UNIVERSITAT DE BARCELONA

Assessment of the lipidomic effects of environmental pollutants on exposed organisms using chemometric and analytical methods

Eva Gorrochategui Matas

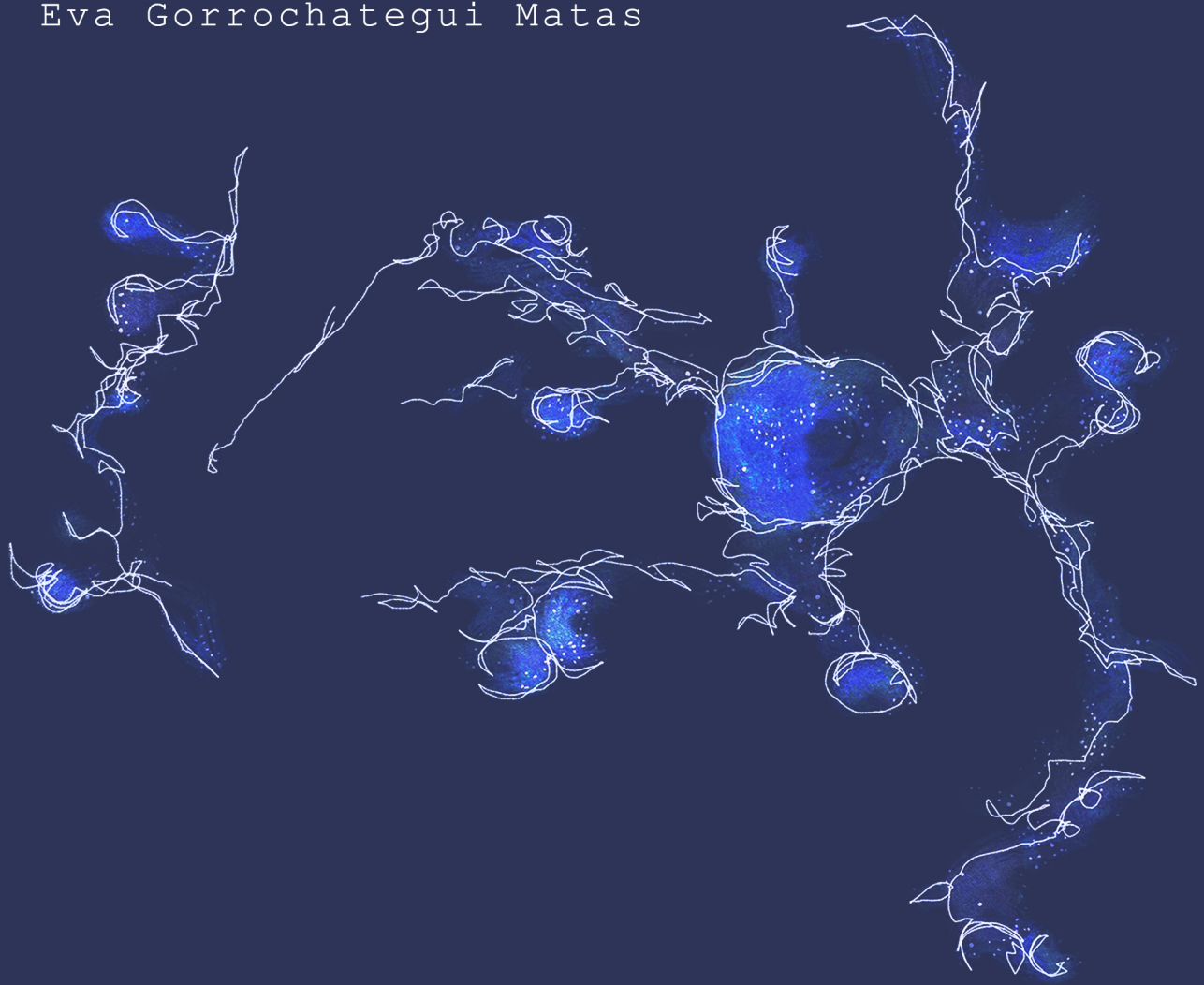
UNIVERSITAT DE BARCELONA

Assessment of the lipidomic effects of environmental pollutants on exposed organisms using chemometric and analytical methods

Eva Gorrochategui Matas

ASSESSMENT OF THE LIPIDOMIC EFFECTS OF ENVIRONMENTAL POLLUTANTS ON EXPOSED ORGANISMS USING CHEMOMETRIC AND ANALYTICAL METHODS

Eva Gorrochategui Matas



UNIVERSITAT DE BARCELONA

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Doctoral programme: “Química Analítica i Medi Ambient (HDK15)”

ASSESSMENT OF THE LIPIDOMIC EFFECTS OF ENVIRONMENTAL POLLUTANTS ON EXPOSED ORGANISMS USING CHEMOMETRIC AND ANALYTICAL METHODS

A Thesis submitted for the degree of

Doctor in Analytical Chemistry by:

Eva Gorrochategui Matas

Supervisors:

Dr. Romà Tauler Ferré
Department of Environmental Chemistry
Institute of Environmental Assessment and Water Research (IDAEA)
Council for Scientific Research (CSIC)

Dr. Sílvia Lacorte Bruguera
Department of Environmental Chemistry
Institute of Environmental Assessment and Water Research (IDAEA)
Council for Scientific Research (CSIC)

Tutor:

Dr. Mª Teresa Galceran i Huguet
Department of Analytical Chemistry University of Barcelona (UB)

Barcelona, March of 2018

Dr. Romà Tauler Ferré, Professor of the Department of Environmental Chemistry of the Institute of Environmental Assessment and Water Research, and **Dr. Sílvia Lacorte Bruguera**, Professor of the same Department,

STATE THAT:

the current PhD report entitled "*Assessment of the lipidomic effects of environmental pollutants on exposed organisms using chemometric and analytical methods*" has been elaborated under our supervision by Ms. **Eva Gorrochategui Matas** in the Department of Environmental Chemistry of the Institute of Environmental Assessment and Water Research, and also that all the results presented in this manuscript are consequence of the research work of the hereby mentioned doctoral student.

And in order to make it certain, we sign the current certificate.

Barcelona, March of 2018

Dr. Romà Tauler Ferré **Dr. Sílvia Lacorte Bruguera**

Als meus pares i l'Anna,

"Aldapeko
sagarraren
adarraren puntan
puntaren puntan
txoria zegoen kantari.

Xiru-liruli, xiru-liruri,
nork dantzatuko ote du
soinutxo hori?"

Oskorri

AGRAÏMENTS

En el moment d'escriure els agraïments, fer la vista enrere i recordar tots aquells amb qui he tingut el plaer de viure aquests anys de tesi, he quedat sorpresa del número de persones que en algun moment o altre m'han ofert un cop de mà, el seu recolzament o una petita (però molt necessària) empenta per a seguir endavant. Això fa que em senti molt afortunada, agraïda i feliç.

Voldria començar agraint als meus directors de tesi, el Dr. Romà Tauler Ferré i la Dra. Sílvia Lacorte Bruguera, no només l'oportunitat de realitzar aquesta tesi, sinó també els coneixements compartits, la guia, el suport i la confiança en les meves capacitats. També voldria agrair a la meva tutora de tesi, la Dra. Maria Teresa Galceran i Huguet la seva ajuda prestada per a la realització d'aquest manuscrit.

A tots els companys del grup CHEMAGEB, amb qui he compartit tants moments aquests últims quatre anys, que s'han convertit en la meva segona família. Començant per la Carma (la "mare" d'aquesta "família"), el Joaquim (el "cap pensant"), el Cristian (l'alegria del grup), l'Stefan (la tranquil·litat en persona), el Marc (l'encarregat de posar-nos al dia de tot el que passa), la Miriam (a qui deixem el nostre lloc!), continuant amb tots aquells companys que ara ja no hi són però que han format part d'aquest grup (Amrita, Mireia, Marta, Yahia, Igor...) i acabant amb els meus companys de batalla, els que van començar amb mi aquesta etapa i que també l'acaben ara. A ells en particular els hi vull fer un especial agraïment. Meri, Francesc, Núria, Elba, Elena i Víctor, més que uns companys de tesi, puc dir que de vosaltres m'emporto uns amics. Gràcies per tots els moments de confidències, escoltant-nos els uns als altres, de les pauses del cafè, també pels sopars, cervesetes, viatges i per les injeccions de positivisme que ens hem fet en moments complicats. En aquests quatre anys que hem compartit, hem après molt els uns dels altres; en definitiva, hem crescut junts. Gràcies.

També voldria fer un especial agraïment a la Dra. Cinta Porte Visa, per la seva ajuda i guia que tan útils m'han estat en el món dels cultius cel·lulars i de la toxicologia i també a la recent Dra. Elisabet Pérez-Albaladejo, qui em va ajudar per primer cop a treballar amb línies cel·lulars. També vull agrair a la Dra. Josefina Casas Brugulat l'ajut que m'ha prestat en el camp de la lipidòmica.

I would also like to acknowledge Dr. Francis L. Martin for his guidance and support during my research stay in Lancaster University. It was a pleasure to work in your group and to learn from your experience and from your knowledge.

No podia oblidar d'incloure en aquests agraïments a totes aquelles persones que ocupen un lloc molt important en la meva vida. Les "sarandongues", Mariona, Anais, Elisabet, Laura i Anna, que sempre heu estat amb mi, des de ben petita, fent-me costat. El Manel, qui ha compartit amb mi molts moments, ha estat el meu suport incondicional i la persona que m'ha fet tirar endavant, a qui li he d'agrair tantíssimes coses, entre d'altres, que pugui estar escrivint això en aquests moments. També vull agrair el suport d'altres persones, algunes a qui fa poc temps que conec, però que també han estat amb mi en aquesta etapa de tesi: Sara, Alec, tabalers, la gran família bluesera que he conegut aquest darrer any i tots aquells que també esteu en la meva ment tot i que aquí no us mencioni perquè no acabaria mai.

Finalment, i no per això menys important, voldria fer un enorme agraïment a la meva família per tota l'ajuda i el recolzament que m'han donat, no només en aquests quatre anys de tesi, sinó durant tota la meva vida. Als meus pares, que han estat i seran un exemple a seguir per a mi en tots els aspectes, a qui respecto i estimo moltíssim. A la meva germana Anna, la meva altra meitat, la meva millor amiga i confident, qui m'ha donat suport i m'ha ajudat en tot moment. També a les meves tietes, que han fet de "segones mares" i sempre m'han donat un cop de mà, a tots els altres tiets/es i cosins/es, i als meus avis i àvies, tant els que encara hi són com els que ja no, però que sempre m'acompanyen.

Gràcies.

CONTENTS

ABSTRACT

Lipidomics is a subset of metabolomics, which in turn, is one of the categorical platforms that constitute omics. Omic sciences aim at the study of the abundance and (or) structural characterization of a broad range of molecules (*e.g.*, lipids in the case of lipidomics) in organisms under distinct scenarios. In the environmental field, omic studies aim at the evaluation of the alterations that organisms might suffer after exposure to environmental stressors such as chemical pollutants, leading to *exposomics*. Thus, lipidomic data can be used to acquire fundamental understanding of the interaction between organisms and external environmental stimuli. Distinct analytical techniques can be used to obtain lipidomic data. Among them, liquid chromatography coupled to mass spectrometry (LC-MS) is one of the most powerful technologies due to its ability in the analysis of low molecular weight compounds in biological systems. However, LC-MS data sets are challenging to analyse because of their very large size and complexity and for this reason chemometric methods are proposed to reveal the information contained in LC-MS lipidomic (and metabolomic) data sets as much as possible.

This Thesis has dealt, on the one hand, with the development of an untargeted LC-MS data analysis strategy based on the use of powerful chemometric tools, with the ultimate goal of facilitating lipidomic studies and demonstrating the usefulness of *Chemometrics* in this field. The developed strategy covers most of the steps involved in the data analysis process: data storage and conversion, import, compression, normalization, scaling and transformation, peak resolution and biomarker detection. However, the main contributions of the present Thesis are related with the steps of data compression and resolution. Distinct data compression strategies have been compared such as the classical *binning* procedure or the time/mass *windowing* and a more recent strategy based on the search of the *regions of interest* or ROI, being the latter, the one that has provided better results. Concerning data resolution, the performance of Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) has been evaluated showing excellent results with LC-MS lipidomic data sets. Also, different methods for variable selection (biomarker screening) have been compared including the

classical ANOVA statistical test followed by a multiple comparisons test and a chemometric method: Partial Least Squares-Discriminant Analysis (PLS-DA), using the variables importance in projection (VIP). Moreover, the study of the contribution of the factors underlying multifactorial experimental designs and the evaluation of the sources of variance has been performed by ANOVA-simultaneous component analysis (ASCA). A step-by-step protocol of the developed LC-MS data analysis strategy and a description of the corresponding methodology have been provided so that researchers in lipidomics (and metabolomics) can use them to analyse their own data without the requirement of external data analysis software.

On the other hand, this Thesis has aimed to generate lipidomic data from LC-MS techniques and from attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy and Surface Enhanced Raman Spectroscopy (SERS) combined with chemometric methods. The obtained lipidomic data have been used to provide knowledge of the effects of some widespread environmental pollutants on human and environmental model-biosystems. The selected environmental pollutants in this Thesis include some perfluoroalkylated substances (PFASs), tributyltin (TBT), different forms of carbon-based nanoparticles (CBNs) and some proautophagic drugs. Moreover, the biological systems that have been exposed to these environmental stressors comprise three cell lines (a *human* placental chroriocarcinoma cell line (JEG-3), a human glioblastoma cell line (T98G) and a *Xenopus laevis* kidney epithelial cell line (A6)) and a model organism (zebrafish, *Danio rerio*). In order to extract more information of the effects of the chemicals on the exposed biological systems, some toxicological assays have also been performed. Overall, this Thesis pretends to provide a useful contribution to the untargeted lipidomic research.

RESUM

La lipidòmica és una subdisciplina de la metabolòmica; aquesta última considerada una de les branques fonamentals de l'òmica. Les ciències òmiques tenen com a objectiu principal l'estudi de l'abundància i (o) de la caracterització estructural d'un gran grup de molècules (entre elles els lípids en el cas de la lipidòmica) en organismes que es troben sota diversos escenaris. En el camp mediambiental, els estudis òmics tenen la finalitat d'avaluar les alteracions que els organismes poden patir com a conseqüència de l'exposició a factors presents en el medi ambient, com poden ser els productes químics, donant lloc a l'*exposòmica*. Per tant, les dades obtingudes en estudis lipidòmics poden ser utilitzades per a obtenir un coneixement bàsic sobre els fenòmens que tenen lloc en la interacció entre els organismes i el seu entorn. Existeixen diverses tècniques analítiques per a l'obtenció d'aquestes dades. Entre totes elles, la cromatografia de líquids acoblada a l'espectrometria de masses (LC-MS) s'ha convertit en una de les tècniques estrella en estudis lipidòmics gràcies a la seva capacitat d'anàlisi de compostos de baix pes molecular en sistemes biològics. No obstant, les dades obtingudes amb aquesta tècnica son molt difícils d'analitzar degut a les seves grans dimensions i a la seva complexa naturalesa. Per aquest motiu, l'ús de mètodes quimiomètrics que permetin extreure i interpretar aquestes dades lipidòmiques (i metabolòmiques) resulta necessari.

Aquesta Tesi ha perseguit, per una banda, el desenvolupament d'una metodologia d'anàlisi no dirigida de dades LC-MS fent ús d'eines quimiomètriques, per tal de facilitar els estudis lipidòmics i demostrar la utilitat de la *Quimiometria* en el camp de l'òmica. L'estratègia d'anàlisi de dades que s'ha desenvolupat cobreix la major part de les etapes del procés: emmagatzematge i conversió de les dades, import, compressió, normalització, correcció per escalars, resolució del perfil cromatogràfic i detecció de possibles biomarcadors. Ara bé, la major contribució d'aquesta Tesi està relacionada amb les etapes de compressió i resolució de les dades. Diverses estratègies de compressió de dades s'han comparat en aquesta Tesi, incloent la clàssica estratègia del *binning* o la divisió del cromatograma en petites finestres de temps i/o de massa (*windowing*) i una aproximació més recent basada en la recerca de

regions d'interès (ROI), éssent la última d'aquestes la que ha generat millors resultats. Pel que fa a la resolució del perfil cromatogràfic, en aquesta Tesi s'ha emprat el mètode de resolució multivariant de corbes per mínims quadrats alternats (MCR-ALS) i se n'ha avaluat la seva eficàcia amb aquest tipus de dades, que ha resultat ésser molt bona. A part, diferents mètodes per a la selecció de variables (*screening* de biomarcadors) s'han examinat comparativament, incloent el clàssic test ANOVA seguit d'un test de comparacions múltiples i un mètode quimiomètric: l'anàlisi discriminant per mínims quadrats parcials (PLS-DA), fent ús de les variables importants en la projecció (VIP). Per altra banda, la contribució dels diferents factors lligats al disseny experimental multifactorial així com la contribució de les diverses fonts de variància s'han estudiat mitjançant el mètode d'ANOVA amb l'anàlisi simultani de components (ASCA). En aquesta part de la Tesi s'ha elaborat un protocol on s'inclou una descripció detallada del procés de tractament de dades així com una explicació de les bases d'aquesta metodologia, per tal que altres científics que treballin en el camp de la lipidòmica (i/o de la metabolòmica) puguin tractar les seves dades sense necessitat d'utilitzar un programa extern si així ho desitgen.

D'altra banda, el segon propòsit d'aquesta Tesi ha estat la generació de dades lipidòmiques a partir de tècniques LC-MS i de l'espectroscòpia IR (ATR-FTIR) i Raman (SERS) i l'ús de mètodes quimiomètrics. Les dades lipidòmiques obtingudes han estat utilitzades per a generar nou coneixement sobre els efectes d'alguns dels contaminants ambientals més presents en el mediambient en sistemes model humans i mediambientals. Els contaminants ambientals escollits en aquesta Tesi inclouen alguns compostos perfluorats (PFASs), el tributilestany (TBT), diverses formes de nanopartícules de carboni (CBNs) i alguns compostos proautofàgics. Els sistemes biològics que s'han exposat a aquests contaminants inclouen tres línies cel·lulars (línia de cèl·lules de carcinoma de placenta humana (JEG-3), línia de cèl·lules humanes de tumor cerebral glioblastoma (T98G) i línia de cèl·lules epitelials de ronyó de granota *Xenopus laevis* (A6)) i un organisme model (zebrafish, *Danio rerio*). Aquests estudis s'han acompanyat d'assaigs toxicològics per a obtenir més informació dels efectes dels contaminants en aquests sistemes biològics model. En conjunt, aquesta Tesi pretén aportar un valor afegit a la recerca no-dirigida en lipidòmica.

ABBREVIATIONS AND ACRONYMS

AIF	All ion fragmentation
ANN	Artificial neural networks
APCI	Atmospheric pressure chemical ionization
ASCA	ANOVA-simultaneous component analysis
ATCC	American type culture collection
ATR-FTIR	Attenuated total reflection Fourier-transform infrared
BPA	Bisphenol A
CBNs	Carbon-based nanoparticles
CCX	Celecoxib
CE	Capillary electrophoresis
CEs	Cholesterol esters
Cer	Ceramide
CNT	Carbon nanotube
COX-2	Cyclooxygenase isozyme
C-SWCNT	Carbonyl single-walled carbon nanotubes
CWT	Continuous wavelet transform
DAG	Diacylglycerol
DOE	Design of experiments
EDC	Endocrine disruptive chemicals
ESI	Electrospray ionization
FA	Fatty Acyls
FLIP	FLICE inhibitory protein
GC	Gas chromatography
GC-MS	Gas chromatography coupled to mass spectrometry
GL	Glycerolipids
GP	Glycerophospholipids
GTE	ɣ-tocotrienol
HGP	Human Genome Project
^{1}H-NMR	Proton nuclear magnetic resonance
HRMS	High-resolution mass spectrometry
ILCNC	International Lipid Classification and Nomenclature Committee
IR	Infrared spectroscopy
IT	Ion trap
IV	Intravenous
KEGG	Kyoto encyclopedia of genes and genomes
LATs	Lysophospholipid acyltransferases
LC	Liquid chromatography
LC-ESI-MS/MS	Liquid chromatography- electrospray ionization- tandem mass spectrometry
LC-MS	Liquid chromatography coupled to mass spectrometry
LDA	Linear discriminant analysis
LDL	Low-density lipoprotein
LLE	Liquid-liquid extraction
LMPD	LIPID MAPS Proteome Database
LMSD	LIPID MAPS Structure Database
LRMS	Low-resolution mass spectrometry
LRMS/MS	Low-resolution tandem mass spectrometry
LV	Latent variable
MAG	Monoacylglycerol
MALDI	Matrix-assisted laser desorption/ionization
MCR-ALS	Multivariate curve resolution-alternating least squares
MeOH	Methanol
MOM	Model Organism Metabolomes
MS	Mass spectrometry
MWCNT	Multi-walled carbon nanotube
m/z	Mass-to-charge

NLS	Neutral-loss scanning
NMR	Nuclear magnetic resonance
NP	Nanoparticles
NSAID	Nonsteroidal anti-inflammatory drug
OBI-warp	Ordered bijective interpolated warping
OPLS	Orthogonal partial least squares
Orbitrap	Orbital ion trap
PC	Principal component
PCA	Principal component analysis
PCB	Polychlorinated biphenils
PFASs	Perfluoroalkylated substances
PFBA	Perfluorobutanoate acid
PFBS	Perfluorobutanesulfonate
PFCAs	Perfluorinated carboxylic acids
PFDoA	Perfluorododecanoate acid
PFHxA	Perfluorohexanoate acid
PFHxS	Perfluorohexanesulfonate
PFNA	Perfluorononanoate acid
PFOA	Perfluorooctanoate acid
PFOS	Perfluorooctanesulfonate
PFSAs	Perfluorinated sulfonic acids
PFTOHs	Fluorotelomer alcohols
PIS	Precursor ion scanning
PK	Polyketides
PLS	Partial least squares
PLS-DA	Partial least squares-discriminant analysis
PPARγ	Peroxisome proliferator-activated receptor gamma
PQN	Probabilistic quotient normalization
PR	Prenol lipids
PTFE	Polytetrafluoroethylene
PUFAs	Polyunsaturated fatty acids
PVDF	Polyvinylidene fluoride
PXD	Phenoxodiol
QLIT	Quadrupole/linear ion trap
QqQ	Triple quadrupole
Q-TOF-MS	Hybrid quadrupole time-of-flight mass spectrometers
ROI	Regions of interest
ROS	Reactive oxygen species
RV	Resveratrol
SFC	Supercritical fluid chromatography
SIM	Selected ion monitoring
SL	Saccharolipids
SNR_{Thr}	Signal-to-noise ratio threshold
SOM	Self-organizing maps
SP	Sphingolipids
SRM	Selected reaction monitoring
ST	Sterol lipids
SWCNT	Single-walled carbon nanotube
TAG	Triacylglycerol
TBME	Tert-butyl methyl ether
TBT	Tributyltin
TLC	Thin-layer chromatography
TOF	Time-of-flight
UHPLC	Ultra high-performance liquid chromatography
VAST	Variable stability scaling
VIP	Variable importance in projection
WWTP	Waste water treatment plant
XCMS	Various forms (X) of chromatography mass spectrometry
XIAP	X-linked inhibitor of apoptosis

OBJECTIVES AND STRUCTURE OF THE THESIS

OBJECTIVES

Every year thousands of new synthesized chemicals are released and accumulate in the environment causing unknown effects on living organisms and ecological systems. Living organisms react to environmental pressures in their surroundings and generate a response that can be followed by monitoring their constituent low molecular weight molecules, such as metabolites (metabolomics) or lipids (lipidomics). Recent advances in high-throughput omic analytical techniques, mainly referring to mass spectrometry methods coupled to chromatographic techniques such as liquid chromatography (LC-MS), enable the identification and quantification of many of these molecules in a particular biological system. However, the generated LC-MS data sets are not easy to analyse due to their large size and complex nature, making data processing one of the most challenging aspects in omic studies. Chemometric and multivariate data analysis methods are powerful tools to facilitate many of the steps involved in analysing the data generated in LC-MS lipidomic (and metabolomic) studies. Considering this challenging situation, the two main objectives of this Thesis were:

- To develop an LC-MS data analysis strategy for untargeted metabolomics/lipidomics based on chemometric methods, which covers most of the steps involved in the process and provides an alternative approach to the already existing data analysis platforms and software.

- To study the effects of some environmental pollutants on the lipids of exposed model biological systems by using analytical techniques and chemometric analysis.

These two general objectives can be broken down into the following specific objectives:

Chemometric objectives:

- To study distinct ways of data conversion and import for the most important LC-MS vendors (*i.e.*, Agilent, AB Sciex, Thermo Fischer, Waters and Bruker). To evaluate distinct data compression strategies for LC-MS omic data sets in terms of degree of compression achievable and loss of mass spectral resolution associated. To study distinct ways of data normalization to correct for sample size, analytical noise and experimental bias. To evaluate the performance of Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) on chromatographic peak resolution and to comparatively analyse the performance of other feature detection methodologies respect to MCR-ALS in terms of peak shaping and time alignment requirements. To comparatively examine the operation of distinct post-processing chemometric methods intended to detect potential biomarkers such as classical statistical test ANOVA, multiple comparisons test and partial least squares-discriminant analysis (PLS-DA) with the corresponding variables important in projection (VIPs). To evaluate the effects of the distinct factors underlying the experimental design by ANOVA-simultaneous component analysis (ASCA).

Analytical objectives:

- To develop an UHPLC-MS method for targeted lipid profiling in human placental choriocarcinoma (JEG-3) cells and to comparatively evaluate the performance of this method when using a TOF instrument respect to an Orbitrap. To optimize lipid extraction conditions and to elaborate a home-made database containing the analysed lipid species, as a preliminary step to set the basis for further untargeted lipidomic evaluation.
- To evaluate the capacity of attenuated total reflection Fourier-transform infrared (ATR-FTIR) and Raman spectroscopies to perform analyses of biological systems in lipidomic studies.

Omic and toxicological objectives:

- To study the cytotoxic effects of perfluoroalkylated substances (PFASs) and their capacity to inhibit aromatase activity on human placental choriocarcinoma JEG-3 cells. To evaluate the capacity of cells to uptake the chemicals at low doses and to define the non-lethal concentrations that can be used in omic experiments. To study the capacity of a mixture of PFASs and of TBT to disrupt the lipidome of JEG-3 cells. To study the effects of five inhibitors of *Des1* (*i.e.*, celecoxib, phenoxodiol, resveratrol, γ-tocotrienol and XM462) on the lipids of human glioblastoma T98G cells.
- To study the effects of PFASs on the metabolites and lipids of *Xenopus laevis* A6 kidney epithelial cells by ATR-FTIR spectroscopy, microscopy and chemometric analysis and their effects on A6 cell population. To study the effects of four different forms of carbon-based nanoparticles (CBNs) on the lipids of three distinct tissues (*i.e.*, gonads, intestinal tracts and brain) of zebrafish (*Danio rerio*) and their capacity to produce DNA methylation.

STRUCTURE

This Thesis is presented in three major sections. The first introductory section includes an overview of omic sciences and their classification, of the analytical techniques and data analysis strategies that can be used in lipidomic studies and of toxicology for environmental risk assessment. Moreover, the model biological systems and the environmental pollutants studied in this Thesis are presented. In the second section, the scientific articles together with their background, results discussion and specific conclusions are presented. Finally, the general conclusions and references are included in the third section of the Thesis. The contents of these major sections are in turn organized in five parts that are described below.

Chapter 1 *Introduction*

Background and state-of-the-art of omic sciences and analytical instrumentation and data analysis procedures (both targeted and untargeted) used in lipidomic (and metabolomic) studies are reviewed. Also, some toxicological approaches for environmental risk assessment are presented.

The model biological systems and the environmental pollutants studied in this Thesis are then introduced. Finally, the chemometric methods developed in the omics field and applied in this Thesis are described.

Chapter 2 *Novel data analysis approaches for metabolomics/lipidomics*

At the beginning of this chapter, a comparative evaluation of different state-of-the-art data analysis strategies for omic studies is presented with the scientific review entitled "*Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow*", published in the journal Trends in Analytical Chemistry, 82 (2016), 425-442. In the same scientific publication, a proposal of a novel LC-MS data analysis strategy for untargeted metabolomics/lipidomics based on the use of chemometric tools is presented and also discussed in this chapter. This strategy is based on the search of regions of interest (ROI) to compress the data and the use of MCR-ALS to resolve the chromatographic profiles. A detailed step-by-step protocol of this data analysis strategy entitled "*A protocol for LC-MS metabolomic data processing using chemometric tools*" was published in the online journal Nature Protocol Exchange (2015), doi:10.1038/protex.2015.102 and it is also included in this chapter. In addition to these two publications, a scientific article containing a more detailed description of the proposed data analysis methodology entitled "*ROIMCR: a powerful data analysis strategy for LC-MS based metabolomic data sets*", submitted recently for publication is also included. Finally, this chapter concludes with a comparison and discussion of the results achieved in the three scientific articles together with the specific conclusions that can be drawn from them.

Chapter 3 *LC-MS lipidomic studies using human cell lines*

In this chapter, the effects of some environmental chemical pollutants on the lipidome of human cells are investigated. These pollutants include PFASs, TBT

and some proautophagic drugs. First, the effects of PFASs are deeply studied on the human placental choriocarcinoma cell line (JEG-3). The cytotoxicity of PFASs and their capacity to inhibit aromatase activity (assessed using the toxicity assays described in the introductory section of the chapter), and a preliminary evaluation of the lipid disruption that these chemicals cause on human cells is included in the scientific article entitled "*Perfluorinated chemicals: Differential toxicity, inhibition of aromatase activity and alteration of cellular lipids in human placental cells*", published in the journal Toxicology and Applied Pharmacology, 277 (2014), 124-130. Further evaluation of the effects of PFASs and also of the obesogen TBT on the lipids of these cells is performed using both targeted and untargeted data analysis strategies. In the targeted approach, the performance of two high-resolution mass spectrometers is compared and a home-made database of lipid species is elaborated. This targeted lipidomic approach is included in the scientific article entitled "*Characterization of complex lipid mixtures in contaminant exposed JEG-3 cells using liquid chromatography and high-resolution mass spectrometry*" published in the journal Environmental Science and Pollution Research, 21 (20) 2014, 11907-11916. In the untargeted approach, also described in this chapter, a chemometric strategy based on the use of *binning* and *time windowing* for data compression and MCR-ALS for peak resolution is developed. This strategy is subsequently presented in the scientific article entitled "*Chemometric strategy for untargeted lipidomics: Biomarker detection and identification in stressed human placental cells*", published in the journal Analytica Chimica Acta, 854 (2015), 20-33. In addition, the effects of some proautophagic drugs are investigated in a glioma cell line and are included in the scientific article entitled "*Chemometric evaluation of glioma cell lipidome in response to proautophagic drugs*", recently submitted for publication. Finally, this chapter concludes with a dissection of the results achieved in these four scientific articles together with the specific conclusions that can be drawn from them.

Chapter 4 *IR and Raman spectroscopy combined with MS and chemometrics for environmental omic assessment*

In this chapter, IR and Raman spectroscopy and MS spectrometry are used to generate environmental omic data. In contrast to chapter 3, the effects of the contaminants are not studied on human cells but in an environmental model organism (zebrafish) and in an environmental model cell line (*Xenopus laevis* A6 cell line). On the one hand, ATR-FTIR is used to evaluate the effects of four individual PFASs on *Xenopus laevis A6 kidney epithelial cells*. This study entitled "*Perfluoroalkylated substance effects in Xenopus laevis A6 kidney epithelial cells determined by ATR-FTIR spectroscopy and chemometric analysis*" was published in the journal Chemical Research in Toxicology, 29 (5) (2016), 924-932, and it is first presented in this chapter. On the other hand, Raman spectroscopy together with MS spectrometry are used to evaluate the effects of CBNs on zebrafish (*Danio rerio*). The latter study entitled "*Diet-sourced carbon-based nanoparticles induce lipid alterations in tissues of zebrafish (Danio rerio) with genomic hypermethylation changes in brain*" was published in the journal Mutagenesis, 00 (2016), 1-13 and it is also hereby presented. This chapter concludes with a comparative discussion of the results achieved in these two scientific articles together with the specific conclusions that can be drawn from them.

Conclusions This section reports the general conclusions that result from this Thesis.