg, T., Athanasiadou, M., Athanassiadis, I., Rahm, S., Bergman & Wehler, E. K. (2002) Identification of hydroxylated PCB metabolites and other phenolic halogenated pollutants in human blood plasma. *Archives of Environmental Contamination and Toxicology,* 42**,** 105-117.

Idborg-Björkman, H., Edlund, P.-O., Kvalheim, O. M., Schuppe-Koistinen, I. & Jacobsson, S. P. (2003) Screening of biomarkers in rat urine using LC/electrospray ionization-MS and two-way data analysis. *Analytical Chemistry,* 75**,** 4784-4792.

Ifelebuegu, M. O. (2011) The fate and behavior of selected endocrine disrupting chemicals in full scale wastewater and sludge treatment unit processes. *International Journal of Environmental Science and Technology,* 8**,** 245-254.

Iguchi, T., Watanabe, H. & Katsu, Y. (2006) Application of ecotoxicogenomics for studying endocrine disruption in vertebrates and invertebrates. *Environmental Health Perspectives,* 114.

Izquierdo-García, J. L., Villa, P., Kyriazis, A., del Puerto-Nevado, L., Pérez-Rial, S., Rodriguez, I., Hernandez, N. & Ruiz-Cabello, J. (2011) Descriptive review of current NMR-based metabolomic data analysis packages. *Progress in Nuclear Magnetic Resonance Spectroscopy,* 59**,** 263-270.

James, M. O., Sikazwe, D. N. & Gadagbui, B. K. M. (1998) Isolation of a Pi class glutathione S-transferase (GST) from catfish intestinal mucosa. *Marine Environmental Research,* 46**,** 57-60.

Jandacek, R. J. & Tso, P. (2007) Enterohepatic circulation of organochlorine compounds: A site for nutritional intervention. *The Journal of Nutritional Biochemistry,* 18**,** 163-167.

Jelic, A., Gros, M., Ginebreda, A., Cespedes-Sánchez, R., Ventura, F., Petrovic, M. & Barcelo, D. (2011) Occurrence, partition and removal of pharmaceuticals in sewage water and sludge during wastewater treatment. *Water Research,* 45**,** 1165-1176.

Jensen, P. H., Juhler, R. K., Nielsen, N. J., Hansen, T. H., Strobel, B. W., Jacobsen, O. S., Nielsen, J. & Hansen, H. C. B. (2008) Potato glycoalkaloids in soil-optimising liquid chromatography–time-of-flight mass spectrometry for quantitative studies. *Journal of Chromatography A,* 1182**,** 65-71.

Jeon, H.-K., Chung, Y. & Ryu, J.-C. (2006) Simultaneous determination of benzophenone-type UV filters in water and soil by gas chromatography–mass spectrometry. *Journal of Chromatography A,* 1131**,** 192-202.

Jiang, J. Q., Yin, Q., Zhou, J. L. & Pearce, P. (2005) Occurrence and treatment trials of endocrine disrupting chemicals (EDCs) in wastewaters. *Chemosphere,* 61**,** 544-550.

Joachim, K. (2006) Current challenges and developments in GC–MS based metabolite profiling technology. *Journal of Biotechnology,* 124**,** 312-322.

Jobling, S., Beresford, N., Nolan, M., Rodgers-Gray, T., Brighty, G. C., Sumpter, J. P. & Tyler, C. R. (2002) Altered sexual maturation and gamete production in wild roach (*Rutilus rutilus*) living in rivers that receive treated sewage effluents. *Biology of Reproduction,* 66**,** 272-281.

Jobling, S., Nolan, M., Tyler, C. R., Brighty, G. & Sumpter, J. P. (1998) Widespread sexual disruption in wild fish. *Environmental Science & Technology,* 32**,** 2498-2506.

Jobling, S., Reynolds, T., White, R., Parker, M. G. & Sumpter, J. P. (1995) A variety of environmentally persistent chemicals, including some phthalate plasticizers, are weakly estrogenic. *Environmental Health Perspectives,* 103**,** 582-7.

Jobling, S., Sumpter, J. P., Sheahan, D., Osborne, J. A. & Matthiessen, P. (1996) Inhibition of testicular growth in rainbow trout (*Oncorhynchus mykiss*) exposed to estrogenic alkylphenolic chemicals. *Environmental Toxicology and Chemistry,* 15**,** 194-202.

Jobling, S., Williams, R., Johnson, A., Taylor, A., Gross-Sorokin, M., Nolan, M., Tyler, C. R., van Aerle, R., Santos, E. & Brighty, G. (2005) Predicted Exposures to Steroid Estrogens in U.K. Rivers Correlate with Widespread Sexual Disruption in Wild Fish Populations. *Environ Health Perspect,* 114.

Johnson, A. C., Acreman, M. C., Dunbar, M. J., Feist, S. W., Giacomello, A. M., Gozlan, R. E., Hinsley, S. A., Ibbotson, A. T., Jarvie, H. P., Jones, J. I., Longshaw, M., Maberly, S. C., Marsh, T. J., Neal, C., Newman, J. R., Nunn, M. A., Pickup, R. W., Reynard, N. S., Sullivan, C. A., Sumpter, J. P. & Williams, R. J. (2009) The British river of the future: How climate change and human activity might affect two contrasting river ecosystems in England. *Science of the Total Environment,* 407**,** 4787-4798.

Johnson, A. C., Belfroid, A. & Di Corcia, A. (2000) Estimating steroid oestrogen inputs into activated sludge treatment works and observations on their removal from the effluent. *Science of the Total Environment,* 256**,** 163-173.

Johnson, A. C., Jürgens, M. D., Williams, R. J., Kümmerer, K., Kortenkamp, A. & Sumpter, J. P. (2008) Do cytotoxic chemotherapy drugs discharged into rivers pose a risk to the environment and human health? An overview and UK case study. *Journal of Hydrology,* 348**,** 167-175.

Johnson, C. H., Patterson, A. D., Idle, J. R. & Gonzalez, F. J. (2012) Xenobiotic metabolomics: Major impact on the metabolome. *Annual Review Pharmacology Toxicology,* 52**,** 37-56.

Johnson, C. H., Patterson, A. D., Krausz, K. W., Lanz, C., Kang, D. W., Luecke, H., Gonzalez, F. J. & Idle, J. R. (2011) Radiation metabolomics. 4. UPLC-ESI-QTOFMS-based metabolomics for urinary biomarker discovery in gamma-irradiated rats. *Radiation Research,* 175**,** 473-484.

Johnson, E. G. & Nier, A. O. (1953) Angular aberrations in sector shaped electromagnetic lenses for focusing beams of charged particles *Physical Review,* 91**,** 10-17.

Jones, L. L., McDonald, D. A. & Borum, P. R. (2010) Acylcarnitines: Role in brain. *Progress in Lipid Research,* 49**,** 61-75.

Jonsson, G., Stokke, T. U., Cavcic, A., Jørgensen, K. B. & Beyer, J. (2008) Characterization of alkylphenol metabolites in fish bile by enzymatic treatment and HPLC-fluorescence analysis. *Chemosphere,* 71**,** 1392-1400.

Joss, A., Zabczynski, S., Göbel, A., Hoffmann, B., Löffler, D., McArdell, C. S., Ternes, T. A., Thomsen, A. & Siegrist, H. (2006) Biological degradation of pharmaceuticals in municipal wastewater treatment: Proposing a classification scheme. *Water Research,* 40**,** 1686-1696.

K.L. Rule, V. R. E., P.J. Vikesland (2005) *Environ. Sci. Technol.,* 39**,** 3176.

Kallio, J.-M., Lahti, M., Oikari, A. & Kronberg, L. (2010) Metabolites of the Aquatic Pollutant Diclofenac in Fish Bile. *Environmental Science & Technology,* 44**,** 7213-7219.

Kao, L. R. & Birnbaum, L. S. (1986) Disposition of *O*-benzyl-*p*-chlorophenol in male rats. *Journal of Toxicology and Environmental Health,* 18**,** 441-458.

Karas, M., Bachmann, D., Bahr, U. & Hillenkamp, F. (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes,* 78**,** 53-68.

Karas, M. & Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry,* 60**,** 2299-2301.

Kasprzyk-Hordern, B., Dinsdale, R. M. & Guwy, A. J. (2008) The occurrence of pharmaceuticals, personal care products, endocrine disruptors and illicit drugs in surface water in South Wales, UK. *Water Research,* 42**,** 3498-3518.

Katsiadaki, I., Morris, S., Squires, C., Hurst, M. R., James, J. D. & Scott, A. P. (2006) Use of the three-spined stickleback (*Gasterosteus aculeatus*) as a sensitive *in vivo* test for detection of environmental antiandrogens. *Environmental Health Perspectives,* 114.

Kennedy, C. J., Gill, K. A. & Walsh, P. J. (1989) Thermal modulation of benzo[a]pyrene metabolism by the gulf toadfish, *Opsanus beta*. *Aquatic Toxicology,* 15**,** 331-343.

Kenny, L. C., Broadhurst, D., Brown, M., Dunn, W. B., Redman, C. W. G., Kell, D. B. & Baker, P. N. (2008) Detection and identification of novel metabolomic biomarkers in Preeclampsia. *Reproductive Sciences,* 15**,** 591-597.

Kim, H. K., Wilson, E. G., Choi, Y. H. & Verpoorte, R. (2010) Metabolomics: A tool for anticancer lead-finding from natural products. *Planta Medica,* 76**,** 1094,1102.

Kim, K.-B., Chung, M., Um, S., Oh, J., Kim, S., Na, M., Oh, H., Cho, W.-S. & Choi, K. (2008) Metabolomics and biomarker discovery: NMR spectral data of urine and hepatotoxicity by carbon tetrachloride, acetaminophen, and D-galactosamine in rats. *Metabolomics,* 4**,** 377-392.

Kind, T. & Fiehn, O. (2006) Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC bioinformatics,* 7**,** 234.

Kind, T. & Fiehn, O. (2007) Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC bioinformatics,* 8**,** 105.

Kiparissis, Y., Metcalfe, T. L., Balch, G. C. & Metcalfe, C. D. (2003) Effects of the antiandrogens, vinclozolin and cyproterone acetate on gonadal development in the Japanese medaka (*Oryzias latipes*). *Aquatic Toxicology,* 63**,** 391-403.

Kobayashi, K., Kimura, S. & Oshima, Y. (1984) Sulfate conjugation of various phenols by liver-soluble fraction of goldfish. *Bulletin of the Japanese Society of Scientific Fisheries,* 50**,** 833-837.

Koh, Y. K. K., Chiu, T. Y., Boobis, A. R., Scrimshaw, M. D., Bagnall, J. P., Soares, A., Pollard, S., Cartmell, E. & Lester, J. N. (2009) Influence of operating parameters on the biodegradation of steroid estrogens and nonylphenolic compounds during biological wastewater treatment processes. *Environmental Science & Technology,* 43**,** 6646-6654.

Kolpin, D. W., Furlong, E. T., Meyer, M. T., Thurman, E. M., Zaugg, S. D., Barber, L. B. & Buxton, H. T. (2002) Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999-2000: A national reconnaissance. *Environmental Science & Technology,* 36**,** 1202-1211.

Kristian Hovde, L. (2011) Multivariate methods in metabolomics – From pre-processing to dimension reduction and statistical analysis. *TrAC Trends in Analytical Chemistry,* 30**,** 827-841.

Krogh, K. A., Halling-Sørensen, B., Mogensen, B. B. & Vejrup, K. V. (2003) Environmental properties and effects of nonionic surfactant adjuvants in pesticides: A review. *Chemosphere,* 50**,** 871-901.

Kunz, P. & Fent, K. (2006) Multiple hormonal activities of UV filters and comparison of *in vivo* and *in vitro* estrogenic activity of ethyl-4-aminobenzoate in fish. *Aquatic Toxicology,* 79**,** 305-324.

Kwintkiewicz, J., Nishi, Y., Yanase, T. & Giudice, L. C. (2009) Peroxisome proliferator-activated receptor-γ mediates bisphenol A inhibition of FSH-stimulated IGF-1, aromatase, and estradiol in human granulosa cells. *Environmental Health Perspectives,* 118.

Laetz, C. A., Baldwin, D. H., Collier, T. K., Hebert, V., Stark, J. D. & Scholz, N. L. (2008) The synergistic toxicity of pesticide mixtures: Implications for risk assessment and the conservation of endangered pacific salmon. *Environmental Health Perspectives,* 117.

Lafaye, A., Junot, C., Gall, B. R.-L., Fritsch, P., Tabet, J.-C. & Ezan, E. (2003) Metabolite profiling in rat urine by liquid chromatography/electrospray ion trap mass spectrometry. Application to the study of heavy metal toxicity. *Rapid Communications in Mass Spectrometry,* 17**,** 2541-2549.

Langenheim, J. H. (Ed.) (2003) *Plant Resins: Chemistry, Evolution, Ecology, and Ethnobotany,* USA, Timber Press.

Lara-Martín, P. A., Gómez-Parra, A. & González-Mazo, E. (2008) Sources, transport and reactivity of anionic and non-ionic surfactants in several aquatic ecosystems in SW Spain: A comparative study. *Environmental Pollution,* 156**,** 36-45.

Lara-Martín, P. A., Gómez-Parra, A., Köchling, T., Sanz, J. L. & González-Mazo, E. (2007) Monitoring the primary biodegradation of linear alkylbenzene sulfonates

and their coproducts in anoxic sediments using liquid chromatography–mass spectrometry. *Environmental Science & Technology,* 41**,** 3580-3586.

Lara-Martín, P. A., Gómez-Parra, A., Sanz, J. L. & González-Mazo, E. (2010) Anaerobic degradation pathway of linear alkylbenzene sulfonates (LAS) in sulfate-reducing marine sediments. *Environmental Science & Technology,* 44**,** 1670-1676.

Larsson, D. G. J., Adolfsson-Erici, M., Parkkonen, J., Pettersson, M., Berg, A. H., Olsson, P. E. & Förlin, L. (1999) Ethinyloestradiol — An undesired fish contraceptive? *Aquatic Toxicology,* 45**,** 91-97.

Latta, R. K., Fiander, H., Ross, N. W., Simpson, C. & Schneider, H. (1993) Toxicity of bile acids to colon cancer cell lines. *Cancer Letters,* 70**,** 167-173.

Lee, H. B. (1999) Review of analytical methods for the determination of nonylphenol and related compounds in environmental samples. *Water Quality Research Journal of Canada,* 34**,** 33-35.

Lee, R., West, D., Phillips, S. M. & Britz-McKibbin, P. (2010) Differential metabolomics for quantitative assessment of oxidative stress with strenuous exercise and nutritional intervention: Thiol-specific regulation of cellular metabolism with N-acetyl-l-cysteine pretreatment. *Analytical Chemistry,* 82**,** 2959-2968.

Lemke, T. L., Williams, D. A., Roche, V. F. & Zito, S. W. (2007) Foye's principles of medicinal chemistry; Sixth edition. *Lippincott Williams & Wilkins, US*.

Levsen, K., Schiebel, H.-M., Behnke, B., Dötzer, R., Dreher, W., Elend, M. & Thiele, H. (2005) Structure elucidation of phase II metabolites by tandem mass spectrometry: an overview. *Journal of Chromatography A,* 1067**,** 55-72.

Lin, C. Y., Viant, M. R. & Tjeerdema, R. S. (2006) Metabolomics: Methodologies and applications in the environmental sciences. *Journal of Pesticide Science,* 31**,** 245-251.

Lindon, J. C., Holmes, E. & Nicholson, J. K. (2006) Metabonomics techniques and applications to pharmaceutical research & development. *Pharmaceutical Research,* 23**,** 1075-1088.

Lindon, J. C., Holmes, E. & Nicholson, J. K. (2007) Metabonomics in pharmaceutical R&D. *Federation of European Biochemical Societies (FEBS) Journal,* 274**,** 1140-1151.

Lindström, A., Buerge, I. J., Poiger, T., Bergqvist, P.-A., Müller, M. D. & Buser, H.-R. (2002) Occurrence and environmental behavior of the bactericide triclosan and its methyl derivative in surface waters and in wastewater. *Environmental Science & Technology,* 36**,** 2322-2329.

Loo, J. A., Udseth, H. R. & Smith, R. D. (1989) Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. *Analytical Biochemistry,* 179**,** 404-412.

Lutz, U., Bittner, N., Lutz, R. W. & Lutz, W. K. (2008) Metabolite profiling in human urine by LC–MS/MS: Method optimization and application for glucuronides from dextromethorphan metabolism. *Journal of Chromatography B,* 871**,** 349-356.

Lyndall, J., Fuchsman, P., Bock, M., Barber, T., Lauren, D., Leigh, K., Perruchon, E. & Capdevielle, M. (2010) Probabilistic risk evaluation for triclosan in surface water, sediments, and aquatic biota tissues. *Integrated Environmental Assessment and Management,* 6**,** 419-440.

MacKenzie, D. A., Defernez, M., Dunn, W. B., Brown, M., Fuller, L. J., de Herrera, S. R. M. S., Günther, A., James, S. A., Eagles, J., Philo, M., Goodacre, R. & Roberts, I. N. (2008) Relatedness of medically important strains of *Saccharomyces cerevisiae* as revealed by phylogenetics and metabolomics. *Yeast,* 25**,** 501-512.

Madalinski, G., Godat, E., Alves, S., Lesage, D., Genin, E., Levi, P., Labarre, J., Tabet, J.-C., Ezan, E. & Junot, C. (2008) Direct introduction of biological samples into a LTQ-orbitrap hybrid mass spectrometer as a tool for fast metabolome analysis. *Analytical Chemistry,* 80**,** 3291-3303.

Maguire, R. (1999) Review of the persistence of nonylphenol and nonylphenolethoxylates in aquatic environments. *Water Quality Research of Canada,* 34**,** 37-78.

Mahadevan, S., Shah, S. L., Marrie, T. J. & Slupsky, C. M. (2008) Analysis of metabolomic data using support vector machines. *Analytical Chemistry,* 80**,** 7562-7570.

Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. & Zagulin, V. A. (1973) The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Physics JETP,* 37**,** 45.

Mann, M., Meng, C. K. & Fenn, J. B. (1989) Interpreting mass spectra of multiply charged ions. *Analytical Chemistry,* 61**,** 1702-1708.

Marshall, A. G., Hendrickson, C. L. & Jackson, G. S. (1998) Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrometry Reviews,* 17**,** 1-35.

Martínez Bueno, M. J., Agüera, A., Gómez, M. J., Hernando, M. D., García-Reyes, J. F. & Fernández-Alba, A. R. (2007) Application of liquid chromatography/quadrupole-linear ion trap mass spectrometry and time-of-flight mass spectrometry to the determination of pharmaceuticals and related contaminants in wastewater. *Analytical Chemistry,* 79**,** 9372-9384.

Mattauch, J. (1936) Double-focusing mass spectrograph and the masses of N^{15} and O^{18}. *Physical Review,* 50**,** 617–623.

Matthijs, E., Holt, M. S., Kiewiet, A. & Rijs, G. B. J. (1999) Environmental monitoring for linear alkylbenzene sulfonate, alcohol ethoxylate, alcohol ethoxy sulfate, alcohol sulfate, and soap. *Environmental Toxicology and Chemistry,* 18**,** 2634-2644.

Maurer, M., Escher, B. I., Richle, P., Schaffner, C. & Alder, A. C. (2007) Elimination of β-blockers in sewage treatment plants. *Water Research,* 41**,** 1614-1622.

Mayer, B. (Ed.) (2011) *Bioinformatics for Omics Data Methods and Protocols,* Austria, Humana Press.

Mazéas, O. & Budzinski, H. (2005) Solid-phase extraction and purification for the quantification of polycyclic aromatic hydrocarbon metabolites in fish bile. *Analytical and Bioanalytical Chemistry,* 383**,** 985-990.

McClellan, J. E., Quarmby, S. T. & Yost, R. A. (2002) Parent and neutral loss monitoring on a quadrupole ion trap mass spectrometer: Screening of acylcarnitines in complex mixtures. *Analytical Chemistry,* 74**,** 5799-5806.

McDonnell, G. E. (Ed.) (2007) *Antisepsis, Disinfection, and Sterilization: Types, Action, and Resistance,* USA, ASM Press.

McGurk, K. A., Remmel, R. P., Hosagrahara, V. P., Tosh, D. & Burchell, B. (1996) Reactivity of mefenamic acid 1-O-acyl glucuronide with proteins *in vitro* and *ex vivo*. *Drug Metabolism and Disposition,* 24**,** 842-849.

Mehinto, A. C. (2009) Impacts of the Human Pharmaceutical Diclofenac in the Aquatic Environment. University of Exeter.

Meinertz, J. R., Stehly, G. R., Hubert, T. D. & Bernardy, J. A. (1999) Liquid chromatographic determination of benzocaine and N-acetylbenzocaine in the edible fillet tissue from rainbow trout. *Journal of Chromatography A,* 855**,** 255-260.

Meriläinen, P. & Oikari, A. (2008) Exposure assessment of fishes to a modern pulp and paper mill effluents after a black liquor spill. *Environmental Monitoring and Assessment,* 144**,** 419-435.

Metz, T. O., Page, J. S., Baker, E. S., Tang, K., Ding, J., Shen, Y. & Smith, R. D. (2008) High resolution separations and improved Ion production and transmission in metabolomics. *Trends in Analytical Chemistry,* 27**,** 205-214.

Mezcua, M., Gómez, M. J., Ferrer, I., Aguera, A., Hernando, M. D. & Fernández-Alba, A. R. (2004) Evidence of 2,7/2,8-dibenzodichloro-*p*-dioxin as a photodegradation product of triclosan in water and wastewater samples. *Analytica Chimica Acta,* 524**,** 241-247.

Millington, D., Norwood, D., Kodo, N., Roe, C. & Inoue, F. (1989) Application of fast atom bombardment with tandem mass spectrometry and liquid

chromatography/mass spectrometry to the analysis of acylcarnitines in human urine, blood, and tissue. *Analytical Biochemistry,* 180**,** 331-339.

Mills, G. & Moolenaar, W. (2003) The emerging role of lysophosphatidic acid in cancer. *Nature Reviews. Cancer,* 3**,** 582-591.

Miracle, A. L. & Ankley, G. T. (2005) Ecotoxicogenomics: Linkages between exposure and effects in assessing risks of aquatic contaminants to fish. *Reproductive Toxicology,* 19**,** 321-326.

Möder, M., Kießling, A., Löster, H. & Brüggemann, L. (2003) The pattern of urinary acylcarnitines determined by electrospray mass spectrometry: a new tool in the diagnosis of diabetes mellitus. *Analytical and Bioanalytical Chemistry,* 375**,** 200-210.

Mommsen, T. P. & Moon, T. W. (Eds.) (2005) The Netherlands, Elsevier B.V.

Moolenaar, W. H. (2000) Development of our current understanding of bioactive lysophospholipids. *Annals of the New York Academy of Sciences,* 905**,** 1-10.

Moran, M. J., Zogorski, J. S. & Squillace, P. J. (2005) MTBE and gasoline hydrocarbons in ground water of the United States. *Ground Water,* 43**,** 615-627.

Moran, M. J., Zogorski, J. S. & Squillace, P. J. (2006) Chlorinated solvents in groundwater of the United States. *Environmental Science & Technology,* 41**,** 74-81.

Morrison, N., Bearden, D., Bundy, J., Collette, T., Currie, F., Davey, M., Haigh, N., Hancock, D., Jones, O., Rochfort, S., Sansone, S.-A., Štys, D., Teng, Q., Field, D. & Viant, M. (2007) Standard reporting requirements for biological samples in metabolomics experiments: Environmental context. *Metabolomics,* 3**,** 203-210.

Muir, D. C. G. & Howard, P. H. (2006) Are There Other Persistent Organic Pollutants? A Challenge for Environmental Chemists†. *Environmental Science & Technology,* 40**,** 7157-7166.

Mulder, G. J. & Scholtens, E. (1978) The availability of inorganic sulphate in blood for sulphate conjugation of drugs in rat liver *in vivo*. (^{35}S)Sulphate incorporation into harmol sulphate. *The Biochemical Journal,* 172**,** 247-251.

Müller, M. T., Zehnder, A. J. B. & Escher, B. I. (1999) Membrane toxicity of linear alcohol ethoxylates. *Environmental Toxicology and Chemistry,* 18**,** 2767-2774.

Murakami, T., Ohoku, K., Yumoto, R., Yoshii, M., Une, M., Kuramoto, T., Hoshita, T. & Yata, N. (2000) Enhancing effect of 5 alpha-cyprinol sulfate on mucosal membrane permeability to sodium ampicillin in rats. *Eurpean Journal Pharmaceutical and Biopharmaceuticals,* 49**,** 111-117.

Murphy, R. C. (Ed.) (2002) *Mass Spectrometry of Phospholipids: Tables of Molecular and Product Ions,* Denver Colorado, USA, Illuminati Press.

Myers, D. S., Ivanova, P. T., Milne, S. B. & Brown, H. A. (2011) Quantitative analysis of glycerophospholipids by LC–MS: Acquisition, data handling, and interpretation. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* 1811**,** 748-757.

Narukawa, J., Inoue, H., Kato, S. & Yokota, H. (2004) Glucuronidation of 1-naphthol and excretion into the vein in perfused rat kidney. *Drug Metabolism and Disposition,* 32**,** 758-761.

Nicholson, J. K., Lindon, J. C. & Holmes, E. (1999) 'Metabonomics': Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica,* 29**,** 1181-1189.

Nicholson, J. K. & Wilson, I. D. (2003) Understanding 'Global' systems biology: Metabonomics and the continuum of metabolism. *Nature Reviews Drug Discovery,* 2**,** 668-676.

NICNAS (2004) National industrial chemicals notification and assessment scheme: Tinosan HP100. *Australia*.

Nikolaou, A., Meric, S. & Fatta, D. (2007) Occurrence patterns of pharmaceuticals in water and wastewater environments. *Analytical and Bioanalytical Chemistry,* 387**,** 1225-1234.

Nomura, D. K., Dix, M. M. & Cravatt, B. F. (2010) Activity-based protein profiling for biochemical pathway discovery in cancer. *Nature Reviews Cancer,* 10**,** 630-638.

Nordstrom, A., O'Maille, G., Qin, C. & Siuzdak, G. (2006) Nonlinear data alignment for UPLC-MS and HPLC-MS based metabolomics: Quantitative analysis of endogenous and exogenous metabolites in human serum. *Analytical Chemistry* 78.

Oikari, A. & Kunnamo-Ojala, T. (1987) Tracing of xenobiotic contamination in water with the aid of fish bile metabolites: A field study with caged rainbow trout (*Salmo gairdneri*). *Aquatic Toxicology,* 9**,** 327-341.

Oliver, S. G., Winson, M. K., Kell, D. B. & Baganz, F. (1998) Systematic functional analysis of the yeast genome. *Trends in Biotechnology,* 16**,** 373-378.

Orvos, D. R., Versteeg, D. J., Inauen, J., Capdevielle, M., Rothenstein, A. & Cunningham, V. (2002) Aquatic toxicity of triclosan. *Environmental Toxicology and Chemistry,* 21**,** 1338-1349.

P. Canosa, S. M., I. Rodr´ıguez, E. Rub´ı, R. Cela, M. G´omez (2005) *Anal. Bioanal. Chem.,* 383**,** 1119.

Pan, Z. & Raftery, D. (2007) Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Analytical and Bioanalytical Chemistry,* 387**,** 525-527.

Pang, K. S., Schwab, A. J., Goresky, C. A. & Chiba, M. (1994) Transport, binding, and metabolism of sulfate conjugates in the liver. *Chemico-Biological Interactions,* 92**,** 179-207.

Pasikanti, K. K., Ho, P. C. & Chan, E. C. Y. (2008) Gas chromatography/mass spectrometry in metabolic profiling of biological fluids. *Journal of Chromatography B,* 871**,** 202-211.

Pasquali, M., Monsen, G., Richardson, L., Alston, M. & Longo, N. (2006) Biochemical findings in common inborn errors of metabolism. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics,* 142C**,** 64-76.

Pauli-Magnus, C. & Meier, P. J. (2006) Hepatobiliary transporters and drug-induced cholestasis. *Hepatology,* 44**,** 778-787.

Pawitan, Y., Michiels, S., Koscielny, S., Gusnanto, A. & Ploner, A. (2005) False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics,* 21**,** 3017-3024.

Pedersen, R. T. & Hill, E. M. (2002) Tissue Distribution and Depuration of 4-tert-Octylphenol Residues in the Cyprinid Fish, Scardinius erythrophthalmus. *Environmental Science & Technology,* 36**,** 3275-3283.

Penttinen, O.-P. (1995) Chlorophenols in aquatic environments: Structure-activity correlations. *Annales Zoologici Fennici,* 32**,** 287-294.

Pérez, S. & Barceló, D. (2007) Fate and occurrence of X-ray contrast media in the environment. *Analytical and Bioanalytical Chemistry,* 387**,** 1235-1246.

Peterman, R. M. (1990) Statistical Power Analysis can Improve Fisheries Research and Management. *Canadian Journal of Fisheries and Aquatic Sciences,* 47**,** 2-15.

Pettersson, M., Adolfsson-Erici, M., Parkkonen, J., Forlin, L. & Asplund, L. (2006) Fish bile used to detect estrogenic substances in treated sewage water. *Science of the Total Environment,* 366**,** 174-186.

Pexa, A., Boeger, R., Henle, T., Schwedhelm, E. & Deussen, A. (2008) Effects of moderate hyperhomocysteinaemia induced by 4 weeks methionine-enriched diet on metabolite profile and mesenteric artery function in rats. *British Journal of Nutrition,* 99**,** 993-9.

Plomley, J. B., Crozier, P. W. & Taguchi, V. Y. (1999) Characterization of nonyl phenol ethoxylates in sewage treatment plants by combined precursor ion scanning and multiple reaction monitoring. *Journal of Chromatography A,* 854**,** 245-257.

Pluskal, T., Nakamura, T., Villar-Briones, A. & Yanagida, M. (2010) Metabolic profiling of the fission yeast *S. pombe*: Quantification of compounds under different temperatures and genetic perturbation. *Molecular BioSystems,* 6**,** 182-98.

Pritchard, J. B., Anderson, J. B., Rall, D. P. & Guarino, A. M. (1980) Comparative hepatic and renal handling of phenol red and indocyanine green by cyclostome, elasmobranch and teleost fish. *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology,* 65**,** 99-104.

Purdom, C. E., Hardiman, P. A., Bye, V. V. J., Eno, N. C., Tyler, C. R. & Sumpter, J. P. (1994) Estrogenic effects of effluents from sewage treatment works. *Chemistry and Ecology,* 8**,** 275-285.

Quinn, B., Gagné, F. & Blaise, C. (2009) Evaluation of the acute, chronic and teratogenic effects of a mixture of eleven pharmaceuticals on the cnidarian, *Hydra attenuata*. *Science of the Total Environment,* 407**,** 1072-1079.

Quinn, B. P., Booth, M. M., Delfino, J. J., Holm, S. E. & Gross, T. S. (2003) Selected resin acids in effluent and receiving waters derived from a bleached and unbleached kraft pulp and paper mill. *Environmental Toxicology and Chemistry,* 22**,** 214-218.

Quinones, M. P. & Kaddurah-Daouk, R. (2009) Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. *Neurobiology of Disease,* 35**,** 165-176.

Rajapakse, N., Silva, E. & Kortenkamp, A. (2002) Combining xenoestrogens at levels below individual no-observed-effect concentrations dramatically enhances steroid hormone action. *Environmental Health Perspectives,* 110**,** 917–921.

Ramadan, Z., Jacobs, D., Grigorov, M. & Kochhar, S. (2006) Metabolic profiling using principal component analysis, discriminant partial least squares, and genetic algorithms. *Talanta,* 68**,** 1683-1691.

Ramautar, R., Mayboroda, O. A., Somsen, G. W. & de Jong, G. J. (2011) CE-MS for metabolomics: Developments and applications in the period 2008–2010. *Electrophoresis,* 32**,** 52-65.

Reemtsma, T., Weiss, S., Mueller, J., Petrovic, M., González, S., Barcelo, D., Ventura, F. & Knepper, T. P. (2006) Polar pollutants entry into the water cycle by municipal wastewater: A european perspective. *Environmental Science & Technology,* 40**,** 5451-5458.

Reiss, R., Mackay, N., Habig, C. & Griffin, J. (2002) An ecological risk assessment for triclosan in lotic systems following discharge from wastewater treatment plants in the United States. *Environmental Toxicology and Chemistry,* 21**,** 2483-2492.

Richter, C. A., Birnbaum, L. S., Farabollini, F., Newbold, R. R., Rubin, B. S., Talsness, C. E., Vandenbergh, J. G., Walser-Kuntz, D. R. & vom Saal, F. S. (2007) In vivo effects of bisphenol A in laboratory rodent studies. *Reproductive Toxicology,* 24**,** 199-224.

Ridley, W. P., Sewall, M. R. & Dietrich, M. W. (1986) Metabolism and disposition of *O*-benzyl-*p*-chlorophenol in male rats. *Journal of Toxicology and Environmental Health,* 18**,** 267-283.

Ritter, L., Solomon, K., Sibley, P., Hall, K., Keen, P., Mattu, G. & Linton, B. (2002) Sources, pathways, and relative risks of contaminants in surface water and groundwater: A perspective prepared for the Walkerton inquiry. *Journal of toxicology and environmental health. Part A,* 65**,** 1-142.

Rivera, R. & Chun, J. (Eds.) (2008) *Biological effects of lysophospholipids 'Reviews of Physiology Biochemistry and Pharmacology',* Berlin Heidelberg, Springer.

Robert, J. S., John, P. H. & Eric, S. (1997) Statistical Power Analysis in Wildlife Research. *The Journal of Wildlife Management,* 61**,** 270-279.

Roberts, J. F., Marshall, S. J. & Roberts, D. W. (2007) Aquatic toxicity of ethoxylated and propoxylated alcohols to *Daphnia magna*. *Environmental Toxicology and Chemistry,* 26**,** 68-72.

Robertson, D. G. (2005) Metabonomics in toxicology: A review. *Toxicological Sciences,* 85**,** 809–822.

Robertson, D. G., Reily, M. D. & Baker, J. D. (2006) Metabonomics in pharmaceutical discovery and development. *Journal of Proteome Research,* 6**,** 526-539.

Rodgers-Gray, T. P., Jobling, S., Kelly, C., Morris, S., Brighty, G., Waldock, M. J., Sumpter, J. P. & Tyler, C. R. (2000) Exposure of juvenile roach (*Rutilus rutilus*) to treated sewage effluent induces dose-dependent and persistent disruption in gonadal duct development. *Environmental Science & Technology,* 35**,** 462-470.

Romesburg, H. C. (1981) Wildlife Science: Gaining Reliable Knowledge. *The Journal of Wildlife Management,* 45**,** pp. 293-313.

Rosal, R., Rodríguez, A., Perdigón-Melón, J. A., Petre, A., García-Calvo, E., Gómez, M. J., Agüera, A. & Fernández-Alba, A. R. (2010) Occurrence of emerging pollutants in urban wastewater and their removal through biological treatment followed by ozonation. *Water Research,* 44**,** 578-588.

Rostkowski, P., Horwood, J., Shears, J. A., Lange, A., Oladapo, F. O., Besselink, H. T., Tyler, C. R. & Hill, E. M. (2011) Bioassay- directed identification of novel antiandrogenic compounds in bile of fish exposed to wastewater effluents. *Environmental Science & Technology,* 45**,** 10660-10667.

Routledge, E. J., Sheahan, D., Desbrow, C., Brighty, G. C., Waldock, M. & Sumpter, J. P. (1998) Identification of estrogenic chemicals in STW effluent. 2. *In vivo* responses in trout and roach. *Environmental Science & Technology,* 32**,** 1559-1565.

Routledge, E. J. & Sumpter, J. P. (1996) Estrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen. *Environmental Toxicology and Chemistry,* 15**,** 241-248.

Roux, A., Lison, D., Junot, C. & Heilier, J.-F. (2011) Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review. *Clinical Biochemistry,* 44**,** 119-135.

Roy, D. & Liehr, J. G. (1999) Estrogen, DNA damage and mutations. *Mutation Research,* 424**,** 107-15.

Ryan, D. & Robards, K. (2006) Metabolomics: The greatest omics of them all? *Analytical Chemistry,* 78**,** 7954-7958.

Samuelsson, L. M., Förlin, L., Karlsson, G., Adolfsson-Erici, M. & Larsson, D. G. J. (2006) Using NMR metabolomics to identify responses of an environmental estrogen in blood plasma of fish. *Aquatic Toxicology,* 78**,** 341-349.

Sana, T. R., Waddell, K. & Fischer, S. M. (2008) A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome. *Journal of Chromatography B,* 871**,** 314-321.

Santos, E. M., Paull, G. C., Van Look, K. J. W., Workman, V. L., Holt, W. V., van Aerle, R., Kille, P. & Tyler, C. R. (2007) Gonadal transcriptome responses and physiological consequences of exposure to oestrogen in breeding zebrafish (*Danio rerio*). *Aquatic Toxicology,* 83**,** 134-142.

Sarasquete, C. & Segner, H. (2000) Cytochrome P4501A (CYP1A) in teleostean fishes. A review of immunohistochemical studies. *Science of the Total Environment,* 247**,** 313-332.

Scherpenisse, P. & Bergwerff, A. A. (2007) Determination of residues of tricaine in fish using liquid chromatography tandem mass spectrometry. *Analytica Chimica Acta,* 586**,** 407-410.

Schummer, C., Delhomme, O., Appenzeller, B. M. R., Wennig, R. & Millet, M. (2009) Comparison of MTBSTFA and BSTFA in derivatization reactions of polar compounds prior to GC/MS analysis. *Talanta,* 77**,** 1473-1482.

Scott, K. D., Fåhræus-Van Ree, G. E. & Parrish, C. C. (2002) Sex differences in hepatic lipids of toxaphene-exposed juvenile yellowtail flounder (*Pleuronectes ferrugineus* Storer). *Ecotoxicology and Environmental Safety,* 51**,** 168-176.

Seki, M., Yokota, H., Maeda, M., Tadokoro, H. & Kobayashi, K. (2003) Effects of 4-nonylphenol and 4-tert-octylphenol on sex differentiation and vitellogenin induction in medaka (*Oryzias latipes*). *Environmental Toxicology and Chemistry,* 22**,** 1507-1516.

Sheahan, D. A., Brighty, G. C., Daniel, M., Jobling, S., Harries, J. E., Hurst, M. R., Kennedy, J., Kirby, S. J., Morris, S., Routledge, E. J., Sumpter, J. P. & Waldock, M. J. (2002a) Reduction in the estrogenic activity of a treated sewage effluent discharge to an english river as a result of a decrease in the concentration of industrially derived surfactants. *Environmental Toxicology and Chemistry,* 21**,** 515-519.

Sheahan, D. A., Brighty, G. C., Daniel, M., Kirby, S. J., Hurst, M. R., Kennedy, J., Morris, S., Routledge, E. J., Sumpter, J. P. & Waldock, M. J. (2002b) Estrogenic activity measured in a sewage treatment works treating industrial inputs

containing high concentrations of alkylphenolic compounds—A case study. *Environmental Toxicology and Chemistry,* 21**,** 507-514.

Shulaev, V. (2006) Metabolomics technology and bioinformatics. *Briefings in Bioinformatics,* 7**,** 128-139.

Silva, E., Rajapakse, N. & Kortenkamp, A. (2002) Something from "Nothing" – eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. *Environmental Science & Technology,* 36**,** 1751-1756.

Singh, Y., Cooke, J. B., Hinton, D. E. & Miller, M. G. (1996) Trout liver slices for metabolism and toxicity studies. *Drug Metabolism and Disposition,* 24**,** 7-14.

Siuzdak, G. & Lewis, J. K. (1998) Applications of mass spectrometry in combinatorial chemistry. *Biotechnology and Bioengineering (Combinatorial Chemistry),* 61**,** 127-134.

Skillman, A. D., Nagler, J. J., Hook, S. E., Small, J. A. & Schultz, I. R. (2006) Dynamics of 17α-Ethynylestradiol exposure in rainbow trout (Oncorhynchus mykiss): Absorption, tissue distribution, and hepatic gene expression pattern. *Environmental Toxicology and Chemistry,* 25**,** 2997-3005.

Smedsgaard, J. & Nielsen, J. (2005) Metabolite profiling of fungi and yeast: From phenotype to metabolome by MS and informatics. *Journal of Experimental Botany,* 56**,** 273-286.

Smith, M. D. & Hill, E. M. (2006) Profiles of short chain oligomers in roach (*Rutilus rutilus*) exposed to waterborne polyethoxylated nonylphenols. *Science of the Total Environment,* 356**,** 100-111.

Soares, A., Guieysse, B., Jefferson, B., Cartmell, E. & Lester, J. N. (2008) Nonylphenol in the environment: A critical review on occurrence, fate, toxicity and treatment in wastewaters. *Environment International,* 34**,** 1033-1049.

Sohoni, P. & Sumpter, J. (1998) Several environmental oestrogens are also anti-androgens. *Journal of Endocrinology,* 158**,** 327-339.

Southam, A. D., Lange, A., Hines, A., Hill, E. M., Katsu, Y., Iguchi, T., Tyler, C. R. & Viant, M. R. (2011) Metabolomics reveals target and off-target toxicities of a model organophosphate pesticide to roach (*Rutilus rutilus*): Implications for biomonitoring. *Environmental Science & Technology,* 45**,** 3759-3767.

Stalling, D. L. & Hogan, J. W. (1978) Preparation, separation and identification of TMS derivatives of hydroxylated PCBS and chlorophenols. *Bulletin of Environmental Contamination and Toxicology,* 20**,** 35-43.

Standley, L. J., Rudel, R. A., Swartz, C. H., Attfield, K. R., Christian, J., Erickson, M. & Brody, J. G. (2008) Wastewater-contaminated groundwater as a source of endogenous hormones and pharmaceuticals to surface water ecosystems. *Environmental Toxicology and Chemistry,* 27**,** 2457-2468.

Statistics: Metabolomics Fiehn Lab. http://fiehnlab.ucdavis.edu/staff/kind/Statistics/Concepts/OPLS-PLSDA/.

Stevens, J. L., Northcott, G. L., Stern, G. A., Tomy, G. T. & Jones, K. C. (2003) PAHs, PCBs, PCNs, organochlorine pesticides, synthetic musks, and polychlorinated *n*-alkanes in U.K. sewage sludge: Survey results and implications. *Environmental Science & Technology,* 37**,** 462-467.

Stuart, M., Lapworth, D., Crane, E. & Hart, A. (2012) Review of risk from potential emerging contaminants in UK groundwater. *Science of the Total Environment,* 416**,** 1-21.

Stump, M. J., Fleming, R. C., Gong, W.-H., Jaber, A. J., Jones, J. J., Surber, C. W. & Wilkins, C. L. (2002) Matrix-assisted laser desorption mass spectrometry. *Applied Spectroscopy Reviews,* 37**,** 275-303.

Sumner, L., Amberg, A., Barrett, D., Beale, M., Beger, R., Daykin, C., Fan, T., Fiehn, O., Goodacre, R., Griffin, J., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A., Lindon, J., Marriott, P., Nicholls, A., Reily, M., Thaden, J. & Viant, M. (2007) Proposed minimum reporting standards for chemical analysis. *Metabolomics,* 3**,** 211-221.

Sumner, L. W., Mendes, P. & Dixon, R. A. (2003) Plant metabolomics: Large-scale phytochemistry in the functional genomics era. *Phytochemistry,* 62**,** 817-836.

Sumpter, J. P. (2003) Endocrine disruption in wildlife: The future? *Pure and Applied Chemistry,* 35**,** 2355–2360.

Swisher, R. D. (Ed.) (1987) *Surfactants Biodegradation: Omega and Beta Oxidation of LAS,* New Yourk, Marcell Dekker, Inc.

Tanaka, Y., Higashi, T., Rakwal, R., Wakida, S.-i. & Iwahashi, H. (2007) Quantitative analysis of sulfur-related metabolites during cadmium stress response in yeast by capillary electrophoresis–mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis,* 44**,** 608-613.

Taylor, N., Weber, R., Southam, A., Payne, T., Hrydziuszko, O., Arvanitis, T. & Viant, M. (2009) A new approach to toxicity testing in *Daphnia magna*: Application of high throughput FT-ICR mass spectrometry metabolomics. *Metabolomics,* 5**,** 44-58.

Teijon, G., Candela, L., Tamoh, K., Molina-Díaz, A. & Fernández-Alba, A. R. (2010) Occurrence of emerging contaminants, priority substances (2008/105/CE) and heavy metals in treated wastewater and groundwater at Depurbaix facility (Barcelona, Spain). *Science of the Total Environment,* 408**,** 3584-3595.

Terasaki, M. & Makino, M. (2008) Determination of chlorinated by-products of parabens in swimming pool water. *International Journal of Environmental Analytical Chemistry,* 88**,** 911-922.

Terasaki, M., Makino, M. & Tatarazako, N. (2009) Acute toxicity of parabens and their chlorinated by-products with *Daphnia magna* and *Vibrio fischeri* bioassays. *Journal of Applied Toxicology,* 29**,** 242-247.

Ternes, T. A. & Hirsch, R. (2000) Occurrence and behavior of X-ray contrast media in sewage facilities and the aquatic environment. *Environmental Science & Technology,* 34**,** 2741-2748.

Ternes, T. A., Joss, A. & Siegrist, H. (2004) Peer reviewed: Scrutinizing pharmaceuticals and personal care products in wastewater treatment. *Environmental Science & Technology,* 38**,** 392A-399A.

Terzic, S., Senta, I., Ahel, M., Gros, M., Petrovic, M., Barcelo, D., Müller, J., Knepper, T., Martí, I., Ventura, F., Jovancic, P. & Jabucar, D. (2008) Occurrence and fate of emerging wastewater contaminants in Western Balkan Region. *Science of the Total Environment,* 399**,** 66-77.

Theodoridis, G., Gika, H. G. & Wilson, I. D. (2008) LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. *TrAC Trends in Analytical Chemistry,* 27**,** 251-260.

Thomas, L. & Juanes, F. (1996) The importance of statistical power analysis: an example from Animal Behaviour. *Animal Behaviour,* 52**,** 856-859.

Thomas, R. E. & Kotchevar, A. T. (2010) Comparative in vitro metabolism of chloroxylenol, chlorophene, and triclosan with rat, mouse, and human hepatic microsomes. *Toxicological and Environmental Chemistry,* 92**,** 1735-1747.

Thorpe, K. L., Cummings, R. I., Hutchinson, T. H., Scholze, M., Brighty, G., Sumpter, J. P. & Tyler, C. R. (2003) Relative potencies and combination effects of steroidal estrogens in fish. *Environmental Science & Technology,* 37**,** 1142-1149.

Thorpe, K. L., Hutchinson, T. H., Hetheridge, M. J., Scholze, M., Sumpter, J. P. & Tyler, C. R. (2001) Assessing the biological potency of binary mixtures of environmental estrogens using vitellogenin induction in juvenile rainbow trout (*Oncorhynchus mykiss*). *Environmental Science & Technology,* 35**,** 2476-2481.

Tohge, T. & Fernie, A. R. (2010) Combining genetic diversity, informatics and metabolomics to facilitate annotation of plant gene function. *Nature Protocols,* 5**,** 1210-1227.

Tolls, J., Haller, M., Seinen, W. & Sijm, D. T. H. M. (2000) LAS bioconcentration: Tissue distribution and effect of hardness−implications for processes. *Environmental Science & Technology,* 34**,** 304-310.

Tong, Z. & James, M. O. (2000) Purification and characterization of hepatic and intestinal phenol sulfotransferase with high affinity for benzo[a]pyrene phenols from channel catfish, *Ictalurus punctatus*. *Archives of Biochemistry and Biophysics,* 376**,** 409-419.

Torkhovskaya, T., Ipatova, O., Zakharova, T., Kochetova, M. & Khalilov, E. (2007) Lysophospholipid receptors in cell signaling. *Biochemistry (Moscow),* 72**,** 125-131.

Trygg, J., Holmes, E. & Lundstedt, T. (2007) Chemometrics in metabonomics. *Journal of Proteome Research,* 6**,** 469-479.

Trygg, J. & Wold, S. (2002) Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics,* 16**,** 119-128.

Tulp, M. T. M., Sundström, G., Martron, L. B. J. M. & Hutzinger, O. (1979) Metabolism of chlorodiphenyl ethers and irgasan DP 300. *Xenobiotica,* 9**,** 65-77.

Tyler, C., Lange, A., Paull, G., Katsu, Y. & Iguchi, T. (2007) The roach (*Rutilus rutilus*) as a sentinel for assessing endocrine disruption. *Environmental Sciences,* 14**,** 235-253.

Tyler, C. R., Spary, C., Gibson, R., Santos, E. M., Shears, J. & Hill, E. M. (2005) Accounting for differences in estrogenic responses in rainbow trout (*Oncorhynchus mykiss*: Salmonidae) and roach (*Rutilus rutilus*: Cyprinidae) exposed to effluents from wastewater treatment works. *Environmental Science & Technology,* 39**,** 2599-2607.

Urakami, K., Zangiacomi, V., Yamaguchi, K. & Kusuhara, M. (2010) Quantitative metabolome profiling of Illicium anisatum by capillary electrophoresis time-of-flight mass spectrometry. *Biomedical Research (Tokyo, Japan),* 31**,** 161-163.

Van der Oost, R., Beyer, J. & Vermeulen, N. P. E. (2003) Fish bioaccumulation and biomarkers in environmental risk assessment: A review. *Environmental Toxicology and Pharmacology,* 13**,** 57-149.

Van Oostdam, J., Donaldson, S. G., Feeley, M., Arnold, D., Ayotte, P., Bondy, G., Chan, L., Dewaily, É., Furgal, C. M., Kuhnlein, H., Loring, E., Muckle, G., Myles, E., Receveur, O., Tracy, B., Gill, U. & Kalhok, S. (2005) Human health implications of environmental contaminants in Arctic Canada: A review. *Science of the Total Environment,* 351–352**,** 165-246.

Vasskog, T., Anderssen, T., Pedersen-Bjergaard, S., Kallenborn, R. & Jensen, E. (2008) Occurrence of selective serotonin reuptake inhibitors in sewage and receiving waters at Spitsbergen and in Norway. *Journal of Chromatography A,* 1185**,** 194-205.

Veldhoen, N., Skirrow, R. C., Osachoff, H., Wigmore, H., Clapson, D. J., Gunderson, M. P., Van Aggelen, G. & Helbing, C. C. (2006) The bactericidal agent triclosan modulates thyroid hormone-associated gene expression and disrupts postembryonic anuran development. *Aquatic Toxicology,* 80**,** 217-227.

Verge, C., Moreno, A., Bravo, J. & Berna, J. L. (2001) Influence of water hardness on the bioavailability and toxicity of linear alkylbenzene sulphonate (LAS). *Chemosphere,* 44**,** 1749-1757.

Verliefde, A., Cornelissen, E., Amy, G., Van der Bruggen, B. & van Dijk, H. (2007) Priority organic micropollutants in water sources in Flanders and the Netherlands and assessment of removal possibilities with nanofiltration. *Environmental Pollution,* 146**,** 281-289.

Viant, M. R. (2007) Metabolomics of aquatic organisms: The new 'omics' on the block. *Marine Ecology Progress Series,* 332**,** 301-306.

Viant, M. R., Bearden, D. W., Bundy, J. G., Burton, I. W., Collette, T. W., Ekman, D. R., Ezernieks, V., Karakach, T. K., Lin, C. Y., Rochfort, S., Ropp, J. S. d., Teng, Q., Tjeerdema, R. S., Walter, J. A. & Wu, H. (2008) International NMR-based environmental metabolomics intercomparison exercise. *Environmental Science & Technology,* 43**,** 219-225.

Viant, M. R., Bundy, J. G., Pincetich, C. A., de Ropp, J. S. & Tjeerdema, R. S. (2005) NMR-derived developmental metabolic trajectories: An approach for visualizing the toxic actions of trichloroethylene during embryogenesis. *Metabolomics,* 1**,** 149-158.

Viant, M. R., Pincetich, C. A., Hinton, D. E. & Tjeerdema, R. S. (2006) Toxic actions of dinoseb in medaka (*Oryzias latipes*) embryos as determined by *in vivo* ^{31}P NMR, HPLC-UV and ^{1}H NMR metabolomics. *Aquatic Toxicology,* 76**,** 329-342.

Viant, M. R., Rosenblum, E. S. & Tjeerdema, R. S. (2003) NMR-based metabolomics: A powerful approach for characterizing the effects of environmental stressors on organism health. *Environmental Science & Technology,* 37**,** 4982-4989.

Vieno, N., Tuhkanen, T. & Kronberg, L. (2007) Elimination of pharmaceuticals in sewage treatment plants in Finland. *Water Research,* 41**,** 1001-1012.

Villas-Boas, S. G., Roessner, U., Hansen, M. A. E., Smedsgaard, J. & Nielsen, J. (Eds.) (2007) *Metabolome Analysis: An Introduction,* Canada, John Wiley & Sons.

Vineis, P., Khan, A., Vlaanderen, J. & Vermeulen, R. (2009) The impact of new research technologies on our understanding of environmental causes of disease: the concept of clinical vulnerability. *Environmental Health,* 8**,** 54.

Viravaidya, K., Sin, A. & Shuler, M. L. (2004) Development of a microscale cell culture analog to probe naphthalene toxicity. *Biotechnology Progress,* 20**,** 316-323.

Vom Saal, F. S. & Hughes, C. (2005) An extensive new literature concerning low-dose effects of bisphenol A shows the need for a new risk assessment. *Environmental Health Perspectives,* 113.

Vrana, B., Allan, I. J., Greenwood, R., Mills, G. A., Dominiak, E., Svensson, K., Knutsson, J. & Morrison, G. (2005) Passive sampling techniques for monitoring pollutants in water. *TrAC Trends in Analytical Chemistry,* 24**,** 845-868.

Vulliet, E. & Cren-Olivé, C. (2011) Screening of pharmaceuticals and hormones at the regional scale, in surface and groundwaters intended to human consumption. *Environmental Pollution,* 159**,** 2929-2934.

Wachtmeister, C. A., Förlin, L., Arnoldsson, K. C. & Larsson, J. (1991) Fish bile as a tool for monitoring aquatic pollutants: Studies with radioactively labelled 4,5,6-trichloroguaiacol. *Chemosphere,* 22**,** 39-46.

Wagner, S., Scholz, K., Sieber, M., Kellert, M. & Voelkel, W. (2007) Tools in metabonomics: An integrated validation approach for LC-MS metabolic profiling of mercapturic acids in human urine. *Analytical Chemistry,* 79**,** 2918-2926.

Wang, H., Tso, V. K., Slupsky, C. M. & Fedorak, R. N. (2010) Metabolomics and detection of colorectal cancer in humans: A systematic review. *Future Oncology,* 6**,** 1395-1406.

Wang, J., Reijmers, T., Chen, L., Van Der Heijden, R., Wang, M., Peng, S., Hankemeier, T., Xu, G. & Van Der Greef, J. (2009) Systems toxicology study of doxorubicin on rats using ultra performance liquid chromatography coupled with mass spectrometry based metabolomics. *Metabolomics,* 5**,** 407-418.

Wang, L., Falany, C. N. & James, M. O. (2004) Triclosan as a substrate and inhibitor of 3'-phosphoadenosine 5'-phosphosulfate-sulfotransferase and UDP-glucuronosyl transferase in human liver fractions. *Drug Metabolism Disposition,* 32**,** 1162-1169.

Want, E. J., Nordström, A., Morita, H. & Siuzdak, G. (2006) From exogenous to endogenous: The inevitable imprint of mass spectrometry in metabolomics. *Journal of Proteome Research,* 6**,** 459-468.

Water UK (2006) Wastewater treatment and recycling.

Watkinson, A. J., Murby, E. J., Kolpin, D. W. & Costanzo, S. D. (2009) The occurrence of antibiotics in an urban watershed: From wastewater to drinking water. *Science of the Total Environment,* 407**,** 2711-2723.

Welshons, W., Thayer, K., Judy, B., Taylor, J., Curran, E. & vom Saal, F. (2003) Large effects from small exposures. I. Mechanisms for endocrine-disrupting chemicals with estrogenic activity. *Environmental Health Perspectives,* 111**,** 994-1006.

Wen, Z., Tallman, M. N., Ali, S. Y. & Smith, P. C. (2007) UDP-glucuronosyltransferase 1A1 is the principal enzyme responsible for etoposide glucuronidation in human liver and intestinal microsomes: Structural characterization of phenolic and alcoholic glucuronides of etoposide and estimation of enzyme kinetics. *Drug Metabolism and Disposition,* 35**,** 371-380.

Werner, A. F., Taulli, T. A., Michael, P. R. & Williams, M. A. (1983) Estimation and verification of the environmental fate of *O*-benzyl-*p*-chlorophenol. *Archives of Environmental Contamination and Toxicology,* 12**,** 569-575.

Werner, J. J., McNeill, K. & Arnold, W. A. (2005) Environmental photodegradation of mefenamic acid. *Chemosphere,* 58**,** 1339-1346.

Wheeler, T. F., Heim, J. R., LaTorre, M. R. & Janes, A. B. (1997) Mass spectral characterization of *p*-nonylphenol isomers using high-resolution capillary GC-MS. *Journal of Chromatographic Science,* 35**,** 19-30.

Wick, A., Fink, G., Joss, A., Siegrist, H. & Ternes, T. A. (2009) Fate of beta blockers and psycho-active drugs in conventional wastewater treatment. *Water Research,* 43**,** 1060-1074.

Wiklund, S., Johansson, E., Sjstrm, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., Gottfries, J., Moritz, T. & Trygg, J. (2008) Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry,* 80**,** 115-122.

Wild, C. P. (2012) The exposome: From concept to utility. *International Journal of Epidemiology***,** 1-9.

William Allwood, J., Ellis, D. I., Heald, J. K., Goodacre, R. & Mur, L. A. J. (2006) Metabolomic approaches reveal that phosphatidic and phosphatidyl glycerol phospholipids are major discriminatory non-polar metabolites in responses by *Brachypodium distachyon* to challenge by *Magnaporthe grisea*. *The Plant Journal,* 46**,** 351-368.

Wilson, I. D., Nicholson, J. K., Castro-Perez, J., Granger, J. H., Johnson, K. A., Smith, B. W. & Plumb, R. S. (2005) High resolution "Ultra Performance" liquid chromatography coupled to oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. *Journal of Proteome Research,* 4**,** 591-598.

Winning, H., Larsen, F. H., Bro, R. & Engelsen, S. B. (2008) Quantitative analysis of NMR spectra with chemometrics. *Journal of Magnetic Resonance,* 190**,** 26-32.

Wishart, D. S. (2008) Metabolomics: Applications to food science and nutrition research. *Trends in Food Science & Technology,* 19**,** 482-493.

Wold, S., Sjöström, M. & Eriksson, L. (2001) PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems,* 58**,** 109-130.

Wolf, C. J., Takacs, M. L., Schmid, J. E., Lau, C. & Abbott, B. D. (2008) Activation of mouse and human peroxisome proliferator–activated receptor alpha by perfluoroalkyl acids of different functional groups and chain lengths. *Toxicological Sciences,* 106**,** 162-171.

Woo, H. M., Kim, K. M., Choi, M. H., Jung, B. H., Lee, J., Kong, G., Nam, S. J., Kim, S., Bai, S. W. & Chung, B. C. (2009) Mass spectrometry based metabolomic approaches in urinary biomarker study of women's cancers. *Clinica Chimica Acta,* 400**,** 63-69.

Workentine, M. L., Harrison, J. J., Weljie, A. M., Tran, V. A., Stenroos, P. U., Tremaroli, V., Vogel, H. J., Ceri, H. & Turner, R. J. (2010) Phenotypic and metabolic profiling of colony morphology variants evolved from *Pseudomonas fluorescens* biofilms. *Environmental Microbiology,* 12**,** 1565-1577.

Wu, H., Xue, R., Tang, Z., Deng, C., Liu, T., Zeng, H., Sun, Y. & Shen, X. (2010) Metabolomic investigation of gastric cancer tissue using gas chromatography/mass spectrometry. *Analytical Bioanalytical Chemistry,* 396**,** 1385-95.

Wu, Z., Gao, W., Phelps, M. A., Wu, D., Miller, D. D. & Dalton, J. T. (2004) Favorable effects of weak acids on negative-ion electrospray ionization mass spectrometry. *Analytical Chemistry,* 76**,** 839-847.

Xiang, X., Han, Y., Neuvonen, M., Laitila, J., Neuvonen, P. J. & Niemi, M. (2010) High performance liquid chromatography–tandem mass spectrometry for the determination of bile acid concentrations in human plasma. *Journal of Chromatography B,* 878**,** 51-60.

Xiao, J. F., Zhou, B. & Ressom, H. W. (2012) Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC Trends in Analytical Chemistry,* 32**,** 1-14.

Xie, B., Gong, T., Gao, R., Liu, J., Zuo, J., Wang, X. & Zhang, Z. (2009) Development of rat urinary HPLC-UV profiling for metabonomic study on Liuwei Dihuang Pills. *Journal of Pharmaceutical and Biomedical Analysis,* 49**,** 492-497.

Xuan, B. H. N., Thi, T. X. N., Nguyen, S. T., Goldfarb, D. S., Stokes, M. B. & Rabenou, R. A. (2003) Ichthyotoxic ARF after fish gallbladder ingestion: A large case series from Vietnam. *American Journal of Kidney Diseases : The Official Journal of the National Kidney Foundation,* 41**,** 220-224.

Yadav, J., Lawrence, D., Nuck, B., Federle, T. & Reddy, C. A. (2001) Biotransformation of linear alkylbenzene sulfonate (LAS) by *Phanerochaete chrysosporium*: oxidation of alkyl side-chain. *Biodegradation,* 12**,** 443-453.

Yeh, Y.-H. & Hwang, D.-F. (2001) High-performance liquid chromatographic determination for bile components in fish, chicken and duck. *Journal of Chromatography B: Biomedical Sciences and Applications,* 751**,** 1-8.

Yeh, Y. H., Wang, D. Y., Deng, J. F., Chen, S. K. & Hwang, D. F. (2002) Short-term toxicity of grass carp bile powder, 5alpha-cyprinol and 5alpha-cyprinol sulfate in rats. *Comparative Biochemistry and Physiology. Toxicology & Pharmacology,* 131**,** 1-8.

Ying, G.-G. (2006) Fate, behavior and effects of surfactants and their degradation products in the environment. *Environment International,* 32**,** 417-431.

Ying, G.-G., Williams, B. & Kookana, R. (2002) Environmental fate of alkylphenols and alkylphenol ethoxylates—A review. *Environment International,* 28**,** 215-226.

Ying, G. G. & Kookana, R. S. (2002) Endocrine disruption: An australian perspective. *Australian Water Association Journal Water,* 29**,** 42-45.

Yu, K., Sheng, G., Sheng, J., Chen, Y., Xu, W., Liu, X., Cao, H., Qu, H., Cheng, Y. & Li, L. (2007) A Metabonomic investigation on the biochemical perturbation in liver failure patients caused by hepatitis B virus. *Journal of Proteome Research,* 6**,** 2413-2419.

Zaharie, M. (2006) Occurrences of pharmaceutical and personal care products as micropollutants in rivers from Romania. *Chemosphere,* 64**,** 1808-1817.

Zang, S., Lian, B., Wang, J. & Yang, Y. (2010) Biodegradation of 2-naphthol and its metabolites by coupling *Aspergillus niger* with *Bacillus subtilis*. *Journal of Environmental Sciences,* 22**,** 669-674.

Zhang, Y., Geißen, S.-U. & Gal, C. (2008) Carbamazepine and diclofenac: Removal in wastewater treatment plants and occurrence in water bodies. *Chemosphere,* 73**,** 1151-1161.

Zollner, G., Marschall, H.-U., Wagner, M. & Trauner, M. (2006) Role of nuclear receptors in the adaptive response to bile acids and cholestasis: Pathogenetic and therapeutic considerations. *Molecular Pharmaceutics,* 3**,** 231-251.

Zollner, G. & Trauner, M. (2008) Mechanisms of cholestasis. *Clinics Liver Disease,* 12**,** 1-26.

Zorita, S., Mårtensson, L. & Mathiasson, L. (2009) Occurrence and removal of pharmaceuticals in a municipal sewage treatment system in the south of Sweden. *Science of the Total Environment,* 407**,** 2760-2770.

Zuccato, E., Castiglioni, S., Bagnati, R., Chiabrando, C., Grassi, P. & Fanelli, R. (2008) Illicit drugs, a novel group of environmental contaminants. *Water Research,* 42**,** 961-968.

APPENDICES

Chapter three appendices

Appendix 3.1: Performance parameters of principal component analyses (PCA) for the comparison of trout exposed to river water and effluent (bile extracts).

Ionization mode	Groups	Multivariate method	Principal Components	R^2X	Q^2
+ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PCA	5	0.418	0.277
-ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PCA	6	0.426	0.258
+ESI	C_{10},C_{14},E_{10},E_{14}	PCA	4	0.415	0.246
-ESI	C_{10},C_{14},E_{10},E_{14}	PCA	2	0.305	0.227
+ESI	C_{10},C_{21},E_{10},E_{21}	PCA	3	0.387	0.277
-ESI	C_{10},C_{21},E_{10},E_{21}	PCA	4	0.401	0.259
+ESI	C_{14},C_{21},E_{14},E_{21}	PCA	2	0.256	0.097
-ESI	C_{14},C_{21},E_{14},E_{21}	PCA	2 (1)	0.251	0.090
+ESI	C_{10},E_{10}	PCA	2	0.36	0.247
-ESI	C_{10},E_{10}	PCA	2 (1)	0.334	0.218
+ESI	C_{14},E_{14}	PCA	2 (2)	0.368	-0.035
-ESI	C_{14},E_{14}	PCA	2 (2)	0.358	-0.050
+ESI	C_{21},E_{21}	PCA	2 (2)	0.287	0.012
-ESI	C_{21},E_{21}	PCA	2 (2)	0.293	0.002

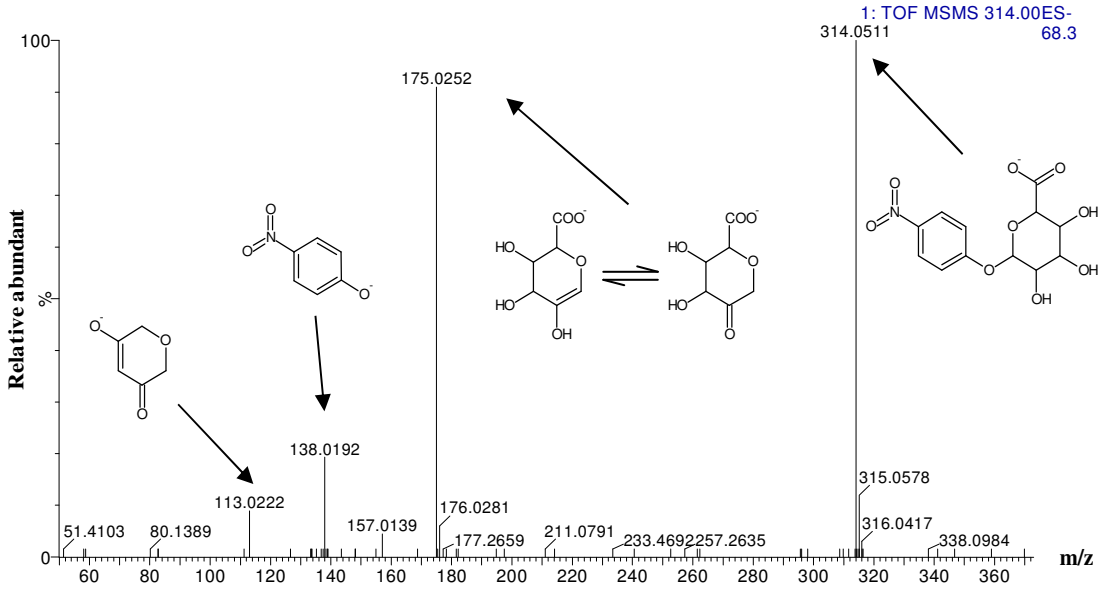
+/-ESI: positive/negative electrospray ionization; C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control); E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; R^2X: variation explained by the models; Q^2: cumulative variation predicted by the models. Discriminating components that are not significant are shown in brackets.

Appendix 3.2: Performance parameters of multivariate discriminant models for the comparison of trout exposed to river water and effluent (bile extracts).

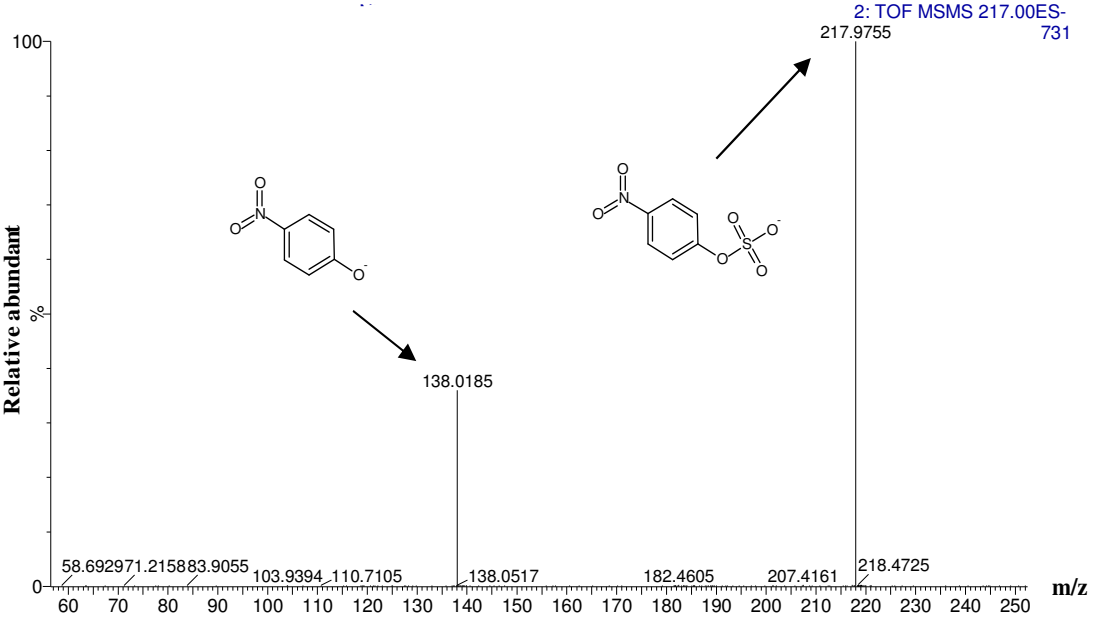
Ionization mode	Groups	Multivariate method	Number of projections	R^2X	R^2Y	Q^2	External Validation
+ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PLS-DA	9	0.471	0.978	0.776	100%
-ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PLS-DA	9	0.455	0.978	0.777	89%
+ESI	C_{10},C_{14},E_{10},E_{14}	PLS-DA	5	0.414	0.984	0.865	90%
-ESI	C_{10},C_{14},E_{10},E_{14}	PLS-DA	5	0.392	0.987	0.866	95%
+ESI	C_{10},C_{21},E_{10},E_{21}	PLS-DA	6	0.443	0.990	0.868	100%
-ESI	C_{10},C_{21},E_{10},E_{21}	PLS-DA	6	0.431	0.990	0.874	100%
+ESI	C_{14},C_{21},E_{14},E_{21}	PLS-DA	6	0.421	0.990	0.792	100%
-ESI	C_{14},C_{21},E_{14},E_{21}	PLS-DA	6	0.416	0.987	0.781	100%
+ESI	C_{10},E_{10}	PLS-DA	2 (1)	0.324	0.999	0.987	100%
-ESI	C_{10},E_{10}	PLS-DA	2 (1)	0.306	0.999	0.985	100%
+ESI	C_{14},E_{14}	PLS-DA	2	0.358	0.990	0.800	100%
-ESI	C_{14},E_{14}	PLS-DA	2	0.352	0.995	0.826	100%
+ESI	C_{21},E_{21}	PLS-DA	3	0.322	0.999	0.820	100%
-ESI	C_{21},E_{21}	PLS-DA	3	0.331	0.999	0.809	100%
+ESI	C_{10},E_{10}	OPLS-DA	1+1 (1)	0.324	0.999	0.986	-
-ESI	C_{10},E_{10}	OPLS-DA	1+1 (1)	0.306	0.999	0.985	-
+ESI	C_{14},E_{14}	OPLS-DA	1+1 (1)	0.358	0.990	0.630	-
-ESI	C_{14},E_{14}	OPLS-DA	1+1 (1)	0.352	0.995	0.651	-
+ESI	C_{21},E_{21}	OPLS-DA	1+1 (1)	0.244	0.989	0.473	-
-ESI	C_{21},E_{21}	OPLS-DA	1+1 (1)	0.257	0.968	0.464	-

+/-ESI: positive/negative electrospray ionization; C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control); E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; R^2Y and Q^2 represent the cumulative variation explained and predicted by the models, respectively. Discriminating components that are not significant are shown in brackets.

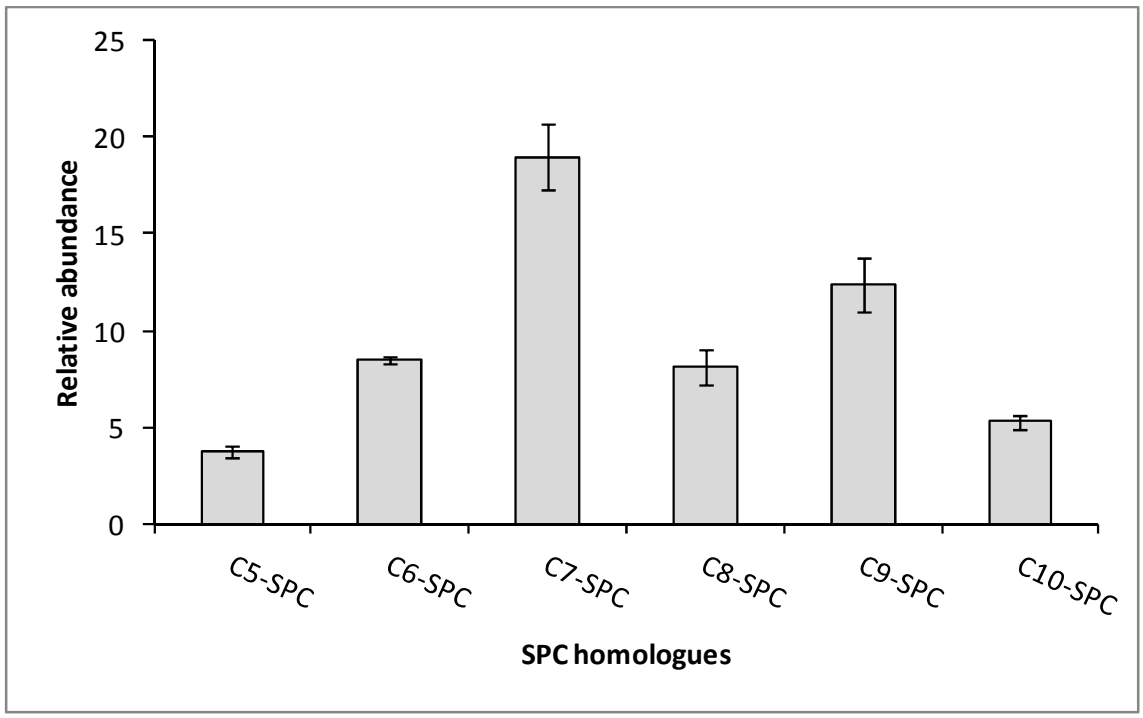
a)



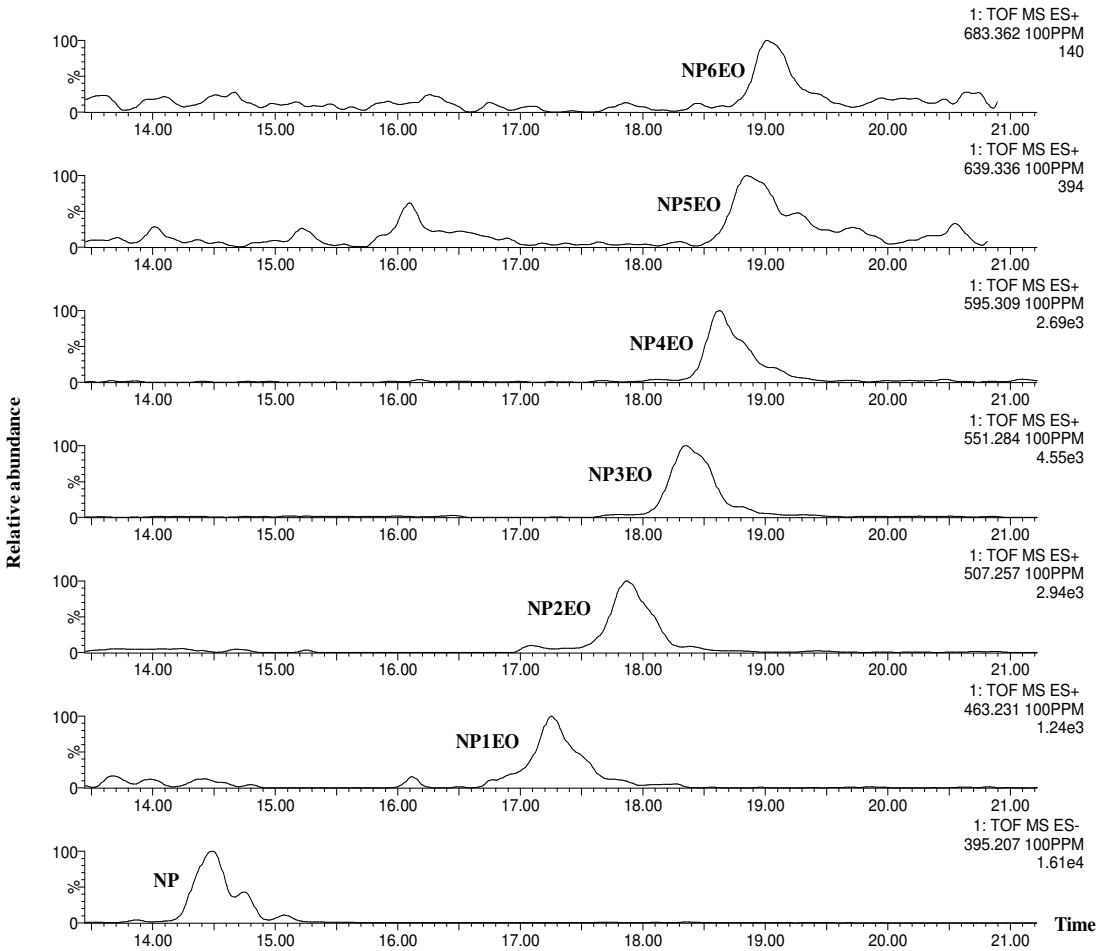
b)



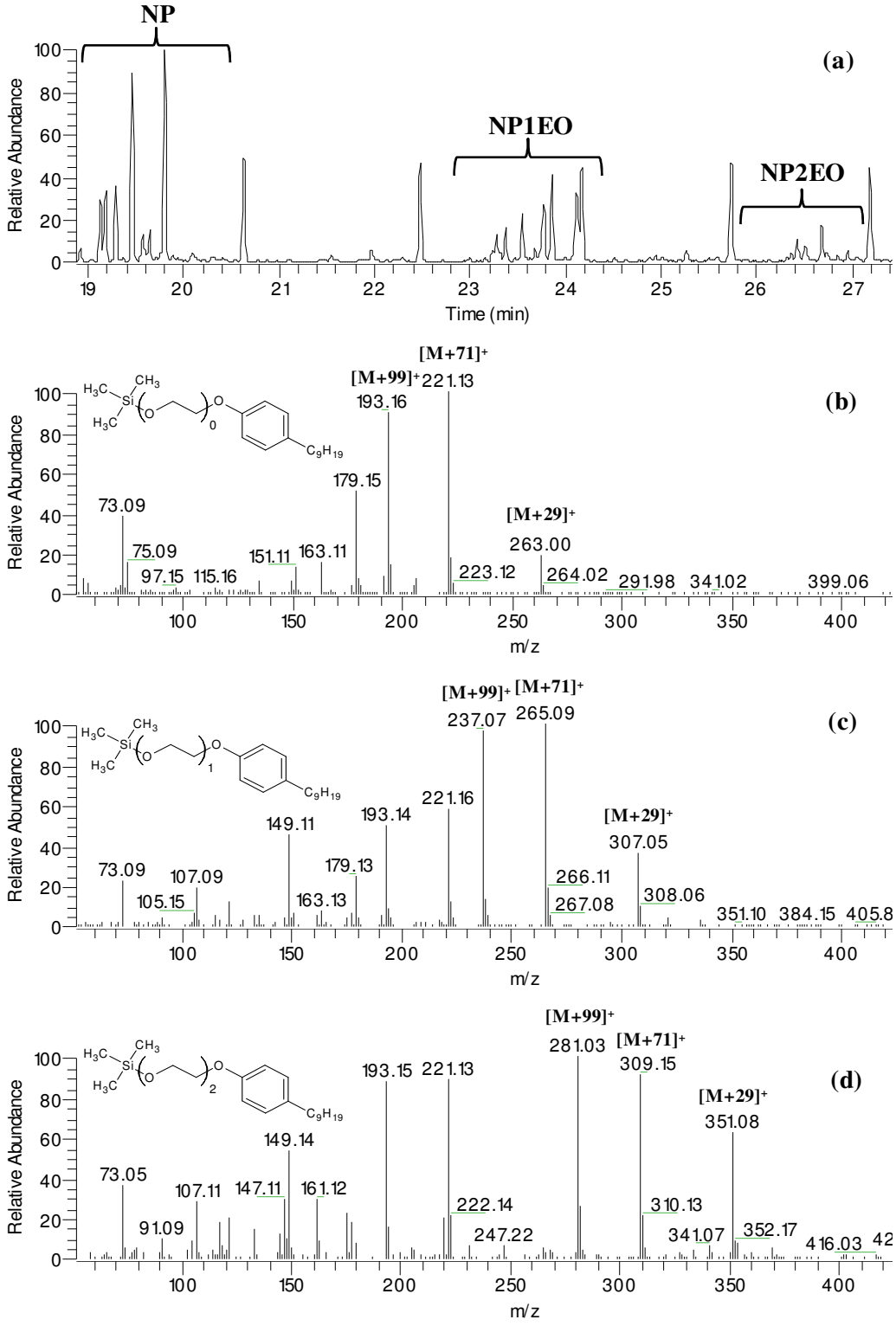
Appendix 3.3: Q-TOFMS mass spectra of the standard compounds a) nitrophenyl glucuronide and b) nitrophenyl sulphate and structures of their fragment ions analyzed in -ESI mode with collision energy of 10 eV.



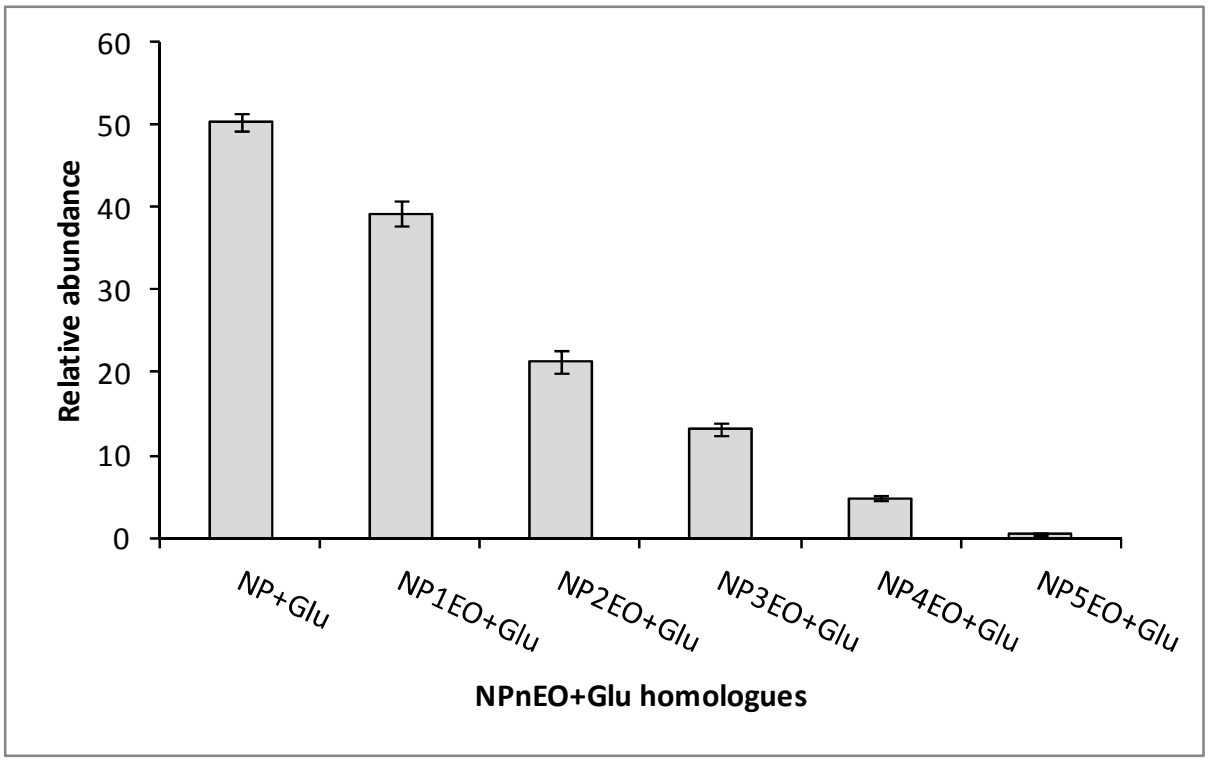
Appendix 3.4: Distribution of the relative concentration for SPC homologues (mean area) in bile extracts of effluent-exposed fish.



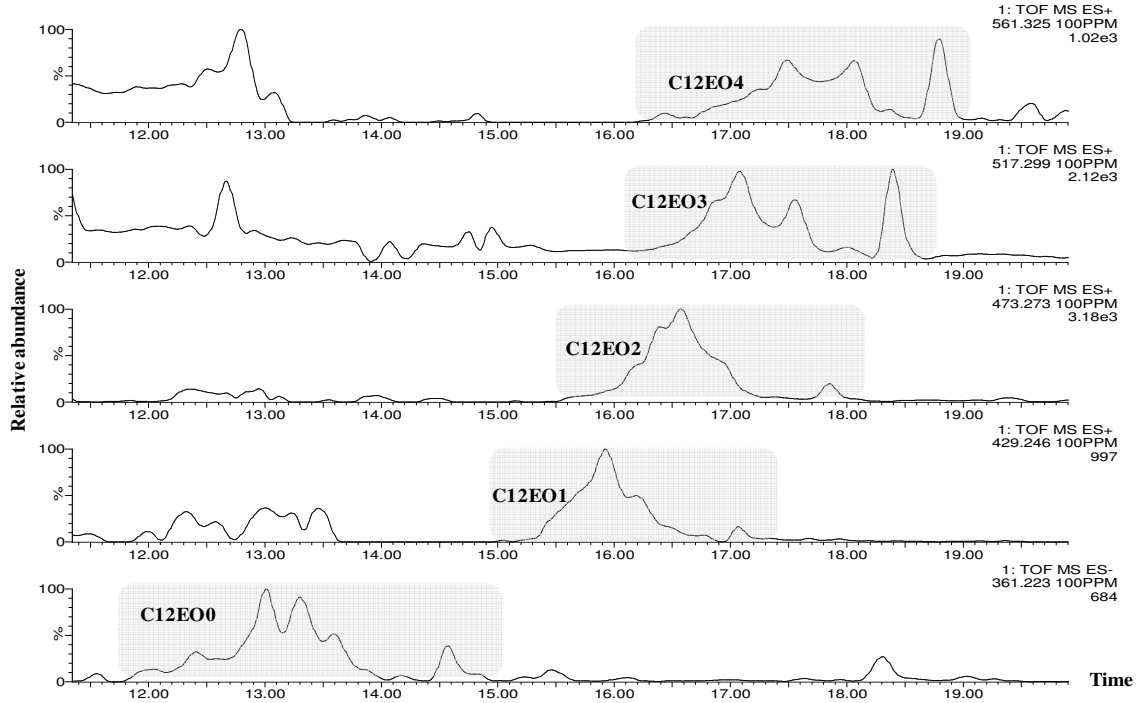
Appendix 3.5: Extracted ion chromatograms for detected of NPEO homologues. All compounds were detected as sodium adducts with the exception of NP, which was considered as the protonated form.



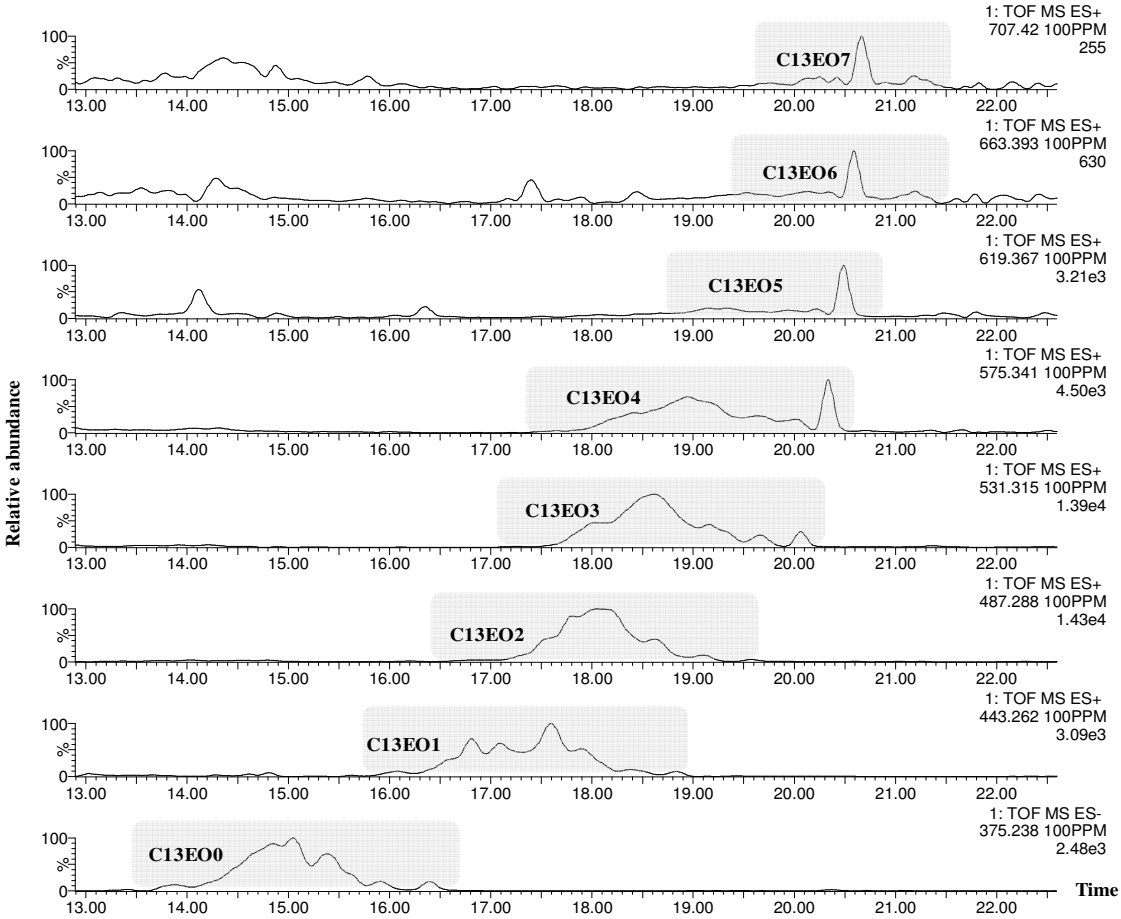
Appendix 3.6: a) GC-MS extracted ion chromatograms for NPEO homologues and relative mass spectra for b) NP, c) NP1EO and d) NP2EO as TMS derivatives.



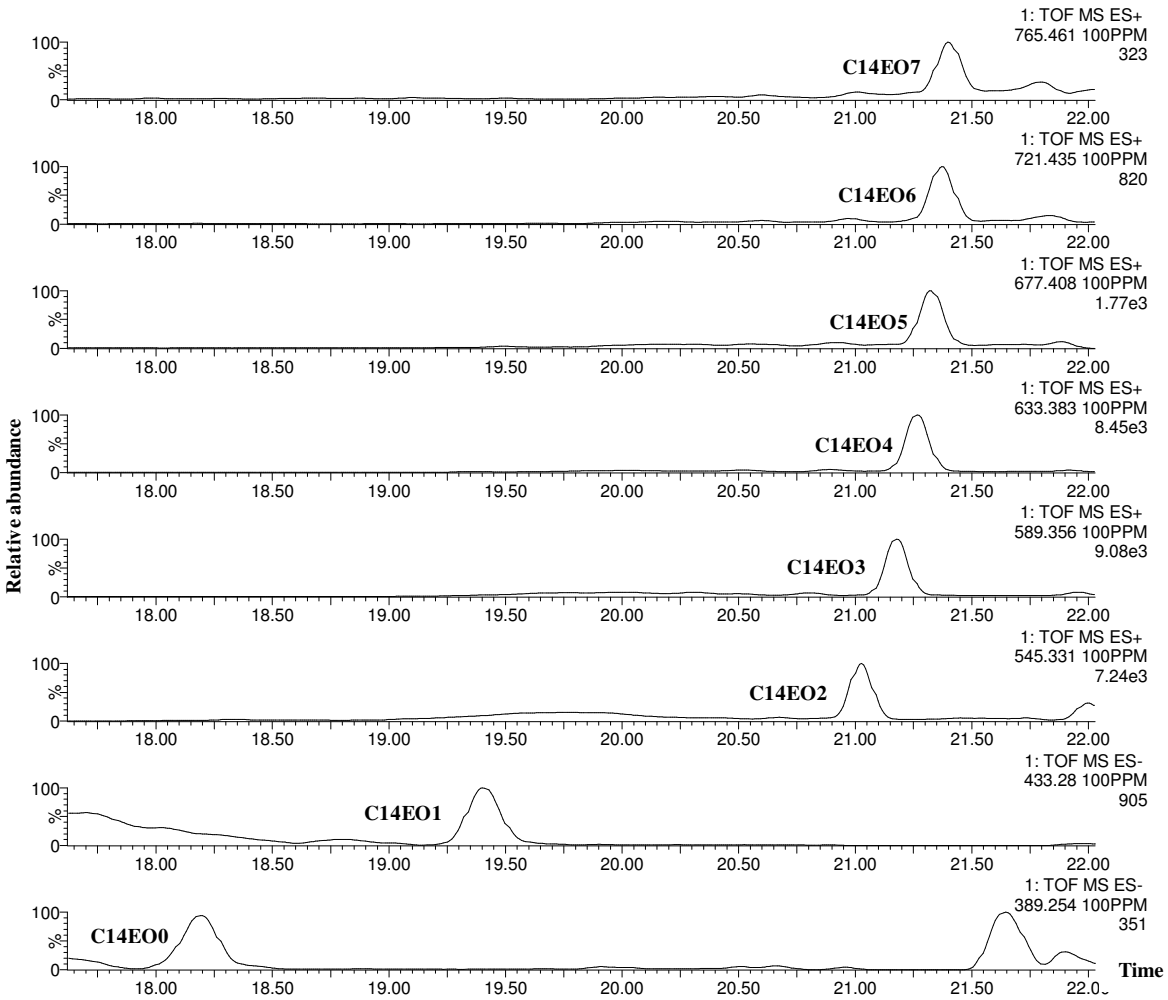
Appendix 3.7: Distribution of the relative abundance as mean area for NPEO homologues as glucuronide conjugates (NPnEO+Glu) in bile extracts of effluent-exposed fish.



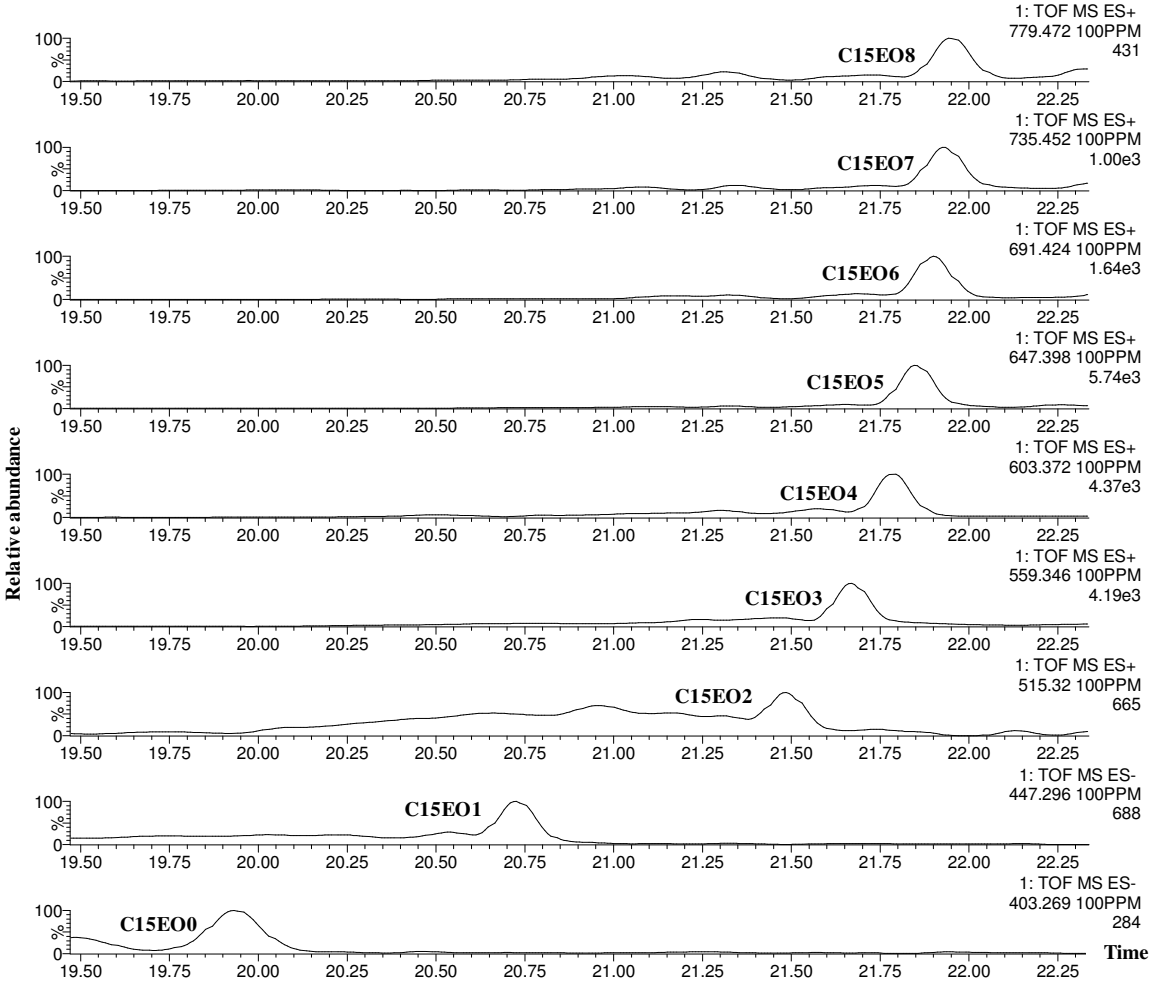
Appendix 3.8: a) Extracted ion chromatograms for C_{12}EO glucuronide homologues. All compounds were detected in +ESI mode as sodium adducts with the exception of dodecanol glucuronide ($C_{12}H_{25}$OH-Glu) detected in –ESI mode as deprotonated form.



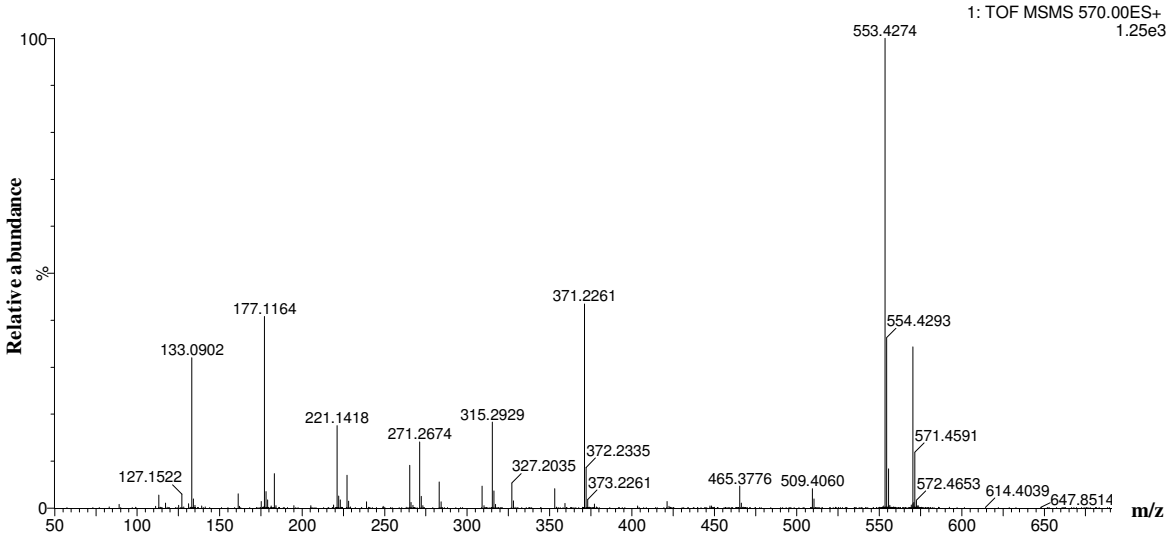
Appendix 3.8: b) Ion chromatograms for $C_{13}EO$ homologues elution. All compounds were extracted as sodium adducts species with exception of tridecanol ($C_{13}H_{27}OH$) as deprotonated form.



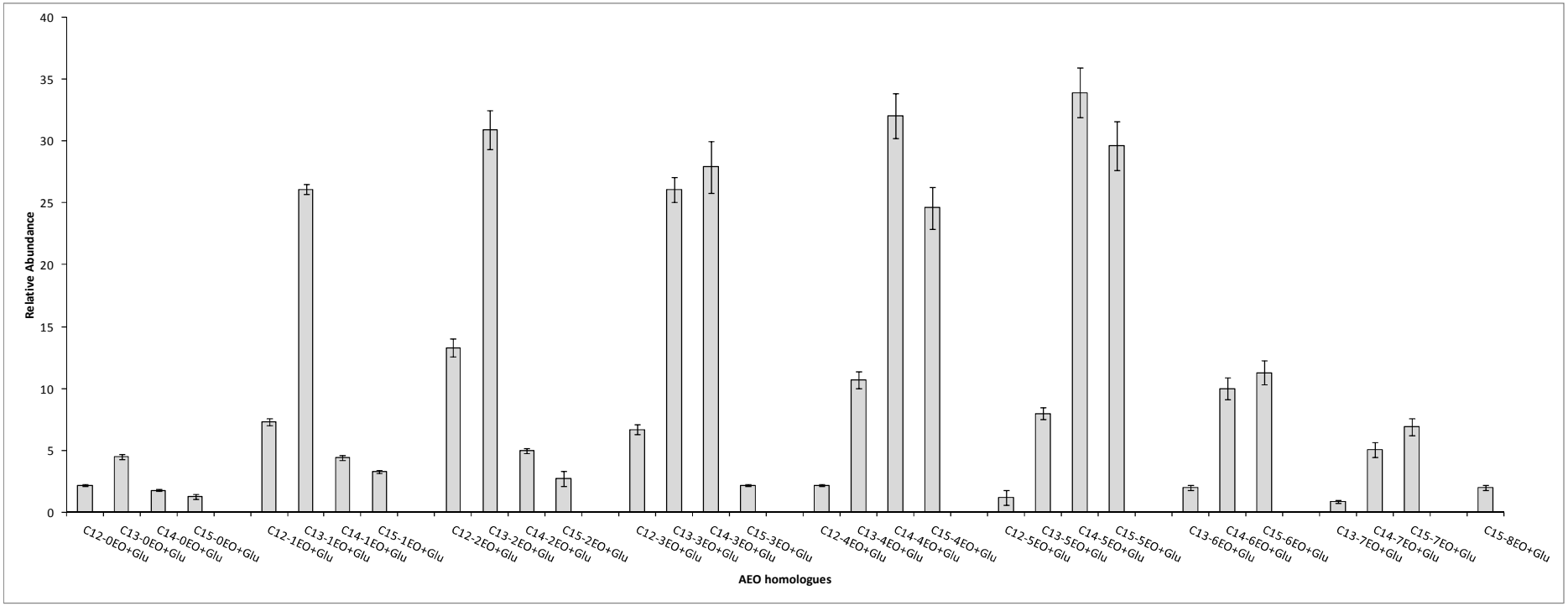
Appendix 3.8: c) Ion chromatograms for $C_{14}EO$ homologues elution. All compounds were extracted as sodium adducts species with exception of tetradecanol ($C_{14}H_{29}OH$) and $C_{14}(EO)_1$ as deprotonated form.



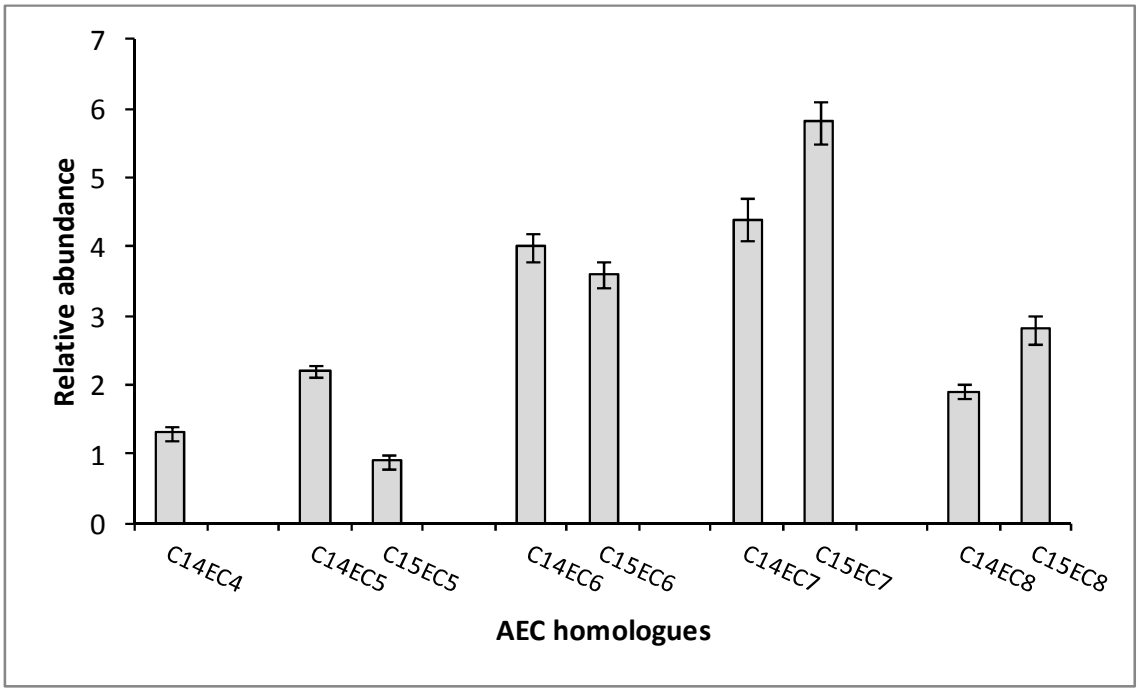
Appendix 3.8: d) Ion chromatograms for $C_{15}EO$ homologues elution. All compounds were extracted as sodium adducts species with exception of ($C_{15}H_{31}OH$) and $C_{15}(EO)_1$ as deprotonated form.



Appendix 3.9: +ESI Q-TOFMS spectrum (CE=20eV) for $C_{13}EO_8$ (technical mixture).



Appendix 3.10: Distribution of the relative abundance mean area for AEO series with EO units ranging from 0 to 8 in bile extracts from effluent-exposed fish.



Appendix 3.11: Distribution of the relative abundance for AEC homologues (mean area) in bile extracts of effluent-exposed fish.

Appendix 3.12a: Unidentified markers of effluent exposure in bile of trout in -ESI mode.

Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Fold change C_{10} vs E_{10}	% Decrease during depuration period		Identity of compound (see legend for explanation)
									4 day	11day	
291.0866	5.83	?				164.0686		0.2±0.0[¶]	69.9[ϕ]	67.5[ϕ]	Unidentified compound
557.0959	5.90	$C_{23}H_{25}O_{14}S$	557.0965	-1.1	0.2	381.0650, 285.0794, 205.0347, 175.0239, 113.0236	$C_{17}H_{18}O_8S$	156.8±10.7	100	100	Glucuronide conjugate of an unidentified LAS metabolite
388.1403	7.41	$C_{20}H_{22}NO_7$	388.1396	1.8	0.8	212.1075,193.0345,175.0241, 113.0188	$C_{14}H_{15}NO$	12.3±2.0[¶]	76.4	93.8	Glucuronide conjugate of unidentified compound (possibly dibenzylhydroxylamine or aminodiphenylethanol)
331.1758	9.07	$C_{16}H_{27}O_7$	331.1757	0.3	0.1	313.1661,175.0237,113.0219	$C_{10}H_{20}O$	263.4±22.4	98.6	100	Glucuronide of putative menthol metabolite
656.3105	9.23	?				480.2789,193.0344,175.0263		140.2±78.0	92.8	100	Glucuronide conjugate of unidentified compound
479.2859	11.69	$C_{23}H_{43}O_{10}$	479.2856	0.6	1.0	461.2734, 419.2577, 175.0251, 157.0148,113.0243	$C_{17}H_{36}O_4$	367.4±34.1	100	100	Glucuronide conjugate of putative monohydroxy tridecaonl-2EO metabolite
383.2073	12.39	$C_{20}H_{31}O_7$	383.2070	0.3	0.8	365.1959, 248.9621, 193.0356, 175.028, 113.0239	$C_{14}H_{24}O$	41.0±13.3[¶]	92.6	100	Glucuronide conjugate of an unidentified compound
433.1831	13.39	$C_{21}H_{31}O_8Na$	433.1838	-1.6	0.1	257.1518, 175.0245	$C_{15}H_{24}O_2$	7.1±1.8[¶]	18.3	46.5	Unidentified compound; Na adduct
403.2329	13.45	$C_{20}H_{35}O_8$	403.2332	-0.7	0.8	385.2234, 359.2065, 343.2132, 175.0259, 157.0130, 113.0239	$C_{14}H_{28}O_2$	744.1±32.3	94.2	99.3	Glucuronide conjugate of an unidentified compound
449.2178	15.64	$C_{24}H_{33}O_8$	449.2175	0.7	1.6	175.0241,113.0242	$C_{18}H_{26}O_2$	21.4±6.6[¶]	13.3	65.8	Glucuronide conjugate of unidentified compound
491.3222	12.77	$C_{25}H_{47}O_9$	491.3220	0.4	0.6	471.3108, 193.0345, 175.0241, 113.0241	$C_{19}H_{40}O_3$	0.2±0.0[¶]	62.4[ϕ]	80.0[ϕ]	Glucuronide conjugate of unidentified compound (two isomers on LCMS); endogenous
433.2801	13.3	$C_{22}H_{41}O_8$	433.2801	0.0	0.1	415.2705, 257.2474,175.0245, 113.0234	$C_{16}H_{34}O_2$	0.2±0.0[¶]	80.6	99.7	Glucuronide conjugate of unidentified compound

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; fold change calculated from relative concentrations of $[M-H]^-$ in bile of effluent-exposed trout compared with the values in control trout held in river water (10 day exposure); [¶] indicates compound was detected at levels >LOD (LOD=0.0113) in bile from reference fish held in river water; C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; [ϕ] represents % increase in metabolite concentrations after a 4 or 11 day depuration period following effluent exposure.

Appendix 3.12a: (continued) Unidentified markers of effluent exposure in bile of trout in -ESI mode.

Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Fold change C_{10} vs E_{10}	% Decrease during depuration period		Identity of compound (see legend for explanation)
									4 day	11day	
464.2470	13.68	?				377.2278, 343.2282, 205.1225		0.002±0.0[¶]	99.2[ϕ]	99.7[ϕ]	Unidentified compound
425.2171	13.88	$C_{22}H_{33}O_8$	425.2175	-0.9	13	407.2070, 381.1915, 249.1857, 193.0341, 175.0234, 113.0234	$C_{16}H_{26}O_2$	889.3±34.8	74.0	100%	Glucuronide conjugate of putative octylphenol-1EO metabolite
397.1859	14.14	$C_{20}H_{29}O_8$	397.1862	-0.8	2.5	221.1539, 193.0344, 175.0248	$C_{14}H_{22}O_2$	446.2±17.0	97.8	100%	Glucuronide conjugate of unidentified compound
381.1922	14.22	$C_{20}H_{29}O_7$	381.1913	2.4	16	205.1587, 175.0241, 157.0136, 113.0238	$C_{14}H_{22}O$	794.6±22.6	86.5	100	Glucuronide conjugate of unidentified compound
393.1924	14.64	$C_{21}H_{29}O_7$	393.1913	2.8	10	217.1597, 175.0241, 113.0234	$C_{15}H_{22}O$	573.1±32.7	98.1	99.2	Glucuronide conjugate of an unidentified compound
407.2071	14.71	$C_{22}H_{31}O_7$	407.207	0.2	2.9	231.1754, 175.0245, 113.0237	$C_{16}H_{24}O$	1151.1±35.3	71.9	100	Glucuronide conjugate of unidentified compound
389.2169	16.13	$C_{19}H_{33}O_8$	389.2175	-1.5	1.1	213.1854, 193.0353, 175.0244	$C_{13}H_{26}O_2$	174.6±6.8	96.4	100	Glucuronide conjugate of putative tridecanoic acid
479.3374	17.5	$C_{28}H_{47}O_6$	479.3373	0.2	0.7	351.2017, 317.2126, 245.1331, 162.0272		0.6±0.2	99.5[ϕ]	99.5[ϕ]	Unidentified compound
542.3336	20.31	$C_{28}H_{48}NO_9$	542.3329	1.3	0.4	175.0271, 113.0310	$C_{22}H_{41}NO_3$	234.2±16.8	91.9	100	Glucuronide conjugate of unidentified metabolite
610.2809	21.12	?				552.3074, 481.2363, 327.2330, 299.1394		0.4±0.1[¶]	21.3[ϕ]	50.2[ϕ]	Unidentified compound
585.3431	24.13	$C_{34}H_{49}O_8$	585.3427	0.7	0.2	409.3108, 277.1228, 175.0246, 113.0241	$C_{28}H_{42}O_2$	0.3±0.1[¶]	52.1[ϕ]	75.3[ϕ]	Glucuronide conjugate of unidentified compound

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; fold change calculated from relative concentrations of $[M-H]^-$ in bile of effluent-exposed trout compared with the values in control trout held in river water (10 day exposure); [¶] indicates compound was detected at levels >LOD (LOD=0.0113) in bile from reference fish held in river water; C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; [ϕ] represents % increase in metabolite concentrations after a 4 or 11 day depuration period following effluent exposure.

Appendix 3.12b: Unidentified markers of effluent exposure in bile of trout in +ESI mode.

Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Fold change C_{10} vs E_{10}	% Decrease during depuration period		Identity of compound (see legend for explanation)
									4 day	11day	
341.1816	4.35	$C_{14}H_{29}O_9$	341.1812	1.2	1.7	301.1434		400.9±24.7	94.7	100	Unidentified compound
317.1578	5.13	$C_{13}H_{26}O_7Na$	317.1576	0.6	0.5	259.1545, 245.1386		307.1±18.2	100	100	Unidentified compound
287.1475	5.32					215.1286		493.7±23.9	100	100	Unidentified compound
509.1564	5.85	$C_{29}H_{26}O_7Na$	509.1576	-2.4	0.1	333.1260	$C_{23}H_{18}O$	0.2±0.0[¶]	71.8[ϕ]	68.3[ϕ]	Glucuronide conjugate of unidentified compound
565.1577	6.58	?				389.1220		409.0±0.0	100	100	Glucuronide conjugate of unidentified compound
468.1426	8.09	?				292.1101, 275.1057, 246.0948		1.6±0.8[¶]	47.2	77.6	Glucuronide conjugate of unidentified compound(two isomers on LCMS)
233.1905	8.67	$C_{16}H_{25}O$	233.1905	0.0	0.8	191.1439		1005.9±42.3	42.4	82.3	Unidentified Compound (isomers on LCMS)
343.1885	9.16	$C_{19}H_{28}O_4Na$	343.1885	0.0	0.6	285.1672		14.4±1.8[¶]	95.5	99.2	Unidentified compound
449.2148 *444.2598*	13.88	$C_{22}H_{34}O_8Na$ *$C_{22}H_{38}NO_8$*	449.2151 *444.2597*	-0.7 *0.2*	1.2 *6.1*	273.1841, 251.2003, 233.1903, 199.0210 *251.2010, 233.1900*	$C_{16}H_{26}O_2$	622.3±139.8	29.6	77.1	Glucuronide conjugate of putative octylphenol-1EO metabolite; Na adduct (isomers on LCMS) *NH_4-adduct*
378.1705	13.89	?	378.1805	0.0	7.8	336.1688, 290.1619, 186.0282, 152.0406		7.1±3.4[¶]	100	100	Unidentified compound
473.2175	13.93	$C_{24}H_{34}O_8Na$	473.2175	0.0	7.0	297.1830, 199.0221	$C_{18}H_{26}O_2$	6.9±0.8[¶]	33.4[ϕ]	44.2	Glucuronide conjugate of unidentified compound
287.2378	14.17	$C_{20}H_{31}O$	287.2375	1.0	1.4	269.2365	$C_{20}H_{30}O$	374.5±14.4	100	100	Unidentified compound
548.2391	14.96	$C_{30}H_{34}N_3O_7$	548.2397	-0.5	2.2	372.2076	$C_{24}H_{25}N_3O$	383.8±19.8	100	100	Glucuronide conjugate of unidentified compound
257.1905	15.59	$C_{18}H_{25}O$	257.1905	0.0	1.9	229.1993	$C_{18}H_{24}O$	7.1±1.2[¶]	11.3[ϕ]	46.5	Unidentified compound
243.2114	15.95	?				228.1906, 213.1662, 173.1368		450.9±20.7	26.1	59.6	Unidentified compound

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; fold change calculated from relative concentrations of $[M+H]^+$ or $[M+Na]^+$ in bile of effluent-exposed trout compared with the values in control trout held in river water (10 day exposure); [¶] indicates compound was detected at levels >LOD (LOD=0.0064) in bile from reference fish held in river water; C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; [ϕ] represents % increase in metabolite concentrations after a 4 or 11 day depuration period following effluent exposure.

Appendix 3.12b: (continued) Unidentified markers of effluent exposure in bile of trout in +ESI mode.

Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Fold change C_{10} vs E_{10}	% Decrease during depuration period		Identity of compound (see legend for explanation)
									4 day	11day	
503.2833	18.14	$C_{23}H_{44}O_{10}Na$	503.2832	0.2	0.2	327.2519	$C_{17}H_{36}O_4$	1195.5±112.6	100	100	Glucuronide conjugate of unidentified ethoxylate compound (3EO)
498.3281		*$C_{23}H_{48}NO_{10}$*	*498.3281*	*0.6*	*1.0*	*305.2690, 151.0977, 133.0871*					*NH_4-adduct*
447.2958	19.7	$C_{23}H_{43}O_8$	447.2958	0.0	1.1	289.2746, 265.0950		977.4±37.7	97.2	100	Unidentified compound
518.3239	21.42	?				459.2377, 313.2710, 146.9847, 104.1075		0.3±0.0[¶]	68.6[ϕ]	69.2[ϕ]	Unidentified compound
573.3608	23.41	$C_{28}H_{54}O_{10}Na$	573.3608	0.0	2.5	397.3311	$C_{22}H_{46}O_4$	13.3±6.1[¶]	100	100	Glucuronide conjugate of unidentified ethoxylate compound
568.4051		*$C_{28}H_{58}NO_{10}$*	*568.4061*	*-1.8*	*5.0*	*375.3479, 177.1122, 133.0866*					*NH_4-adduct*
631.404	23.99	$C_{31}H_{60}O_{11}Na$	631.4033	1.1	2.8	455.3712	$C_{25}H_{52}O_5$	616.5±36.4	64.6	97.7	Glucuronide conjugate of unidentified compound; Na adduct
626.4476		*$C_{31}H_{64}NO_{11}$*	*626.4479*	*-0.5*	*0.2*	*433.3883, 415.2560, 177.1121, 133.0866*					*NH_4-adduct*

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; fold change calculated from relative concentrations of $[M+H]^+$ or $[M+Na]^+$ in bile of effluent-exposed trout compared with the values in control trout held in river water (10 day exposure); [¶] indicates compound was detected at levels >LOD (LOD=0.0064) in bile from reference fish held in river water; C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; [ϕ] represents % increase in metabolite concentrations after a 4 or 11 day depuration period following effluent exposure.

Chapter four appendices

Appendix 4.1: Performance parameters of principal component analyses (PCA) for the comparison of trout exposed to river water and effluent (plasma extracts).

Ionization mode	Groups	Multivariate method	Principal Components	R^2X	Q^2
+ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PCA	3	0.230	0.126
-ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PCA	3	0.206	0.104
+ESI	C_{10},C_{14},E_{10},E_{14}	PCA	2	0.186	0.092
-ESI	C_{10},C_{14},E_{10},E_{14}	PCA	2	0.178	0.079
+ESI	C_{10},C_{21},E_{10},E_{21}	PCA	2	0.196	0.114
-ESI	C_{10},C_{21},E_{10},E_{21}	PCA	2	0.174	0.090
+ESI	C_{14},C_{21},E_{14},E_{21}	PCA	2 (1)	0.215	0.066
-ESI	C_{14},C_{21},E_{14},E_{21}	PCA	2 (1)	0.199	0.038
+ESI	C_{10},E_{10}	PCA	2 (1)	0.207	0.074
-ESI	C_{10},E_{10}	PCA	2 (1)	0.210	0.073
+ESI	C_{14},E_{14}	PCA	2 (2)	0.337	-0.015
-ESI	C_{14},E_{14}	PCA	2 (2)	0.337	-0.028
+ESI	C_{21},E_{21}	PCA	2 (2)	0.263	-0.006
-ESI	C_{21},E_{21}	PCA	2 (2)	0.262	-0.028

+/-ESI: positive/negative electrospray ionization; C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control); E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; R^2X: variation explained by the models; Q^2: cumulative variation predicted by the models. Discriminating components that are not significant are shown in brackets.

Appendix 4.2: Performance parameters of multivariate discriminant models for the comparison of trout exposed to river water and effluent (plasma extracts).

Ionization mode	Groups	Multivariate method	Number of projections	R^2X	R^2Y	Q^2	External Validation
+ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PLS-DA	10	0.383	0.984	0.772	66%
-ESI	C_{10},C_{14},C_{21},E_{10},E_{14},E_{21}	PLS-DA	11	0.395	0.987	0.724	78%
+ESI	C_{10},C_{14},E_{10},E_{14}	PLS-DA	7	0.357	0.993	0.851	100%
-ESI	C_{10},C_{14},E_{10},E_{14}	PLS-DA	6	0.321	0.983	0.796	100%
+ESI	C_{10},C_{21},E_{10},E_{21}	PLS-DA	7	0.356	0.990	0.857	91%
-ESI	C_{10},C_{21},E_{10},E_{21}	PLS-DA	6	0.304	0.981	0.738	100%
+ESI	C_{14},C_{21},E_{14},E_{21}	PLS-DA	6	0.375	0.993	0.739	40%
-ESI	C_{14},C_{21},E_{14},E_{21}	PLS-DA	6	0.375	0.984	0.659	40%
+ESI	C_{10},E_{10}	PLS-DA	3	0.215	0.999	0.863	100%
-ESI	C_{10},E_{10}	PLS-DA	2	0.163	0.988	0.790	100%
+ESI	C_{14},E_{14}	PLS-DA	2	0.320	0.993	0.677	100%
-ESI	C_{14},E_{14}	PLS-DA	2	0.311	0.991	0.653	100%
+ESI	C_{21},E_{21}	PLS-DA	2	0.226	0.994	0.589	100%
-ESI	C_{21},E_{21}	PLS-DA	2	0.188	0.996	0.354	
+ESI	C_{10},E_{10}	OPLS	1+1	0.156	0.988	0.722	
-ESI	C_{10},E_{10}	OPLS	1+1	0.163	0.988	0.815	
+ESI	C_{14},E_{14}	OPLS	1+1 (1)	0.320	0.993	0.513	
-ESI	C_{14},E_{14}	OPLS	1+1 (1)	0.311	0.991	0.549	
+ESI	C_{21},E_{21}	OPLS	1+1 (1)	0.226	0.994	0.202	
-ESI	C_{21},E_{21}	OPLS	1+1 (1)	0.188	0.996	0.065	

+/-ESI: positive/negative electrospray ionization; C_{10}: 10 days river water exposure (control); C_{14}: 14 days river water exposure (control); C_{21}: 21 days river water exposure (control); E_{10}: 10 days effluent exposure, E_{14}: 10 days effluent exposure followed by 4 days depuration, E_{21}: 10 days effluent exposure then 11 days depuration; R^2Y and Q^2 represent the cumulative variation explained and predicted by the models, respectively.

Appendix 4.3a: Unidentified markers of effluent exposure in plasma of trout in -ESI mode.

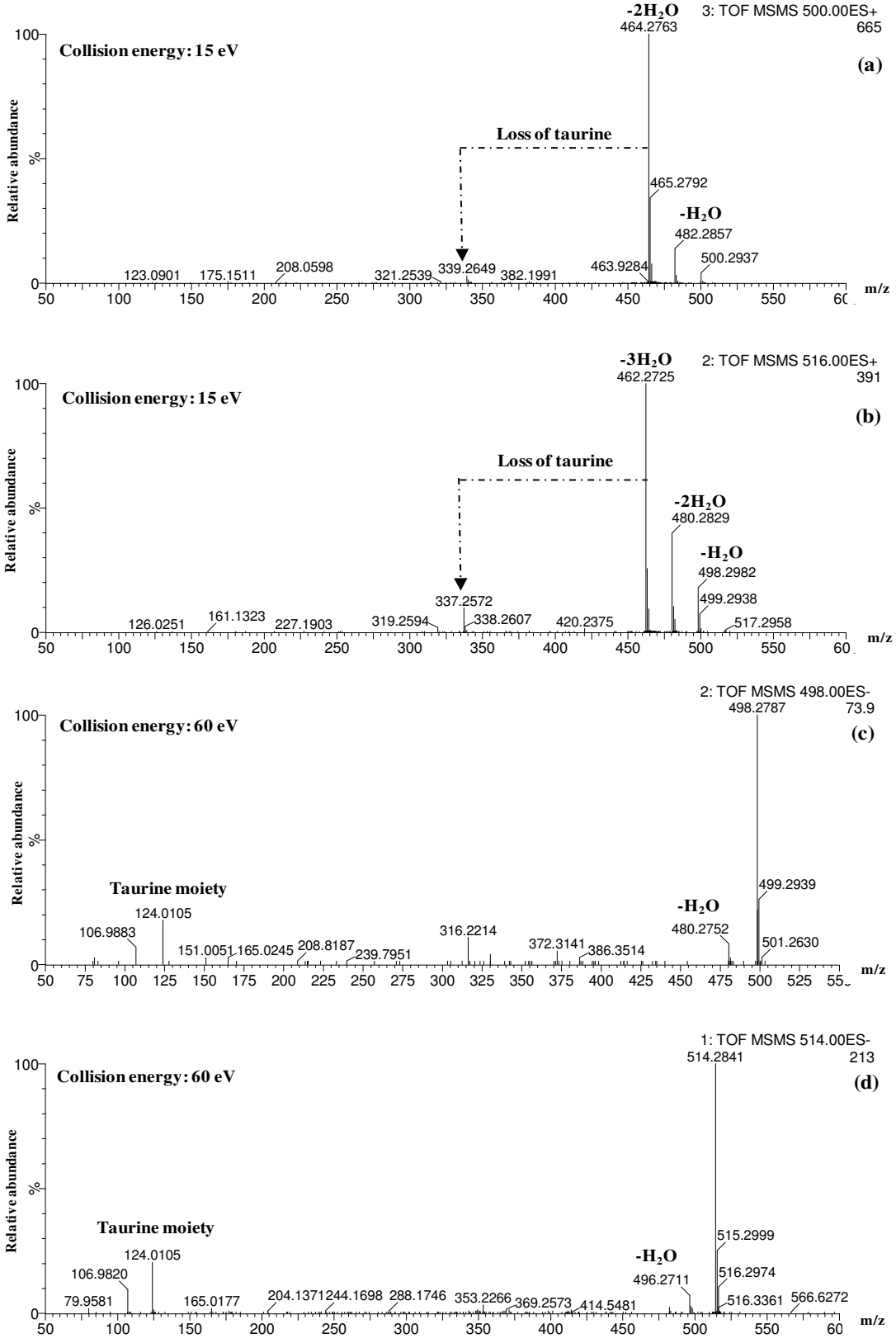
Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Fold change C_{10} vs E_{10}	% Decrease during depuration period		Identity of compound
									4 day	11 day	
571.3837	4.39	$C_{31}H_{55}O_9$	571.3846	-1.6	0.1	527.3581	$C_{31}H_{54}O_9$	71.2±20.7	100	100	Unidentified compound
249.0218	8.04	$C_{12}H_9O_4S$	249.0222	-1.6	0.8	169.0654	$C_{12}H_{10}O$	4.2±3.1	100	100	Sulphate conjugate of putative hydroxybiphenyl
229.0561	8.07	$C_6H_{13}O_9$	229.0560	0.4	3.7			69.9±4.2	100	100	Unidentified compound
512.2705	8.34	$C_{26}H_{42}NO_7S$	512.2707	-0.4	7.5	124.0072, 106.9808	$C_{26}H_{43}NO_7S$	6.9±2.1[¶]	72.3	55.6	Taurine conjugate of unidentified bile acid
250.9783	8.77	$C_8H_8O_5SCl$	250.9781	0.8	1.3	171.0216	$C_8H_9O_2Cl$	42.7±6.8	100	100	Sulphate conjugate of putative chlorinated phenol
327.1633	11.55	$C_{17}H_{27}O_4S$	327.1630	0.9	0.5	183.0112	$C_{17}H_{28}O_4S$	49.7±4.1	100	100	Putative monohydroxy C_{11}-LAS
482.2920	12.52	?						20.8±2.8	100	100	Unidentified compound
409.9714	14.18	$C_{10}H_{14}NO_{10}SCl_2$	409.9715	-0.2	1.7	330.0133	$C_{10}H_{14}NO_7Cl_2$	28.3±2.3	100	100	Sulphate conjugate of unidentified chlorinated compound
347.9540	15.75							24.4±3.2[¶]	33.8[ϕ]	11.5[ϕ]	Unidentified compound
397.9537	18.43	?				295.1954, 277.1877		2.1±0.1[¶]	3.2[ϕ]	22.1	Unidentified compound
497.9463	20.28	?				255.2321, 153.9952		2.3±0.1[¶]	0.0	20.1	Unidentified compound
262.8394	21.33	?						48.2±2.3	100	100	Unidentified halogenated compound
450.9278	21.69	?				414.9528, 345.9497, 281.9895, 255.2363		12.8±8.8[¶]	87.7	93.1	Unidentified compound
569.9671	22.41	?				511.9587, 482.9446, 418.9709, 218.9920, 168.9868		60.9±9.2	14.6[ϕ]	37.8	Unidentified compound
321.2068	22.58	$C_{19}H_{29}O_4$	321.2066	0.6	4.2	219.1744		46.4±6.0	100	100	Unidentified compound
329.2695	24.57	$C_{19}H_{37}O_4$	329.2692	0.9	1.7	281.2480	$C_{19}H_{38}O_4$	66.8±5.8	100	100	Unidentified compound
373.2960	24.63	$C_{21}H_{41}O_5$	373.2954	1.6	3.3			36.4±1.4	100	100	Unidentified compound

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; fold change calculated from relative concentrations of $[M-H]^-$ in plasma of effluent-exposed trout compared with the values in control trout held in river water (10 day exposure); [¶] indicates compound was detected at levels >LOD (LOD=0.0169) in plasma from reference fish held in river water; C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; [ϕ] represents % increase in metabolite concentrations after a 4 or 11 day depuration period following effluent exposure.

Appendix 4.3b: Unidentified markers of effluent exposure in plasma of trout in +ESI mode.

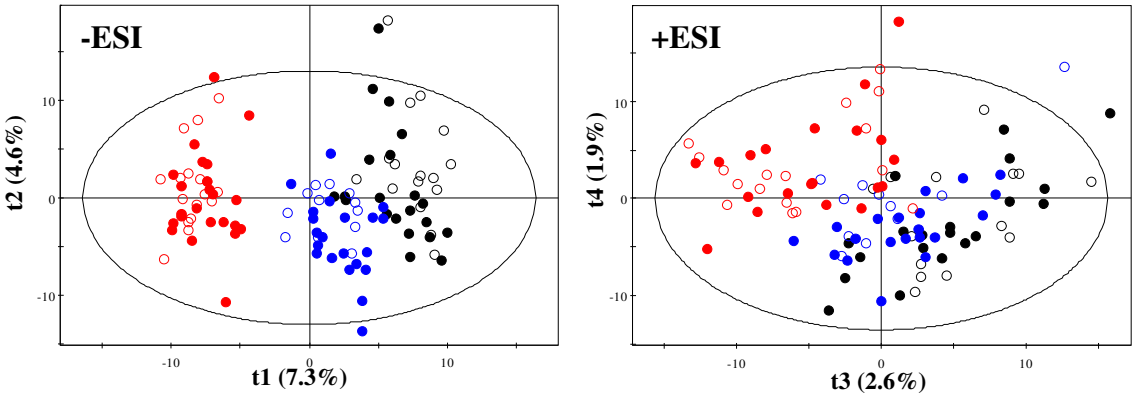
Observed ion (*m/z*)	RT	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula[§]	Fold change C_{10} vs E_{10}	% Decrease during depuration period		Identity of compound
									4 day	11 day	
287.2014	4.30	$C_{19}H_{27}O_2$	287.2011	1.0	3.4	147.1137	$C_{19}H_{28}O_2$	76.6±18.4	91.1	100	Steroid like
349.1902	4.30	?						4.9±1.5[¶]	50.7	68.1	Unidentified compound
573.3974	4.31	$C_{29}H_{58}O_9Na$	573.3974	-0.9	3.2	373.2820, 314.1804	$C_{29}H_{58}O_9$	211.5±45.5	96.2	99.5	Unidentified compound
915.0370	4.10	?						0.8±0.1[¶]	58.6	74.7	
801.2529	4.27	?						0.4±0.1[¶]	67.3[ϕ]	58.1[ϕ]	Phospholipid like
961.3008	4.27	?						0.3±0.1[¶]	73.1[ϕ]	68.0[ϕ]	Phospholipid like
710.7023	4.69	?						162.3±51.0	79.1	100	Phospholipid like
799.4440	5.18	?				790.3388, 719.3889, 618.3337, 583.3423, 470.2602, 339.2231		2.9±1.3[¶]	86.6	100	Phospholipid like
849.7238	5.61	?						0.03±0.00[¶]	89.3[ϕ]	94.1[ϕ]	Phospholipid like
679.9800	5.63	?						0.05±0.00[¶]	83.6[ϕ]	78.1[ϕ]	Possibly fragmented from 849.7238
566.8147	5.63	?						0.01±0.00[¶]	98.8[ϕ]	98.2[ϕ]	Possibly fragmented from 849.7238
445.3036	5.39	?				314.2055, 245.1862, 201.1239, 173.1292		323.1±138.9	100	100	Unidentified compound
327.0796	20.28	?				251.0654, 215.0405, 228.1049		32.1±1.3	100	91.5	Unidentified compound
330.3380	21.58	?						4.9±1.9	100	100	Unidentified compound
284.3318	22.34	?				155.1180		31.3±3.3[¶]	37.6	45.4	Unidentified compound
344.3165	23.13	$C_{20}H_{42}NO_3$	344.3165	0.0	2.2	326.3024, 283.2635	$C_{20}H_{43}NO_3$	64.2±8.5	100	100	Unidentified compound
251.0477	23.28	?						8.3±5.9	100	100	Unidentified compound

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit; [#]Q-TOFMS fragments within ± 5ppm; [§]aglycone formula; fold change calculated from relative concentrations of $[M+H]^+$ or $[M+Na]^+$ in plasma of effluent-exposed trout compared with the values in control trout held in river water (10 day exposure); [¶] indicates compound was detected at levels >LOD (LOD=0.0054) in plasma from reference fish held in river water; C_{10}: 10 days river water exposure (control); E_{10}: 10 days effluent exposure; [ϕ] represents % increase in metabolite concentrations after a 4 or 11 day depuration period following effluent exposure.



Appendix 4.4: Q-TOFMS spectra for the standard compounds: a) and c) taurochenodeoxycholic acid and b) and d) taurocholic acid in both ESI modes.

Chapter five appendices



Appendix 5.1: Principal component analyses (PCA) scores plots of the chemical profiles of plasma samples from roach exposed to either control river water or wastewater effluent in both +/-ESI modes. Control (C)= river water exposure where n=16 female, n=21 male. E_{50}=exposure to 50% effluent where n=12 female, n=21 male. E_{100}= exposure to 100% effluent exposures where n=17 female n=23 male. Female roach symbols :(○), (○) and (○) represent C, E_{50} and E_{100} respectively. Male roach symbols: (●), (●) and (●) symbols represent C, E_{50} and E_{100}, respectively. The percentages of explained variation (R^2Y) modelled for the two selected components (t1, t2) in –ESI and (t3, t4) in +ESI mode are displayed on the related axes. Outliers were detected in both –ESI (4 outliers) and +ESI (3 outliers).

Appendix 5.2: Performance parameters of principal component analyses (PCA) for the comparison of roach exposed to river water, 50% effluent and 100% effluent.

Ionization mode	Gender	Groups	Multivariate method	Principal Components	R^2X	Q^2
+ESI	F & M	C, E_{50}, E_{100}	PCA	6	0.396	0.295
-ESI	F & M	C, E_{50}, E_{100}	PCA	4	0.179	0.094
+ESI	F	C, E_{50}, E_{100}	PCA	3	0.384	0.268
-ESI	F	C, E_{50}, E_{100}	PCA	2 (1)	0.158	0.050
+ESI	M	C, E_{50}, E_{100}	PCA	4	0.388	0.266
-ESI	M	C, E_{50}, E_{100}	PCA	2	0.132	0.060
+ESI	F	C, E_{50}	PCA	2 (1)	0.380	0.251
-ESI	F	C, E_{50}	PCA	2 (2)	0.151	-0.020
+ESI	M	C, E_{50}	PCA	2 (1)	0.358	0.255
-ESI	M	C, E_{50}	PCA	2 (2)	0.127	-0.003
+ESI	F	C, E_{100}	PCA	2	0.358	0.255
-ESI	F	C, E_{100}	PCA	2 (1)	0.200	0.060
+ESI	M	C, E_{100}	PCA	2 (1)	0.342	0.243
-ESI	M	C, E_{100}	PCA	2 (1)	0.166	0.066
+ESI	F & M	C	PCA	2 (1)	0.349	0.245
-ESI	F & M	C	PCA	2 (2)	0.141	0.016
+ESI	F & M	E_{50}	PCA	2 (1)	0.388	0.283
-ESI	F & M	E_{50}	PCA	2 (2)	0.127	-0.032
+ESI	F & M	E_{100}	PCA	2	0.377	0.293
-ESI	F & M	E_{100}	PCA	2 (2)	0.139	0.018

+/-ESI: positive/negative electrospray ionization; C: 28 days river water exposure (control); E_{50}: 28 days 50% effluent exposure; E_{100}: 28 days 100% effluent exposure; F: female; M: male; R^2X: variation explained by the models; Q^2: cumulative variation predicted by the models. Discriminating components that are not significant are shown in brackets.

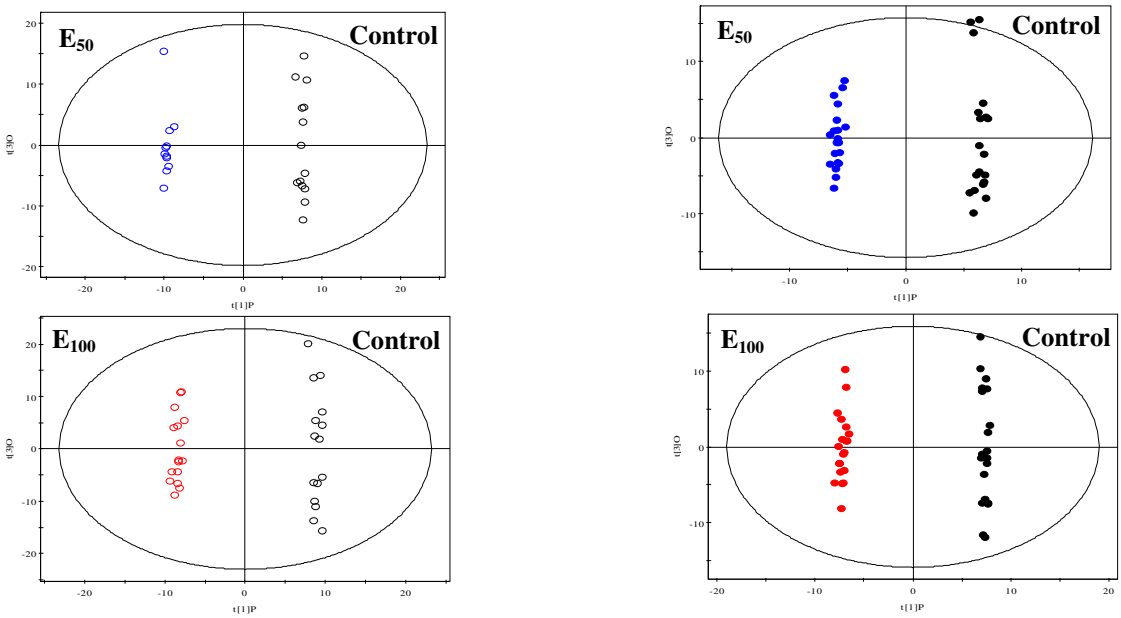
Appendix 5.3: Performance parameters of multivariate discriminant models for the comparison of roach exposed to river water, 50% effluent and 100% effluent.

Ionization mode	Gender	Groups	Multivariate method	Number of projections	R^2X	R^2Y	Q^2	External Validation
+ESI	F & M	C, E_{50}, E_{100}	PLS-DA	9	0.407	0.965	0.651	38%
-ESI	F & M	C, E_{50}, E_{100}	PLS-DA	4	0.146	0.680	0.433	73%
+ESI	F	C, E_{50}, E_{100}	PLS-DA	5	0.419	0.995	0.855	74%
-ESI	F	C, E_{50}, E_{100}	PLS-DA	4	0.210	0.991	0.800	100%
+ESI	M	C, E_{50}, E_{100}	PLS-DA	6	0.416	0.996	0.849	88%
-ESI	M	C, E_{50}, E_{100}	PLS-DA	5	0.200	0.995	0.882	100%
+ESI	F	C, E_{50}	PLS-DA	3	0.409	0.997	0.889	100%
-ESI	F	C, E_{50}	PLS-DA	2	0.132	0.994	0.778	100%
+ESI	M	C, E_{50}	PLS-DA	4	0.406	0.998	0.913	100%
-ESI	M	C, E_{50}	PLS-DA	3	0.145	0.999	0.854	100%
+ESI	F	C, E_{100}	PLS-DA	3	0.387	0.997	0.833	100%
-ESI	F	C, E_{100}	PLS-DA	2 (1)	0.167	0.998	0.900	100%
+ESI	M	C, E_{100}	PLS-DA	3	0.367	0.995	0.919	100%
-ESI	M	C, E_{100}	PLS-DA	2	0.136	0.996	0.921	100%
+ESI	F	C, E_{50}	OPLS-DA	1+2 (1)	0.413	0.998	0.418	
-ESI	F	C, E_{50}	OPLS-DA	1+1 (1)	0.132	0.994	0.680	
+ESI	M	C, E_{50}	OPLS-DA	1+2 (2)	0.379	0.996	0.383	
-ESI	M	C, E_{50}	OPLS-DA	1+1 (1)	0.106	0.992	0.734	
+ESI	F	C, E_{100}	OPLS-DA	1+2 (1)	0.395	0.997	0.730	
-ESI	F	C, E_{100}	OPLS-DA	1+1 (1)	0.270	0.993	0.912	
+ESI	M	C, E_{100}	OPLS-DA	1+2 (1)	0.364	0.998	0.701	
-ESI	M	C, E_{100}	OPLS-DA	1+1 (1)	0.136	0.996	0.919	
+ESI	F & M	C	OPLS-DA	1+1 (1)	0.329	0.917	0.287	
-ESI	F & M	C	OPLS-DA	1+1	0.107	0.973	0.460	
+ESI	F & M	E_{50}	OPLS-DA	1+1 (1)	0.378	0.963	0.273	
-ESI	F & M	E_{50}	OPLS-DA	1+1	0.091	0.994	0.406	
+ESI	F & M	E_{100}	OPLS-DA	1+2	0.383	0.994	0.289	
-ESI	F & M	E_{100}	OPLS-DA	1+1	0.102	0.987	0.515	

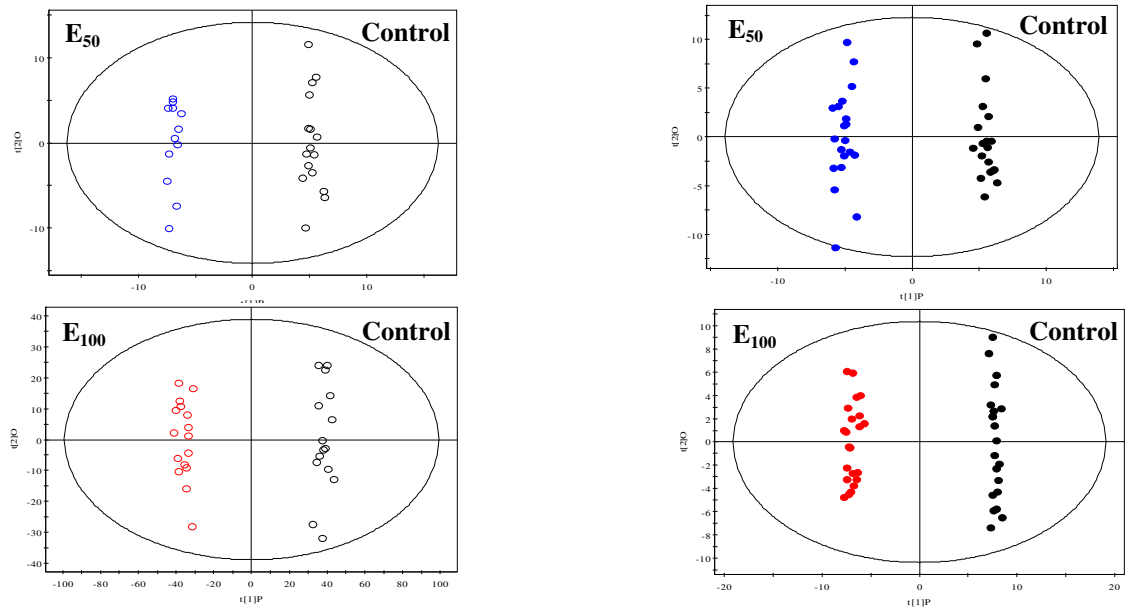
+/-ESI: positive/negative electrospray ionization; C: 28 days river water exposure (control); E_{50}: 28 days 50% effluent exposure; E_{100}: 28 days 100% effluent exposure; F: female; M: male; FC: female control; MC: male control; FE_{100}: female 100% effluent; ME_{100}: male 100% effluent; R^2X and R^2Y: variation explained by the models; Q^2: cumulative variation predicted by the models. Discriminating components that are not significant are shown in brackets.

Positive mode

Females *Males*



Negative mode



Appendix 5.4: OPLS-DA scores plots of the chemical profiles of plasma samples from roach exposed either to wastewater effluent or control river water. The samples were profiled in both +/-ESI modes by UPLC-TOF-MS. Control (C): river water exposure where n=16 females, n=21 males. E_{50}: exposure to 50% effluent where n=12 females, n=21 males. E_{100}: exposure to 100% effluent exposures where n=17 females, n=23 males. Female roach symbols: (○), (○) and (○) represent C, E_{50} and E_{100}, respectively. Male roach symbols: (●), (●) and (●) represent C, E_{50} and E_{100}, respectively.

Appendix 5.5: list of discriminatory variables between female and male in plasma from roach exposed to river water, 50% or 100% effluent.

Ionization mode	*m/z* value of marker ion	RT (min)	Groups					
			FC	MC	FE_{50}	ME_{50}	FE_{100}	ME_{100}
+ESI	386.2161	3.63	↑					
+ESI	456.2281	3.63	↑					
+ESI	481.2425	5.74	↑					
+ESI	465.2487	7.26			↑		↑	
+ESI	425.1378	12.28			↑			
+ESI	395.1846	12.54			↑			
+ESI	310.2746	14.72				↑		
+ESI	532.3298	15.05					↑	
-ESI	496.3018	16.20		↑				
+ESI	466.3301	16.24						↑
+ESI	513.2925	16.57			↑			
+ESI	401.3373	17.09		↑				
+ESI	583.2663	17.09			↑			
-ESI	498.2916	17.20		↑				
+ESI	547.2992	17.49		↑				
+ESI	512.3724	17.82		↑				
+ESI	562.3480	18.08				↑		
-ESI	524.3327	18.20						↑
+ESI	504.3399	18.74						↑
-ESI	556.3585	19.04		↑				
-ESI	464.2957	19.21		↑				
+ESI	277.1805	19.34				↑		
-ESI	436.2821	20.37						↑
-ESI	470.3464	20.71		↑				
-ESI	537.3248	20.88					↑	
-ESI	550.3489	20.88						↑
+ESI	283.6567	21.12	↑					
-ESI	479.2813	22.04					↑	
-ESI	508.3376	22.21		↑				
-ESI	534.3545	22.55		↑				
-ESI	423.3420	23.71	↑					

+/-ESI: positive/negative electrospray ionization; RT: retention time (minute); *m/z*: mass to charge ratio; ↑ elevated markers; C: 28 days river water exposure (control); E_{50}: 28 days 50% effluent exposure; E_{100}: 28 days 100% effluent exposure; F: female; M: male; FC: female control; MC: male control; FE_{50}: female 50% effluent; ME_{50}: male 50% effluent; FE_{100}: female 100% effluent; ME_{100}: male 100% effluent.

Appendix 5.6a: Unidentified markers of effluent exposure in plasma of roach in +ESI mode.

Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula	Gender	Control vs 50% effluent		Control vs 100% effluent		Putative identity
									Fold change	*p*-value	Fold change	*p*-value	
361.1837	3.9	$C_{15}H_{30}O_8Na$	361.1838	-0.3	0.6		$C_{15}H_{30}O_8$	F	0.4[¶]	1.10E-02	0.3[¶]	6.00E-03	
								M	0.6[¶]	4.25E-01	0.5[¶]	3.49E-01	
271.1160	4.8	$C_{11}H_{20}O_6Na$	271.1158	0.7	1.0		$C_{11}H_{20}O_6$	F	0.3[¶]	7.50E-02	5.9[¶]	3.38E-07	
								E	1.4[¶]	5.80E-01	4.8[¶]	2.27E-08	
331.2095	5.3	$C_{15}H_{32}O_6Na$	331.2097	-0.6	0.4		$C_{15}H_{32}O_6$	F	4.3[¶]	2.00E-03	5.6[¶]	7.36E-06	
								M	2.3[¶]	2.00E-04	3.8[¶]	1.69E-06	
453.1868	5.4	?						F	0.5[¶]	2.50E-02	0.3[¶]	1.00E-03	
								M	0.5[¶]	4.82E-01	0.3[¶]	2.96E-01	
387.2381	6.6	$C_{20}H_{37}O_7$	387.2383	-0.5	1.3		$C_{20}H_{36}O_7$	F	ND	1.00E+00	90.9	1.43E-04	
								M	5.4	4.88E-01	46.4	3.00E-03	
313.1992	7.3	$C_{15}H_{30}O_5Na$	313.1991	0.3	0.1		$C_{15}H_{30}O_5$	F	8.2	1.69E-01	89.5	1.40E-05	
								M	8.0	1.07E-01	49.6	1.00E-03	
366.3007	11.3	$C_{22}H_{40}NO_3$	366.3008	-0.3	0.8	348.2901, 274.2509	$C_{22}H_{39}NO_3$	F	205.8	1.21E-07	684.9	2.25E-10	
								M	228.6	9.34E-11	804.3	2.60E-11	
274.2751	12.1	$C_{16}H_{36}NO_2$	274.2746	1.8	0.2		$C_{16}H_{35}NO_2$	F	22.8[¶]	2.00E-03	70.1[¶]	3.54E-09	Possibly C16-sphinganine
								M	8.7	1.00E-02	41.1	4.22E-09	
200.2380	12.7	$C_{13}H_{30}N$	200.2378	1.0	0.7			F	55.4	2.40E-04	76.7	1.23E-08	
								M	43.8	9.95E-07	82.9	7.09E-09	
853.2957	13.0	?						F	0.8[¶]	5.40E-01	0.2[¶]	5.36E-05	fragment from 959.8316
								M	1.3[¶]	9.90E-02	0.4[¶]	1.00E-03	
959.8316	13.0	?						F	0.3[¶]	8.00E-03	0.06[¶]	1.86E-05	Phospholipid like
								M	0.7[¶]	3.55E-01	0.5[¶]	9.60E-02	
344.2799	15.6	$C_{19}H_{38}NO_4$	344.2801	-0.6	0.4		$C_{19}H_{37}NO_4$	F	1.0[¶]	8.22E-01	0.6[¶]	1.00E-03	Acyl carnitine like
								M	1.4[¶]	1.44E-01	0.9[¶]	5.71E-01	
302.3061	16.7	$C_{18}H_{40}NO_2$	302.3059	0.7	0.7	284.2947	$C_{18}H_{39}NO_2$	F	154.0	1.17E-05	283.7	1.63E-07	
								M	211.8	1.34E-11	441.2	2.16E-11	
540.3669	16.8	?				441.8062, 184.0734		F	6.8[¶]	7.00E-03	11.1[¶]	4.03E-07	
								M	2.0[¶]	2.40E-02	5.1[¶]	8.20E-07	
492.3089	18.5	$C_{24}H_{47}NO_7P$	492.3090	-0.2	0.9	474.2973,184.0739	$C_{24}H_{46}NO_7P$	F	30.3[¶]	1.00E-03	51.5[¶]	3.54E-09	LysoPC like
								M	4.3[¶]	2.00E-03	16.2[¶]	7.43E-12	
228.2688	18.6	$C_{15}H_{34}N$	228.2691	-1.3	1.3			F	162.6	2.95E-05	217.0	1.23E-06	
								M	171.7	2.09E-08	279.6	2.41E-12	

m/z: mass to charge ratio; RT: retention time (min); Δppm: part per million difference between the theoretical and observed mass; i-fit: isotope fit. [#] fragments were obtained from Q-TOFMS or high collisional dissociation fragmentation in full scan mode within ±5ppm. Fold Mean fold change calculated from relative concentrations of the compound in plasma from roach exposed to either 50% or 100% effluent for 28 days compared with the levels in control roach held in river water (n=16-20 fish for each gender). [¶]Concentrations above the limit of detection were detected in plasma of the control roach. Lyso PC: lysophosphatidylcholine; F: female; M: male. Non-normally distributed data (determined as such by the Kolmogorov-Smirnov test) were log transformed prior to statistical analysis; *p*-value was calculated from *t*-test between control and the effluent exposure samples, and values below the BH threshold are statistically significant with a false discovery rate of <5% (Benjamini and Hochberg (BH) threshold was 4.0×10^{-02}).

Appendix 5.6a: (continued) Unidentified markers of effluent exposure in plasma of roach in +ESI mode.

Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula	Gender	Control vs 50%effluent		Control vs 100%effluent		Putative identity
									Fold change	*p*-value	Fold change	*p*-value	
328.3215	18.9	$C_{20}H_{42}NO_2$	328.3216	-0.3	0.9		$C_{20}H_{41}NO_2$	F	21.9	7.00E-03	74.2	2.83E-06	
								M	17.0	9.00E-03	97.8	7.52E-09	
242.2850	20.0	$C_{16}H_{36}N$	242.2848	0.8	0.3			F	15.8[¶]	2.59E-07	15.5[¶]	3.54E-09	
								M	22.5[¶]	7.43E-12	35.4[¶]	7.43E-12	
641.2710	20.5	?						F	1.6[¶]	9.00E-03	1.9[¶]	9.87E-06	
								M	1.3[¶]	3.70E-02	1.9[¶]	2.05E-05	
313.2369	20.6	?						F	34.4[¶]	4.01E-06	117.9[¶]	3.54E-09	
								M	6.0[¶]	2.29E-04	33.7[¶]	7.43E-12	
330.3373	20.7	$C_{20}H_{44}NO_2$	330.3372	0.3	0.6		$C_{20}H_{43}NO_2$	F	163.2	3.63E-07	210.3	1.25E-08	
								M	177.2	3.15E-11	256.6	1.73E-12	
374.3637	20.8	$C_{22}H_{48}NO_3$	374.3634	0.8	0.1		$C_{22}H_{47}NO_3$	F	65.8	4.70E-06	93.2	7.83E-07	
								M	68.4	1.06E-08	106.9	4.08E-10	
356.3530	21.7	$C_{22}H_{46}NO_2$	356.3529	0.3	0.3	338.3415	$C_{22}H_{45}NO_2$	F	112.1	1.29E-07	137.8	3.54E-09	
								M	84.3	3.72E-12	159.9	3.72E-12	
627.2927	21.8	?						F	3.1[¶]	8.00E-02	6.4[¶]	1.18E-05	Fragment from 683.2527
								M	1.7[¶]	1.56E-01	3.0[¶]	2.00E-03	
667.2849	21.8	?						F	1.7[¶]	1.00E-02	2.8[¶]	1.80E-08	Fragment from 683.2527
								M	1.1[¶]	4.81E-01	1.7[¶]	1.53E-04	
683.2527	21.8	?						F	1.7[¶]	9.00E-03	2.8[¶]	1.05E-09	
								M	1.1[¶]	4.44E-01	1.6[¶]	1.00E-03	
284.3322	20.7	$C_{19}H_{42}N$	284.3317	1.8	1.2			F	3.7[¶]	9.10E-09	3.7[¶]	4.30E-12	Possibly a fragment
								M	4.8[¶]	1.71E-11	5.5[¶]	7.32E-16	
468.3089	21.1	$C_{22}H_{47}NO_7P$	468.3090	-0.2	0.5	184	$C_{22}H_{46}NO_7P$	F	1.3[¶]	1.20E-02	1.7[¶]	1.05E-04	Lyso PC like
								M	1.1[¶]	3.59E-01	1.2[¶]	1.75E-01	
295.1671	21.2	$C_{18}H_{24}O_2Na$	295.1674	-1.0	0.4		$C_{18}H_{24}O_2$	F	22.0	6.30E-02	244.4	4.24E-08	
								M	17.4	1.07E-01	284.3	3.72E-12	
618.3542	21.2	?						F	1.4[¶]	4.70E-02	2.5[¶]	7.48E-05	
								M	0.9[¶]	5.97E-01	0.8[¶]	5.61E-01	
493.2345	21.4	?						F	1.9[¶]	7.65E-01	16.0[¶]	7.07E-09	
								M	3.2[¶]	2.50E-02	10.8[¶]	1.58E-09	
494.3322	21.5	?						F	1.5[¶]	2.00E-03	1.7[¶]	8.28E-06	
								M	1.3[¶]	1.10E-02	1.5[¶]	1.00E-03	
405.2400	21.5	?						F	ND	1.00E+00	148.1	2.83E-06	
								M	4.5[¶]	3.56E-01	61.9[¶]	1.96E-06	
341.3057	21.5	$C_{21}H_{41}O_3$	341.3056	0.3	1.0		$C_{21}H_{40}O_3$	F	1.4[¶]	5.00E-03	1.7[¶]	2.25E-04	
								M	1.5[¶]	3.70E-02	2.0[¶]	4.29E-04	

Appendix 5.6a: (continued) Unidentified markers of effluent exposure in plasma of roach in +ESI mode.

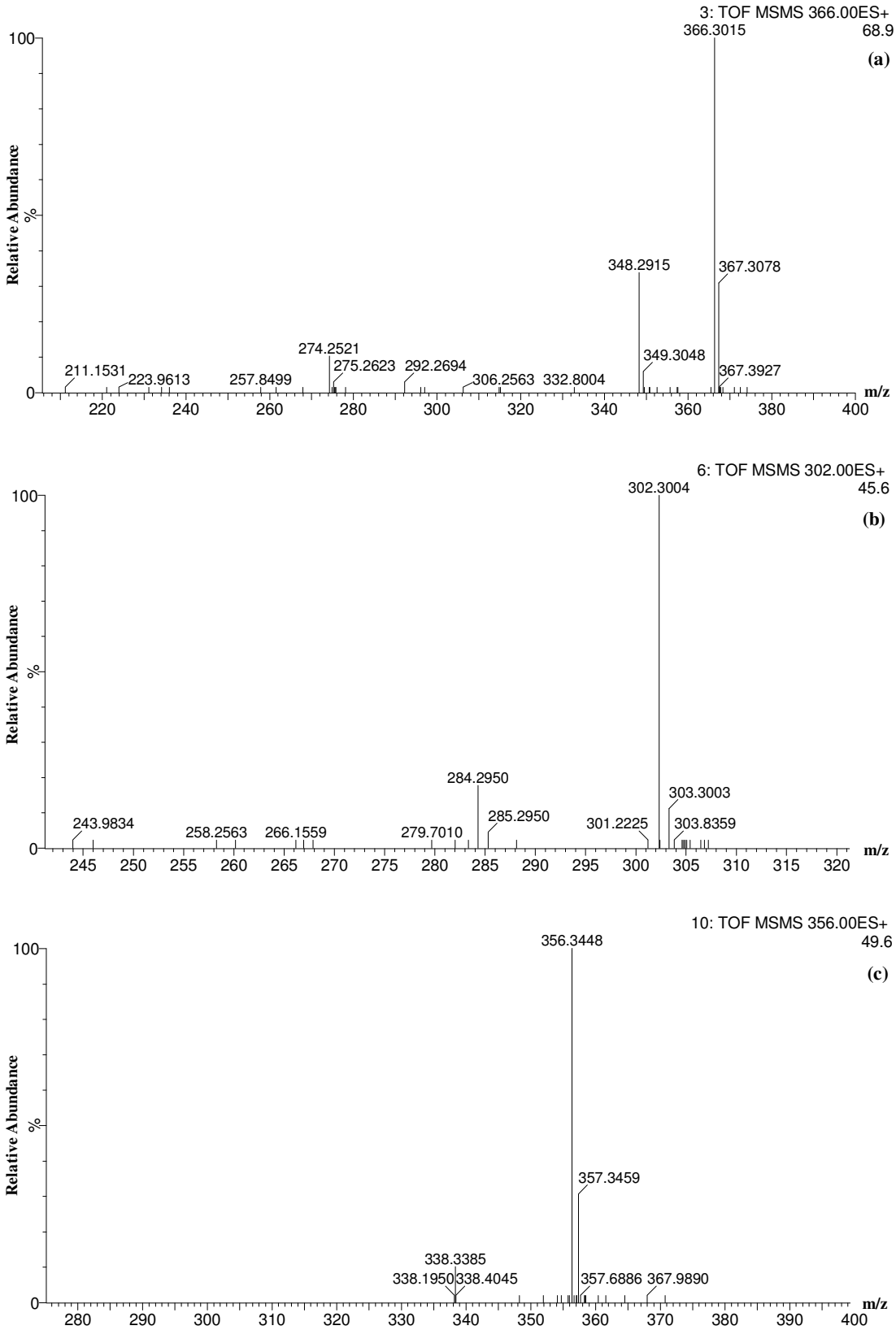
Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula	Gender	Control vs 50%effluent		Control vs 100%effluent		Putative identity
									Fold change	*p*-value	Fold change	*p*-value	
331.2216	21.7	?						F	19.5[¶]	2.59E-07	44.1[¶]	3.54E-09	
								M	2.1[¶]	1.70E-02	7.1[¶]	2.23E-10	
367.3266	22.0	?						F	1.6[¶]	1.30E-02	1.9[¶]	1.95E-04	
								M	1.1[¶]	2.85E-01	1.2[¶]	1.83E-01	
345.2047	21.9	?						F	3.9[¶]	2.59E-07	11.1[¶]	3.54E-09	
								M	3.2[¶]	1.30E-07	10.3[¶]	1.49E-11	
270.2796	22.2	$C_{17}H_{36}NO$	270.2797	-0.4	0.7			F	43.6	3.00E-03	245.9	1.10E-06	Possibly fragment
								M	35.6	3.00E-03	247.1	6.28E-09	
247.2059	22.3	$C_{17}H_{27}O$	247.2062	-1.2	0.5			F	24.9[¶]	1.29E-07	74.9[¶]	3.54E-09	Possibly fragment
								M	27.0	7.43E-12	113.0[¶]	7.43E-12	
355.2845	22.4	$C_{21}H_{39}O_4$	355.2848	-0.8	0.4		$C_{21}H_{38}O_4$	F	3.0[¶]	6.08E-06	6.8[¶]	2.04E-07	
								M	2.5[¶]	3.58E-04	8.2[¶]	1.02E-08	
519.2508	22.4	?						F	2.7[¶]	5.00E-01	26.0[¶]	2.48E-08	
								M	3.1[¶]	1.90E-02	11.6[¶]	6.99E-09	
235.2064	22.4	$C_{16}H_{27}O$	235.2062	0.9	0.4			F	5.8[¶]	2.59E-07	18.0[¶]	3.54E-09	Possibly fragment
								M	8.6[¶]	7.43E-12	41.0[¶]	7.43E-12	
369.2647	22.4	$C_{21}H_{37}O_5$	369.2641	1.6	1.7		$C_{21}H_{36}O_5$	F	2.9[¶]	1.17E-05	4.9[¶]	7.07E-09	
								M	2.0[¶]	2.28E-04	3.8[¶]	1.41E-10	
399.3112	22.3	$C_{23}H_{43}O_5$	399.3110	0.5	1.0		$C_{23}H_{42}O_5$	F	1.6[¶]	2.94E-04	2.9[¶]	6.89E-06	
								M	1.5[¶]	6.00E-03	3.3[¶]	1.28E-07	
259.2065	22.3	$C_{18}H_{27}O$	259.2062	1.2	0.6			F	26.6[¶]	2.59E-07	84.8[¶]	3.54E-09	Possibly fragment
								M	43.1[¶]	7.43E-12	159.0[¶]	7.43E-12	
219.1751	22.4	$C_{15}H_{23}O$	219.1749	0.9	1.3			F	1.6[¶]	2.70E-02	3.2[¶]	2.03E-04	Possibly fragment
								M	1.9[¶]	6.00E-03	1.3[¶]	2.25E-01	
251.0484	22.7	?						F	2.7[¶]	8.65E-05	4.5[¶]	6.08E-11	
								M	1.4[¶]	5.60E-02	3.3[¶]	2.96E-09	
385.1567	22.7	?						F	5.3[¶]	4.00E-03	11.0[¶]	6.20E-07	
								M	2.2[¶]	2.34E-01	12.5[¶]	5.29E-08	
407.2586	23.0	$C_{27}H_{35}O_3$	407.2586	0.0	3.7			F	2.4[¶]	5.92E-01	15.1[¶]	1.79E-07	
								M	4.9[¶]	5.00E-03	20.8[¶]	8.87E-09	
413.3263	23.0	$C_{24}H_{45}O_5$	413.3267	-1.0	0.7			F	1.6[¶]	1.00E-03	2.8[¶]	6.13E-07	
								M	1.5[¶]	1.00E-03	2.8[¶]	6.30E-06	
297.2793	23.0	$C_{19}H_{37}O_2$	297.2794	-0.3	0.2			F	11.0[¶]	4.87E-05	21.6[¶]	1.13E-07	
								M	4.6[¶]	2.64E-05	8.6[¶]	1.29E-08	
353.2689	23.4	$C_{21}H_{37}O_4$	353.2692	-0.8	7.4			F	38.7	2.20E-02	152.6	4.66E-05	
								M	7.0[¶]	2.70E-02	15.0[¶]	8.00E-03	

Appendix 5.6a: (continued) Unidentified markers of effluent exposure in plasma of roach in +ESI mode.

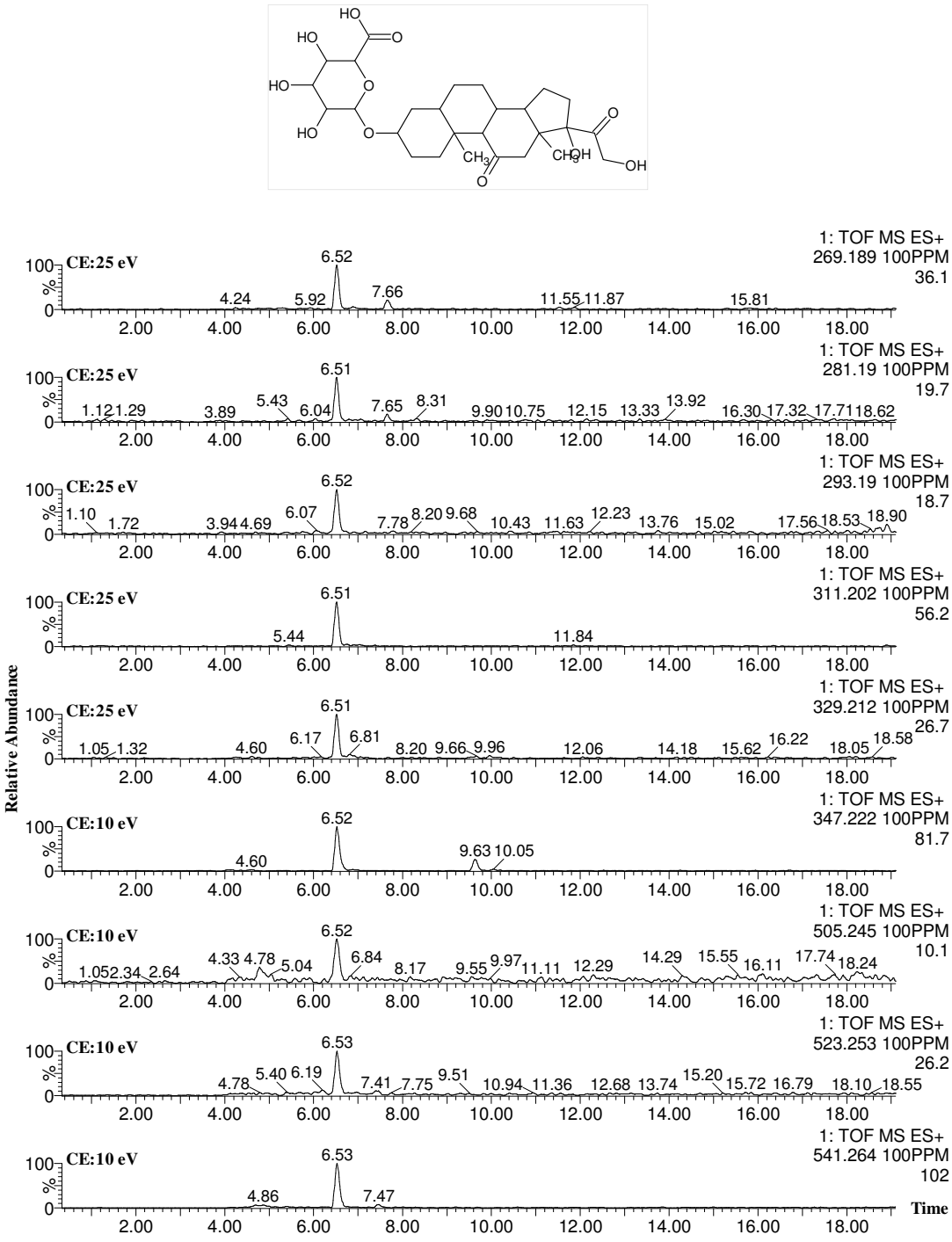
Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula	Gender	Control vs 50% effluent		Control vs 100% effluent		Putative identity
									Fold change	*p*-value	Fold change	*p*-value	
383.3161	23.4	?						F	1.9[¶]	5.42E-05	3.1[¶]	8.33E-07	
								M	1.5[¶]	4.00E-03	3.4[¶]	2.76E-07	
782.5387	23.7	?						F	31.3[¶]	4.87E-05	50.0[¶]	1.61E-07	
								M	3.7[¶]	1.68E-07	5.0[¶]	5.39E-08	
866.5905	23.8	?						F	2.7[¶]	6.50E-02	12.0[¶]	4.30E-06	
								M	1.9[¶]	2.90E-02	6.5[¶]	6.48E-09	
954.6312	24.0	?						F	2.1[¶]	2.20E-02	2.8[¶]	4.24E-08	
								M	1.1[¶]	7.57E-01	2.0[¶]	1.00E-03	
702.5140	24.4	?						F	1.4[¶]	1.80E-02	2.4[¶]	1.52E-07	
								M	1.2[¶]	4.10E-02	1.7[¶]	5.89E-06	

Appendix 5.6b: Unidentified markers of effluent exposure in plasma of roach in -ESI mode.

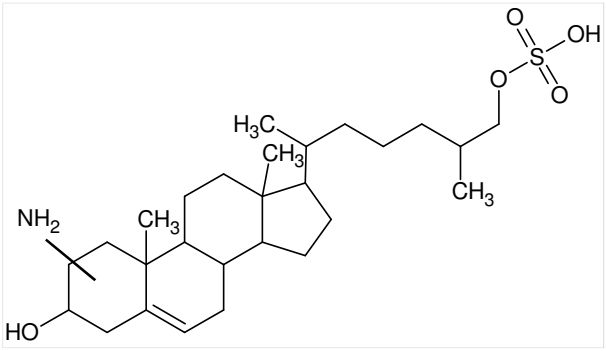
Observed ion (*m/z*)	RT (min)	Putative formula	Theoretical mass of ion	Δ ppm	i-fit	Q-TOFMS fragments[#]	Parent compound formula	Gender	Control vs 50% effluent		Control vs 100% effluent		Putative identity
									Fold change	*p*-value	Fold change	*p*-value	
634.3027	10.0	?						F	1.8[¶]	1.37E-01	3.1[¶]	1.65E-04	
								M	0.7[¶]	1.78E-01	1.8[¶]	3.00E-03	
643.2881	20.5	?						F	1.6[¶]	1.10E-02	3.2[¶]	6.51E-07	
								M	1.5[¶]	5.20E-01	1.5[¶]	2.10E-02	
658.3326	20.8	?						F	1.2[¶]	1.75E-01	1.7[¶]	3.85E-05	Formate adduct
								M	1.1[¶]	8.94E-01	1.1[¶]	5.67E-01	
999.6437	21.0	?						F	1.4[¶]	1.00E-03	2.1[¶]	9.84E-09	Possibly cluster
								M	1.2[¶]	3.25E-01	1.2[¶]	5.40E-02	
621.3051	21.2	?						F	1.2[¶]	5.78E-01	3.4[¶]	2.79E-06	
								M	1.4[¶]	3.97E-01	1.4[¶]	1.35E-01	
987.6355	21.2	?						F	1.8[¶]	1.00E-03	2.6[¶]	2.16E-07	
								M	1.4[¶]	9.00E-03	1.4[¶]	2.00E-03	
466.2929	21.3	$C_{22}H_{45}NO_7P$	466.2934	-1.1	0.0		$C_{22}H_{46}NO_7P$	F	0.9[¶]	4.00E-01	1.2[¶]	6.60E-02	Lyso PC like
								M	1.2[¶]	9.00E-03	1.2[¶]	4.40E-02	
640.3613	21.4	?						F	1.2[¶]	1.00E-01	2.2[¶]	1.93E-05	
								M	0.8[¶]	3.82E-01	0.8[¶]	3.39E-01	
616.3633	21.5	?						F	1.6[¶]	1.00E-03	1.9[¶]	9.17E-06	
								M	1.0[¶]	3.24E-01	1.0[¶]	8.50E-01	
665.2913	21.5	?						F	1.4[¶]	5.00E-03	1.6[¶]	7.58E-05	
								M	1.4[¶]	1.25E-01	1.4[¶]	2.05E-04	
469.2355	21.5	?				301.2174,257.2290,167.0154		F	6.6[¶]	2.69E-01	30.3[¶]	1.71E-09	
								M	2.8[¶]	1.20E-02	9.1[¶]	9.64E-11	
323.1683	21.6	?						F	36.8	6.14E-06	163.0	1.76E-04	
								M	19.8	1.63E-04	112.0	1.33E-16	
277.1805	21.7	$C_{17}H_{25}O_3$	277.1804	0.4	0.7	219.1782	$C_{17}H_{26}O_3$	F	16.2[¶]	6.57E-08	62.9[¶]	1.71E-09	Fatty alcohol like
								M	14.5[¶]	7.43E-12	64.2[¶]	9.94E-13	
845.4563	22.2	?						F	0.5[¶]	3.00E-03	0.3[¶]	1.21E-04	
								M	0.6[¶]	2.60E-01	0.6[¶]	2.52E-01	
337.187	22.5	?						F	93.1	3.29E-08	708.0	1.71E-09	
								M	62.5	8.18E-11	188.0	9.94E-13	
960.5995	24.1	?						F	0.6[¶]	1.60E-02	0.4[¶]	2.81E-06	
								M	1.2[¶]	7.30E-02	1.2[¶]	3.30E-02	
716.5263	25.4	?						F	2.3[¶]	2.41E-01	5.2[¶]	5.24E-05	
								M	0.6[¶]	3.60E-02	0.6[¶]	2.24E-01	
742.5351	25.4	?						F	1.5[¶]	5.80E-02	4.7[¶]	1.00E-03	
								M	0.8[¶]	5.70E-02	0.8[¶]	1.97E-01	
986.8198	27.7	?						F	22.5[¶]	1.30E-02	29.4[¶]	1.23E-07	
								M	3.3[¶]	2.81E-06	3.3[¶]	2.51E-08	

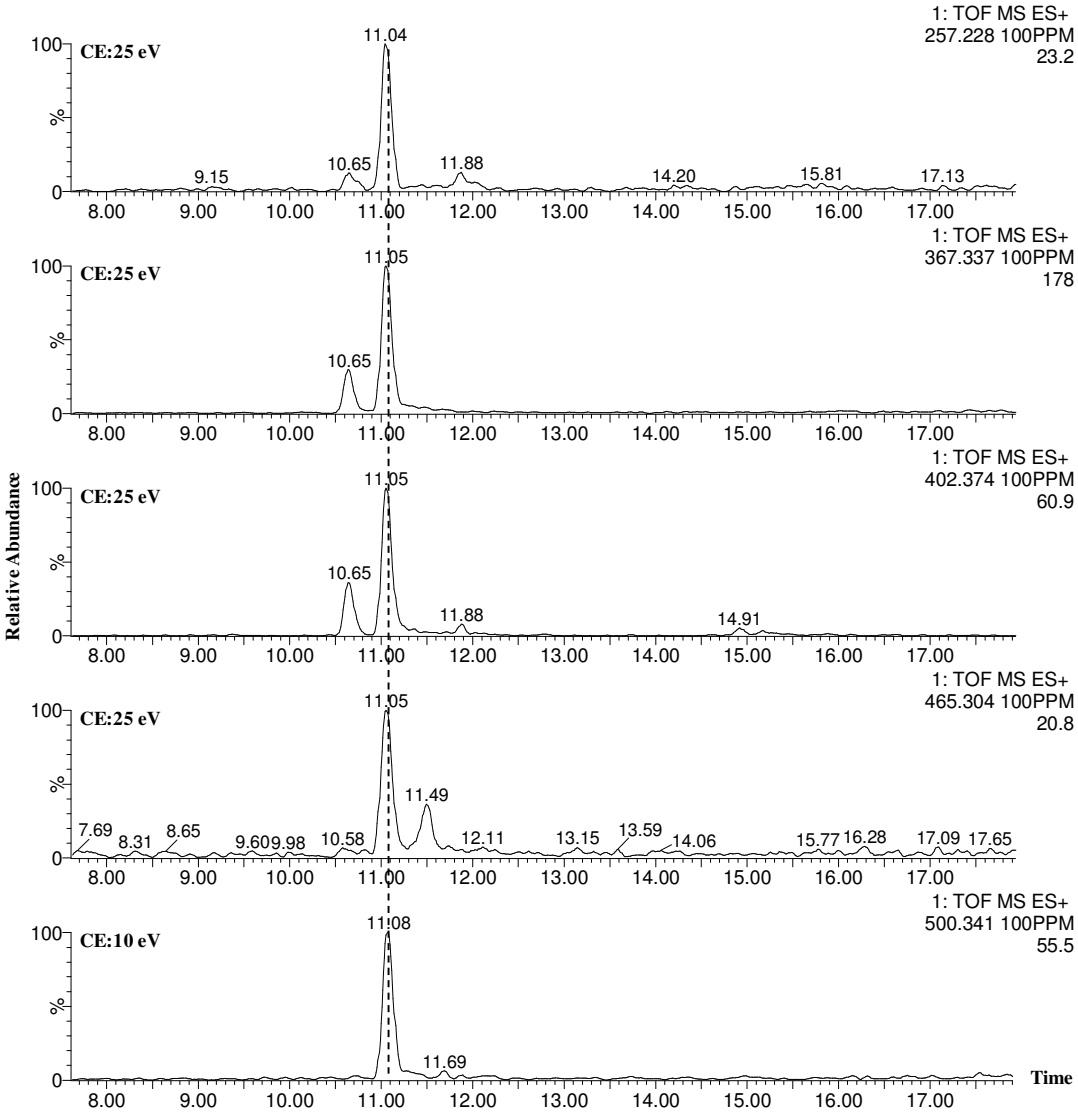


Appendix 5.7: Q-TOFMS mass spectra of three unidentified xenobiotics found in plasma from effluent-exposed roach at collision energy of 20eV.



Appendix 5.8: Ion chromatograms of the in-source fragmented of the putative glucuronide conjugate of tetrahydrocortisone in plasma from effluent-exposed roach in +ESI mode (collision energy = 10 eV and 25 eV).





Appendix 5.9: Ion chromatograms of the in-source fragmented of a bile acid sulphate in plasma from effluent-exposed roach in +ESI mode (collision energy = 10 eV and 25 eV).