Characterization of organic aerosols from ship emissions by high resolution mass spectrometry : Development of new analytical methods and data visualization software.

Maxime Sueur

Normandie Université

UNIVERSITÉ DE ROUEN NORMANDIE

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité **CHIMIE**

Préparée au sein de l'**Université de Rouen Normandie**

Characterization of organic aerosols from ship emissions by high resolution mass spectrometry:Development of new analytical methods and data visualization software.

Présentée et soutenue par

MAXIME SUEUR

Thèse soutenue le 27/03/2024

devant le jury composé de :

M. CRISTIAN FOCSA	Professeur des Universités - UNIVERSITE DE LILLE	Rapporteur du jury
MME VERONIQUE VUITTON	Directeur de Recherche - UNIVERSITE GRENOBLE ALPES	Rapporteur du jury
M. MARK BARROW	Professeur des Universités - Université de Coventry (ROYAUME-UNI)	Membre du jury
MME ANGELICA BIANCO	Chargé de recherche du CNRS - UNIVERSITE CLERMONT AUVERGNE CLERMONT AUVERGNE	Membre du jury
M. JEAN-FRANÇOIS BRIERE	Directeur de Recherche - Université de Rouen Normandie	Membre du jury
M. CHRISTOPHER RÜGER	Maître de Conférences - Université de Rostock (ALLEMAGNE)	Membre du jury
M. CARLOS AFONSO	Professeur des Universités - Université de Rouen Normandie	Directeur de thèse
MME HELENE LAVANANT	Maître de Conférences - Université de Rouen Normandie	Co-directeur de thèse

Thèse dirigée par **CARLOS AFONSO** (CHIMIE ORGANIQUE, BIOORGANIQUE, REACTIVITE, ANALYSE) et **HELENE LAVANANT** (CHIMIE ORGANIQUE, BIOORGANIQUE, REACTIVITE, ANALYSE)

ED 508 NC Normande de Chimie

Laboratoire Cobra — Chimie organique, Bioorganique, Réactivité et Analyse.

Abstract

Organic aerosols (OA), whether of anthropogenic or natural origin, have a major impact on the environment, both in terms of climate changes and health effects. Among the many sources of OA, the maritime transport sector occupies a far from negligible place, on the one hand, because of the annual volume of exhaust emitted and their nature, and on the other because of the proximity of emissions to inhabited areas such as ports or coastlines. Although the composition of these exhausts is monitored by means of routine analyses, it is important to study them in depth in order to fine-tune regulations on marine traffic. However, the analysis of the molecular-scale composition of a mixture as complex as ship emissions requires the use of techniques such as mass spectrometry (MS), and in particular Fourier transform mass spectrometry (FTMS). However, FTMS generates a large amount of data and generally requires the use of processing and visualization software to extract and highlight relevant information. Also, although FTMS provides information on the molecular composition of a sample, this technique does not allow the evaluation of the isomeric diversity, unlike ion mobility spectrometry (IMS). So, with the aim to characterize ship emissions, we adopted three lines of work: the development of open-access software under Python to facilitate the processing and visualization of FTMS data for the characterization of complex mixtures, the targeted characterization of petroporphyrins in naval fuels and their combustion products by Fourier transform ion cyclotron resonance (FTICR) mass spectrometry using electron transfer matrix-assisted laser desorption/ionization (ET-MALDI), and the study of structural modifications caused by photochemical aging on ship emissions using IMS-MS.

Keywords: mass spectrometry, FTICR, ion mobility spectrometry, organic aerosols, particulate matter, software development, Python, ET-MALDI.

Résumé

Les aérosols organiques (OA), qu'ils soient d'origine anthropogène ou naturelle, ont un grand impact sur l'environnement, tant d'un point de vue climatique que sanitaire. Parmi les nombreuses sources d'OA, le secteur du transport maritime occupe une place loin d'être négligeable, d'une part à cause du volume annuel de gaz d'échappement et de leur nature ; d'autre part à cause de la proximité des émissions avec des zones habitées telles que les ports ou les côtes. Même si la composition de ces gaz d'échappement est contrôlée au moyen d'analyses de routine, il est important de l'étudier en profondeur afin d'affiner les réglementations. Cependant, une analyse de la composition à l'échelle moléculaire d'un mélange aussi complexe que les émissions de navire requiert l'utilisation de techniques telle que la spectrométrie de masse (MS), et notamment la spectrométrie de masse à transformée de Fourier (FTMS). Cependant, la FTMS génère une grande quantité de données et nécessite généralement l'utilisation de logiciels de traitement et de visualisation afin d'extraire et de mettre en évidence les informations pertinentes. De plus, bien que la FTMS fournisse des informations sur la composition moléculaire d'un échantillon, cette technique ne permet pas d'en évaluer la diversité isomérique, contrairement à la spectrométrie de mobilité ionique (IMS). Ainsi, afin de caractériser les émissions de navire, nous avons adopté trois axes de travail : le développement d'un logiciel open-access sous python pour faciliter le traitement et la visualisation des données FTMS pour la caractérisation de mélanges complexes, la caractérisation ciblée de pétroporphyrines dans les carburants navals et leur produits de combustion par spectrométrie de masse à résonnance cyclotronique des ions par transformée de Fourier (FTICR) grâce à la désorption/ionisation laser assistée par matrice à transfert d'électrons (ET-MALDI), et l'étude des modification structurelles provoquées par le vieillissement photochimique sur les émissions de navires par IMS-MS.

Mots clés : Spectrométrie de masse, FTICR, spectrométrie de mobilité ionique, aérosols organiques, matière particulaire, développement de logiciel, Python, ET-MALDI.

« *The trick is to never stop looking.*

There's always *another secret. -Kelsier* »

Brandon Sanderson, The Final Empire, The Mistborn Saga

Acknowledgements / Remerciements

I would like to address my first acknowledgements to the members of the jury: Véronique Vuitton, Cristian Focsa, Angelica Bianco, Mark Barrow and Jean-François Briere for accepting to evaluate my work as a PhD student.

Je remercie mon directeur de thèse Carlos Afonso de m'avoir donné une chance alors que j'étais loin d'être un expert en chimie analytique et pour m'avoir aidé à en arriver là. Merci également pour ces bons souvenirs de congrès ainsi que pour vos conseils et vos anecdotes. Je remercie Hélène Lavanant, ma co-directrice, pour ses précieux conseils et pour tout le temps qu'elle a passé à corriger mes posters, papiers et autres présentations. J'ai eu de la chance de pouvoir compter sur vous. Christopher Rüger, thank you for your kindness and your brilliant ideas, you are a great supervisor and an amazing kicker player.

Je remercie Pierre Giusti de m'avoir intégré au laboratoire commun iC2MC dans lequel j'ai pu développer mes compétences de programmeur. Je remercie Marie Hubert-Roux d'avoir résolu presque chacun des problèmes que j'ai pu rencontrer sur les spectros. Sans toi je serai encore en train d'essayer de brancher l'APCI sur le Synapt ! Merci à Isabelle Schmitz pour m'avoir inclus dans les projets Infranalytics, ce qui m'a permis de découvrir d'autres applications de la FTMS. Je remercie également Corinne Loutellier-Bourhis et Mathilde Lauzent pour leur gentillesse et leur bonne humeur.

Merci à mes encadrants de stage d'ingénieur Lokmane Abdelouahed, Chetna Mohabeer et Bechara Taouk pour m'avoir donné le gout de la recherche et m'avoir mis le pied à l'étrier dans ce domaine.

Merci à mes parents, Karine et Laurent, pour votre soutien, pour avoir cru en moi et pour m'avoir permis d'entreprendre cette longue série d'études. Normalement, c'est fini ! Merci à Morgane et Manon, pour ne pas avoir été si casse-pied que ça pour des petites sœurs, je m'en tire plutôt bien, je trouve.

Merci à Franck et Virginie pour votre aide et votre gentillesse. Désolé de vous remercier pour tout ce que vous avez fait pour moi en accaparant votre fille… Merci Jules pour toutes ces fois où tu me laisses gagner aux jeux de cartes, c'est vraiment gentil de ta part ! Merci à tata Catherine sans qui je ne serais probablement pas allé si loin dans mes études.

Florian, merci d'être comme un frère pour moi, ton soutien remonte à bien avant cette thèse et m'a été essentiel. Au passage bon courage pour ta soutenance, espèce de futur docteur va. Et, par pitié, arrête de te faire spoil sur TikTok !

Florentin, alias CrossbarMaster, merci de ne pas m'avoir aidé à sortir de mon addiction pour les jeux vidéo et pour tout ce temps passé en vocal à se faire rouler dessus sur RL, ou autre, parce qu'on fait des commérages au lieu de parler du jeu.

Merci à mes amis de longue date Marie, Quentin, Florine, Baptiste et Caroline pour ces tout merveilleux moments !

Merci Julien d'avoir été un super collègue et ami. Merci pour tes blagues bien beaufs et pour m'avoir fait tomber dans la marmite PyC2MC quand je n'étais qu'un jeune thésard. Merci Oscar de m'avoir ouvert les yeux quant à mon besoin d'investir dans de la boisson énergisante en poudre. Et merci pour ton bureau ! Olivier, c'est grâce toi que j'ai pu postuler à cette (presque comme ton chat) offre de thèse. Merci beaucoup pour cette opportunité et pour tous ces bons moments passés au labo et en dehors. Merci Clément de m'avoir appris plein de choses au début de ma thèse, dont le fait que je n'étais pas très bon aux jeux de cartes. Quand je pense que ça va mieux, heureusement que tu es là pour me péter les chevilles. Merci Estelle, la kalash', pour me péter les chevilles de temps en temps à grands coup de punchlines bien senties, ça fait réfléchir. Décidemment ils se sont bien trouvé ces deux-là. Merci Anaïs pour avoir partagé avec moi les moments de solitudes que sont les formations d'enseignement. Merci Delphine, ou Phine-del de son vrai nom, pour ta bonne humeur permanente et tes photobombs. Merci Guillaume d'avoir le rôle de méchant dans 90% des jeux auxquels tu joues, ça facilite pas mal les choses ! Merci Théo de rire comme ma voiture démarrant par un matin de janvier, ça m'a poussé à la réparer. Merci Aurélien de nous permettre de faire du team-building après chaque pause repas. Merci Sébastien d'être aussi

bon, ou nul, que moi aux sessions escalades. Merci Matthieu d'amener un peu de classe au labo. Merci Gautier pour tes remarques douteuses et ta façon bien à toi de voir le monde. Merci Salomé de rire à mes blagues pourries quand tous les autres préfèrent m'ignorer, c'est sympa. Merci Marie van A pour avoir essayé d'introduire un peu de culture dans les discussions des pauses repas. Thank you Silvia, for your kindness and perseverance. Merci Nathaniel pour avoir apporté le soleil au labo et pour tous nos bons moments. Merci à Jason, alias le « DJI », pour veiller au respect des règles de nos jeux, mais aussi pour rappeler que finalement le FT n'est qu'un détecteur. Merci aussi de m'aider à me moquer de nos collègues sur le départ avec des memes de qualité.

Merci Charlotte, alias bébou, de m'avoir fait découvrir des expressions étranges et un peu insensées venues d'Yvetôt, merci de partager ton rire si communicatif (et que j'imite à la perfection). Merci de me rendre sourd d'une oreille aussi. Merci également de participer si activement à notre carnet de citations. Merci pour ces réels trop meugnooooons. Plus sérieusement, merci pour tous ces bons moments et ces fous rires. Je suis très heureux d'avoir fait ta connaissance et de pouvoir continuer à travailler avec toi !

La meilleure pour la fin, merci Chloé de me supporter au quotidien, ce n'est pas rien ! Merci de me soutenir et de m'apporter tout le bonheur dont j'ai besoin et plus encore. Merci de t'occuper de moi à chaque fois que j'ai un bobo, ce qui arrive assez souvent l'air de rien. Merci de me pousser à devenir meilleur parce que c'est carrément un job à plein temps. Merci de me faire rire tous les jours que l'on passe ensemble même si parfois c'est parce que tu te moques de moi. Enfin pour reprendre les paroles de quelqu'un qui s'exprime bien mieux que moi, « *Hopelessly, I'll love you endlessly/ Hopelessly, I'll give you everything/ But I won't give you up* ». Je t'aime.

Table of Content

Abbreviations

ACOS: Average carbon oxidation state

AMS: Aerosol mass spectrometry

APCI: Atmospheric pressure chemical ionization

APLI: Atmospheric pressure laser ionization

APPI: Atmospheric pressure photoionization

BBOA: Biomass burning organic aerosol

BC: Black carbon

BrC: Brown carbon

CCN: Cloud condensation nuclei

CCS: Collision cross section

CCSexp: Experimental CCS

cIMS: Cyclic ion mobility spectrometry

CV: Coefficient of variation

DBE: Double bond equivalent

DCM: Dichloromethane

DFT: Density functional theory

DMA: Differential mobility analyzer

DTIMS: Drift tube ion mobility spectrometry

EC: Elemental carbon

ECA: Emission control area

EGR: Exhaust gas recirculation

EHSS: Exact hard sphere scattering

EIM: Extracted ion mobilogram

ELVOC: Extreme low-volatile organic compounds

ESI: Electrospray ionization

ET-MALDI: Electron-transfer matrix-assisted laser desorption/ionization

FAIMS: Field asymmetric ion mobility spectrometry

FID: Free induction decay

FSC: Fuel sulfur content

FTICR: Fourier transform ion cyclotron resonance

FTMS: Fourier transform mass spectrometry

FWHM: Full-width half maximum

GHG: Greenhouse gases

HFO: Heavy fuel oil

HOA: Hydrocarbon-like organic aerosol

iC2MC: International laboratory for complex matrices molecular characterization

ICR: Ion Cyclotron resonance

IE: Ionization energy

IMS: Ion mobility spectrometry

IPCC: Intergovernmental panel on climate change

IR: Infrared

IVOC: Intermediate-volatile organic compounds

LNG: Liquified natural gas

LPG: Liquified petroleum gas

LV(-OOA): Low-volatile (oxidized organic aerosol)

LVOC: Low-volatile organic compounds

MALDI: Matrix-assisted laser desorption-ionization

MARPOL: Marine pollution convention

MDO: Marine diesel oil

MeOH: Methanol

MGO: Marine gas oil

MS: Mass spectrometry

NECA: Nitrogen emission control area

NOx: Nitrogen oxides

OA: Organic aerosol

OC: Organic carbon

OM: Organic matter

OPS: On-shore power supply

PA: Projection approximation

PAM: Potential aerosol mass

PE: Polyethylene

PEAR: Photochemical emission aging flowtube reactor

PM : Particulate matter

POA: Primary organic aerosol

RF: Radiative forcing

RF: Radiofrequency

SECA: Sulphur emission control area

SOA: Secondary organic aerosol

SOx: Sulfur oxides

SPE: Solid-phase extraction

SRIG: Stacked-ring ion guide

SV(-OOA): Semi-volatile (oxidized organic aerosol)

SVOC: Semi-volatile organic compounds

t_D: Drift time

TIMS: Trapped ion mobility spectrometry

TM: Trajectory method

TOCA: Thermal-optical carbon analyzer

TOF: Time of flight

TWIMS: Traveling wave ion mobility spectrometry

UHRMS: Ultrahigh-resolution mass spectrometry

ULTRHAS: Ultrafine particles from transportation – health assessment of sources

UV: Ultraviolet

VOC: Volatile organic compounds

V-UV: Vacuum UV

WESP: Wet electrostatic precipitator

YLL/YOLL: Years of life loss

List of Figures

List of Tables

General Introduction

Organic aerosol emissions play an important role in both short- and long-term climate change by modifying the interactions of clouds with light[1, 2]. Organic aerosols also affect living beings, causing several diseases by their introduction into the respiratory and blood systems [3-5]. There is a wide variety of aerosols, from natural particle emissions due to sea spray formed by bursting droplets at the sea-air interface [6] to anthropogenic emissions caused by air, sea, and land traffic. Furthermore, the composition and morphology of an aerosol is subjected to changes with time as it ages in the atmosphere [7].

Among the many sources of organic aerosol emissions, we focused on ship exhaust, which is as important as car emissions in terms of volume emitted per year [8-10]. Moreover, ship emissions are of particular interest not only because of the quantity emitted but also because of their composition, as the fuels used for shipping have a high sulfur content [11]. Furthermore, an important proportion of these emissions occur while ships are in transit close to coastlines or at berth in port areas [12]. As a consequence of this proximity to inhabited areas, the health impact of ship exhaust is also significant in areas where the marine traffic is dense, such as the Baltic sea [13].

Even though numerous efforts are being made to reduce the impact of shipping on health and the environment, such as replacing the heavy fuels currently used with carbon-neutral fuels [14] or using various exhaust gas treatment techniques [15], ship emissions are still the subject of many studies, particularly in terms of their composition. Indeed, a way to reduce adverse effects of ship emissions is to monitor their composition by means of regulations that limit, for example, the amount of sulfur allowed in the fuel [16]. The compliance with these regulations by shipping companies is ensured by elemental analysis and particulate matter physical characterization. However, more in-depth chemical characterizations can be performed using high-resolution mass spectrometry (MS). Mass spectrometry provides information on the composition of ship emissions at the molecular level, making it the technique of choice for identifying problematic compounds. Nonetheless, MS alone does not give information on the isomeric diversity of a sample. However, this can be achieved by hyphenating ion mobility spectrometry (IMS) to MS.

The analysis of such highly complex mixture by mass spectrometry generate an important quantity of information. The quantity of data obtained in MS, especially ultrahigh-resolution MS (UHRMS), can be overwhelming for the user and often requires the use of advanced data processing and visualization software to process and extract meaningful information.

The main objective of this work is to carry out the in-depth characterization of ship emissions. This objective was divided into the following three axes: the development of a software aiming at facilitating data processing and visualization, the characterization of ship fuels and emissions by UHRMS, and the structural study of aged ship emissions by IMS-MS.

The first chapter of this thesis describes organics aerosols, from their formation to their harmful effects with a focus on the specificity of ship emissions. As for the second chapter, reports the experiment setup used for ship emission collection as well as the analytical techniques used in this work: mass spectrometry and ion mobility.

Chapters three and four summarize in the form of published articles the development of our data visualization software and the targeted characterization by UHRMS of petroporphyrins in ship fuels and their emissions. The fifth chapter focuses on the initial results obtained in the structural study of aged ship emissions by IMS-MS.

References:

1. Andreae, M.O., *The dark side of aerosols.* Nature, 2001. **409**(6821): p. 671-672.
2. Boucher, O., et al., *Clouds and aerosols*, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker, et al., Editors. 2013, Cambridge University Press: Cambridge, UK. p. 571-657.
3. Pope, C.A., et al., *Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution.* Circulation, 2004. **109**(1): p. 71-77.
4. Gauderman, W.J., et al., *The Effect of Air Pollution on Lung Development from 10 to 18 Years of Age.* New England Journal of Medicine, 2004. **351**(11): p. 1057-1067.

5. Bernstein, J.A., et al., *Health effects of air pollution.* Journal of Allergy and Clinical Immunology, 2004. **114**(5): p. 1116-1123.

6. Lewis, E.R. and S.E. Schwartz, *Sea Salt Aerosol Production: Mechanisms, Methods, Measurements and Models—A Critical Review*. Geophysical Monograph Series. 2004: American Geophysical Union.

7. Kroll, J.H. and J.H. Seinfeld, *Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere.* Atmospheric Environment, 2008. **42**(16): p. 3593-3624.

8. *IMO 2014: Third IMO GHG Study*. Available from: https://www.imo.org/en/ourwork/environment/pages/greenhouse-gas-studies-2014.aspx.

9. Liu, H., et al., *Health and climate impacts of ocean-going vessels in East Asia.* Nature Climate Change, 2016. **6**(11): p. 1037-1041.

10. Serra, P. and G. Fancello, *Towards the IMO's GHG Goals: A Critical Overview of the Perspectives and Challenges of the Main Options for Decarbonizing International Shipping.* Sustainability, 2020. **12**(8): p. 3220.

11. Winnes, H. and E. Fridell, *Particle Emissions from Ships: Dependence on Fuel Type.* Journal of the Air & Waste Management Association, 2009. **59**(12): p. 1391-1398.

12. Chang, C.-C. and C.-M. Wang, *Evaluating the effects of green port policy: Case study of Kaohsiung harbor in Taiwan.* Transportation Research Part D: Transport and Environment, 2012. **17**(3): p. 185-189.

13. Jonson, J.E., et al., *Model calculations of the effects of present and future emissions of air pollutants from shipping in the Baltic Sea and the North Sea.* Atmospheric Chemistry and Physics, 2015. **15**(2): p. 783-798.

14. Manigandan, S., et al., *Comparative study of nanoadditives TiO2, CNT, Al2O3, CuO and CeO2 on reduction of diesel engine emission operating on hydrogen fuel blends.* Fuel, 2020. **262**: p. 116336.

15. Karjalainen, P., et al., *Real-world particle emissions and secondary aerosol formation from a diesel oxidation catalyst and scrubber equipped ship operating with two fuels in a SECA area.* Environmental Pollution, 2022. **292**: p. 118278.

16. *IMO 2020*. Available from: https://www.imo.org/en/MediaCentre/HotTopics/Pages/Sulphur-2020.aspx.

Chapter I. Atmospheric aerosols and specificity of ship emissions

1. Atmospheric Aerosols

The term aerosol refers to a suspension of particles, whether liquid or solid, in gas. More precisely, in atmospheric sciences, the term designates the particulate phase also called particulate matter (PM). PM has an aerodynamic diameter (D_{ae}) in the range of 10^{-9} to 10^{-4} m [1]. The lower limit is determined by the size of molecules and molecular clusters, while the upper limit corresponds to the size at which large particles and droplets tend to sediment rapidly. PM is often divided into three categories using its D_{ae}. Indeed, particles with $D_{ae} < 0.1$ µm ($PM_{0.1}$), $D_{ae} < 2.5$ µm ($PM_{2.5}$) and $D_{ae} < 10$ µm (PM_{10}) are called ultrafine, fine, and coarse fraction, respectively [2]. However, the aerodynamic parameter is not the only parameter that determines the shape and morphology of aerosol particles, as they also depend on the aerosol source [3].

1.1 Primary Sources and Sinks

Primary sources of aerosol emissions can be grouped into two types: natural and anthropogenic sources. The most important natural sources of aerosol are sea spray formed by bursting droplets at the sea-air interface[4] and soil erosion. Nevertheless, other natural phenomena can influence the concentration of particles in the atmosphere such as volcanic activity, wildfires, and emissions from the biosphere (spores, pollen, and microorganisms like viruses). Anthropogenic sources themselves can be separated into mobile and stationary sources. The mobile sources include traffic on sea, air, railroad, and land while the stationary sources refer to domestic and large-scale energy generation, but also to cooking and industrial operations [5]. Regarding mass, atmospheric aerosols mainly originate from natural sources, except in urban areas where anthropogenic sources predominate [6]. The lifetime of PM in the atmosphere varies between hours and weeks and depends on its properties such as size, structure, and composition [6, 7]. Airborne particles are removed from the atmosphere in two deposition mechanisms (Figure I-1). The first deposition mechanism, called wet deposition can itself be divided into two types of scavenging. In fact, aerosol particles can function as cloud condensation nuclei (CCN), gathering the surrounding humidity and initiating cloud formation.

Then, when the cloud precipitates, the particles within it fall down to the Earth's surface; a phenomenon also known as in-cloud scavenging. Furthermore, raindrops can also scavenge other particles during precipitation, which is known as below-cloud scavenging. The second, less relevant, deposition mechanism is the dry deposition. For smaller aerosols, dry deposition occurs when particles impact a larger aerosol, or any obstacles preventing them from following the streamlines of the gas flow, which is usually wind. For larger aerosols, it refers to gravitational sedimentation. Consequently, aerosol deposition strongly depends on meteorological conditions such as the amount of precipitation and wind velocity. These conditions can also affect the composition of aerosol, especially through the aging process described below.

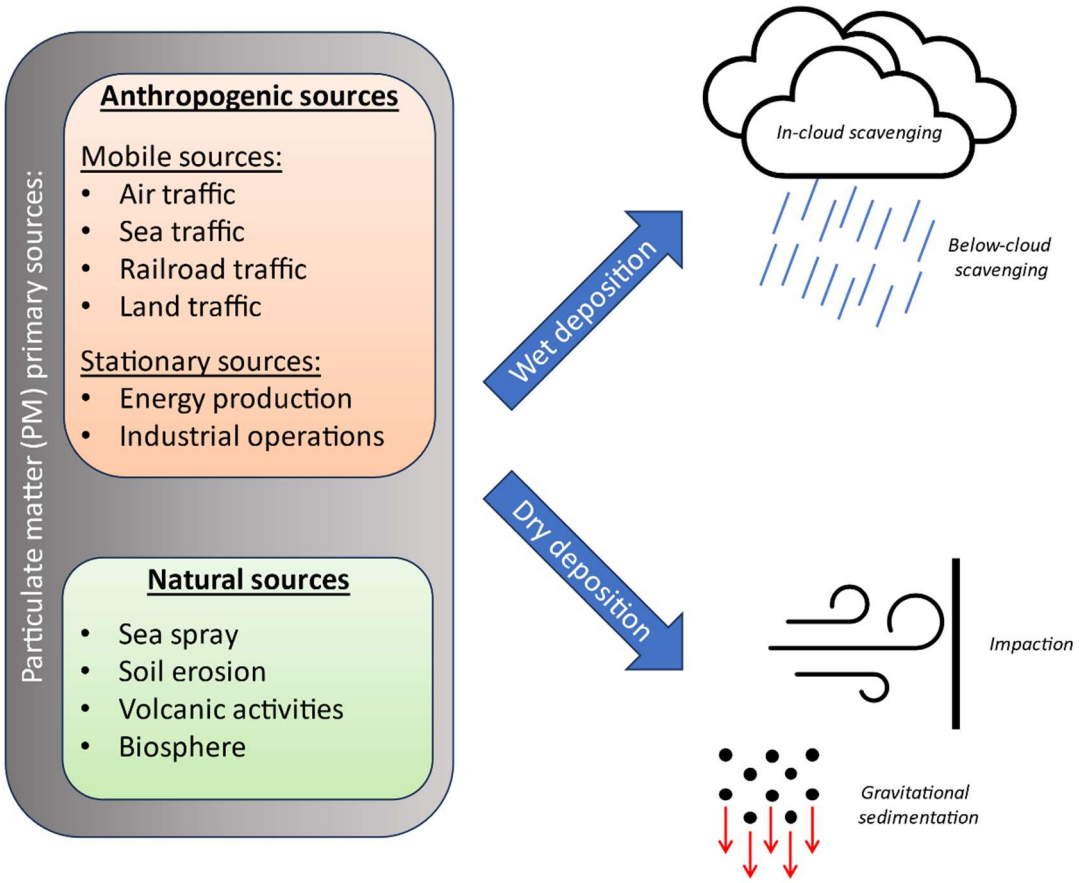


Figure I-1 Principal sources and sinks of organic aerosol.

1.2 Chemical Composition

The chemical composition of an aerosol depends on the sources it originates from and the atmospheric processing (ageing). Hence, at different sites, the composition of the ambient

aerosol may vary depending on the presence of nearby anthropogenic sources and the aerosol aging state. However, its constituents can be broken down into two main categories: inorganic and organic constituents. The inorganic compounds are mainly nitrate (NO_3^+), sulfate (SO_4^{2-}), ammonium (NH_4^+), chloride (Cl^-), and metal cations (alkali metals/ alkaline earth metals). Figure I-2 shows the typical composition of an urban aerosol.

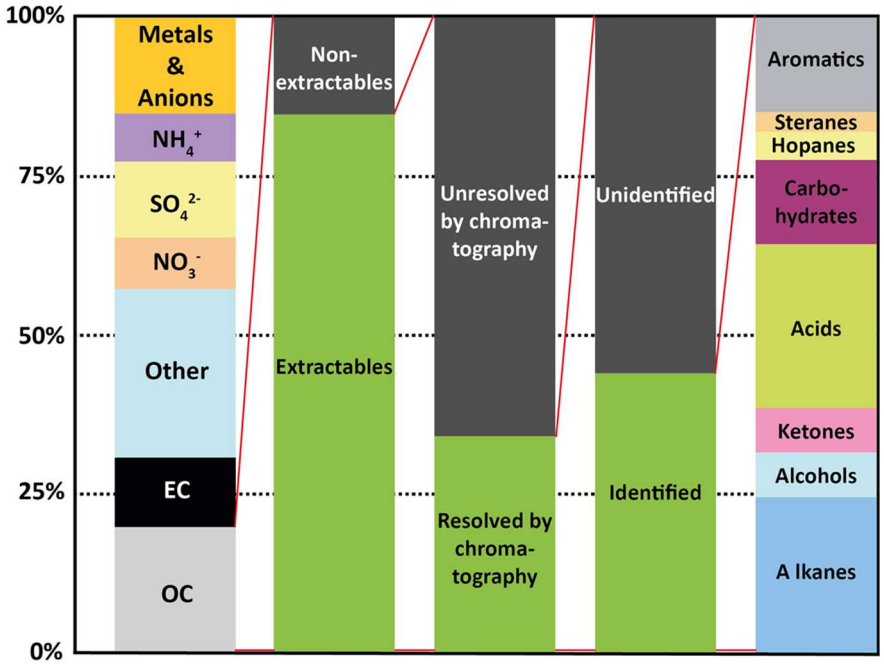


Figure I-2: Urban aerosol composition, adapted from [8]

The carbon content of an aerosol is divided into two categories. On the one hand, there is organic carbon (OC) which can be quantified by techniques, such as thermal-optical carbon analysis (TOCA). OC, when multiplied by an empirical factor depending on the aerosol's source, can be converted to organic aerosol (OA) also referred to as organic matter (OM). This organic matter contains hydrogen and heteroelements like oxygen, nitrogen, and sulfur [9]. On the other hand, there is elemental carbon (EC) which consists of almost pure graphite-like carbon in the form of soot [10]. The OC/EC ratio depends on the aerosol source and is an important bulk parameter for classification and source apportionment.[11]

Other terms are commonly used when referring to OA carbon content: black carbon (BC) and brown carbon (BrC or colored organic carbon). These terms are used when optical methods are used to characterize aerosol composition. Hence, by definition, EC is determined

by thermal-optical methods, while BC is determined by optical methods, but they substantially refer to the same entity as illustrated in Figure I-3.

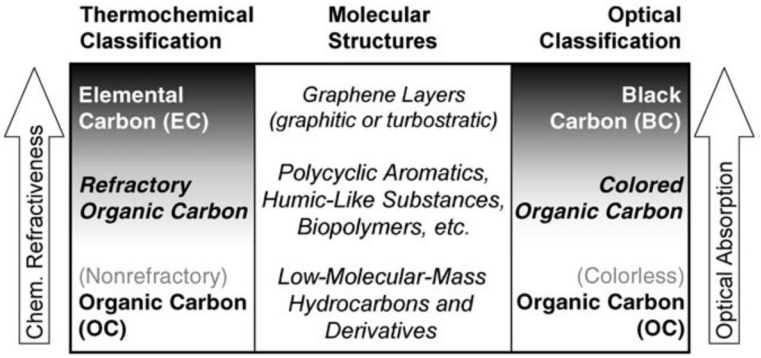


Figure I-3: Thermochemical and optical classifications of carbon in organic aerosol [12]

1.3 Organic Aerosol Classification

OA can be further be classified using several bulk and molecular parameters such as their emission source, volatility, oxidation state, or even their main compound classes. Historically, OA was first classified as follows: hydrocarbon-like organic aerosol (HOA), biomass-burning organic aerosol (BBOA), and oxidized organic aerosol (OOA). Furthermore, HOA and BBOA are combined into the primary organic aerosol (POA) category for they are emitted by primary sources as opposed to OOA which are mostly associated with secondary organic aerosol (SOA) [13]. Actually, high oxygen content is strongly correlated with photochemical processing occurring during aerosol aging [14, 15] (See I.1.4). In 2006, Donahue *et al.* proposed a physical classification of OA based on their specific volatility ranges determined from saturation vapor pressure and the correlated gas-particle partitioning [16]. In 2012, the group of Donahue *et al.* added a second dimension to this classification: the oxygen content, often represented by the oxygen-to-carbon ratio or the average carbon oxidation state (ACOS or $\overline{OS_C}$) [17]. This leads to the classification presented on Figure I-4. On this representation, five volatility domains are defined for the OA: volatile (VOC), intermediate-volatile (IVOC), semi-volatile (SVOC), low-volatile (LVOC), and extreme low-volatile organic compounds (ELVOC). Furthermore, the categories from the first classification are also represented: HOA, BBOA, and both low-volatile (LV-) and semi-volatile (SV-) OOA; in addition to the more volatile "vapor" categories.

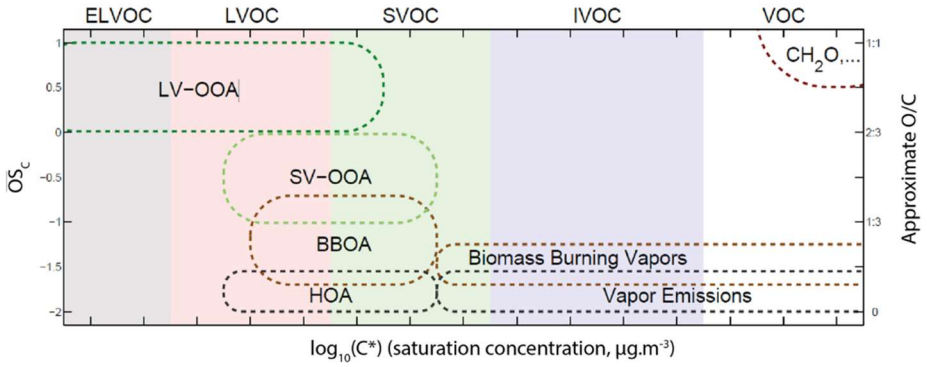


Figure I-4: Organic aerosol classification based on their volatility and oxygen content. LV-OOA: Low-volatile oxidized organic aerosol. SV-OOA: Semi-volatile oxidized organic aerosol. BBOA: Biomass burning organic aerosol. HOA: Hydrocarbon-like organic aerosol. (Adapted from [17])

1.4 Aerosol Aging

Organic aerosols released into the atmosphere are likely to be chemically transformed, or aged, following photochemical reactions with airborne oxidizing agents such as ozone (O_3), $NO_3^{\bullet}$ radical or $OH^{\bullet}$ radicals [18-20]. The contribution of each species is correlated to the solar radiation, thus to the daytime. During the day, $OH^{\bullet}$ radicals are the main drivers of atmospheric photochemical aging as ozone absorbs UV light and decomposes to $OH^{\bullet}$ radicals in presence of water following Equations I.1-I.2.

$$O_3 + h\nu\ (\lambda < 320nm) \rightarrow O_2 + O(^1D) \quad \text{I.1}$$

$$O(^1D) + H_2O \rightarrow 2\ OH^{\bullet} \quad \text{I.2}$$

The necessary ozone is produced by the photolysis of NO_2 and the degradation of VOCs [21, 22]. However, in absence of UV light, no excited oxygen ($O(^1D)$) can be formed; hence, ozone reacts with NO_2 to form nitrate radicals ($NO_3^{\bullet}$). Contrary to the other two oxidizing agent, oxidation with O_3 is more likely to take place if its reaction partner presents a structural specificity. Indeed, ozone requires a double bond to form a reaction intermediate called ozonide [23]. Thus, it can't directly oxidize saturated compounds present in the POA, contrary to OH radicals, but still affects the oxidation product distribution [24]. Oxidization of OA constituent leads to functionalization [25]. The increase in functionalization induces a decrease in volatility leading to nucleation or condensation on existing particles. The resulting aerosol,

formed by gas-to-particle conversion is called secondary organic aerosol (SOA). The typical atmospheric oxidation of VOC is shown in Figure I-5. On the other hand, gain in particle mass by heterogeneous oxidation reaction is slow and depends on the particle composition and morphology [26, 27] and refers to aged-POA. As SOA and aged-POA are mainly composed of carbonyls, acids, organic sulphates, and nitrates [28, 29]; they exhibit a higher functionality and polarity than POA; thus, they might act as more effective cloud condensation nuclei.[30]

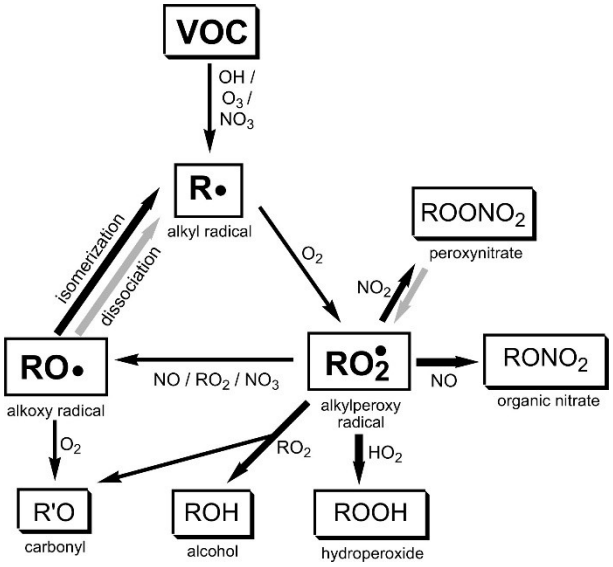


Figure I-5: Most important atmospheric oxidation pathways of volatile organic carbon [19]

The aging process of OA can be experimentally simulated. For this purpose, two approaches are primarily used. The first technique is a batch reactor called “smog chamber”, which consists of a Teflon bag, UV lamps mimicking the solar radiation, and an external ozone stream. The main inconveniences of such a batch reactor are the particle and vapor losses to the walls of the reactor [31] that have to be corrected by mathematical models and simulations [32]. One way to diminish the importance of this phenomenon is to increase the size of the smog chamber, this would allow for longer ageing time equivalents, while maintaining acceptable wall losses [33]. The other experimental approach for simulating aerosol ageing is based on the use of a plug-flow reactor. The two main reactors deployed are called the potential aerosol mass (PAM) and the photochemical emission aging flow-tube reactor (PEAR) [34-36]. Both consist in a steel cylinder equipped with UV lamps and an external O_3 and water vapor flow to allow $OH^{\bullet}$ radical formation. The design of such instruments allows to mimic

extensive photochemical aging of several equivalent days, while the residence time within the reactor is only around one minute [37].

1.5 Effects on Climate and Health

One of the main driving forces of organic aerosol research is the study of their impact on our environment, especially how they contribute to climate change or provoke health issues.

1.5.1. Effects on Climate

The climate impact of OA constituents is classified either as direct or indirect with respect to their radiative forcing (RF) on the climate. RF is the difference in the energy flux of solar and terrestrial radiation caused by an anthropogenic or natural change in atmospheric composition, Earth's surface properties, or solar activity [12]. This quantity is measured in Wm^{-2}. Positive RF such as terrestrial radiation absorption by clouds or greenhouse gases, also called greenhouse effect, lead to the warming of the atmosphere; whereas negative RF, such as the scattering or reflection of solar radiation, lead to a cooling effect to the atmosphere [38]. Direct effects are caused by the absorption or reflection of light by the aerosol particles, while the indirect effects result from their tendency to act as CCN, thus forming clouds interacting with solar and terrestrial radiations. Concerning OA constituents, colorless particles tend to reflect and scatter solar radiation, leading to a direct negative RF; nevertheless, BC and BrC fractions exhibit a high albedo, *i.e.,* they absorb more light than they reflect, leading to a direct positive RF. This direct positive RF is, for most cases, greater than the negative RF of colorless particles, leading to a warming effect [39]. On the other hand, clouds reflect and scatter solar radiations, producing a negative RF. Furthermore, sulphate ions, that also compose organic aerosol, have a highly negative RF [40].

1.5.2. Effects on Health

Numerous studies have proven that airborne particulate matter and other air pollutants are related to severe health effects, including enhanced mortality, cardiovascular, respiratory, and allergic diseases [41-43]. The first evidence of a correlation between poor air

quality and adverse health effects was the “London smog” event of 1952, in which several thousand deaths were attributed to the deplorable air quality [44]. More recently, in 2022, WHO was estimated that 6.7 million deaths worldwide were imputable to air pollution [45]. Air pollution in inhabited areas has various sources, from traffic to industrial processes, and its toxicity depends on the shape, chemical composition, and size of the emitted particles. In fact, the size of PM directly influences how far it can penetrate into the human respiratory system, as shown in Figure I-6 adapted from J. Heyder’s work [46]. It is shown that it is mainly the particles with a diameter comprised between 1 and 2.5 µm and below 0.1 µm, or $PM_{2.5}$ and $PM_{0.1}$, that reaches the alveolar region in which air-blood exchanges take place. Especially, ultrafine particles ($PM_{0.1}$) are suspected to be sufficiently small to enter the blood circulation or be transported to the brain directly via olfactory nerves [47, 48]. On the other hand, coarse particles are shown to be mainly deposited in the pharynx and larynx area. In addition to the particle size, chemical composition of PM is also key factor for the toxicity. As a matter of fact, compounds like PAH and trace metals found in combustion aerosol are known for acute and chronic toxicity [49, 50].

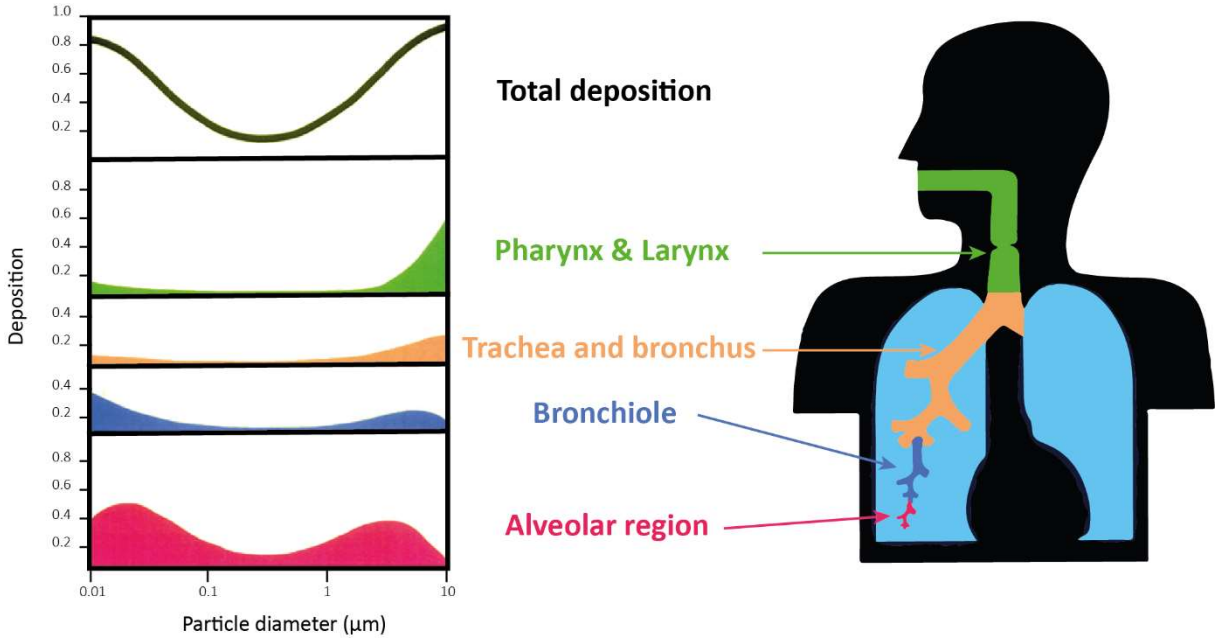


Figure I-6: Particles size distribution in the human respiratory system. Small particles (<0.1 µm) and large particles (>1µm) are more likely to be deposited in the respiratory system than particles with an intermediate size. [46]

2. Ship emissions

2.1 Introduction

As a consequence of globalization, shipping is the transport mode used to move 80% to 90% of the traded goods around the globe, which is equivalent to approximately 10 billion tons transported goods annually [51]. In addition, recent studies predict a demand growth of 40% for seaborne trade by 2050 [52]. The marine transport sector consequently contributes significantly to global air pollution in a non-negligible way. Actually, shipping emissions were estimated to account for the emission of 3% the global greenhouse gases (GHG), 13% of the nitrogen oxides (NO_x) and around 12% of the sulfur oxides (SO_x) [52-54]. Furthermore, approximately 70% of the total shipping emissions are released less than 400 km distance to the coastline and inhabited areas [55]. Thus, among the various sources of anthropogenic OA emissions, the study of shipping emissions effects on health and environment presents an important interest [56].

2.2 Impacts on Health and Environment

Corbett *et al*. [57] estimated in 2007 that marine transport emissions contributed to around 60 000 annual deaths, which represented 3 to 8% of the total deaths related to $PM_{2.5}$ emissions [58]. Ship-induced air pollution has a stronger effect near major trade routes. In fact, Liu *et al.* [54] studied the health impact of ocean-going vessel in East Asia, which is one of the major trade routes. They evidenced that ship-related emissions not only affect shores and coastlines but also extend their influence up to four hundred kilometers from the coast. They estimated that ship emissions led to 24 000 premature deaths annually in East Asia and mainly in mainland China and Japan.

Also, Broome *et al.* [59] studied the mortality effect of ship-related PM emission in Sydney greater metropolitan region which is also nearby an important trade route. They estimated in 2011 that 220 years of life were lost (YLL or YOLL) due to $PM_{2.5}$ emitted by marine transport. In the case of this study, 220 YLL can be translated to a loss of life expectancy of 0.803 days per person per year. This value is calculated exclusively on ship-related air pollution which in this case accounts for less than 2% of the region-wide $PM_{2.5}$ concentration.

Jonson *et al.* [60] estimated, in 2010, that ship emission in the North sea and the Baltic sea regions were responsible for a total of 69 189 YLL distributed among the different countries as shown in Table I-1 Furthermore, Barregard *et al.* [61] estimated that, in 2014, around 1 500 cases of ischemic heart disease and stroke cases were due to $PM_{2.5}$ emissions from shipping in the Baltic sea region.

Table I-1: Estimated total years of life lost (YLL) in 2010 due to shipping emission in the North sea and the Baltic sea regions. [60]

Country	YLL in 2010
Belgium	4061
The Netherlands	4665
Germany	26071
Great Britain	11716
Denmark	660
Norway	657
Sweden	928
Finland	629
Poland	18085
Latvia	562
Lithuania	986
Estonia	169

In addition to their adverse impact on health, emissions from the marine sector also exhibit a relevant climate impact.

For example, in shipping emission, NO_x mainly refers to NO and NO_2 [62]. In fact, NO is formed by the combustion of airborne nitrogen when the engine is working at high temperature. Then, once in the atmosphere, part of the NO is oxidized to NO_2. These nitrogen oxides pose a threat to human health as they can cause respiratory diseases [60] but also to the environment, as they are the cause of phenomena such as acidic rain [63] and water eutrophication [64]. In addition, the emission of NO_x or other ozone precursors from shipping activities lead to ozone formation that perturbs the concentration of atmospheric hydroxyl radical and, consequently, the lifetime of the methane emitted by other sector of activity. This reduction of airborne methane concentration induces a direct negative RF, *i.e.* a cooling effect of the global surface temperature [65].

Ship emissions are also composed of GHG and BC that induce a direct positive RF, *i.e.* a warming effect. This warming effect can be quantified by a factor called the global warming potential (GWP). This value represents the quantity of energy absorbed by one ton of emission over a given period of time relative to the energy absorbed by the emission of one ton of CO_2. The main GHG emitted by ships are carbon dioxide, methane and nitrous oxide. Table 2 gathers the GWP of the mentioned GHG estimated by the intergovernmental panel on climate change (IPCC) in 2014 [66] and the GWP of BC [67], both over 20 and 100 years.

Table I-2: Global warming potential (GWP) of common greenhouse gases [67]

Pollutant	GWP over 20 years	GWP over 100 years
CO_2	1	1
CH_4	72	28
N_2O	289	265
Black carbon	3200	900

These data are used to calculate the total equivalent of GHG emissions as exemplified in the study of Styhre *et al.* [68] in which they calculated the GHG emissions caused by ships in four different ports: the ports of Gothenburg, Long Beach, Osaka and Sydney. In this study, they calculated not only the total CO_2 equivalent emissions, but they also show that the 'at berth' mode, *i.e.* when the ship is docked but its auxiliary engine are still running, contributed to about half of the in-ports emissions. This contribution can be reduced by different measure. One of them is to provide electrical on-shore power supply (OPS), thus allowing the ship to turn its auxiliary engine off without disruption of on-board services. However OPS might not always be beneficial in terms of air pollution, especially in countries where the electricity is mainly produced by burning fossil resources without exhaust treatment [69]. In addition, OPS for cruise ships might be highly beneficial in touristic ports due to the high electricity demand of those ships and the need of an efficient energy supply while at berth to maintain on-ship activities. However, OPS for domestic cruise ships isn't considered reasonable economically speaking as the cost of OPS are too high compared to the short stay at berth of such ships [70]. Another way to reduce the contribution of at-berth GHG emissions basically consist in reducing the turnaround time for at-berth ships. However, while appearing as an 'evident' solution, optimizing turnaround time relies on an overall efficiency increase of ports as discussed by Johnson *et al*. [71]. Other methods to generally reduce shipping emission exists and are presented later (See section I.2.4).

Nevertheless, even by taking GHG emissions into account, the overall effect on climate of ship emissions translates to a long term negative RF, contrary to the other main transportation modes as illustrated by Fuglestvedt *et al.* [72]. This negative RF is mainly due to the indirect effect of shipping emissions on clouds properties as illustrated by Eyring *et al.* [73]. Indeed, in their 2009 study, they showed that ship emissions are a cause of modification of the cloud reflectivity. This can be locally identified on satellite images as the 'ship tracks' (or ship plumes) formed as the ship is travelling. In fact, Hobbs *et al.* [74] showed that approximately 12% of exhaust particles above a cargo ship acted as CCN as, when emitted, the SO_2 is oxidized into SO_4 which acts as an efficient CCN [66]. However, this effect on reflectivity also depends on the background aerosol composition. For instance, as shown by Taylor *et al.* [75], if the background aerosol contains a high CCN concentration issued from continental emissions, it is most likely unaffected by ship emissions. Contrary to a cleaner maritime air that would be subjected to an increase in droplet concentration of a factor of two and a reduction in droplet size, increasing the cloud's reflectivity. It is this increased reflectivity that contributes to the indirect negative RF of ship tracks. Ship emissions also affect clouds on a larger scale after the plume 's dispersion. Nonetheless, Liu *et al.* [54] stated that, in the long term, the RF of a given ship emission would eventually become positive given the high airborne lifetime of CO_2 compared to the other components.

2.3 Regulations

With the aim to reduce the impacts of marine transport on global health and climate, regulations are set up by the international maritime organization (IMO) through the marine pollution convention (MARPOL). IMO also defined sensitive regions in which its regulations are even stricter in order to further decrease the emission of target pollutants: SO_x, NO_x and particulate matter. They are called nitrogen emission control area (NECA), sulfur emission control area (SECA) or emission control area (ECA) whether they aim to reduce nitrogen emissions, sulfur emissions or both nitrogen, sulfur and PM emissions respectively. ECA are shown on Figure I-7.

Figure I-7: Global map of existing (dark blue) and potential (light blue) emission control areas (IFPEN 2018)

2.3.1. Regulation of Nitrogen-Containing Pollutants

Rising concerns about the harmful effects on health and climate of ship emissions have led to several limitation policies aiming to enhance global air quality. Emissions from shipping are regulated globally by the Annex VI "Regulations for the Prevention of Air Pollution from Ships" [76] to the MARPOL convention of the IMO. Considering NO_x emission, IMO set up the tier-based regulations[77] as follow:

- Tier 1, applicable on ships constructed after 2000 imposed a reduction of NO_x emissions by 10% compared to ships built before 2000.
- Tier 2 applies on ships built after 2011 and imposes stricter regulation of 15% relative to tier 1 on NO_x emissions.
- Tier 3 is only effective in NECA and is 75% stricter than tier 2.

The impact of such regulations has been studied and estimated. Especially, the case of the greater North Sea region (North Sea, Baltic Sea and the English Channel) were studied with different points of view and considerations. In 2016, Matthias et al [78] predicted that consequences of the Tier 3 NO_x emission regulation on future shipping emissions depend critically on the projected growth of transported volume, the increase in ship number and the share of new ships in the future fleet. In their study, they concluded that NECA would reduce the emission of particulate sulfur and $PM_{2.5}$ but not reduce the emissions of O_3 and particulate

nitrates by 2030 as ship number would increase too fast. However, other studies on the benefit of establishing the NECA in the Baltic Sea, considering a more moderate increase in the number of ships, were published. Indeed, for example, in 2015 Jonson et al [60] estimated that the creation of a NECA in the Baltic Sea would significantly reduce the emissions of NO_x by 2030. Furthermore, in their study, Kalli *et al*. [79] consider the emissions separately for every kind of ship, also considering their respective expected traffic growth and fleet renewal. Their result corroborates the strong decrease in NO_x shipping emissions due to the establishment of a NECA in the greater North Sea region in 2016, *i.e.* by 11% in 2020 and by 79% in 2040. In a more recent study, *i.e.* after the establishment of the NECA, Karl *et al*. [77] estimated that the Baltic sea NECA would lead to 80% less NO_x emissions by 2040.

Regulations on NO_x emissions are based on the engine, on the ship and the year of its construction. Hence, they are targeting ship constructors more than ship operators, contrary to SO_x regulations which are focused upon fuel.

2.3.2. Regulation of Sulfur-Containing Pollutants

Regulations on SO_x emissions are based on fuel sulfur content (FSC). The legislative enforcement created by the IMO in 2020, states that the FSC of shipping fuel is limited to 0.5% (m/m) instead of the previous 3.5% (m/m). A regulation even more strict in SECA, in which the FSC is limited to 0.1% (m/m) [80].

The benefits of these regulations on air quality have been predicted by several studies. Winebrake *et al.* [81] considered three realistic FSC levels: one unregulated (2.7%) and two corresponding to the IMO 2020 policy inside and outside SECA (0.5% and 0.1% respectively). Briefly, they estimated that no control on FSC would lead to 87 000 premature deaths per year linked to $PM_{2.5}$ emissions and that the limits of 0.5% and 0.1% near coastlines would respectively decrease this number substantially by 33 500 and 43 500. They also estimated that a global limit of 0.5%, which corresponds to IMO 2020 new policy, would decrease the annual premature deaths linked to ship emissions by 41 200. In their study, Barregard *et al.* [61] that the implementation of the Baltic Sea SECA reduced by one third the number of premature deaths in 2016, *i.e.* a reduction of more than 1 000 cases. They also estimated that

the cases of stroke and ischemic heart disease due to $PM_{2.5}$ emissions from marine traffic were reduced by one third in 2016 due to the increased regulations. Broome *et al.* [59] studied the impact that a SECA, *i.e.* a limit of 0.1% on FSC, in the region of Sydney would have on the local health. They estimated that a SECA comprising the area within 300 km of the coast would reduce the ship-related ship emissions by 56% and lead to 920 life years gained, *i.e.* a diminution of YLL, over a span of 20 years.

2.4 Air Pollution Reduction Strategies

There are two main approaches to comply with the above-mentioned legal requirements. First, there are the "primary" strategies that concerns the engine itself, this encompasses the use of alternative fuel, adaptions to the combustion process, and exhaust gas recirculation (EGR) [82]. Then, there are the "secondary" methods that aim at improving an already installed engine, such as scrubbers and selective catalytic reduction (SCR) [83]. Here, we will focus on the use of alternative fuels and EGR as primary methods and SCR and scrubbers as secondary methods.

2.4.1. Alternative Fuels

The high proportion of sulphates is a specificity of shipping emissions. In fact, fuels used for marine transport mainly are heavy fuel oils (HFO) with a high fuel sulfur content (FSC). The use of such fuel dates back to the oil crisis of the 1970s and the resulting increase in the price of crude oil middle distillates, like marine diesel oil (MDO) or marine gas oil (MGO). This incited transport companies to use a lower but cheaper cut of crude oil: the heavy fuel oils (or residual fuel oils or bunker fuel). The use of MGO or MDO instead of HFO would greatly reduce shipping emission [84]. Nevertheless, alternative fuels are being investigated to achieve even better results in terms of pollution reduction. These fuels are divided into three classes[85]:

- Low-carbon fuel: liquified natural gas (LNG), liquified petroleum gas (LPG), alcohols
- Carbon-neutral fuels (or bio-sourced fuels): biodiesel or biomethanol,
- Carbon-free fuels: hydrogen and ammonia

Hereafter, we will develop the cases of LNG, methanol, biodiesel, hydrogen, and ammonia.

2.4.1.1. Liquefied Natural Gas

The main advantage to switch from HFO to LNG for marine transportation is that PM, BC, NO_X and CO_2 emissions are reduced by 93%, 97%, 92% and 18% respectively [86]. However, the emission of another GHG is greatly increased: methane, which is the main component of LNG, is able to leak from the engine during the combustion process. The emission of the so-called unburnt fuel may counter the benefits of the reduction of CO_2 emissions [87]. Indeed, CH_4 has a global warming potential over 100 years (GWP) more than 25 times greater than CO_2 [88]. Furthermore, most of the international ports lack the infrastructures that are mandatory to the use of LNG as marine fuel [89], including LNG storage facilities.

2.4.1.2. Methanol and Bio-methanol

Methanol is a sulfur-free fuel with lower exhaust emissions [90]. It can be produced from fossil sources (fossil fuel-based methanol) as well as renewable sources (bio-methanol) [91] and can be used either pure [92] or in blends in dual fuel engines [93]. However, it also has drawbacks such as: the emission of CO_2, a low solubility in diesel fuel; corrosive properties; and a high toxicity to human health [90].

2.4.1.3. Biodiesel

Biodiesel is a renewable fuel prepared from vegetable oil, animal fat or other biomass sources [94]. It is similar to ordinary diesel regarding its physical properties and can be fully mixed with it. Biodiesel also has the same energetic potential than ordinary diesel while emitting less pollutants [89]. The use of such mixture can reduce the emissions of SO_x, PM, hydrocarbons and carbon monoxide [95]. Nevertheless, biodiesel is an oxygenated fuel, thus it emits more NO_x than ordinary diesel fuels under high engine load conditions [96]. This excessive NO_x emission can be reduced using additives like (bio-)alcohols [97], antioxidants [98] or nanoparticles [99]. However, its large-scale applications are slowed by several drawbacks. Indeed, the physical and chemical properties of the biodiesel greatly depend on

the biomass feedstock and chemical reactions used for its production [100]. Additionally, its production cost is greater than the one of fossil fuel-based diesel [101].

2.4.1.4. Carbon-free Ship Fuel Alternatives: Hydrogen and Ammonia

Hydrogen can be used in internal combustion engines [102] or in fuel cells [103]. Even though hydrogen is nowadays principally produced through a CO_2 intensive process, it is also possible to produce it through carbon neutral processes [90]. Hydrogen is a suitable carbon-free energy carrier, but its use presents several disadvantages such as its cost [104], its low volumetric energy density, and its storage/transport conditions [102]. Also, hydrogen is odorless and invisible which induces safety issues [105]. Furthermore, hydrogen is able to penetrate through tanks [106] leading to the formation of an explosive mixture with the atmospheric oxygen.

On the other hand, ammonia is easier to transport and store than hydrogen and also presents a higher volumetric energy density [90]. Similarly, to hydrogen, it can be used either in internal combustion engines [107] or fuel cells [108]. Nonetheless, ammonia is also a colorless toxic gas that can cause severe skin burns and respiratory issues [89]. Thus, also requiring supplementary safety measures.

Nonetheless, most of the existing ships are not able to rely on hydrogen or ammonia alone and must use a co-combustion with another carbon-containing fuel [109]. The emissions from this type of co-combustion are yet largely unknown.

2.4.2. Exhaust Gas Recirculation

Exhaust gas recirculation consists in partially using exhaust gas as the engine air intake. One of its effects is to lower the combustion temperature as the air is more concentrated in higher heat capacity species, *i.e.* CO_2 and H_2O, while oxygen becomes less concentrated in the engine intake manifold. This leads to less NO_x emissions, that are formed at high temperature. Nonetheless, EGR also decrease combustion rate. This translates to higher fuel consumption and soot emissions. Furthermore, especially when using HFO, the high sulfur content of the

exhaust gas can cause corrosion. However, the disadvantages can be balanced by tuning the ratio of recirculated exhaust gas to avoid unacceptably high fuel consumption and soot emission [83]. To overcome the issues of acid recirculating gases and soot emissions, a scrubber can be coupled with the EGR unit [110].

2.4.3. Selective Catalytic Reduction

Another efficient way to reduce NO_x emissions but as a secondary method is the use of selective catalytic reduction. In this case, NO_x are catalytically reduced to nitrogen and water by adding ammonia or another reducing agent, *i.e.* urea or H_2, to the exhaust gas [111]. This technique requires the use of metal oxides as catalysts [89], *i.e.* V_2O_5-WO_3 or TiO_2, and can achieve high conversion rate [112]. However, this technique is subjected to catalyst poisoning [113]. As a secondary method, it is relatively easy to install on a ship, however it also has drawbacks such as the large occupied space [114] or an increase in hydrocarbon and CO emissions [115].

2.4.4. Scrubbers

In a scrubber unit, engine exhaust is sprayed with water to remove pollutants, especially SO_2 from the gas phase (See Figure I-8). In their study, Karjalainen *et al.* [116] used a scrubber on a ship operated in realistic conditions and measured that the scrubber removed 99% of SO_2, making it a very effective unit for sulfur control. Additionally, Yang *et al.* [117] showed that while the NO_x emissions were not reduced by the scrubber during cruising, they decreased largely when the ship was at berth. They also shown that a scrubber unit is mostly ineffective for CO and CO_2 removal. Nevertheless, scrubbers are shown by Fridell *et al*. [118] to be effective for particles removal as they measured a reduction of the particulate matter mass of 75% after the unit. Furthermore, they also evidenced that scrubbers were more effective for the abatement of volatile particles. This study also showed that the reduction of particulate matter is almost equivalent to switching from HFO to MGO. Nonetheless, in their study [119, 120] Jeong *et al.* showed that a scrubber might not remove inhalable particles, therefore requiring the use of a wet electrostatic precipitator (WESP) downstream. In their experiments, the combination of the two removed particles emissions almost entirely.

A scrubber unit has two operating modes: closed-loop and open-loop. Closed-loop refers to the closed circuit used for the water circulation in the unit. In such mode, the water is treated and recycled back to the unit. Thus, this mode requires an on-board water treatment unit using bases. In the open-loop mode however, the water is pumped directly from the sea and only run once through the unit. However, the wastewater is afterwards discharged into the sea [121]. The release, for open-loop operations, of such polluted water in the hydrosphere poses ecological problems [119, 120] while closed-loop operations allow for water treatment once the ship is at berth.

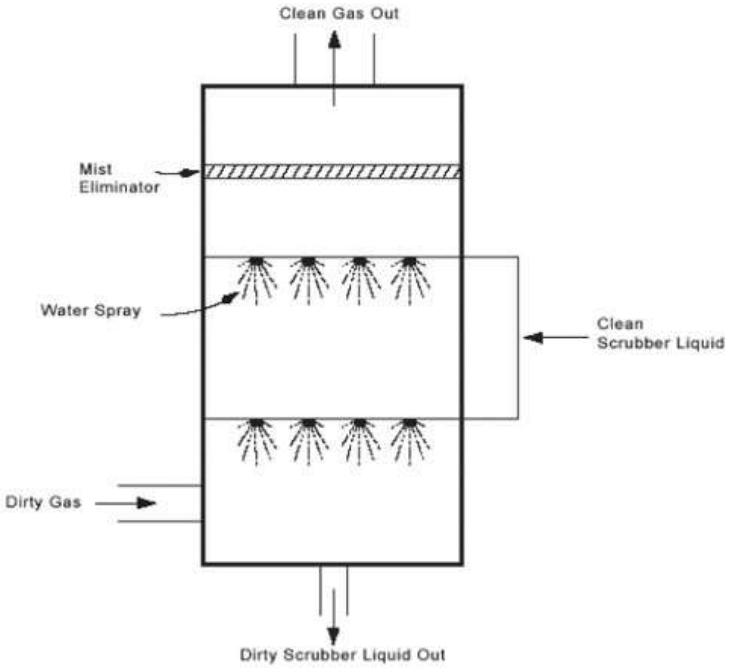


Figure I-8: Schematic representation of a wet scrubber unit [122].

3. Conclusion

Air pollution regulations for ship emissions are monitored on a routine-scale to ensure shipping companies complies with them. However, the study of the in-depths and long-terms consequences of air pollution by ships requires a more thorough investigation. The approach used in this work is to characterize the composition of ship emissions, especially particulate matter, at the molecular level to identify compounds that could be subject to future

regulations. This can be achieved by using high-resolution mass spectrometry that allows to dive in the molecular composition of the studied sample. This technique can be coupled to another powerful instrumental approach called ion mobility spectrometry to provide information on the structure of the molecules identified by mass spectrometry and thus give information on the isomeric complexity. In the chapter II of this thesis, both mass spectrometry and ion mobility spectrometry will be presented as used in this work, as well as the setup used to collect the studied aerosol samples.

4. References

1. Kulkarni, P., P.A. Baron, and K. Willeke, eds. *Aerosol Measurement - Principles, Techniques, and Applications, Third Edition*. 2011.

2. Schnelle-Kreis, J., et al., *Atmosphärische Aerosole: Quellen, Vorkommen, Zusammensetzung.* Chemie in unserer Zeit, 2007. **41**(3): p. 220-230.

3. Zhang, R., et al., *Morphology and property investigation of primary particulate matter particles from different sources.* Nano Research, 2018. **11**(6): p. 3182-3192.

4. Lewis, E.R. and S.E. Schwartz, *Sea Salt Aerosol Production: Mechanisms, Methods, Measurements and Models—A Critical Review*. Geophysical Monograph Series. 2004: American Geophysical Union.

5. Fuzzi, S., et al., *Particulate matter, air quality and climate: lessons learned and future needs.* Atmos. Chem. Phys., 2015. **15**(14): p. 8217-8299.

6. Seinfeld, J.H. and S.N. Pandis, *Atmospheric chemistry and physics: from air pollution to climate change*. 2016: John Wiley & Sons.

7. Williams, J., et al., *Application of the variability-size relationship to atmospheric aerosol studies: estimating aerosol lifetimes and ages.* Atmospheric Chemistry and Physics, 2002. **2**(2): p. 133-145.

8. Zimmermann, R., *Organische Verbindungen in Feinstaub.* Nachrichten aus der Chemie, 2006. **54**(6): p. 676-680.

9. Chow, J.C., et al., *Mass reconstruction methods for PM(2.5): a review.* Air Qual Atmos Health, 2015. **8**(3): p. 243-263.

10. Andreae, M.O. and A. Gelencsér, *Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols.* Atmospheric Chemistry and Physics, 2006. **6**(10): p. 3131-3148.

11. Pio, C., et al., *OC/EC ratio observations in Europe: Re-thinking the approach for apportionment between primary and secondary organic carbon.* Atmospheric Environment, 2011. **45**(34): p. 6121-6132.

12. Poschl, U., *Atmospheric aerosols: composition, transformation, climate and health effects.* Angew Chem Int Ed Engl, 2005. **44**(46): p. 7520-40.

13. Jimenez, J.L., et al., *Evolution of Organic Aerosols in the Atmosphere.* 2009. **326**(5959): p. 1525-1529.

14. Volkamer, R., et al., *Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected.* Geophysical Research Letters, 2006. **33**(17).

15. Herndon, S.C., et al., *Correlation of secondary organic aerosol with odd oxygen in Mexico City.* Geophysical Research Letters, 2008. **35**(15).

16. Donahue, N.M., et al., *Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics.* Environmental Science & Technology, 2006. **40**(8): p. 2635-2643.

17. Donahue, N.M., et al., *A two-dimensional volatility basis set – Part 2: Diagnostics of organic-aerosol evolution.* Atmospheric Chemistry and Physics, 2012. **12**(2): p. 615-634.

18. Atkinson, R., *Atmospheric chemistry of VOCs and NOx.* Atmospheric Environment, 2000. **34**(12): p. 2063-2101.

19. Kroll, J.H. and J.H. Seinfeld, *Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere.* Atmospheric Environment, 2008. **42**(16): p. 3593-3624.

20. Chacon-Madrid, H.J. and N.M. Donahue, *Fragmentation vs. functionalization: chemical aging and organic aerosol formation.* Atmospheric Chemistry and Physics, 2011. **11**(20): p. 10553-10563.

21. Ehhalt, D.H., *Photooxidation of trace gases in the troposphere Plenary Lecture.* Physical Chemistry Chemical Physics, 1999. **1**(24): p. 5401-5408.

22. Bufalini, J.J., et al., *Ozone formation potential of organic compounds.* 1976. **10**(9): p. 908-912.

23. Criegee, R., *Die Ozonolyse.* Chemie in unserer Zeit, 1973. **7**(3): p. 75-81.

24. Zhang, X., et al., *Role of ozone in SOA formation from alkane photooxidation.* Atmospheric Chemistry and Physics, 2014. **14**(3): p. 1733-1753.

25. Kroll, J.H., et al., *Heterogeneous Oxidation of Atmospheric Organic Aerosol: Kinetics of Changes to the Amount and Oxidation State of Particle-Phase Organic Carbon.* The Journal of Physical Chemistry A, 2015. **119**(44): p. 10767-10783.

26. Ringuet, J., et al., *Reactivity of polycyclic aromatic compounds (PAHs, NPAHs and OPAHs) adsorbed on natural aerosol particles exposed to atmospheric oxidants.* Atmospheric Environment, 2012. **61**: p. 15-22.

27. Hennigan, C.J., et al., *Chemical and physical transformations of organic aerosol from the photo-oxidation of open biomass burning emissions in an environmental chamber.* Atmospheric Chemistry and Physics, 2011. **11**(15): p. 7669-7686.

28. Farmer, D.K., et al., *Response of an aerosol mass spectrometer to organonitrates and organosulfates and implications for atmospheric chemistry.* Proceedings of the National Academy of Sciences, 2010. **107**(15): p. 6670-6675.

29. Ng, N.L., et al., *Changes in organic aerosol composition with aging inferred from aerosol mass spectra.* Atmospheric Chemistry and Physics, 2011. **11**(13): p. 6465-6474.

30. Trivitayanurak, W. and P.J. Adams, *Does the POA–SOA split matter for global CCN formation?* Atmospheric Chemistry and Physics, 2014. **14**(2): p. 995-1010.

31. Trump, E.R., et al., *Wall effects in smog chamber experiments: A model study.* Aerosol Science and Technology, 2016. **50**(11): p. 1180-1200.

32. Wang, N., et al., *Particle wall-loss correction methods in smog chamber experiments.* Atmospheric Measurement Techniques, 2018. **11**(12): p. 6577-6588.

33. Czech, H., et al., *Wildfire plume ageing in the Photochemical Large Aerosol Chamber (PHOTO-LAC).* Environmental Science: Processes & Impacts, 2023.

34. Kang, E., et al., *Introducing the concept of Potential Aerosol Mass (PAM).* Atmos. Chem. Phys., 2007. **7**(22): p. 5727-5744.

35. Keller, A. and H. Burtscher, *A continuous photo-oxidation flow reactor for a defined measurement of the SOA formation potential of wood burning emissions.* Journal of Aerosol Science, 2012. **49**: p. 9-20.

36. Ihalainen, M., et al., *A novel high-volume Photochemical Emission Aging flow tube Reactor (PEAR).* Aerosol Science and Technology, 2019. **53**(3): p. 276-294.

37. Leskinen, J., et al., *Photochemical Aging Induces Changes in the Effective Densities, Morphologies, and Optical Properties of Combustion Aerosol Particles.* Environmental Science & Technology, 2023. **57**(13): p. 5137-5148.

38. Houghton, J.E.T., et al., *Climate Change 2001: The Scientific Basis*. 2001. p. 881.

39. Andreae, M.O., *The dark side of aerosols.* Nature, 2001. **409**(6821): p. 671-672.

40. Boucher, O., et al., *Clouds and aerosols*, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker, et al., Editors. 2013, Cambridge University Press: Cambridge, UK. p. 571-657.

41. Pope, C.A., et al., *Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution.* Circulation, 2004. **109**(1): p. 71-77.

42. Gauderman, W.J., et al., *The Effect of Air Pollution on Lung Development from 10 to 18 Years of Age.* New England Journal of Medicine, 2004. **351**(11): p. 1057-1067.

43. Bernstein, J.A., et al., *Health effects of air pollution.* Journal of Allergy and Clinical Immunology, 2004. **114**(5): p. 1116-1123.

44. Bell, M.L. and D.L. Davis, *Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution.* 2001. **109**(suppl 3): p. 389-394.

45. WHO. [cited 2024; Available from: https://www.who.int/data/gho/data/themes/air-pollution?lang=en.

46. Heyder, J., *Deposition of inhaled particles in the human respiratory tract and consequences for regional targeting in respiratory drug delivery.* Proc Am Thorac Soc, 2004. **1**(4): p. 315-20.

47. Oberdörster, G., et al., *Translocation of Inhaled Ultrafine Particles to the Brain.* Inhalation Toxicology, 2004. **16**(6-7): p. 437-445.

48. Nemmar, A., et al., *Passage of Inhaled Particles Into the Blood Circulation in Humans.* Circulation, 2002. **105**(4): p. 411-414.

49. Uski, O., et al., *Effect of fuel zinc content on toxicological responses of particulate matter from pellet combustion in vitro.* Science of The Total Environment, 2015. **511**: p. 331-340.

50. Claxton, L.D., *The history, genotoxicity, and carcinogenicity of carbon-based fuels and their emissions. Part 3: Diesel and gasoline.* Mutation Research/Reviews in Mutation Research, 2015. **763**: p. 30-85.

51. Schnurr, R.E.J. and T.R. Walker, *Marine Transportation and Energy Use ☆*, in *Reference Module in Earth Systems and Environmental Sciences*. 2019, Elsevier.

52. Serra, P. and G. Fancello, *Towards the IMO's GHG Goals: A Critical Overview of the Perspectives and Challenges of the Main Options for Decarbonizing International Shipping.* Sustainability, 2020. **12**(8): p. 3220.

53. *IMO 2014: Third IMO GHG Study*. Available from: https://www.imo.org/en/ourwork/environment/pages/greenhouse-gas-studies-2014.aspx.

54. Liu, H., et al., *Health and climate impacts of ocean-going vessels in East Asia.* Nature Climate Change, 2016. **6**(11): p. 1037-1041.

55. Chang, C.-C. and C.-M. Wang, *Evaluating the effects of green port policy: Case study of Kaohsiung harbor in Taiwan.* Transportation Research Part D: Transport and Environment, 2012. **17**(3): p. 185-189.

56. Contini, D. and E. Merico, *Recent Advances in Studying Air Quality and Health Effects of Shipping Emissions.* Atmosphere, 2021. **12**(1): p. 92.

57. Corbett, J.J., et al., *Mortality from Ship Emissions: A Global Assessment.* Environmental Science & Technology, 2007. **41**(24): p. 8512-8518.

58. Cohen, A.J., et al., *The Global Burden of Disease Due to Outdoor Air Pollution.* Journal of Toxicology and Environmental Health, Part A, 2005. **68**(13-14): p. 1301-1307.

59. Broome, R.A., et al., *The mortality effect of ship-related fine particulate matter in the Sydney greater metropolitan region of NSW, Australia.* Environment International, 2016. **87**: p. 85-93.

60. Jonson, J.E., et al., *Model calculations of the effects of present and future emissions of air pollutants from shipping in the Baltic Sea and the North Sea.* Atmospheric Chemistry and Physics, 2015. **15**(2): p. 783-798.

61. Barregard, L., et al., *Impact on Population Health of Baltic Shipping Emissions.* International Journal of Environmental Research and Public Health, 2019. **16**(11): p. 1954.

62. Wright, A.A., *Exhaust emissions from combustion machinery*. Vol. 13. 2000. 212-226.

63. Boningari, T. and P.G. Smirniotis, *Impact of nitrogen oxides on the environment and human health: Mn-based materials for the NOx abatement.* Current Opinion in Chemical Engineering, 2016. **13**: p. 133-141.

64. Stipa, T., et al., *Emissions of NOx from Baltic shipping and first estimates of their effects on air quality and eutrophication of the Baltic Sea*. 2009.

65. Myhre, G., et al., *Radiative forcing due to changes in ozone and methane caused by the transport sector.* Atmospheric Environment, 2011. **45**(2): p. 387-394.

66. *Clouds and Aerosols*, in *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. Intergovernmental Panel on Climate, Editor. 2014, Cambridge University Press: Cambridge. p. 571-658.

67. Bond, T.C., et al., *Bounding the role of black carbon in the climate system: A scientific assessment.* Journal of Geophysical Research: Atmospheres, 2013. **118**(11): p. 5380-5552.

68. Styhre, L., et al., *Greenhouse gas emissions from ships in ports – Case studies in four continents.* Transportation Research Part D: Transport and Environment, 2017. **54**: p. 212-224.

69. Winkel, R., et al., *Shore Side Electricity in Europe: Potential and environmental benefits.* Energy Policy, 2016. **88**: p. 584-593.

70. Black, J. and L. Styhre. *Environmental conflicts in port cities: A case study of Port Jackson and Port Botany in Metropolitan Sydney*. in *World Conference on Transport Research, Special Interest Group Meeting, Antwerp*. 2015.

71. Johnson, H. and L. Styhre, *Increased energy efficiency in short sea shipping through decreased time in port.* Transportation Research Part A: Policy and Practice, 2015. **71**: p. 167-178.

72. Fuglestvedt, J., et al., *Climate forcing from the transport sectors.* 2008. **105**(2): p. 454-458.

73. Eyring, V., et al., *Transport impacts on atmosphere and climate: Shipping.* Atmospheric Environment, 2010. **44**(37): p. 4735-4771.

74. Hobbs, P.V., et al., *Emissions from Ships with respect to Their Effects on Clouds.* Journal of the Atmospheric Sciences, 2000. **57**(16): p. 2570-2590.

75. Taylor, J.P., et al., *Effects of Aerosols on the Radiative Properties of Clouds.* Journal of the Atmospheric Sciences, 2000. **57**(16): p. 2656-2670.

76. *MARPOL Annex IV 2008a*. Available from: http://www.marpoltraining.com/MMSKOREAN/MARPOL/Annex_IV/index.htm.

77. Karl, M., et al., *Impact of a nitrogen emission control area (NECA) on the future air quality and nitrogen deposition to seawater in the Baltic Sea region.* Atmos. Chem. Phys., 2019. **19**(3): p. 1721-1752.

78. Matthias, V., et al., *The impact of shipping emissions on air pollution in the greater North Sea region – Part 2: Scenarios for 2030.* Atmospheric Chemistry and Physics, 2016. **16**(2): p. 759-776.

79. Kalli, J., et al., *Atmospheric emissions of European SECA shipping: long-term projections.* WMU Journal of Maritime Affairs, 2013. **12**(2): p. 129-145.

80. *IMO 2020*. Available from: https://www.imo.org/en/MediaCentre/HotTopics/Pages/Sulphur-2020.aspx.

81. Winebrake, J.J., et al., *Mitigating the Health Impacts of Pollution from Oceangoing Shipping: An Assessment of Low-Sulfur Fuel Mandates.* Environmental Science & Technology, 2009. **43**(13): p. 4776-4782.

82. Lamas, M. and C. Rodriguez, *Emissions from marine engines and NOx reduction methods.* Journal of maritime research, 2012. **9**(1): p. 77-81.

83. Lion, S., I. Vlaskos, and R. Taccani, *A review of emissions reduction technologies for low and medium speed marine Diesel engines and their potential for waste heat recovery.* Energy Conversion and Management, 2020. **207**: p. 112553.

84. Winnes, H. and E. Fridell, *Particle Emissions from Ships: Dependence on Fuel Type.* Journal of the Air & Waste Management Association, 2009. **59**(12): p. 1391-1398.

85. Wang, Y., et al., *A review of low and zero carbon fuel technologies: Achieving ship carbon reduction targets.* Sustainable Energy Technologies and Assessments, 2022. **54**: p. 102762.

86. Peng, W., et al., *Comprehensive analysis of the air quality impacts of switching a marine vessel from diesel fuel to natural gas.* Environmental Pollution, 2020. **266**: p. 115404.

87. Lindstad, E., et al., *Decarbonizing Maritime Transport: The Importance of Engine Technology and Regulations for LNG to Serve as a Transition Fuel.* Sustainability, 2020. **12**(21): p. 8793.

88. Solomon, S., *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*. Vol. 4. 2007: Cambridge university press.

89. Zhao, J., et al., *Progress of ship exhaust gas control technology.* Science of The Total Environment, 2021. **799**: p. 149437.

90. Xing, H., et al., *Alternative fuel options for low carbon maritime transportation: Pathways to 2050.* Journal of Cleaner Production, 2021. **297**: p. 126651.

91. Svanberg, M., et al., *Renewable methanol as a fuel for the shipping industry.* Renewable and Sustainable Energy Reviews, 2018. **94**: p. 1217-1228.

92. Leng, X., et al., *A preliminary numerical study on the use of methanol as a Mono-Fuel for a large bore marine engine.* Fuel, 2022. **310**: p. 122309.

93. Paulauskiene, T., M. Bucas, and A. Laukinaite, *Alternative fuels for marine applications: Biomethanol-biodiesel-diesel blends.* Fuel, 2019. **248**: p. 161-167.

94. Issariyakul, T. and A.K. Dalai, *Biodiesel from vegetable oils.* Renewable and Sustainable Energy Reviews, 2014. **31**: p. 446-471.

95. Mohd Noor, C.W., M.M. Noor, and R. Mamat, *Biodiesel as alternative fuel for marine diesel engine applications: A review.* Renewable and Sustainable Energy Reviews, 2018. **94**: p. 127-142.

96. Devarajan, Y., B.K. Nagappan, and D.B. Munuswamy, *Performance and emissions analysis on diesel engine fuelled with cashew nut shell biodiesel and pentanol blends.* Korean Journal of Chemical Engineering, 2017. **34**(4): p. 1021-1026.

97. Nour, M., et al., *Adding n-butanol, n-heptanol, and n-octanol to improve vaporization, combustion, and emission characteristics of diesel/used frying oil biodiesel blends in <scp>DICI</scp> engine.* Environmental Progress & Sustainable Energy, 2021. **40**(3).

98. Manikandan, K., J. Chandran, and R. Ganesh, *Study on NOx and smoke emission reduction techniques in biodiesel fuelled research engine.* International Journal of Ambient Energy, 2020. **41**(14): p. 1604-1607.

99. Venu, H., et al., *Influence of Al2O3nano additives in ternary fuel (diesel-biodiesel-ethanol) blends operated in a single cylinder diesel engine: Performance, combustion and emission characteristics.* Energy, 2021. **215**: p. 119091.

100. Santos, B.S. and S.C. Capareda, *A Comparative Study on the Engine Performance and Exhaust Emissions of Biodiesel from Various Vegetable Oils and Animal Fat.* Journal of Sustainable Bioenergy Systems, 2015. **05**(03): p. 89-103.

101. Gebremariam, S.N. and J.M. Marchetti, *Economics of biodiesel production: Review.* Energy Conversion and Management, 2018. **168**: p. 74-84.

102. Manigandan, S., et al., *Comparative study of nanoadditives TiO2, CNT, Al2O3, CuO and CeO2 on reduction of diesel engine emission operating on hydrogen fuel blends.* Fuel, 2020. **262**: p. 116336.

103. Wang, Y., et al., *Materials, technological status, and fundamentals of PEM fuel cells – A review.* Materials Today, 2020. **32**: p. 178-203.

104. Crabtree, G.W. and M.S. Dresselhaus, *The Hydrogen Fuel Alternative.* MRS Bulletin, 2008. **33**(4): p. 421-428.

105. Mazloomi, K. and C. Gomes, *Hydrogen as an energy carrier: Prospects and challenges.* Renewable and Sustainable Energy Reviews, 2012. **16**(5): p. 3024-3033.

106. Su, Y., et al., *Review of the Hydrogen Permeability of the Liner Material of Type IV On-Board Hydrogen Storage Tank.* 2021. **12**(3): p. 130.

107. Rehbein, M.C., et al., *Mixtures of Ammonia and Organic Solvents as Alternative Fuel for Internal Combustion Engines.* Energy & Fuels, 2019. **33**(10): p. 10331-10342.

108. Jeerh, G., M. Zhang, and S. Tao, *Recent progress in ammonia fuel cells and their potential applications.* Journal of Materials Chemistry A, 2021. **9**(2): p. 727-752.

109. Dimitriou, P. and T. Tsujimura, *A review of hydrogen as a compression ignition engine fuel.* International Journal of Hydrogen Energy, 2017. **42**(38): p. 24470-24486.

110. Hansen, J.P., et al., *Reduction of SO2, NOx and Particulate Matters from Ships with Diesel Engines*. 2014: Miljøstyrelsen.

111. Kim, H.-S., et al., *Current Catalyst Technology of Selective Catalytic Reduction (SCR) for NOx Removal in South Korea.* 2020. **10**(1): p. 52.

112. Tân, P.H. and N.M. Tri, *Solution to Reduce Air Environmental Pollution from Ships* TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation, 2015. **9**(2): p. 257-261.

113. Lisi, L. and S. Cimino *Poisoning of SCR Catalysts by Alkali and Alkaline Earth Metals*. Catalysts, 2020. **10**, DOI: 10.3390/catal10121475.

114. Lu, X., P. Geng, and Y. Chen, *NOx Emission Reduction Technology for Marine Engine Based on Tier-III: A Review.* Journal of Thermal Science, 2020. **29**(5): p. 1242-1268.

115. Ezhilnilavan, R. and A. Kajavali, *A comparative study of six-cylinder direct injection compression ignition engine fueled with Argemone Mexicana bio-diesel with and without selective catalytic reduction.* Energy Sources, Part A: Recovery, Utilization, and Environmental Effects: p. 1-18.

116. Karjalainen, P., et al., *Real-world particle emissions and secondary aerosol formation from a diesel oxidation catalyst and scrubber equipped ship operating with two fuels in a SECA area.* Environmental Pollution, 2022. **292**: p. 118278.

117. Yang, J., et al., *Controlling emissions from an ocean-going container vessel with a wet scrubber system.* Fuel, 2021. **304**: p. 121323.

118. Fridell, E. and K. Salo, *Measurements of abatement of particles and exhaust gases in a marine gas scrubber.* 2016. **230**(1): p. 154-162.

119. Endres, S., et al., *A New Perspective at the Ship-Air-Sea-Interface: The Environmental Impacts of Exhaust Gas Scrubber Discharge.* 2018. **5**.

120. Koski, M., C. Stedmon, and S. Trapp, *Ecological effects of scrubber water discharge on coastal plankton: Potential synergistic effects of contaminants reduce survival and feeding of the copepod Acartia tonsa.* Marine Environmental Research, 2017. **129**: p. 374-385.

121. Tran, T.A., *Research of the Scrubber Systems to Clean Marine Diesel Engine Exhaust Gases on Ships.* Journal of Marine Science: Research & Development, 2017. **07**(06).

122. EPA. [cited 2024; Available from: https://www.epa.gov/air-emissions-monitoring-knowledge-base/monitoring-control-technique-wet-scrubber-particulate-0.

Chapter II. Sampling and Instrumentation

1. Ship engine and particle sampling setup

In this work, realistic ship emissions particle samples were studied. These PM samples were obtained during several measurement campaigns that took place at the University of Rostock, Germany. For the campaigns, an experimental ship diesel engine was used to burn different marine fuels under variable operating conditions to mimic real-life conditions of a ship engine's use. The two types of fuel used in the campaigns are HFO and MGO, *i.e.* a heavy fuel with high sulfur content and a cleaner fuel complying with the IMO 2020 regulations on FSC. The engine load, *i.e.* the power output of the engine, was an important parameter as it allowed to simulate different running states of the ship that usually depend on its distance from the coast: a low engine load is used for maneuvering operations in harbors while a high engine load is used when the ship is cruising in open sea. A typical setup of one of the measurement campaigns is described in the work of Jeong *et al.* [1] and shown in Figure II-1.

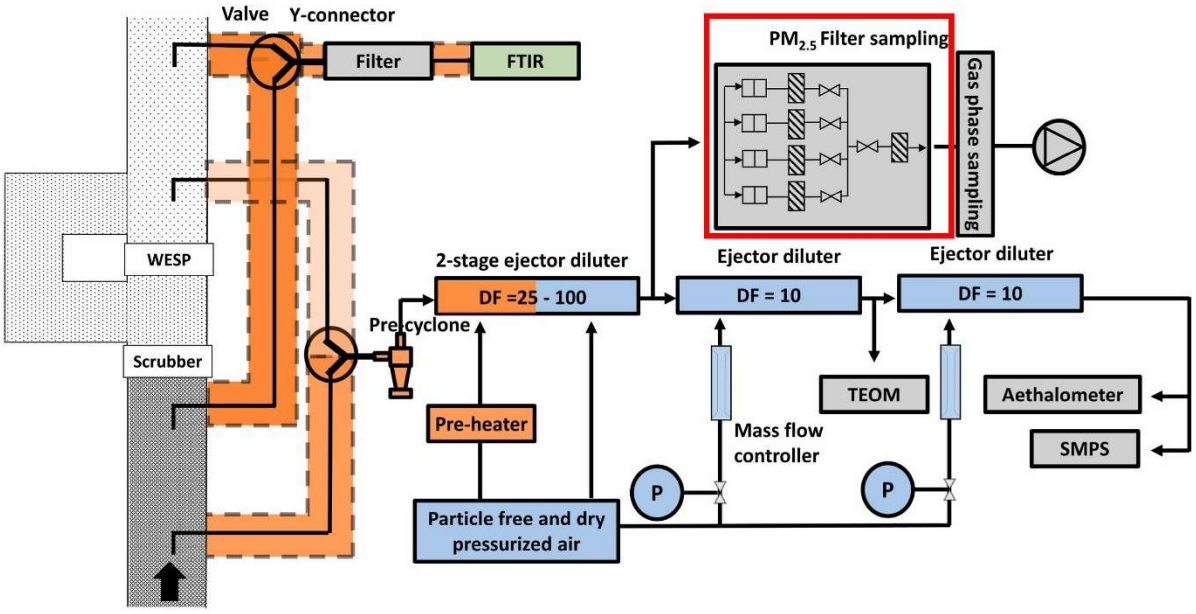


Figure II-1: Scheme of the sampling setup used for a measurement campaign at the University of Rostock. DF: dilution factor, FTIR: Fourier transform infrared spectroscopy, P: pressure regulator, SMPS: Scanning mobility particle sizer, TEOM: Tapered element oscillating microbalance, WESP: wet electrostatic precipitator.

In the figure, the exhaust comes from the engine at the black arrow at the bottom left corner. The setup used for PM sampling is highlighted by the red rectangle and consists of four filter holders connected to a pump that allows a continuous and stable exhaust flow to pass

through the filters. Furthermore, a flowmeter is also connected to this setup to monitor the experiment. The PM was collected on pre-baked quartz fiber filters and stored under -20°C until analysis. The particulate matter is then studied by mass spectrometry, either as is or after a solid-liquid extraction step allowing it to target a specific category of compounds, such as polar or apolar species.

2. Mass spectrometry

2.1 General Principle

Mass spectrometry is an analytical chemistry technique used in numerous research fields for its selectivity, sensitivity, and precision. This technique involves the ionization of molecules and the separation of the formed ions according to their mass-to-charge ratio (*m/z*). The versatility of mass spectrometry lies in the fact that it directly measures an intrinsic propriety of the studied molecule: the molecular mass. Mass spectrometry can be used for identification, structural analysis, or quantification of molecules.

The first mass spectrometer, *i.e.* the first instrument capable of separating ions along their *m/*z ratio, was built by Joseph John Thomson in the early 1900s [2-4]. Since then, this technique has undergone huge developments, with improvements in sensitivity, resolution, and mass accuracy allowing the analysis of complex samples.

As shown in Figure II-2, a mass spectrometer can be divided into three parts: an ion source, an analyzer, and a detector. Depending on its nature, its properties, and the aim of the study, the sample can either be directly introduced into the instrument or using a coupling with a separative technique such as liquid chromatography, gas chromatography, etc... The ion source plays a key role in mass spectrometry as it allows to form the ions to be detected but also conditions what type of molecules will be ionized in the sample. The mass analyzer is the heart of the instrument as it allows the separation of the generated ions along their *m/z* ratio using a magnetic and/or electrical field. Finally, the detector produces an electrical signal that is a function of the number of ions detected for a given *m/z*. This signal is then amplified and computed to yield a mass spectrum.

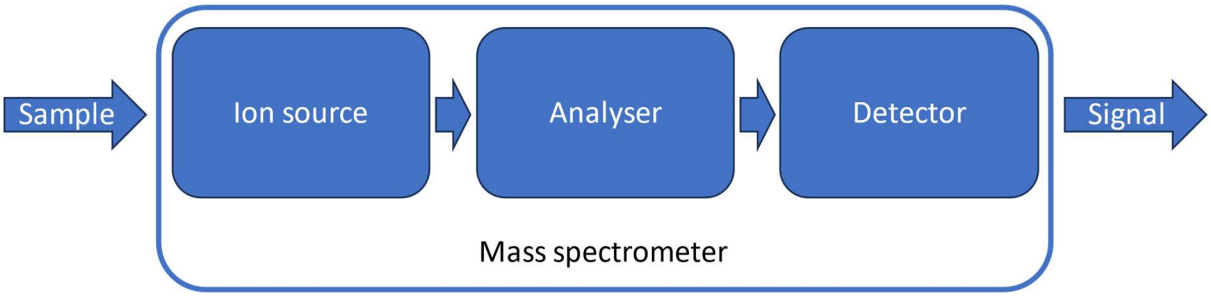


Figure II-2: Schematic diagram of a mass spectrometer

During this thesis, we mainly used four ion sources: electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), atmospheric pressure photoionization (APPI), and matrix-assisted laser desorption-ionization (MALDI). The analyzers used in our work are the time-of-flight (TOF) and the Fourier transform ion cyclotron resonance (FTICR)

2.2 Ion Sources

As stated above, ionization is a key step for MS analysis. During this step, analyte molecules are charged, either positively or negatively, and transferred to the gas phase following several mechanisms and yielding different types of ions. Indeed, the analytes species [M] can be ionized by electron capture [$M^{-\bullet}$] or electron ejection [$M^{+\bullet}$], resulting in radical anions and cations, respectively. Analytes can also be protonated [$M+H]^+$ or deprotonated [$M-H]^-$ Besides, the analytes can form adducts with ions such as sodium [$M+Na]^+$ or chloride [$M+Cl]^-$. As the ionization mechanism and consequently its selectivity depends on the chosen ion source, the latter must be chosen among the great diversity of sources according to the properties of the sample and the goal of the study. The main properties to be discriminated by the ionization step are the polarity, the aromaticity, the acidity (or basicity), and the molecular mass. The selectivity of the most common ion sources along the polarity and

molecular mass of the analyte is shown in Figure II-3. The selectivity of these sources, as well as their principle, is discussed below.

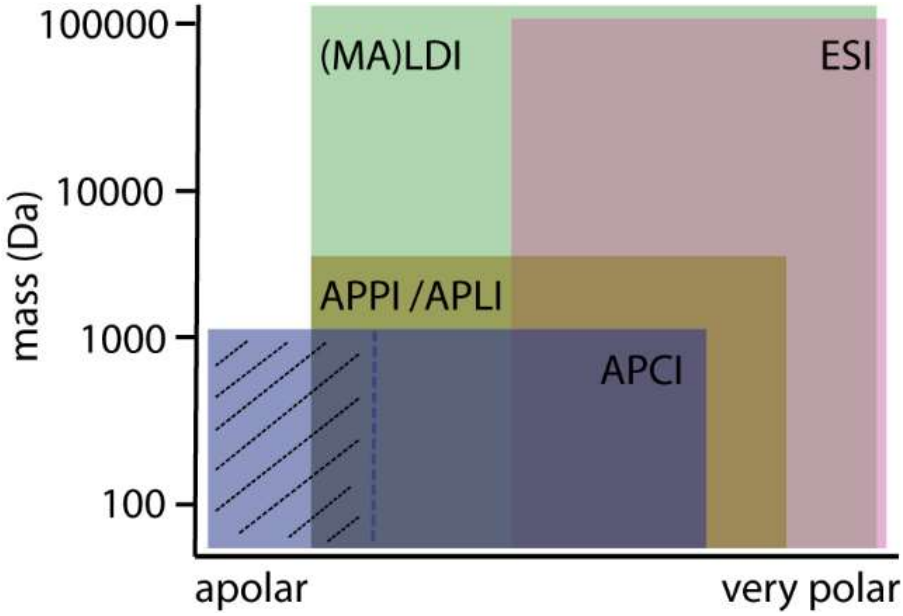


Figure II-3: Selectivity of the most-used ion sources. The striped area shows the extended selectivity of APCI when using hydrocarbonated solvents. APCI: atmospheric pressure chemical ionization, APPI: atmospheric pressure photoionization, APLI: atmospheric pressure laser ionization, MALDI: matrix-assisted laser desorption/ionization, ESI: electrospray ionization. Adapted from our submitted review (Appendix 2).

2.2.1. Electrospray ionization (ESI)

Though it was conceptualized in 1917 by John Zeleny [5], ESI was only coupled to MS several years later by Malcolm Dole [6] in 1968 and then by John B. Fenn [7, 8] in the mid-1980s. An analog of ESI, named solvated ion extraction at atmospheric pressure (known by the Russian acronym ERIAD), was, in parallel, developed by Lidija N. Gall [9] This ionization technique operates at atmospheric pressure and is thus qualified as an atmospheric pressure ion source (API). ESI is also considered a soft ionization technique that leads to almost no fragmentation of the analyte molecules and allows for the formation of multicharged ions. These two specificities make ESI a technique of choice for the study of fragile macromolecules such as proteins. Nevertheless, ESI is also commonly used for the characterization of small polar molecules.

In ESI, the analytes in solution are introduced via a capillary tube, as shown in Figure II-4. The spraying nozzle and the counter electrode are subjected to a high potential difference

in the range of the kilovolt. The mechanisms and underlying phenomena in ESI are not all completely known and still debated [10]. Briefly, the high electric field at the spraying tip causes an accumulation of charges on the liquid's surface at the edge of the nozzle which leads to repulsion. The surface tension of the liquid counterbalances the repulsion force forming a Taylor cone at the edge of the nozzle, pointing towards the counter electrode. At the tip of the Taylor cone, electrostatic forces overcome the tension surface forces leading to the formation of a charged droplets spray. A counter-current flow of inert and heated gas (typically N_2) leads to the progressive evaporation of the solvent in the charged droplets whose size progressively decreases until they reach the Rayleigh limit at which the electrostatic repulsion overcomes the tension surface. At the Rayleigh limit, a Coulomb fission occurs, forming even smaller charged droplets [11]. In one interpretation, known as the charged residue model [12] this process is repeated until the obtention of "naked" charged analytes in the gas phase that are guided to the mass analyzer. Depending on the polarity of the high voltage tension between the nozzle and the counter electrode, ESI produces cations (positive mode) or anions (negative mode) that can be singly or multiply charged.

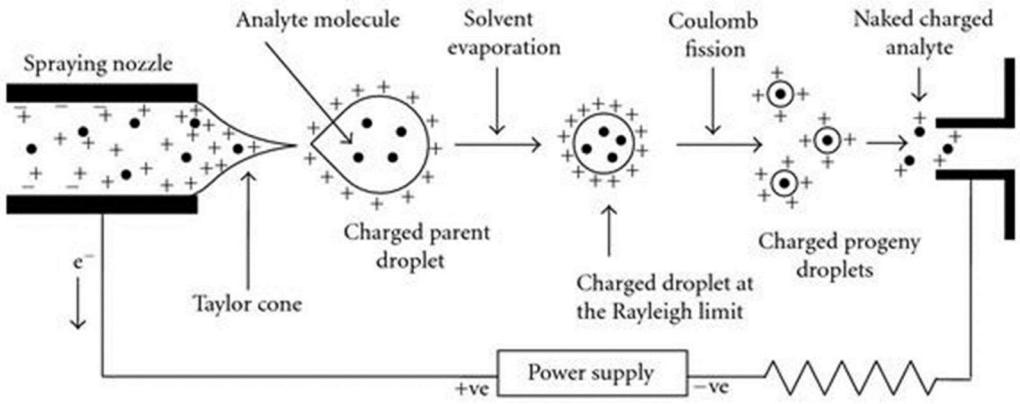


Figure II-4: Schematic representation of an ESI source in positive mode [13].

In ESI, the ionization mostly takes place in solution or in nanodroplets due to an acid-base equilibrium between the analytes and the solvent [14]. Thus, electrospray ionization exhibits a selectivity towards basic/acidic molecules and the use of a protic solvent (*e.g.* H_2O, MeOH) increases its sensitivity [15]. Furthermore, it is possible to add species, *e.g.* formic acid or ammonium hydroxide, to shift the acid-base equilibrium and increase the sensitivity for

either acidic or basic compounds [16]. ESI is prone to form adduct ions such as $[M+Na]^+$ or $[M+Cl]^-$ as soon as salts are present in the solution; hence, it is possible to promote the formation of such ions by adding salts such as sodium acetate. If the ESI source can ionize a wide range of molecules in terms of mass, it is however only applicable to soluble molecules and is selective of polar compounds as represented on Figure II-3 [17]. ESI is sometimes prone to ion suppression through matrix effects [18].

2.2.2. Atmospheric pressure chemical ionization (APCI)

APCI was introduced in 1973 by Horning *et al*. [19, 20]. It is based on the CI previously developed by Munson and Field in 1966 [21]. Contrary to ESI, the ionization takes place in the gas phase. Use of ^{63}Ni blade as electron source initially but then, the corona discharge, using a needle, was preferred [22]. A schematic representation of the APCI source is shown in Figure II-5

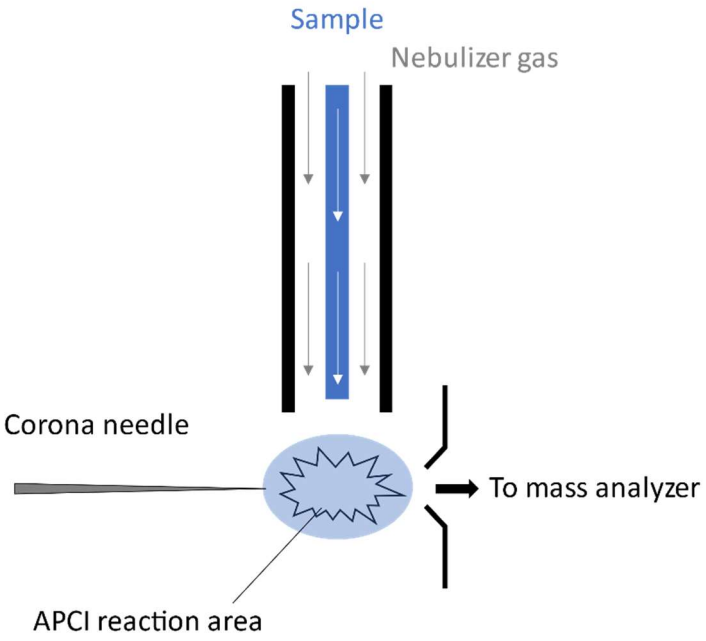


Figure II-5: Scheme of the APCI source

In this source, the solution is vaporized using a heated N_2 flow, then subjected to a corona discharge via a needle that leads to the ionization of the nebulizing gas which transfers its charge to the solvent and analytes molecules [22, 23]. Then the ionization of the nitrogen takes place, as shown in Equations II.1-II.2. In APCI's positive mode, it is possible to form radical

ions [$M^{+\bullet}$] following charge exchange (Equation II.3) in the absence of a protic solvent. However, in the presence of a protic solvent or moisture, the reactive ions of the nitrogen plasma, *i.e.* $N_2^{+\bullet}$, and $N_4^{+\bullet}$, can collide with solvent molecules to form secondary ions (Equation II.4-II.7) to form water clusters. These clusters can then exchange a proton with the analyte to form a protonated ion by following Equation II.8

$$N_2 + e^- \rightarrow N_2^{+\bullet} + 2e^- \quad \text{II.1}$$

$$N_2^{+\bullet} + 2N_2 \rightarrow N_4^{+\bullet} \quad \text{II.2}$$

$$N_2^{+\bullet} + M \rightarrow M^{+\bullet} + N_2 \quad \text{II.3}$$

$$N_4^{+\bullet} + H_2O \rightarrow H_2O^{+\bullet} + 2N_2 \quad \text{II.4}$$

$$H_2O^{+\bullet} + H_2O \rightarrow H_3O^+ + OH^\bullet \quad \text{II.5}$$

$$H_3O^+ + H_2O + N_2 \rightarrow H_3O^+(H_2O) + N_2 \quad \text{II.6}$$

$$H_3O^+(H_2O)_{n-1} + H_2O + N_2 \rightarrow H_3O^+(H_2O)_n + N_2 \quad \text{II.7}$$

$$H_3O^+ + M \rightarrow [M+H]^+ + H_2O \quad \text{II.8}$$

Hence, in APCI's positive mode, it is possible to form either radical ions [$M^{+\bullet}$] or protonated ions $[M+H]^+$, the ratio of abundance between the two depends on the solvent used. Using an aprotic solvent will favor the charge exchange phenomena, yielding mainly radical ions. On the other hand, the use of a protic solvent will favor the proton exchange phenomena if the proton affinity of the analyte is higher than the proton affinity of water.

A different type of cation is produced when small hydrocarbons are used as solvent: $[M-H]^+$. For years, the initial mechanism proposed for the formation of such ions from saturated hydrocarbons was hydride abstraction [24]. However Manheim *et al.* [25] have proven this mechanism wrong. They propose a new ionization mechanism that occurs when a saturated hydrocarbon is protonated by a strong acid formed from atmospheric compounds, *i.e.* N_2H^+ or H_3O^+. This protonated and saturated hydrocarbon then releases a dihydrogen molecule (Equation II.9).

$$H_3O^+ + M \rightarrow [M-H]^+ + H_2 + H_2O \qquad \text{II.9}$$

This variety in ionization mechanisms is the reason why APCI can ionize compounds within a wide range of polarity, but lower masses, as represented in Figure II. 3. APCI is less affected than ESI by ion suppression and matrix effects.

2.2.3. Atmospheric pressure Photoionization (APPI)

APPI was developed by Robb et al in 2000 [26] for a coupling with liquid chromatography. This source presents a selectivity towards aromatics and is even suitable for the analysis of non-polar hydrocarbons [27]. Furthermore, APPI presents a low sensitivity to ion suppression and matrix effects and a very low chemical noise as most of the solvent molecules are not ionized [28]. Similarly to APCI, the sample is nebulized and vaporized using a heated flow of gas, typically N_2. However, the sample is ionized by a vacuum UV (V-UV) photon beam emitted from the source's lamp (Figure II-6). If several lamps can be used, the most common is the krypton lamp allowing to emit photons of 10.0/10.6 eV.

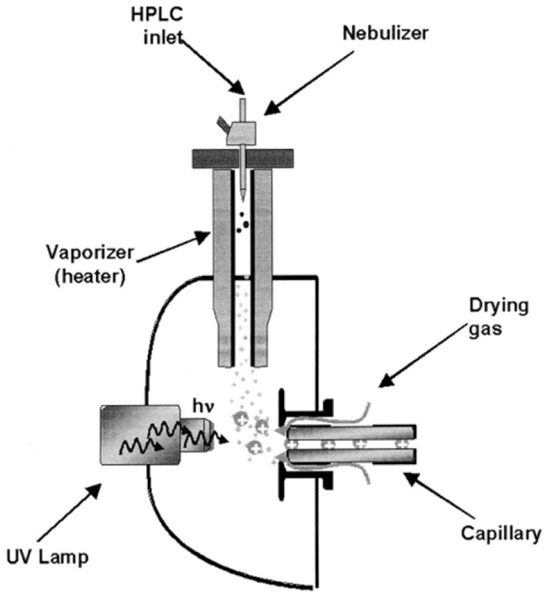


Figure II-6: Scheme of the APPI source [29]

The driving force of APPI is the difference between the energy of the photons and the ionization energy (IE) of the analyte. If the analyte IE is lower than the photon's energy, the analyte is directly ionized (Equation II.10). However, in complex mixtures mainly aromatic

analytes such as PAH or polyphenols are efficiently ionized, which makes direct photoionization a very selective technique towards those compounds.

$$M + h\nu \rightarrow M^{+\bullet} + e^{-} \quad \text{II.10}$$

Furthermore, there exists another ionization path that implies a dopant. In this case, is considered a dopant a species having a rather low IE, *e.g.* toluene IE: 8.83eV. The dopant (D) is easily ionized by the photon beam and can then exchange its charge with the analyte molecules (Equations II.11 and II.12). Hence, radical ions are expected in both direct photoionization and dopant-assisted photoionization.

$$D + h\nu \rightarrow D^{+\bullet} + e^{-} \quad \text{II.11}$$

$$D^{+\bullet} + M \rightarrow D + M^{+\bullet} \quad \text{II.12}$$

In the presence of a protic solvent such as water or methanol, it is possible to obtain protonated ions by proton exchange between the analyte's radical ion and the solvent molecule (S) (Equation II.13). In addition, the solvent molecule can also exchange a proton with the dopant ion and then transfer it to the analyte (Equation II.14 and II.15)

$$M^{+\bullet} + S \rightarrow [M + H]^{+} + [S - H]^{\bullet} \quad \text{II.13}$$

$$D^{+\bullet} + S \rightarrow [D - H]^{\bullet} + [S + H]^{+} \quad \text{II.14}$$

$$[S + H]^{+} + M \rightarrow S + [M + H]^{+} \quad \text{II.15}$$

Thus, the type of form ions depends on the IE and proton affinities of both analyte and solvent.

2.2.4. Laser desorption/ionization (LDI) and matrix-assisted LDI (MALDI)

Laser desorption ionization (LDI) was developed in the 1960s by Fenner *et al.* [30] and Vastola et al [31]. With this source, the ionization is initiated by a pulsed laser, whether ultraviolet (UV) or infrared (IR), irradiating the sample. The laser pulses (with an irradiance

usually comprised between 10^6 and 10^{10} W.cm^{-2}) are focused on a small area of the sample surface, between 0.05 and 0.2 mm in diameter. UV lasers, like the tripled-frequency Nd-YAG laser (λ = 355 nm) used in this work, are more commonly used than IR lasers. The energy absorbed by the sample allows for its desorption and ionization simultaneously, provided that the sample has a significant absorbance at the laser wavelength. The short pulses of the laser prevent thermal degradation of the sample [32].

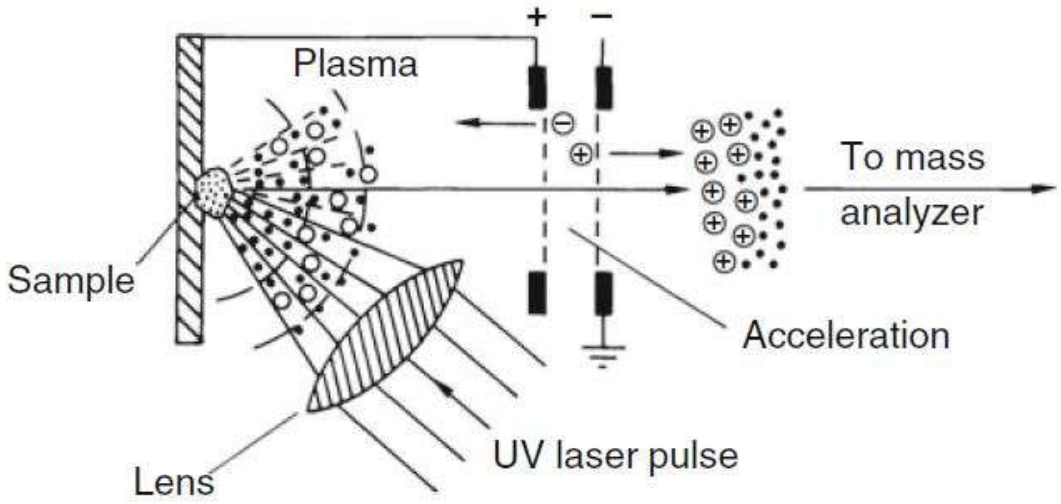


Figure II-7: Scheme of the LDI source [33]

In LDI, the sample, in a condensed state (*i.e.* liquid or solid), is deposited on a conductive plate. Once the deposit is dried, the plate is introduced into the LDI source that operates under vacuum, contrary to the API techniques presented above. LDI can be used both in positive and negative modes. In positive mode, both radical and protonated ions are formed, and it is also possible to obtain adduct ions, *e.g.* $[M+Na]^+$.

LDI is only suitable for the characterization of molecules with a strong absorbance, *i.e.* with chromophore groups, and not for molecules such as aliphatic compounds hydrocarbons that do not absorb UV light. Nevertheless, it is possible to add a strongly light-absorbing matrix, thus moving from LDI to MALDI (matrix-assisted laser desorption ionization). MALDI was developed by Karas *et al.* [34]. In MALDI, the analytes molecules are co-crystalized in an excess

of matrix. Compared to LDI, it is the matrix that absorbs the energy and transfers its charge to the analyte molecules [35]. Around the same period, Tanaka *et al* [36] developed an analogous technique: the soft laser desorption (SLD) which was the first technique allowing to ionize intact proteins, however, it is rarely used nowadays. MALDI allows to cover a wider range of mass than LDI and is suitable for the analysis of fragile biological samples such as neuropeptides [37] or intact proteins [38] due to its softness. It is also possible to use another type of matrix that, instead of exchanging a proton, transfers an electron to the analyte molecule [39]. This MALDI variation called electron transfer MALDI (ET-MALDI) is further discussed in Chapter IV.

2.2.5. Source complementarity

One of the main characteristics of mass spectrometry, its selectivity, is greatly influenced by the ionization source used. Indeed, a specific source can be used to target a selected compound family within our sample. Some ionization sources such as APPI and APCI are less selective and have a broader ionization range. In practice, the use of several sources provides a more exhaustive characterization of the sample. This is especially useful when analyzing complex mixtures by ultrahigh resolution MS as shown in our submitted review found in Appendix 2, where we discuss the ion sources complementarity in the case of complex mixtures characterization by ultrahigh resolution MS. However, if using several ion sources greatly widens the chemical space accessible by MS, it also increases the experiment time and the amount of sample consumed. Hence, there must be a balance between the amount of information needed for the research purpose and the time required to achieve it.

In any case, the ions formed by the sources described above are then guided by electrostatic lenses to the mass analyzer of the instrument. In this work, we used two mass analyzers, the TOF and the FTICR. Both will be described hereafter.

2.3 Time Of Flight Mass Spectrometry

The time-of-flight analyzer was developed by Stephens [40] in 1946. It separates the ions along their velocity inside a vacuum tube of a given length. The ions ejected from the ionization source are accelerated under the effect of an electric field. The potential energy (E_p)

of a charged particle in an electrical field is given by Equation II.16 where *U* is the applied voltage, *q* is the charge of the particle, *z* is the number of charges and *e* is the charge of an electron. When the ion is accelerated in the TOF tube, its potential energy is converted to kinetic energy (E_k) which formula is given in Equation II.17 where *m* is the mass of the ion and *v* is its velocity. As all the ions possess the same potential energy at the entry of the TOF tube, they can be differentiated by their speed within the tube, which depends on their mass and their charge. By assembling Equations II.16 and II.17 and taking *t* as the time of flight and *d* as the tube length, it is possible to calculate the *m/z* ratio of the ion, as shown in Equation II.18. In this equation, the mass *m* is given in kg.

$$E_p = qU = ezU \qquad \text{II.16}$$

$$E_k = \frac{1}{2}mv^2 \qquad \text{II.17}$$

$$\frac{m}{z} = \frac{2eUt^2}{d^2} \qquad \text{II.18}$$

The resolution of this analyzer depends on its length: the longer the tube, the more resolving power. However, in 1994 Mamyrin [41] developed a new TOF with an electrostatic mirror called a reflectron that allows to double the distance traveled by the ions without having to build a tube twice as long. A scheme of this mirror and how it changes the ions pathway compared to a linear TOF is shown in Figure II-8. This setup allows to reach a resolution of up to 100,000 instead of the several thousand provided by a linear TOF. In our work, we used the Synapt G2 from Waters which present a W-geometry reflectron, as seen in Figure II-15, that allows high-resolution analysis.

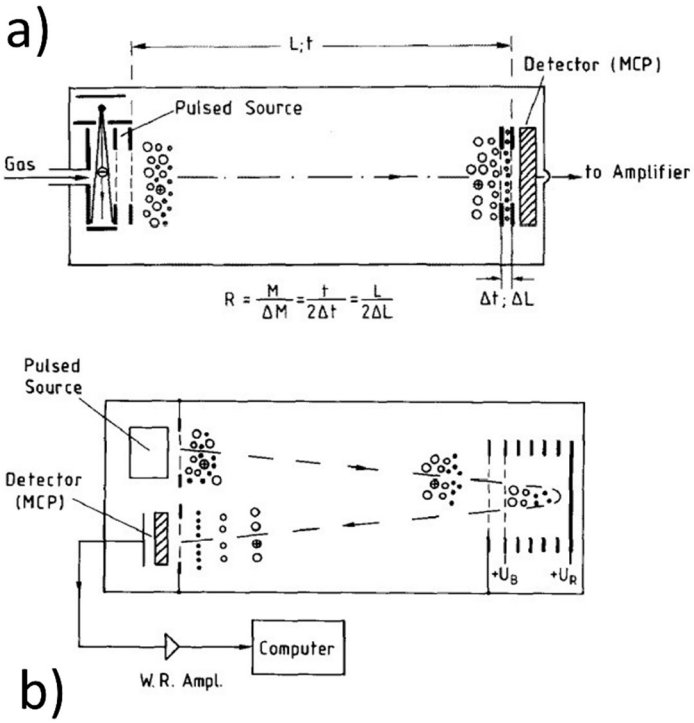


Figure II-8: Scheme of a linear TOF analyzer (a) and an electrostatic mirror or reflectron (b) [41]

2.4 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

The principle of ion cyclotron resonance (ICR) was discovered in 1930 by Lawrence *et al.* while aiming to produce high-speed protons without using high voltages [42]. They used a magnet producing a magnetic field of around 1.5 T without using high voltage to produce high-energy protons (>500 MeV). This study evidenced that it was possible to easily produce proton beams with intensities exceeding those produced from radioactive sources and granted Lawrence a Nobel award in 1939. This work led to the first ICR mass spectrometer built by Sommer *et al.* [43] in 1949. It is however the work of Comisarow and Marshall in 1974 [44, 45] that resulted in the first FTICR. Nowadays, the FTICR mass spectrometer is the instrument providing the best resolution (>10^6), sensitivity, and mass accuracy, allowing to unambiguously attribute molecular formulae, study the isotopic fine structure, and solve the very narrow mass differences observed in complex mixtures analysis such as the C_3/SH_4 split of 3.4 mDa.

2.4.1. Principle

In an FTICR mass spectrometer, the ions are trapped inside a cell (also named Penning trap) by the action of a uniform magnetic field and an electrical field [46, 47]. They are then excited by a radiofrequency pulse giving them the so-called ion cyclotron movement that is a function of their *m/z* ratio. For a better understanding of the technique, its explanation will be divided into three parts that are: the ion cyclotron motion, the trapping of the ions inside the ICR cell, and the excitation and detection of the ions.

2.4.2. Ion Cyclotron Motion

The ICR technique is based on the ion cyclotron motion of the ions inside a uniform magnetic field. Indeed, a particle with a charge *q* and a velocity $\vec{v}$ in a magnetic field $\vec{B}$ will be subjected to a Lorentz force as given in Equation II.19. The motion induced by this force is perpendicular to the magnetic field and its direction depends on the charge of the particle as shown in Figure II-9.

$$F_{Lorentz} = q.v.B = z.e.\omega_c.R.B \text{ with: } \omega_c = \frac{v}{R} = 2\pi\nu_c \quad \text{II.19}$$

$$F_C = m.R.\omega_c^2 \quad \text{II.20}$$

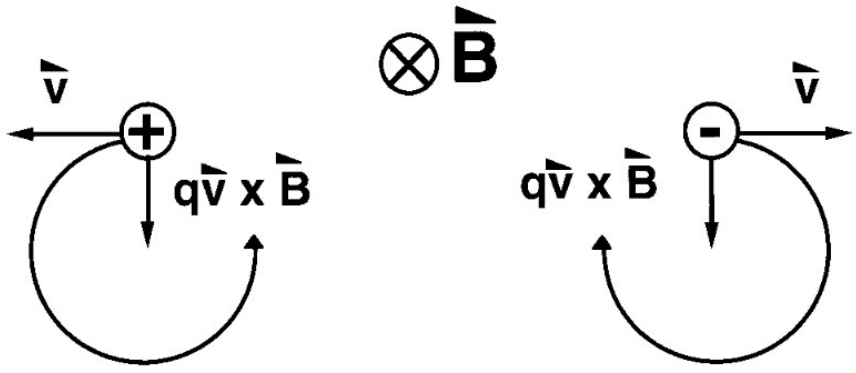


Figure II-9: Ion cyclotron movement of positive ion (left and negative ion (right) [48]

This movement is stable when the Lorentz force is completely counterbalanced by the centrifugal force described in Equation II.20 where *m* is the mass of the ion, ω_c is its angular

velocity and *r* is the ion cyclotron movement radius. Hence, considering that ν_c is the cyclotron movement frequency, at equilibrium, *i.e.* when $F_{Lorentz}$ is equal to F_c, equation II.21 defining the ion cyclotron frequency is obtained. This equation clearly shows the dependency of the ion cyclotron frequency on the *m/z* ratio of the ion and the magnetic field while exhibiting the independence on the ion's velocity. This affords very high performance in terms of resolution as most of the other analyzers are dependent on the kinetic energy of the ions and affected by its dispersion.

$$\nu_c = \frac{z}{m} \times \frac{eB}{2\pi} \qquad \text{II.21}$$

Furthermore, it also shows that the lower the mass of the ions, the higher the frequency. Digitalization of the signal induces a lower limit on the measurable mass by an FTICR. Indeed, as explained by the Nyquist-Shannon sampling theorem, the sampling rate, or acquisition rate, must be at least twice as much as the measured phenomenon frequency which is inversely proportional to the ion mass. For example, measuring a mass of 20 Da, *i.e.* around $3.32.10^{-26}$ kg, with a 12 T FTICR would require a sampling rate of more than 18 MHz.

2.4.3. Ion trapping

Once the ions are formed in the source, they are guided to the ICR cell in which they are trapped by the action of both a magnetic and an electric field. There exist several ICR cell designs, however, we will take here the case of the cylindrical cell, as shown in Figure II-10.

The magnetic field described above only confines the ion in the plane perpendicular to itself (*xy* plane in Figure II-10). Hence, it is necessary to apply a voltage to the trapping plates of the ICR cell to bind the ions in the *z*-direction. This electric field causes the ion trapping motion on the z-axis that has an angular frequency ω_z described in Equation II.22 where V_T is the voltage applied to the plates, *a* is the distance between the trapping plates and α is the trapping scale factor that depends on the cell geometry [49]. Furthermore, this electric field yields an electric force (F_{elec}) that opposes the magnetic force ($F_{Lorentz}$) and decreases the cyclotron movement frequency.

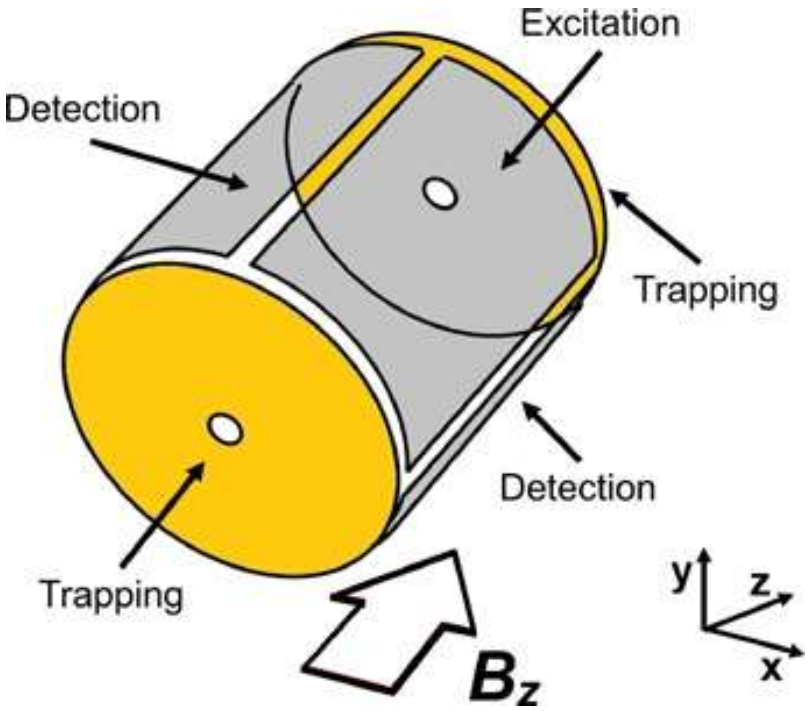


Figure II-10: Scheme of the cylindrical ICR cell [50]

The new sum of the forces applied to the ion is shown in equation II.23 and can be solved as a quadratic equation on ω_c. This equation has two solutions $\omega+$ which are the reduced cyclotron frequency (Equation II.24) and $\omega-$ which is the magnetron frequency (Equation II.25). The magnetron frequency characterizes the movement that consists of the drift of the cyclotron movement center in the *xy* plane. Hence, the ions are subjected to three different periodic movements. Among the three frequencies that define these movements, the trapping frequency and the magnetron frequency are lower than the cyclotron frequency, less than 10 kHz and 100 Hz for the trapping and magnetron frequencies against several kHz to MHz for the cyclotron frequency. Hence, cyclotron frequency is often the only frequency considered [50]. However, this consideration can lead to frequency shifts and peak broadening as the trapping and magnetron movement perturbs the ion's cyclotron movement.

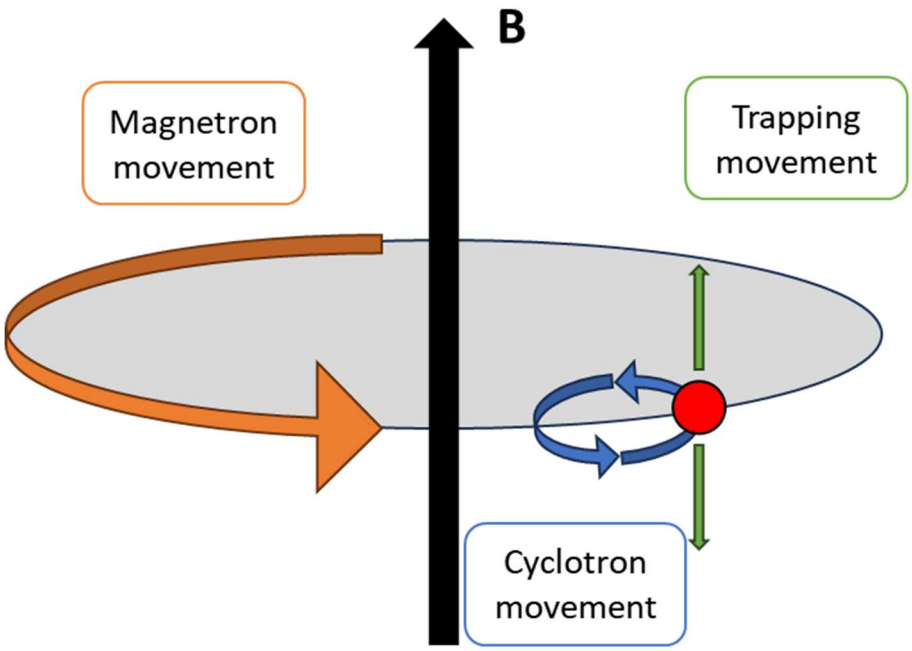


Figure II-11: Ion movements in the ICR cell

$$\omega_z = \sqrt{\frac{2z.e.V.\alpha}{ma^2}} \qquad \text{II.22}$$

$$F_C = F_{Lorentz} - F_{elec} \text{ or } m.R.\omega_c^2 = z.e.\omega_c.R.B - \frac{2z.e.V.\alpha}{ma^2}.R \qquad \text{II.23}$$

$$\omega+ = \frac{\varpi_c}{2} + \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \qquad \text{II.24}$$

$$\omega- = \frac{\varpi_c}{2} - \sqrt{\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}} \qquad \text{II.25}$$

2.4.4. Excitation/detection

Upon entering the ICR cell, the ion's cyclotron movement has a very short radius, and ions sharing the same *m/z* do not move coherently. For the detection of the cyclotron frequency, a radiofrequency (RF) potential is applied to one of the two pairs of the cell's plates that are parallel to the magnetic field (grey Excitation plates in Figure II-10). This RF potential creates an RF electrical field perpendicular to the magnetic field. If the frequency of the

electrical field is equal to the cyclotron frequency of an ion, it absorbs the energy. This leads to an increase in the radii of its cyclotron motion and provides coherent "ion packages" for ions of the same *m/z* ratio. Usually, during the excitation, the applied RF pulse consists of a fast frequency sweep over a broadband called a chirp [45]. This excitation method aims at exciting all the ions within the chirp's frequency range at the same time.

Once the ion packages have a sufficient radius, they can be detected by the other pair of plates in the ICR cell by creating an image current [51]. This current amplitude is amplified and then digitalized along the time domain. The amplitude of the image current depends on the ions' charge. This current is affected by the presence of neutral molecules within the cell that, by collision with the ions, progressively decrease the current until it becomes null. This phenomenon is called free induction decay (FID), its duration (t) directly affects the resolution (R) of the mass spectra following Equation II.26. The FID length depends on experimental parameters, especially on the ultra-high vacuum inside the ICR cell. Indeed, the lower the pressure, the lower the chances of collision between the neutral molecules and the ions.

$$R = \frac{\omega_c . t}{2\pi} \qquad \text{II.26}$$

At the end of the scan, *i.e.* the excitation/detection cycle, the FID is converted from the time domain to the frequency domain by using the Fourier transform. This transformation results in a frequency spectrum that can be converted to a mass spectrum (*m/z vs* intensity).

2.4.5. Instrument

During this work, we used the SolariX 2XR FTICR mass spectrometer (Bruker Daltonics, Bremen, Germany), whose simplified scheme is given in Figure II-12, equipped with a 12 T superconducting magnet. Concerning the ionization, this instrument is compatible with the sources described above. For the ion transmission, it is equipped with a pair of ion funnels and a pair of skimmers, an octupole, a quadrupole (which can be used as a mass filter), and two hexapole: one acts as an ion trap to accumulate the ions before the ICR cell and the other guides the ions into the cell. For the detection part, this instrument is also equipped with a dynamically harmonized ICR cell from Bruker commercially called Paracell™ [52]. This mass spectrometer is capable of achieving a resolution of above 1 000 000.

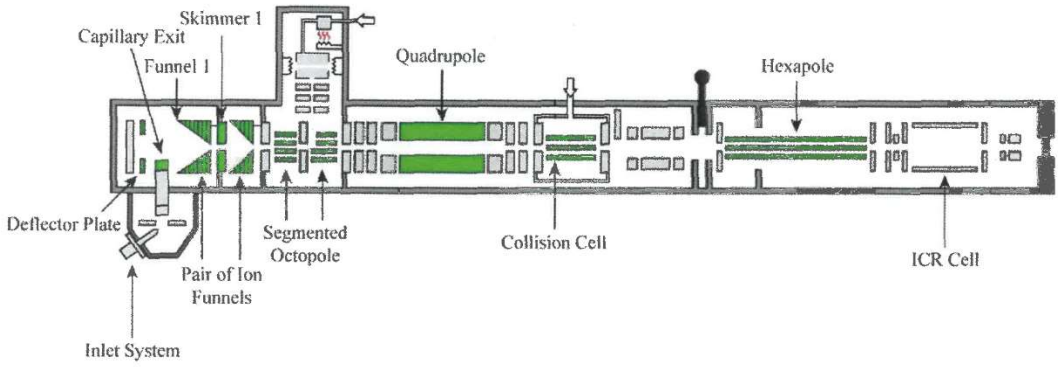


Figure II-12: Scheme of the Solarix 2XR (Bruker Daltonics, Bremen, Germany)

3. Ion Mobility Spectrometry

3.1 Principle

Ion mobility spectrometry (IMS) is a gas-phase ion separation technique. In IMS, the ions are separated using an electrical field in a cell filled with a buffer gas, typically N_2 or He. In IMS, the ions are separated according to their charge and size, or more precisely by the collision they undergo with the buffer gas [53]. The principle of separating ions in an inert gas was first described in 1903 by Langevin *et al*. [54]; since then, this technique has known a widespread growth in terms of instrumentation and applications [55]. Here, the basic concepts of this technique will be described using the most basic IMS system: the drift-tube ion mobility spectrometry (DTIMS). In DTIMS, the ions enter the drift tube and are accelerated by a uniform electrical field while being slowed by collision with the buffer gas filling the cell as shown in Figure II-13.

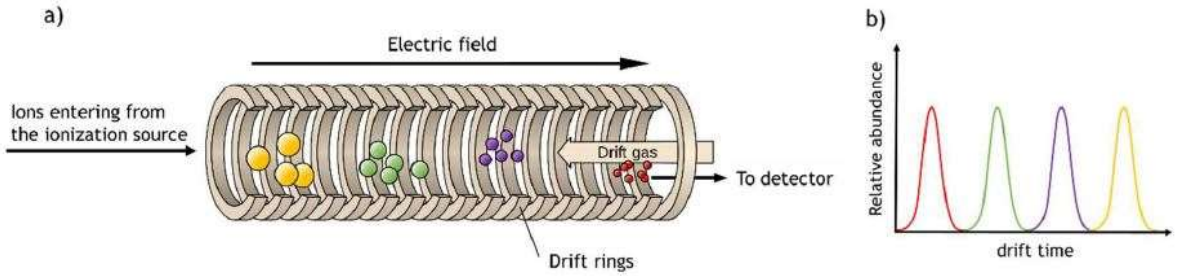


Figure II-13: a) Principle of drift tube ion mobility spectrometry b) Example of ion mobility spectrum or mobilogram [56]

In DTIMS the arrival time, *i.e.* the time that ions take to reach the detector, is measured: it comprises the drift time (t_D), and the time t_0 for the ions to travel between the end of the

ion mobility cell and the detector. The drift time depends on the ion collision cross section (CCS), which is a surface that characterizes the collisions between the buffer gas and the analyte ion. The CCS roughly depend on the size and shape of the ions in the gas phase, but also on the buffer gas. In one gas, the higher the drift time, the larger the ions and the higher the CCS. Indeed, a smaller ion would experience fewer collisions with the buffer gas than a larger one. Furthermore, multicharged ions are expected to have a short drift time compared to a single-charged ion of the same shape as they are more accelerated by the electrical field. In the drift tube, the velocity of an ion (v_d) is linked to its ion mobility (K) and the intensity of the electrical field (E) as shown in Equation II.27. It is also possible to express this relationship using the drift time and the length of the drift tube (L) (See Equation II.28)

$$v_d = K.E \qquad \text{II.27}$$

$$t_d = \frac{L}{K.E} \qquad \text{II.28}$$

The ion mobility gives an indication on the interaction between the ions and the buffer gas and is dependent on their collision frequency. This value depends on the pressure (P), the temperature (T), and the number of molecules per unit volume (N). In order to ease comparisons [57], the mobility can be normalized to the standard number of molecules per unit volume N_0 (which is obtained with standard pressure (P_0 = 760mmHg) and temperature (T_0 = 273.15K) to obtain the reduced ion mobility (K_0) (See Equation II.29). As written by McDaniel and Viehland in their 1984 article [58] "*It must be emphasized that the use of*" equation II.29 "*merely provides a standardization or normalization with respect to the molecular number density; the temperature to which the standard mobility actually refers is the temperature of the gas during the measurement.*"

$$K_0 = K\frac{N}{N_0} = K\frac{P}{P_0}\frac{T_0}{T} \qquad \text{II.29}$$

The reduced mobility K_0 of an ion depends on its size, its charge, and its collision cross section (Ω), the buffer gas and its temperature *T*. Under a low electric field, these parameters are linked by the Mason-Schamp[59] equation (Equation II.30). Where *z* is the charge of the ion, *e* is the elementary charge of the electron, k_B is the Boltzmann constant and μ is the

reduced mass of the ion, calculated with Equation II.31 where m is the of the ion and M is the molecular mass of the buffer gas.

$$\Omega = \frac{3ze}{16N_0} \cdot \frac{1}{K_0} \cdot \sqrt{\frac{2\pi}{\mu k_B T}} \qquad \text{II.30}$$

$$\mu = \frac{mM}{m+M} \qquad \text{II.31}$$

3.2 Traveling Wave Ion Mobility Spectrometry (TWIMS)

There exist several IMS techniques, each with its own advantages and drawbacks. They mainly differ by the nature of the electrical field inside the cell. The main techniques are the following:

- Drift tube ion mobility spectrometry (DTIMS)
- Traveling wave ion mobility spectrometry (TWIMS)
- Trapped ion mobility spectrometry (TIMS)
- Cyclic ion mobility spectrometry (cIMS)
- Field asymmetric ion mobility spectrometry (FAIMS)
- Differential mobility analyzer (DMA)

In this work, we used a single technique: the traveling wave ion mobility spectrometry, which will be described hereafter.

3.2.1. Principle of Traveling Wave Ion Mobility Spectrometry

The traveling wave ion mobility spectrometry, first commercialized by Waters Corporation in 2006 [60], uses a non-uniform and non-constant electrical field that can be described as waves of potential that move along the cell, as shown on Figure II-14. The cell is composed of a series of electrodes that form a stacked-ring ion guide (SRIG). A pulsed voltage is applied to the electrodes to push the ions along the cell in a wave-like motion. It is the

succession of the waves of potential that separates the ions depending on their CCS. Indeed, the ions with a higher CCS tend to roll over the waves of potential due to collisions with the buffer gas while ions of lower CCS are more like to follow, or 'surf', the waves along the cell. In TWIMS it is possible to adjust both the velocity and the height (voltage) of the waves to optimize the separation.

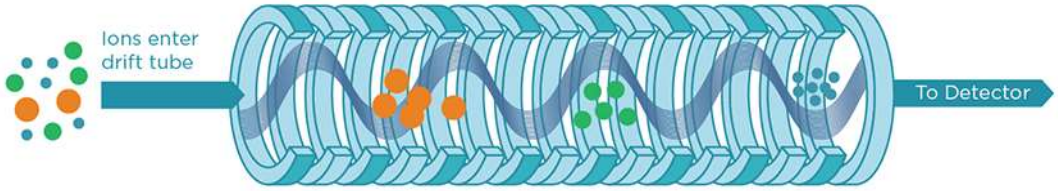


Figure II-14: Principle of traveling wave ion mobility spectrometry.

As opposed to DTIMS, the electric field in TWIMS is neither uniform nor constant. Hence, in TWIMS it is not possible to directly determine the CCS of an ion using the Mason-Schamp equation. Instead, it is necessary to make a CCS calibration using compounds with a known CCS once the cell parameters are optimized for the analyte. This calibration is used to determine two parameters A and B in Equation II.32, empirically derived, where t_d is the ion drift time [59]. To ease the calculations it is possible to use Ω' (reduced CCS) and A', defined in Equation II.33, to simplify Equation II.32 and obtain Equation II.34 which links Ω' and t_d [61].

$$\Omega = \frac{3ze}{16N} \times \sqrt{\frac{1}{\mu}} \times \sqrt{\frac{2\pi}{k_B T}} \times \frac{P_0}{P} \times \frac{T}{T_0} \times A(t_d)^B \quad \text{II.32}$$

$$\Omega' = \frac{\Omega\sqrt{\mu}}{z} \;\&\; A' = \frac{3e}{16N} \times \sqrt{\frac{2\pi}{k_B T}} \times \frac{P_0}{P} \times \frac{T}{T_0} \times A \quad \text{II.33}$$

$$\Omega' = A' t_d^B \quad \text{II.34}$$

3.2.2. *Instrumentation*

The instrument used to perform IMS in this work is the Synapt G2 from Waters Corporation which is schematically represented in Figure II-15. This instrument combines a TOF mass spectrometer with the TWIMS technology (TWIMS-TOF-MS). It is compatible with

several atmospheric pressure ion sources such as ESI, APPI, and APCI that were used in this work. It is provided with three SRIG series: the first one (*trap* in Figure II-15) allowing the accumulation of the ions before entering the second SRIG series which is the ion mobility cell and the third one (*transfer* in Figure II-15) allowing the ion transfer to the TOF analyzer. The TOF analyzer can be used under modes: sensitivity (S), resolution (V), and high-resolution (W) depending on the path taken by the ions. In the S mode, the ion path is linear which allows an enhanced sensitivity but a lower resolution. In V mode the reflectron of the instrument is used to improve the resolution and in the W mode, it is the double reflectron that is used to further increase the resolution (up to 40 000) at the expense of the sensitivity. The resolving power (R_P), defined in Equation II.32, in TWIMS is usually around 40.

$$R_P = \frac{t_D}{FWHM} \qquad 2.35$$

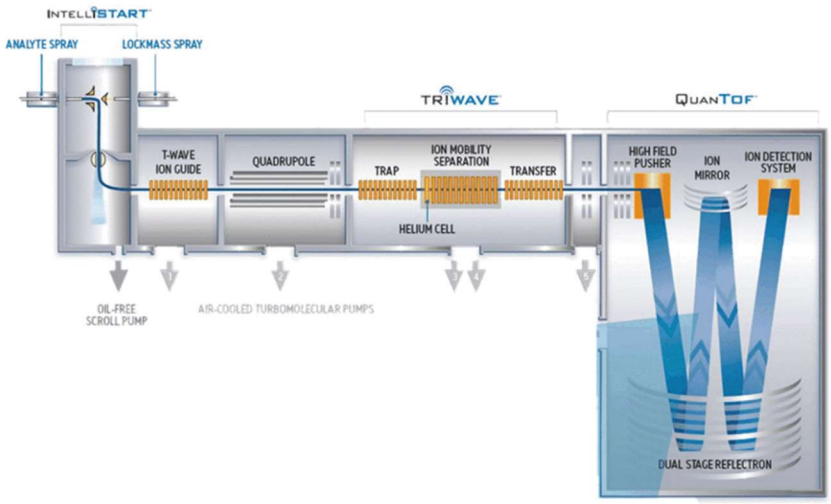


Figure II-15: Schematical representation of the Synapt G2, Waters corporation

In this work, for the IMS calibration, we used the CCS values in N_2 of the ESI Low concentration Tuning Mix which is a solution provided by Agilent (TuneMix).The *m/z* ratio of the TuneMix ions used for calibration and their CCS values is shown in Table II-1.

Table II-1: List of TuneMix ions and their CCS values [62].

m/z	z	CCS(N_2) (Å²)	*m/z*	z	CCS(N_2) (Å²)
118.0857	1	121.3	112.9856	-1	108.23
322.048	1	153.73	431.9823	-1	140.04

622.0273	1	202.96	601.9790	-1	180.77
922.0053	1	243.64	1033.9881	-1	255.34

It is then possible to linearize equation II.34 by applying a logarithmic transformation to obtain a linear calibration curve allowing to determine parameters A' and B as shown in Figure II-16. This calibration step must take place after the optimization of the ion mobility cell parameters.

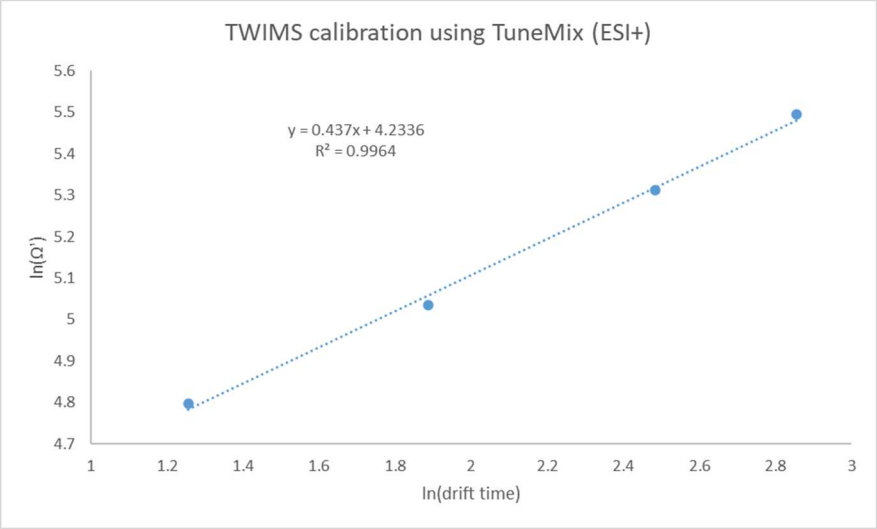


Figure II-16: Calibration curve using TuneMix ions in ESI positive mode.

In practice, the determination of the drift time was performed by extracting the ion mobilogram (EIM) of the ions of the ESI Low concentration Tuning Mix. These EIM were then fitted using a Gauss function with the OriginPro software to determine the drift time which corresponds to the apex of the Gauss fit of the mobility peak (in the case of a single isomer).

4. Conclusion

As presented, FTICR MS and TWIMS-TOF MS are high-end analytical techniques that allow to dive into the molecular composition of a sample, with a focus on its isobaric or isomeric complexity, respectively. This high level of technology is particularly useful for the characterization of complex mixtures such as ship emissions. These techniques generate a large amount of data that needs to be treated and visualized to extract relevant information and reach a conclusion. This is especially true with the characterization of complex mixtures by Fourier transform MS (FTMS). In this case, solutions for both treatment and visualization of

generated data must be used or even created. Chapter III of this thesis focuses on the development of such a solution in the form of an open-source software called Python Tools for Complex Matrices Molecular Characterization, or PyC2MC.

5. References

1. Jeong, S., et al., *Aerosol emissions from a marine diesel engine running on different fuels and effects of exhaust gas cleaning measures.* Environmental Pollution, 2023. **316**: p. 120526.
2. Thomson, J.J., *LIV. Positive rays.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1908. **16**(94): p. 657-691.
3. Thomson, J.J., *XLII. Rays of positive electricity.* The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1910. **19**(111): p. 424-435.
4. Thomson, J.J., *Rays of positive electricity.* 1913. **89**(607): p. 1-20.
5. Zeleny, J., *Instability of Electrified Liquid Surfaces.* Physical Review, 1917. **10**(1): p. 1-6.
6. Dole, M., et al., *Molecular Beams of Macroions.* The Journal of Chemical Physics, 2003. **49**(5): p. 2240-2249.
7. Fenn, J.B., et al., *Electrospray Ionization for Mass Spectrometry of Large Biomolecules.* 1989. **246**(4926): p. 64-71.
8. Yamashita, M. and J.B. Fenn, *Electrospray ion source. Another variation on the free-jet theme.* The Journal of Physical Chemistry, 1984. **88**(20): p. 4451-4459.
9. Verenchikov, A.N., N.V. Krasnov, and V.A. Shkurov, *Electrospray ionization developed by Lidija Gall's group.* International Journal of Mass Spectrometry, 2023. **490**: p. 117067.
10. Prabhu, G.R.D., et al., *Mass spectrometry using electrospray ionization.* Nature Reviews Methods Primers, 2023. **3**(1): p. 23.
11. Konermann, L., *A simple model for the disintegration of highly charged solvent droplets during electrospray ionization.* Journal of the American Society for Mass Spectrometry, 2009. **20**(3): p. 496-506.
12. Pimlott, D.J.D. and L. Konermann, *Using covalent modifications to distinguish protein electrospray mechanisms: Charged residue model (CRM) vs. chain ejection model (CEM).* International Journal of Mass Spectrometry, 2021. **469**: p. 116678.

13. Banerjee, S. and S. Mazumdar, *Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte.* International Journal of Analytical Chemistry, 2012. **2012**: p. 282574.

14. Ehrmann, B.M., T. Henriksen, and N.B. Cech, *Relative importance of basicity in the gas phase and in solution for determining selectivity in electrospray ionization mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2008. **19**(5): p. 719-728.

15. Loo, J.A., et al., *Solvent-induced conformational changes of polypeptides probed by electrospray-ionization mass spectrometry.* 1991. **5**(3): p. 101-105.

16. Wilm, M., *Principles of Electrospray Ionization.* Molecular & Cellular Proteomics, 2011. **10**(7): p. M111.009407.

17. Kruve, A., *Strategies for Drawing Quantitative Conclusions from Nontargeted Liquid Chromatography–High-Resolution Mass Spectrometry Analysis.* Analytical Chemistry, 2020. **92**(7): p. 4691-4699.

18. Furey, A., et al., *Ion suppression; A critical review on causes, evaluation, prevention and applications.* Talanta, 2013. **115**: p. 104-122.

19. Horning, E., et al., *New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure.* Analytical Chemistry (Washington, DC, United States), 1973. **45**(6): p. 936-943.

20. Carroll, D., et al., *Subpicogram detection system for gas phase analysis based upon atmospheric pressure ionization (API) mass spectrometry.* Analytical Chemistry (Washington, DC, United States), 1974. **46**(6): p. 706-710.

21. Munson, M.S.B. and F.-H. Field, *Chemical ionization mass spectrometry. I. General introduction.* Journal of the American Chemical Society, 1966. **88**(12): p. 2621-2630.

22. Andrade, F.J., et al., *Atmospheric Pressure Chemical Ionization Source. 1. Ionization of Compounds in the Gas Phase.* Analytical Chemistry, 2008. **80**(8): p. 2646-2653.

23. Schwemer, T., et al., *Gas Chromatography Coupled to Atmospheric Pressure Chemical Ionization FT-ICR Mass Spectrometry for Improvement of Data Reliability.* Analytical Chemistry, 2015. **87**(24): p. 11957-11961.

24. Tose, L.V., et al., *Analyzes of hydrocarbons by atmosphere pressure chemical ionization FT-ICR mass spectrometry using isooctane as ionizing reagent.* Fuel, 2015. **153**: p. 346-354.

25. Manheim, J.M., et al., *Fragmentation of Saturated Hydrocarbons upon Atmospheric Pressure Chemical Ionization Is Caused by Proton-Transfer Reactions.* Analytical Chemistry, 2020. **92**(13): p. 8883-8892.

26. Robb, D.B., T.R. Covey, and A.P. Bruins, *Atmospheric Pressure Photoionization: An Ionization Method for Liquid Chromatography–Mass Spectrometry.* Analytical Chemistry, 2000. **72**(15): p. 3653-3659.

27. Huba, A.K., K. Huba, and P.R. Gardinali, *Understanding the atmospheric pressure ionization of petroleum components: The effects of size, structure, and presence of heteroatoms.* Science of The Total Environment, 2016. **568**: p. 1018-1025.

28. Zimmermann, R. and L. Hanley, *Photoionization and Photo-Induced Processes in Mass Spectrometry: Fundamentals and Applications*. 2021: John Wiley & Sons.

29. Hanold, K.A., et al., *Atmospheric Pressure Photoionization. 1. General Properties for LC/MS.* Analytical Chemistry, 2004. **76**(10): p. 2842-2851.

30. Fenner, N. and N. Daly, *Laser used for mass analysis.* Review of Scientific Instruments, 1966. **37**(8): p. 1068-1070.

31. Vastola, F.J., R.O. Mumma, and A.J. Pirone, *Analysis of organic salts by laser ionization.* 1970. **3**(1): p. 101-104.

32. Gross, J.H., *Matrix-Assisted Laser Desorption/Ionization*, in *Mass Spectrometry: A Textbook*, J.H. Gross, Editor. 2017, Springer International Publishing: Cham. p. 651-720.

33. Gross, J.H., *Mass Spectrometry*. 2017.

34. Karas, M., D. Bachmann, and F. Hillenkamp, *Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules.* Analytical Chemistry (Washington, DC, United States), 1985. **57**(14): p. 2935-2939.

35. Knochenmuss, R., *The Coupled Chemical and Physical Dynamics Model of MALDI.* 2016. **9**(1): p. 365-385.

36. Tanaka, K., et al., *Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry.* Rapid communications in mass spectrometry, 1988. **2**(8): p. 151-153.

37. Kutz, K.K., J.J. Schmidt, and L. Li, *In Situ Tissue Analysis of Neuropeptides by MALDI FTMS In-Cell Accumulation.* Analytical Chemistry, 2004. **76**(19): p. 5630-5640.

38. Spraggins, J.M., et al., *MALDI FTICR IMS of Intact Proteins: Using Mass Accuracy to Link Protein Images with Proteomics Data.* Journal of the American Society for Mass Spectrometry, 2015. **26**(6): p. 974-985.

39. Ramírez-Pradilla, J.S., C. Blanco-Tirado, and M.Y. Combariza, *Electron-Transfer Ionization of Nanoparticles, Polymers, Porphyrins, and Fullerenes Using Synthetically Tunable α-Cyanophenylenevinylenes as UV MALDI-MS Matrices.* ACS Applied Materials & Interfaces, 2019. **11**(11): p. 10975-10987.

40. Stephens, W.E., *A pulsed mass spectrometer with time disaersion.* Phys. Rev., 1946. **69**: p. 691.

41. Mamyrin, B.A., *Laser assisted reflectron time-of-flight mass spectrometry.* International Journal of Mass Spectrometry and Ion Processes, 1994. **131**: p. 1-19.

42. Lawrence, E.O. and M.S. Livingston, *The Production of High Speed Protons Without the Use of High Voltages.* Physical Review, 1931. **38**(4): p. 834-834.

43. Sommer, H., H.A. Thomas, and J.A. Hipple, *The Measurement of e/M by Cyclotron Resonance.* Physical Review, 1951. **82**(5): p. 697-702.

44. Comisarow, M.B. and A.G. Marshall, *Fourier transform ion cyclotron resonance spectroscopy.* Chemical Physics Letters, 1974. **25**(2): p. 282-283.

45. Comisarow, M.B. and A.G. Marshall, *Frequency-sweep fourier transform ion cyclotron resonance spectroscopy.* Chemical Physics Letters, 1974. **26**(4): p. 489-490.

46. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: A primer.* 1998. **17**(1): p. 1-35.

47. Nikolaev, E.N., Y.I. Kostyukevich, and G.N. Vladimirov, *Fourier transform ion cyclotron resonance (FT ICR) mass spectrometry: Theory and simulations.* 2016. **35**(2): p. 219-258.

48. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: A primer.* Mass Spectrometry Reviews, 1998. **17**(1): p. 1-35.

49. Brown, L.S. and G. Gabrielse, *Geonium theory: Physics of a single electron or ion in a Penning trap.* Reviews of Modern Physics, 1986. **58**(1): p. 233-311.

50. Qi, Y. and P.B. O'Connor, *Data processing in Fourier transform ion cyclotron resonance mass spectrometry.* Mass Spectrom Rev, 2014. **33**(5): p. 333-52.

51. Dass, C., *Fundamentals of contemporary mass spectrometry*. 2007: John Wiley & Sons.

52. Kostyukevich, Y.I., G.N. Vladimirov, and E.N. Nikolaev, *Dynamically Harmonized FT-ICR Cell with Specially Shaped Electrodes for Compensation of Inhomogeneity of the Magnetic Field. Computer Simulations of the Electric Field and Ion Motion Dynamics.* Journal of The American Society for Mass Spectrometry, 2012. **23**(12): p. 2198-2207.

53. Kanu, A.B., et al., *Ion mobility-mass spectrometry.* Journal of Mass Spectrometry, 2008. **43**(1): p. 1-22.

54. Langevin, P., *Recombinaison et mobilites des ions dans les gaz.* Ann. Chim. Phys, 1903. **28**(433): p. 122.

55. May, J.C. and J.A. McLean, *Ion Mobility-Mass Spectrometry: Time-Dispersive Instrumentation.* Analytical Chemistry, 2015. **87**(3): p. 1422-1436.

56. Martin-Gómez, H. and J. Tulla-Puche, *Lasso peptides: chemical approaches and structural elucidation.* Organic & Biomolecular Chemistry, 2018. **16**(28): p. 5065-5080.

57. Clemmer, D.E. and M.F. Jarrold, *Ion Mobility Measurements and their Applications to Clusters and Biomolecules.* 1997. **32**(6): p. 577-592.

58. McDaniel, E.W. and L.A. Viehland, *The transport of slow ions in gases: Experiment, theory, and applications.* Physics Reports, 1984. **110**(5): p. 333-367.

59. Mason, E.A. and H.W. Schamp, *Mobility of gaseous Ions in weak electric fields.* Annals of Physics, 1958. **4**(3): p. 233-270.

60. Pringle, S.D., et al., *An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument.* International Journal of Mass Spectrometry, 2007. **261**(1): p. 1-12.

61. Smith, D.P., et al., *Deciphering Drift Time Measurements from Travelling Wave Ion Mobility Spectrometry-Mass Spectrometry Studies.* 2009. **15**(2): p. 113-130.

62. Stow, S.M., et al., *An Interlaboratory Evaluation of Drift Tube Ion Mobility–Mass Spectrometry Collision Cross Section Measurements.* Analytical Chemistry, 2017. **89**(17): p. 9048-9055.

Chapter III. Processing and Visualization of FTMS Data for Complex Mixtures Molecular Characterization

1. Introduction and Summary of "PyC2MC: an open-source software solution for visualization and treatment of high-resolution mass spectrometry data"

Complex mixtures encompass a wide variety of samples from different research fields such as petroleomics, proteomics, or environmental science. Hence, complex mixtures can be, for example, blood serum, liquid fuel, wastewater, or organic aerosol. This kind of sample carries a tremendous quantity of information at the molecular level that can be useful for a better understanding of biological, industrial, or environmental processes. However, it requires high-end techniques such as Fourier transform mass spectrometers (FTMS) to access this molecular information. Indeed, FTMS, in the form of either FTICR MS or Orbitrap MS, presents sufficient mass-resolving power and accuracy to dive into the molecular composition of such complex organic mixtures. Nonetheless, as mentioned earlier, such high-end techniques generate high amounts of data that need to be efficiently processed and visualized to extract valuable information. In this respect, we developed the software called PyC2MC for "Python tool for complex matrices molecular characterization" within the international laboratory for complex matrices molecular characterization (iC2MC). This tool aims to be user-friendly and to provide the ability to use the most common representations in high resolution mass spectrometry, such as double bond equivalent *vs* carbon number plots or van Krevelen diagrams, while not being limited to a specific brand of instruments. It also provides statistical analysis tools to ease the comparison between samples, replicates or not, to evidence the reproducibility of an analysis and/or to highlight differences at the molecular level. This software was the subject of the article "PyC2MC: an open-source software solution for visualization and treatment of high-resolution mass spectrometry data" that we published and whose full text is available hereafter.

Another aim of this tool is to be used, revised, and improved by the FTMS users' community. For that purpose, we made it open source along with the publication of an article

describing the software. PyC2MC is under continuous development and improvement thanks to the suggestions of fellow FTMS users, within or from outside the original project's team.

2. Full text article: "PyC2MC: an open-source software solution for visualization and treatment of high-resolution mass spectrometry data"

Maxime SUEUR[1,3], Julien F. MAILLARD[1,3], Oscar LACROIX-ANDRIVET[1,2,3], Christopher P. RÜGER[3,4*], Pierre GIUSTI[1,2,3], Hélène LAVANANT[1], Carlos AFONSO[1,3]

[1] Normandie Univ, UNIROUEN, INSA Rouen, CNRS, COBRA, 76000 Rouen, France.

[2] TotalEnergies OneTech R&D, TotalEnergies Research & Technology Gonfreville, BP 27, 76700 Harfleur, France

[3] International Joint Laboratory - iC2MC: Complex Matrices Molecular Characterization, TRTG, BP 27, 76700 Harfleur, France.

[4] Joint Mass Spectrometry Centre, Chair of Analytical Chemistry, University of Rostock, 18059 Rostock, Germany; Interdisciplinary Faculty, Department Life, Light & Matter (LL&M), University of Rostock, 18051 Rostock, Germany.

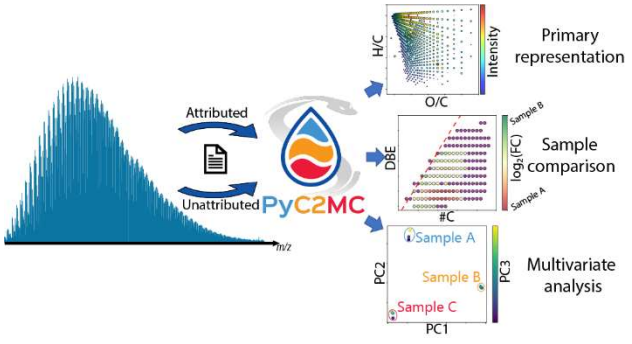
*** corresponding author: christopher.rueger@uni-rostock.de**

Keywords: complex matrices, data visualization, statistical analysis, open access software, Python

ABSTRACT

Complex molecular mixtures are encountered in almost all research disciplines, such as biomedical 'omics, petroleomics, and environmental sciences. State-of-the-art characterization of sample materials related to these fields, deploying high-end instrumentation, allow for gathering humongous quantity of molecular composition data. One established technological platform is ultrahigh-resolution mass spectrometry, *e.g.*, Fourier-transform mass spectrometry (FT-MS). However, the huge amounts of data acquired in FT-MS often result in tedious data treatment and visualization. FT-MS analysis of complex matrices can easily lead to single mass spectra with more than 10,000 attributed unique molecular formulae. Sophisticated software solutions to conduct these treatment and visualization attempts from commercial and non-commercial origins exist. However, existing applications have distinct drawbacks, such as focusing on only one type of graphic representation, being unable to handle large datasets, or not being publicly available. In this respect, we developed a software, within the international complex matrices molecular characterization joint lab (IC2MC), named "python tools for complex matrices molecular characterization" (PyC2MC). This piece of software will be open-source and free to use. PyC2MC is written under python 3.9.7 and relies on well-known libraries such as pandas, NumPy, or SciPy. It is provided with a graphical user interface developed under PyQt5. The two options for execution, 1) user-friendly route with pre-packed executable file or 2) running the main python script through a Python interpreter, ensure a high applicability but also an open characteristic for further development by the community. Both are available on the GitHub platform (https://github.com/iC2MC/PyC2MC_viewer).

GRAPHICAL ABSTRACT



INTRODUCTION

Complex molecular mixtures are encountered in almost all research domains. Their comprehensive chemical description is referred to as 'omics sciences, such as in proteomics, lipidomics, and metabolomics [1], petroleomics [2-5] or in environmental sciences [2, 6-8], to name a few. Omics approaches allow for a massive richness of in-depth information on these complex samples. Consequently, transformations and processes can be studied at the molecular level. Naturally, omics experiments are associated with large numbers of molecular data. In practice, the chemical characterization of complex mixtures requires state-of-the-art analytical instrumentation. One commonly applied technique is high-resolution mass spectrometry. In this category, Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) and Orbitrap MS are the two primary platforms with adequate performance, *i.e.*, resolving power above 10^5 and mass accuracy below 1 ppm [9-13]. Their superior resolving power allows the separation of isobaric constituents, whereas the high mass accuracy enables unique molecular formula attribution to each ion. The attribution of molecular formulas can be performed either using vendor proprietary software, such as DataAnalysis (Bruker Daltonics, US), Xcalibur (Thermo Fisher Scientific, US), and Composer (Sierra Analytics, US) or using specific workflows, such as PetroOrg [14], Peak-by-Peak (Spectroswiss, Switzerland), Attributor [15], OpenMS [16] or mMass [17]. Moreover, numerous specific workflows for given research domains or single laboratory solutions exist. An example is ICBM-OCEAN [18], developed at the University of Oldenburg, which is specialized in dissolved organic matter (DOM) analysis. Results can then usually be recovered in large spreadsheets or software-specific formats. Even though these software workflows allow for robust molecular attribution, most do not permit advanced visualization or comparison capabilities. Hence, in daily routine as a rapid and easily accessible solution, data treatment and visualization are performed based on simple scripts (*e.g.*, R, Python, MATLAB) or even utilizing spreadsheet software like Microsoft Excel or OriginLab Origin. However, when the number of analyses to be processed becomes high, this becomes tedious or even impossible. Complex matrices analyses require the use of fingerprint visualizations that are essential to understand the underlying chemistry. Diagrams such as double bond equivalent [19-21] (DBE) maps, van Krevelen [22, 23], or Kendrick [24, 25] plots are of common use in petroleomics due to the clear representation of a complex organic sample. In the same way, average carbon oxidation state [26], modified

aromatic index [27], and maximum carbonyl ratio [28] representations are often used in environmental sciences. Manually generating these graphs is extremely time-consuming and requires the use of individual coding solutions.

Several open-source software packages are publicly available for the visualization of complex molecular data. Exemplary, DEIMoS (Data Extraction for Integrated Multidimensional Spectrometry) [29], which is a Python package for treating data from hyphenated analytical techniques, such as liquid or gas chromatography coupled to mass spectrometry, can be named. In brief, DEIMoS allows an alignment of the *m/z* information and to visualize the extracted compounds and molecular features. Although it provides two-dimensional visualization options, often missed in vendor software, it does not give any other type of data representation like the aforementioned DBE maps and other utilization of the attributed molecular information. Furthermore, no graphical user interface (GUI) is available, and command-line programming is required. Another recently published software package is Constellation [30]. This program focuses explicitly on systematic trend detection using Kendrick mass defect (KMD) plots. It aims at finding repetitive patterns in a KMD plot obtained from non-attributed data to exhibit homologue rows with the goal of supporting and/or validate molecular formula attribution [31]. Unlike other software solutions, Constellation has the advantage of being a web application where computing takes place on a remote server. Thus, beneficially, it does not require any installation nor extended local computational resources. However, a limit is imposed on the uploaded data size, with a maximum of 5,000 peaks that can be treated at once, and permanent network connection is needed. Even though a number of 5,000 peaks seems high at first, in most FTMS data on complex mixtures, it is not sufficient as it is widespread to recover more than 10,000 attributed species per mass spectra. A recent application is PyKrev [32], which provides several useful visualization options, such as van Krevelen plots (VK) and violin representations, of intersections between samples, *i.e.*, the molecular formulas unique to a sample or common between the selected samples. In addition to these very recent software solutions, established workflows might often use older approaches, such as it is the case for OpenMS [16], published in 2016, which proposes a flexible workflow designer for data treatment and basic visualization. It has been optimized for proteomics and metabolomics and thus is not necessarily adapted to other fields like environmental sciences. Last, the work of the Barrow group needs to be mentioned here, *e.g.*,

KairosMS, published in 2020 [33]. KairosMS is specialized in handling hyphenated mass spectrometric data, including scan-by-scan recalibration, a suite of visualization tools, including DBE, VK, evolution of class intensity over time, and principal component analysis. Even though KairosMS efficiently addresses most of the desired features, it is neither open-source nor available in a public repository.

In this context, we have developed a user-friendly and open-source solution for attributed high-resolution mass spectrometric data visualization. We aimed to process and handle complex organic mixture data with ten-thousands of molecular features. This package, called PyC2MC, is based on Python as a high-level, interpreted language with an intrinsic comprehensive library and compiled with a graphical user interface. The software should be easily utilizable and improvable even by people outside of the initial project. Thus, the primary goal of this work is to provide a robust data visualization tool producing numerous plots as well as statistical analysis. Indeed, inspired by tools such as InteractiVenn [34], we implemented features allowing to print Venn diagrams or to perform clustering or principal component analysis (PCA). Intending to deliver an easy-to-use and extendable application, we use a PyQT-based GUI designed under QtDesigner for its cross-platform functionality and broad usage in the scientific community. For broad applicability to various research areas and types of complex mixtures, numerous parameters derived from the molecular formula have been considered, such as DBE, heteroatoms ratio (O/C, N/C, and others), average carbon oxidation state, aromaticity index or maximum carbonyl ratio.

Here we present our work on PyC2MC by highlighting its main features and capabilities. We will first describe the choices of programming language and libraries in the software construction, as well as the file architecture that input files should follow, exemplified in the used datasets. Then, we will present the workflow and data processing, that allow rapid plotting of basic functions, how specific KMD plots may be built and used or how environmental science variables may be represented. Finally, we will present the basic comparative features and the statistical tools available within PyC2MC.

METHODS

Environment. This code has been written in Python release 3.9.7 (August 2021). It relies on robust and broadly deployed Python libraries, the most important ones being: PyQT5, pandas [35], NumPy [36], SciPy [37], sklearn [37], matplotlib [38] and chemparse [39]. It was developed and primarily tested on a computer embedding Intel Core i5-9500 CPU at 3 GHz, 16 GB RAM, and Intel UHD Graphics 630 GPU, running under Microsoft Windows 10 as the operating system. The software was also beta-tested on various machines ranging from desktop working stations to laptops. The application can be launched directly from the main script in a Python interpreter, like Spyder 5.1.5 [40], or using the command line option. Consequently, individual changes from the programming user base can easily be made and directly tested. For classical usage cases, an executable file compiled using pyinstaller 4.8 [41] enables easy, straightforward exploitation without any programming knowledge required.

Input file architecture. Currently, four file extensions are supported: American standard code (*.asc*), comma-separated values (*.csv*), and in specific cases binary interchange file formats, such as Microsoft Office Excel sheets (*.xls* and *.xlsx)*. Examples of files are given in the supplementary information (SI 1). New formats can also be implemented easily by modifying the “loading_function.py” file.

PyC2MC does not provide a built-in molecular attribution feature, so molecular attribution should be performed using another software, exemplary workflows described below. Nevertheless, the input can be either attributed or non-attributed mass lists, as both data types may be processed to build relevant plots as described afterward. The *.asc* files are used to load non-attributed mass lists and should follow a simple structure with *m/z* values and intensities. The *.csv* files, if not exported from the already compatible vendor software, such as Bruker DataAnalysis or Thermo Fisher Scientific Xcalibur, should only be used for importing attributed data with the hereafter structure: *m/z* ratio, absolute intensity, attribution error (in ppm), molecular formula. Concerning Excel sheets, they are respectively the output format containing attributed peak lists of PetroOrg [14] and a MATLAB user-interface suite called CERES [42] (MATLAB R2022a) used in previous research [4, 42, 43]. All other variables and parameters are calculated further on. Attention should be paid to the order

of columns, and a header should be included; however, not necessarily using the names mentioned above.

Datasets. In this study, several datasets are used to demonstrate the features of the PyC2MC application. Classical fingerprint visualization functions will be illustrated with the help of recently published data obtained from analyzing plastic pyrolysis oil by direct infusion ESI(+) and APCI(+) FTICR MS [44]. The representation of parameters more common to environmental sciences will be performed using a dataset obtained with water-soluble ambient aerosol particles from emissions affected by anthropogenic industrial and wildfire sources analyzed by ESI(+/−) FTICR MS [42]. Specific use of the Kendrick mass defect plot will be shown using data gathered by the selective characterization of petroporphyrins in shipping fuels and their corresponding emission by electron-transfer MALDI FTICR MS [31]. This example highlights the usage of data on complex mixtures containing organic (C, H, N, O, S) and metal-organic (V, Ni) compounds. Finally, aerosol samples obtained under the mimicked atmospheric conditions of an exoplanet will constitute the last dataset to demonstrate the statistical analysis and inter-sample comparison. [45] With compounds having a very high N/C ratio – commonly not detected for natural earth mixtures – this dataset exemplarily represents the utilization of the PyC2MC software for less common complex matrices. Additional information on the selected datasets is found in the supplementary information (SI 2).

RESULTS AND DISCUSSION

Software workflow and data processing functionalities. Figure 1 gives a simplified graphical representation of the PyC2MC workflow. A view of the graphical user interface (GUI) is also presented in Figure 2. A primary input file can on the one hand contain a peak list (*m/z* and intensities without molecular formulae), it will then be identified by the import function as a *mass list*. On the other hand, the input file may be an *attributed* list (with molecular attribution performed with another software), it will then be identified as *attributed*. The loading of files is typically carried out from the "File" menu, and when an input file meets the specific data architecture of *attributed* or non-attributed *mass list* described in the data architecture paragraph (and SI), the import function will automatically recognize and identify it (as *attributed* or *mass list*) and select the appropriate loading and formatting method. For a user-friendly and easy recognition, the data type is transcribed in the *loaded files* section of the GUI using a color code on the file name (blue for non-attributed *mass lists* and pale yellow for *attributed* mass lists) (Figure 2). Peak lists are directly obtained through the user's routine mass spectrometric data treatment software (typically after noise detection, peak picking, creation of centroid *m/z* peak position lists with intensities).

However, for raw peak lists (in PyC2MC referred to as *mass lists*), only several basic functions of PyC2MC are available: Kendrick's mass defect plots and statistical analysis features such as volcano plot and Venn diagram. Indeed, PyC2MC, for now, does not support the attribution of peak lists to molecular formulae. Lists of attributed molecular formula (referred in PyC2MC as *attributed*) are obtained from vendor software or other software workflows, as outlined above. Instead of directly importing them, the data files may be merged by different processes to create a new data file. Most of the comparison and statistical analysis functions rely on input files created by the PyC2MC functionality of merging files, thus creating implicit 3D matrices, also called hypercubes as input [46]. Input files can be merged using one of the merge functions available. In the GUI, the function "Fuse replicates" is used to create one data file from several *attributed* mass lists obtained from analytical replicates (*n*) of the same sample. It results in a .csv file composed of the list of every ion observed in at least X replicate(s), X being an integer number chosen by the user ($1 \leq X \leq n$). Compounds not found in a sufficient number of replicates are discarded. The resulting file contains the molecular formulas, exact masses (*m/z*), arithmetic mean intensities, and individual intensity values of

each replicate. This type of file is subsequently identified as *Fused*. This function is meant for visualizing and exporting a series of replicates as one single averaged data set file and not for comparing different samples. Rather than choosing one replicate for a fingerprint plot, the *Fused* file allows the selection and illustration of robust data from a number of technical replicates.

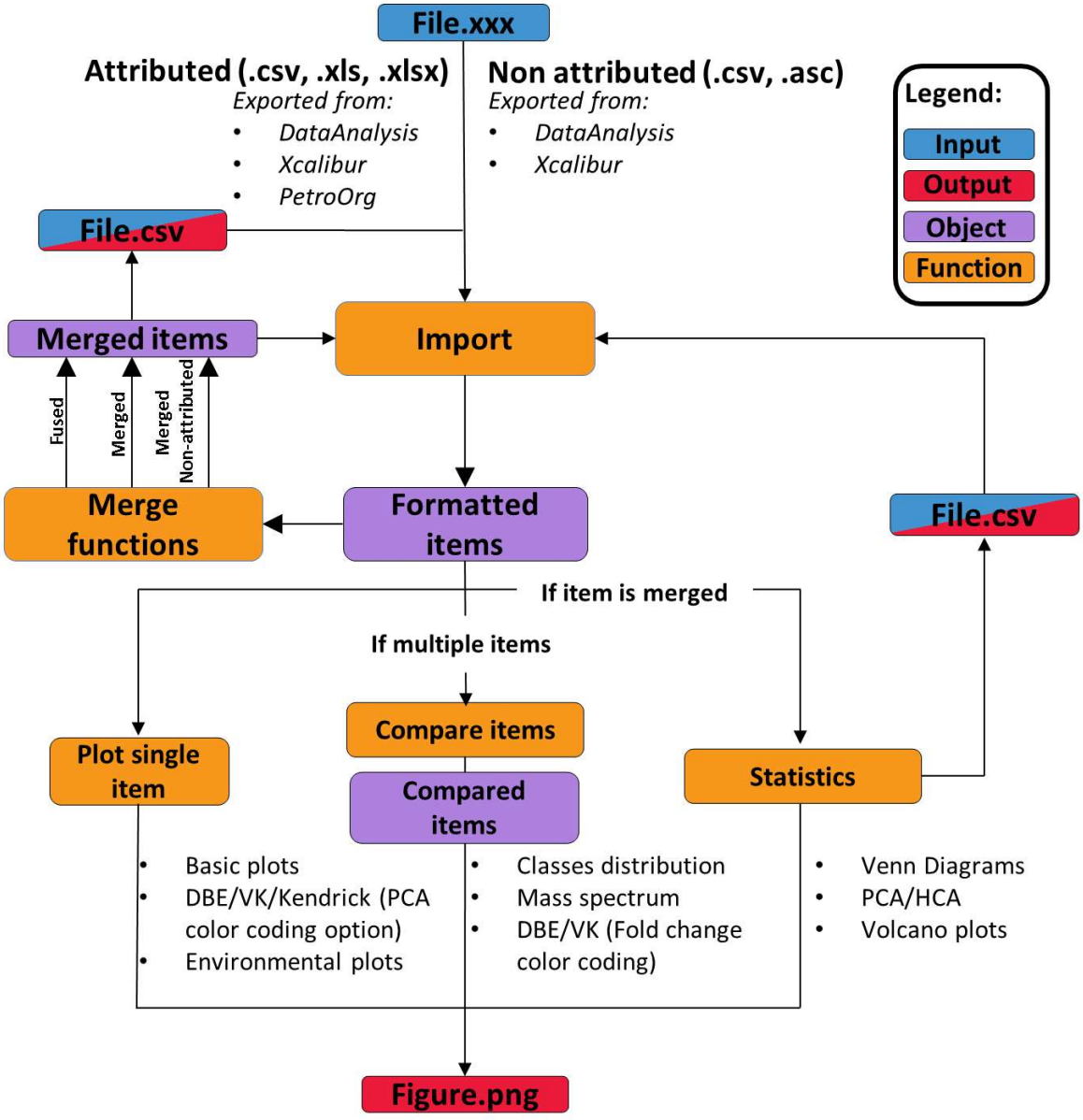


Figure III-1: Workflow diagram of the PyC2MC software. Attributed – list of elemental composition resulting from molecular formula assignment, non-attributed – peak-picked mass spectrometric data containing position as m/z and abundance (I), DBE – double bond equivalents, VK – van Krevelen.

In the “Process” menu, two other merging functions called “Merge files (with attributions)” or “Merge files (without attributions)” were designed for inter-sample comparison. These functions create one file encompassing all entries of the selected files, their intensities in each file, and the other data, such as molecular formula attribution, in the case of attributed data. The resulting files are identified either as *Merged* or as *Merge non-attributed* appearing respectively with a green or grey/violet color.

Then, depending on the type and quantity of loaded items (*Mass list*, *Attributed*, *Fused*, *Merged,* or *Merge non-attributed,* each identified in the Loaded files section of the GUI by a specific color), specific functionalities are enabled, as seen on Figure 1. For example, merged items allow for statistical analysis such as PCA and HCA.

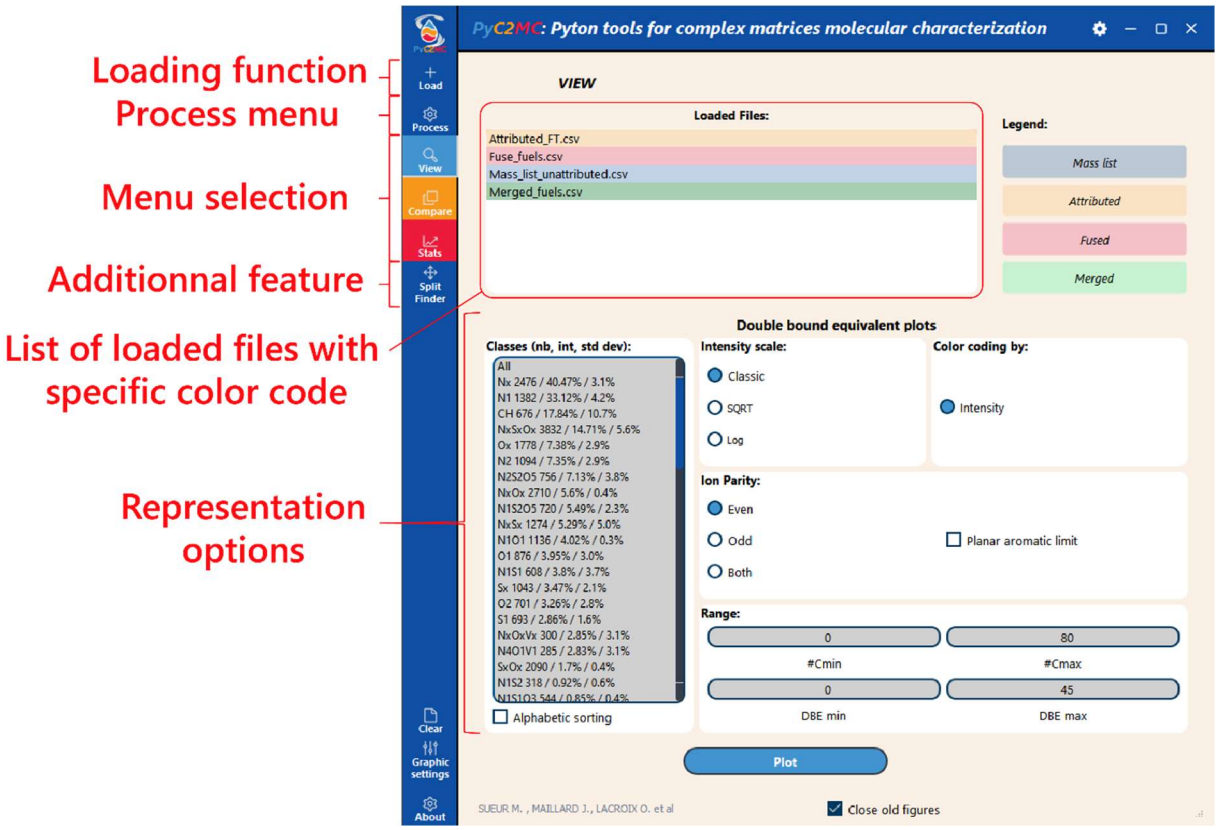


Figure III-2: Graphical user interface (GUI) of the PyC2MC viewer software package.

As stated before, a minimalist GUI is provided within the software and is shown in Figure 2. Aside from combining all functionalities, this user-friendly interface requires no dedicated and lengthy software training. The program functionalities are divided into three parts: “View”, “Compare” and “Stats”, each one corresponding to a menu accessible with the corresponding button on the left side of the GUI. The interface allows the users to directly select the part of the application they want to use, the file(s) they want to visualize among the loaded ones, the type of visualization, and finally, the options of the selected representation, if any.

Fundamental Visualization. Molecular attribution lists from the analysis of complex mixtures are classically visualized by common fingerprint approaches. Different plots may be built utilizing general parameters of the molecular formulae, such as carbon number (#C), unsaturation degree or double bond equivalents (DBE), or elemental ratios (H/C, O/C). Selected examples of these basic representations are represented in Figure 3 and available in the GUI under the “View” menu. This menu encompasses all plots that can be created using only one file containing attributions either from one sample/replicate, or more in the case of a *fused* or *merged* input file. Moreover, PyC2MC allows to create data evaluation representations like error plots: relative error (in ppm) between the theoretical *m/z* ratio calculated with the molecular formula and the observed *m/z* ratio against the observed *m/z* ratio. Using the same information, three variants of the error plot are feasible: the error distribution in a histogram or a boxplot (Figure S2), as well as error distribution versus *m/z* and abundance as the color coding variable, to evidence *m/z*-dependencies. The application can also print the mass spectrum of a selected file. Here, color-coding of signals belonging to a specific compound class, as seen in Figure 3a with the peaks of the O_x and N_xO_y colored respectively in blue and red, allow for accessible insights into complex spectral information. For first insights into the chemical composition and a rapid overview, this functionality can be beneficial, *e.g.*, this can be used to demonstrate the predominance of a compound class of interest or highlight low abundant classes within dominating molecular series.

Another approach to data reduction is to visualize the compound class distribution bar plot (summed relative or absolute abundance of all attributed signals belonging to the same compound class) shown in Figure 3b. In these plots, the data is reduced by summing the

intensities of molecular formulae belonging to the same class. Other measures are also used to present the distribution of other characteristic values such as the DBE or the number of carbon atoms in the molecular formulae (Figure 3c).

PyC2MC can also build diagrams that have become typical for complex matrices analysis such as the DBE versus carbon number maps, van Krevelen plots and Kendrick mass defect plots [47]. In such representations, a color code is usually used to represent the intensity of each species. DBE versus carbon number maps (Figure 3d) allow to bring out repetitive moieties and are useful to highlight alkylated series. The planar limit line (red line in Figure 3d) can be utilized for planar limit-assisted structural interpretation [47]. For these maps, the user can apply a filter to only display data relative to one compound class of interest. The van Krevelen plot consists in plotting the H/C ratio, or the ratio of a heteroelement to carbon, against the ratio of another heteroelement to carbon. Usually, the selected axes are H/C and O/C, giving a representation related to the degree of saturation versus oxygen content of the compounds, exemplarily given in Figure 3e. However, it is also possible to plot an atom number ratio against other parameters, such as *m/z*.

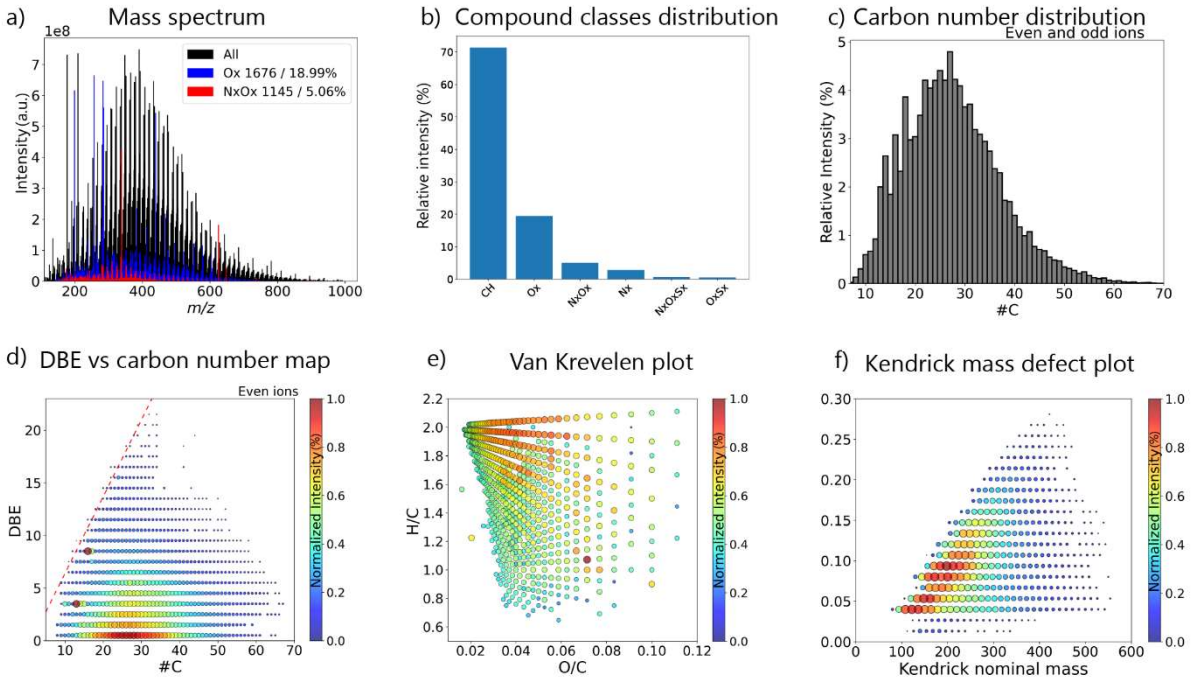


Figure III-3: Basic visualization plotted using the Petroleomics dataset: a) Mass spectrum with overlaid compound classes (Compound class, number of peaks / Relative intensity). b) Compound class distribution. c) Distribution of the number of carbon atoms. d) DBE versus C# map with the aromatic planar limit (red dotted line) showing the maximum DBE value at given C# for aromatic compounds. e) Van Krevelen plot displaying alkylation (H/C) versus oxidation (O/C). f) Kendrick plot with -CH2 as the repetition unit highlighting alkylated series without needing molecular formulas.

Finally, PyC2MC computes for Kendrick mass defect plots as displayed in Figure 3f, based on the calculation of the Kendrick mass defect (KMD) with Equation (1):

$$KMD = Nominal\ Kendrick\ mass - Kendrick\ mass \quad (1)$$

with:

$$Kendrick\ mass\ = observed\ mass\ \times \frac{Repetition\ pattern\ nominal\ mass}{Repetition\ pattern\ exact\ mass} \quad (2)$$

and:

$$Kendrick\ nominal\ mass = rounded\ Kendrick\ mass \quad (3)$$

The Kendrick nominal mass can be obtained from equation (3) using two rounding methods: either round to the closest integer or round to the upper integer (equivalent to rounding the nominal mass to the lower integer). Both methods have their limits [48] and are available within the software. Kendrick mass defect plots consist in plotting KMD versus Kendrick nominal masses, which allows species differing only by *n* repetition pattern(s) to be

displayed on the same horizontal line. Historically and still often used, the repetition pattern is CH_2; in this case, KMD plots highlight alkylated series, just as DBE versus carbon number maps. However, the repeating unit can be changed to exhibit any other interesting particularity of the sample, such as methoxy moieties (CHO), oxidation (O), or any polymeric building units (*e.g.*, polystyrene C_8H_8) [49]. Contrary to most other visualization concepts, KMD plots do not require prior molecular formula attribution and thus can be used with non-attributed mass lists.

KMD extraction for molecular formula validation. KMD plots built from non-attributed mass lists, using only the *m/z* ratio (and the intensity as a color code), may be used with PyC2MC to assist the molecular formula attribution. Indeed, the KMD value of a compound of interest can be computed and highlighted on a KMD plot, allowing one to point out homolog series with the same KMD value (same repetitive feature, such as alkylation CH_2). Non-attributed peaks and compounds sharing the same KMD can be extracted and exported into a *.csv* file and used as input in a molecular attribution software or for further cross-software usage. Classically in KMD workflows, one ion for which the molecular formula has been confirmed, may be used as starting point to deduce the molecular formulae of neighboring compounds in the KMD plots. As the KMD is calculated using a chosen repetition pattern, the user can easily find the signals corresponding to their starting compound plus or minus the chosen repetition unit and manually attribute the molecular formula in their molecular attribution software with this guidance. Figure 4 illustrates this functionality with metal-organic petroporphyrins contained in a complex organic matrix. Knowing the *m/z* ratio of the most intense porphyrin signal in this dataset ($C_{28}H_{28}N_4VO$: *m/z* 487.16972), we were able to identify not only the other porphyrins belonging to the same alkylated series (red dots) but also porphyrins of higher DBE, *i.e.*, with one or more additional carbon atoms. This functionality was particularly useful in this case, as the attribution of petroporphyrins requires validations using the isotopic fine structure [50]. This is impossible in the case of low-intensity signals but easily achievable with this Kendrick homologue row approach.

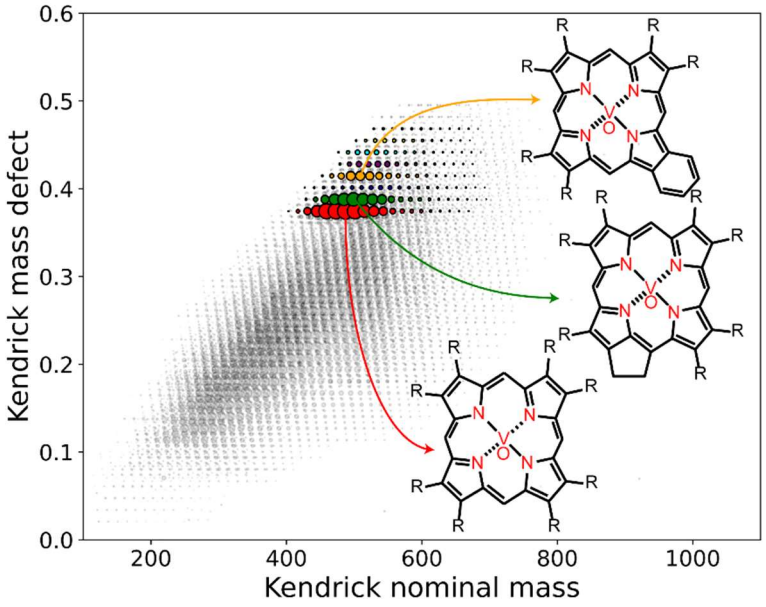


Figure III-4: PyC2MC allows to extract homologue rows on non-attributed data via visualization of KMD (CH_2 – alkylation) series. Here, the challenge of metal-organic compounds (petroporphyrins) within a complex organic matrix (bunker fuel combustion aerosol). This graphical approach facilitates molecular formula attribution or can be used as additional validation and evaluation tool.

Variables related to environmental sciences. Another feature of PyC2MC is the "Environmental science variables" section of the 'View' menu, which computes parameters and variables retrieved from the molecular formula attribution commonly utilized in environmental sciences, such as dissolved organic matter (DOM) or particulate matter (aerosol, PM) research. Namely, the average carbon oxidation state (OSC) [26], the modified aromaticity index (MAI) [27], or the maximum carbonyl ratio (MCR) [28]. Figure 5 is an example of an often-used representation (average carbon oxidation state as function of the carbon atom number or Kroll plot) from which organic aerosol classes can be easily identified and information on the chemical nature of the sample material retrieved. Here, the diagram was plotted using data from an environmental dataset, specifically extracts of ambient particulate matter sampled in areas strongly affected by wildfires [42]. Other helpful visualizations proposed in this context and given by the software are Van Krevelen-type diagrams with a color code corresponding to limit values of either MAI or MCR, see Figure S3.

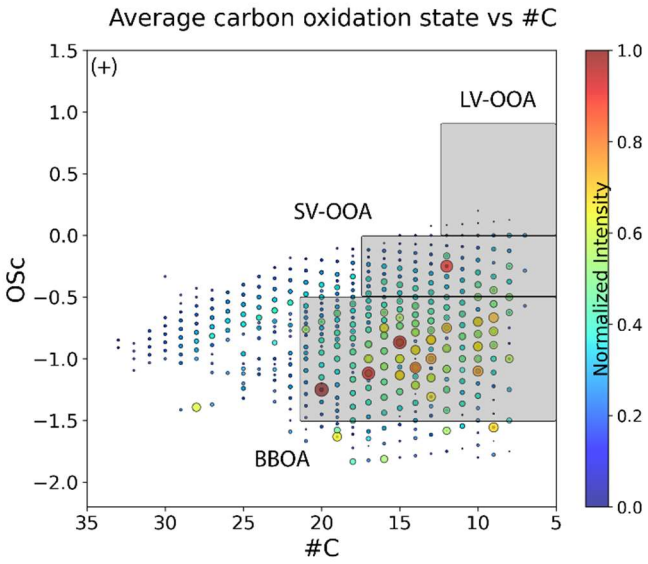


Figure III-*5*: Determining the compound classes within an ambient aerosol sample using the average carbon oxidation state versus carbon number plot (so-called Kroll plot). LV-OOA: Low-volatility oxidized organic aerosol; SV-OOA: Semi-volatile oxidized organic aerosol and BBOA: Biomass burning organic aerosol.

Primary comparative features. For the features in the compare menu, the multiple files to be compared should be loaded in the tabular section at the left center of the interface (Figure 2), and the user should select the files they want to compare among the loaded ones. A merging is then performed automatically using the same principle as the previously mentioned “Merge files (with attributions)” function but no new csv files are created, which lightly reduces computing time, thus allowing a faster comparison. Once these import and preprocessing operations are completed, the user is notified by an indicator and can plot several representations such as chemical class distribution or the DBE distribution (Figure S4) as well as, DBE versus #C maps and van Krevelen plots using a calculated fold change (FC) as color coding. The fold change is derived from equation (4), from the intensity ratio of each peak between two analyses among the merged. According to the petroinformatics principles nicely described by Hur et al. [51], the binary logarithm of the FC (log2[FC]) can be used as color coding to emphasize similarities and differences in the common chemical space based on intensity/abundance.

(4) $$FC = \frac{Peak\ intensity\ in\ sample\ 2}{Peak\ intensity\ in\ sample\ 1}$$

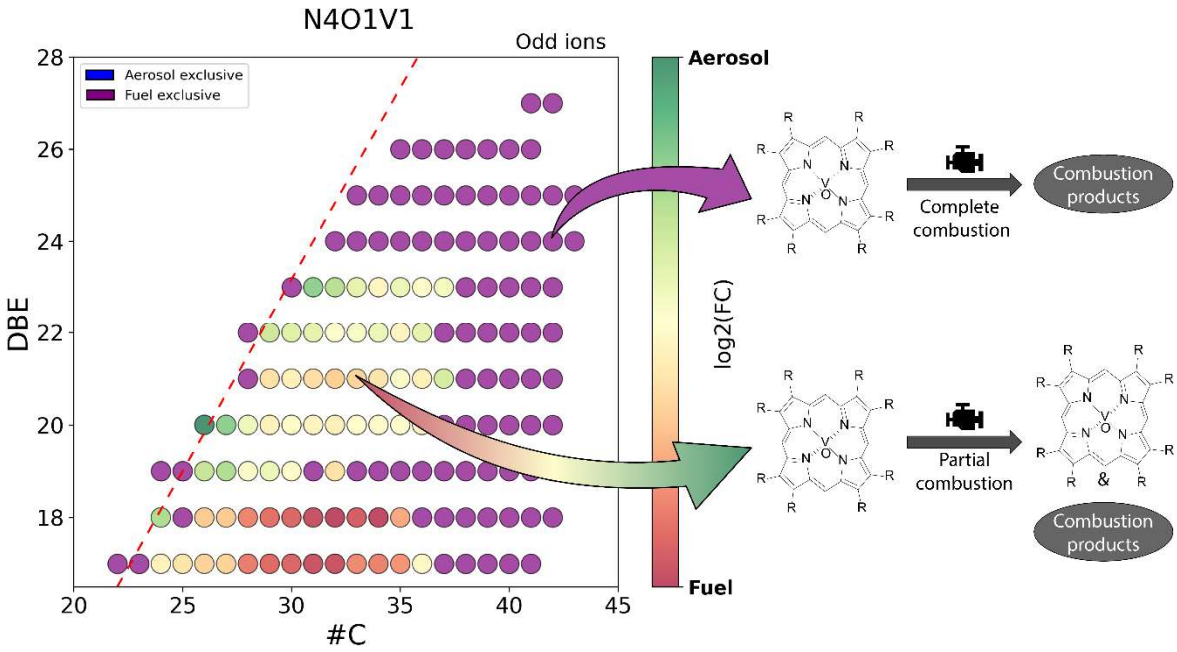


Figure III-6: Exemplary utilization of DBE versus #C fingerprint diagram with color-coded fold change information for insights into the molecular fate of petroporphyrins through the combustion of bunker fuels in a ship diesel engine (feed fuel versus primary combustion aerosol).

This kind of representation graphically summarizes in which sample each molecular feature is given with a higher relative/absolute abundance and to which extent. The petroporphyrins dataset, used in the previous section for extraction of KMD series and improved molecular attribution, was used to plot the DBE vs #C map of extracted metal-organic constituents presented in Figure 6. For this figure, the data from two samples were selected to be compared: a feed bunker ship fuel (heavy fuel oil, fuel sulphur content > 0.5 w-%) and its primary combustion aerosol gathered at the engine exhaust of a research ship diesel engine.[52] The FC color coding in Figure 6, shows that most porphyrins were solely detected

in the feed fuel (purple dots) or detected with lower intensity (red dots) in the corresponding aerosol emission, suggesting an overall total consumption through combustion.

Statistical tools and analysis. Contrary to the “Compare” menu, which allows fast and intuitive comparison, from multiple loaded files, the “Stats” menu allows more complex comparison but from one single file that should result from one of the merge processes: *merged* or *merged non attributed*. Note that the single merged file includes information on each replicate. Thus, the user is able to plot Venn diagrams and volcano plots and compute unsupervised multivariate statistics for dimensionality reduction, such as principal component analysis (PCA) or hierarchical cluster analysis (HCA). PCA and HCA are performed using the Scikit-learn [37] algorithm on the relative intensity (normalized to the summed intensity of each peak) values of each peak in each sample. The user can choose to perform the analysis on every peak in the merged file or only on those common to every sample. In the first case, when a peak is not in a sample, its intensity is set to zero. The calculated PCA loadings of each component can then be exported in a csv file to be used in basic representations under the “View” tab (namely DBE vs #C maps, van Krevelen plots, and average carbon oxidation state vs #C) where the PCA loadings can be used as the color-coding variable [43]. The results of HCA are represented in a dendrogram plot with Euclidian distance as the variable characterizing the dissimilarity between samples (Figure S5). Concerning the Venn diagrams, an .xlsx file can be exported containing the species found in each region of the Venn diagram. Finally, echoing the fold change color-coded plot in the “Compare” section, volcano plots [53] can be built using various color-coding options such as the compound class as displayed in Figure 7a where the oxygen-containing compounds of the aged tholin sample [45] are color-coded in yellow. It is also possible to select another parameter as the color-coding variable *e.g.*, *m/z* ratio, DBE or O/C ratio (Figure 7b).

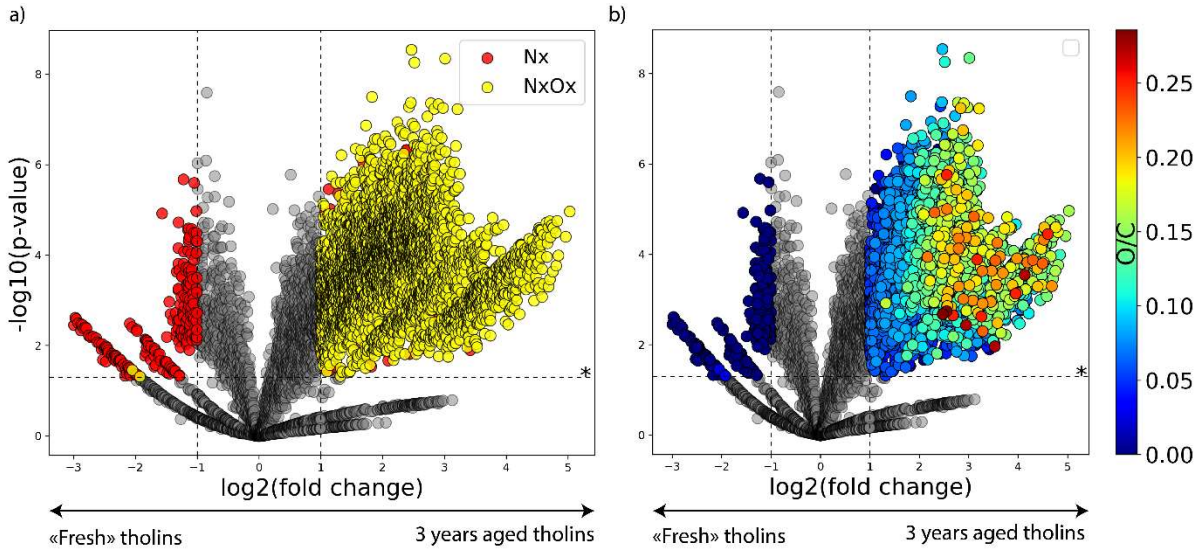


Figure III-7: Volcano plot created via PyC2MC exhibiting the oxidation process of an astrochemical mimicked aerosol sample, so-called Tholins. On the left: "Fresh" tholin sample. On the right: 3 years-old tholin sample. Each sample here consists in the arithmetic mean of 3 technical replicates. Red dots: Species containing only nitrogen. Yellow dots: Species containing nitrogen and oxygen.

Additional feature. Alongside the visualization features, PyC2MC proposes an additional feature named Split Finder, which searches for a given pair of atoms (isotopes, for example) corresponding to a given *Δm/z* within a tolerance. This feature has a separate GUI on which the user can input a *Δm/z* observed on a mass spectrum and a tolerance value (10^{-4} Da by default). This feature will then search in a database [54] comprising the molecular mass with atomic masses and isotopic compositions of each element and return the user a list of the possible atom couples, highlighting the solution with less error. A screenshot of the GUI and the results for a *Δm/z* of 1.003355 and a tolerance of 5.10^{-4} Da, *i.e.,* $^{12}C/^{13}C$ split, is shown in the supplementary information (Figure S6).

CONCLUSION

We developed an open-access program dedicated to the visualization and processing of highly complex high-resolution mass spectrometry data. On the one hand, the program can be handled by users with little to no programming experience, but on the other hand, it is improvable and/or adaptable for the community with knowledge of Python. The capabilities of this application have been demonstrated over a wide range of samples (plastic pyrolysis oil, fuel, ambient air, and Tholins simulating Titan's atmosphere). The intention of this software is to be useful for anyone treating high-resolution mass spectrometry data of complex matrices and, to do so, both standalone application and source code will be available at the following GitHub repository: (https://github.com/iC2MC/PyC2MC_viewer). The software will be continuously updated; thus, we also welcome and encourage any contribution to the repository.

ACKNOWLEDGMENTS

This work has been partially supported by the European Regional Development Fund (ERDF, HN0001343), Labex SynOrg (Grant ANR-11- LABX-0029), Carnot Institute I2C, the Graduate School for Research XL-Chem (Grant ANR-18EURE-0020), the European Union's Horizon 2020 Research Infrastructures program (Grant Agreement 731077), and Région Normandie. Access to the CNRS research infrastructure Infranalytics (FR2054) is gratefully acknowledged. We thank the DFG (ZI 764/28-1) and ANR (ANR-20-CE92-0036) for funding the research project TIMSAC.

SUPPORTING INFORMATION

1. Example files

Even though the developed application is made to load result files produced by specific third-party software, it is also possible to load generic files. In this case, two options are presented to the users: create and load a file containing non-attributed data or create and load a file containing attributed data. For the first option, the user should create a .asc file whit m/z ratio, intensity, and S/N ratio (optionally) as columns. This can be done by copy-pasting a table from spreadsheet software to a text editor and saving this file in the .asc format. No header has to be specified, only the column arrangement matters here. An example is displayed on the left side of Table S1 For the other option, a custom attributed data file should contain the m/z ratio, absolute intensity, attribution error (in ppm), and the molecular formula in a .csv file. A header must be specified for this type of file; however, it is not necessary to use the same names as only the columns arrangement matters. An example is displayed on the right side of table S1.

Non-attributed data file	Attributed data file			
No headers	m/z ratio	Intensity	Error	Sum formula
98.84049 1287542 0.00008	335.182798	16687068	0.368	C23 H27 S
101.38376 834970 0.00008	335.211790	3733477	0.219	C22 H27 N2 O
102.51128 3366713 0.00009	335.224366	2366441	-0.131	C23 H29 N O
102.61279 1008359 0.00007	335.236942	7102445	0.082	C24 H31 O
102.95155 1034841 0.00008	335.240313	6684655	0.075	C21 H35 O S
103.40512 844062 0.00009	335.248175	5672945	0.163	C23 H31 N2
103.69160 1079009 0.00009	335.260751	55652340	0.424	C24 H33 N
103.70885 1023189 0.00007	335.273328	3583993	0.360	C25 H35
103.71896 1190111 0.00009	336.063709	2128115	-0.372	C20 H16 O S2
104.10701 5630360 0.00011	336.084147	3498233	0.267	C23 H14 N S

104.41281 869834 0.00011	336.087518	2377978	0.132	C20 H18 N S2
105.26436 837208 0.00007	336.096723	23154940	0.340	C24 H16 S
106.04990 2359553 0.00013	336.100094	14373118	0.408	C21 H20 S2
106.28434 1115756 0.00008	336.114481	1868035	0.268	C24 H16 O2
106.28457 1231404 0.00008	336.125715	1520390	-0.201	C23 H16 N2 O
106.40151 1157645 0.00011	336.138291	11501834	0.013	C24 H18 N O
106.95061 1028292 0.00009	336.141662	3336459	-0.011	C21 H22 N O S
107.01977 891559 0.00009	336.150867	12230926	0.062	C25 H20 O
107.07981 1231473 0.00009	336.154238	4158735	0.112	C22 H24 O S
109.04053 908808 0.00009	336.162100	20107538	0.220	C24 H20 N2

Table S 1: Tables representing accepted architectures. On the left: The non-attributed data file contains m/z ratio, intensity, and S/N ratio in this order with only a blank space between each piece of information. On the right: The attributed data file is a proper table with a header for each piece of information: m/z ratio, intensity, error, and molecular formula.

2. Datasets details

Hereafter is the tree structure of the data files used in the main body that will also be available on the GitHub repository of the software.

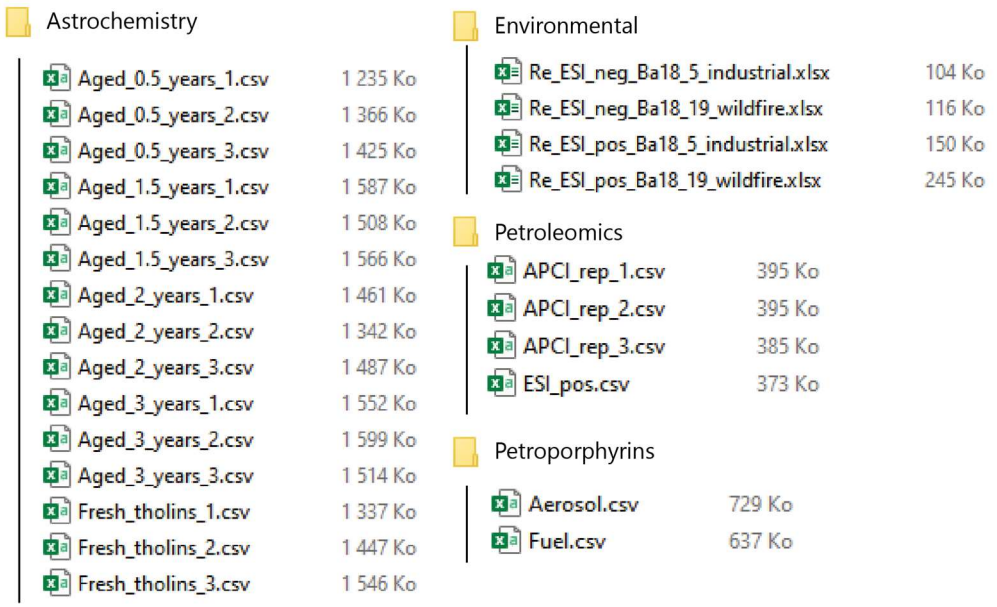


Figure S 1: Tree view of the datasets used

Dataset	Petroleomics	Petroporphyrins	Environmental	Astrochemistry
Analysis type(s)	ESI+ / APCI +	ET-MALDI	ESI +/-	LDI +
Mean number of attributions	4942	7615	1988	9719
Molecular formula boundary	$C_xH_yN_2O_5S_2$	$C_xH_yN_4$ (-VO_3) or (-Ni)	$C_xH_yN_3O_{17}S_2$	$C_xH_yO_2N_{30}$

Table S 2: Details on the datasets used

These datasets were respectively used by Mase et al. [44], Sueur et al. [31], Schneider et al[42] and Maillard et al [45] for their studies.

3. Error plots

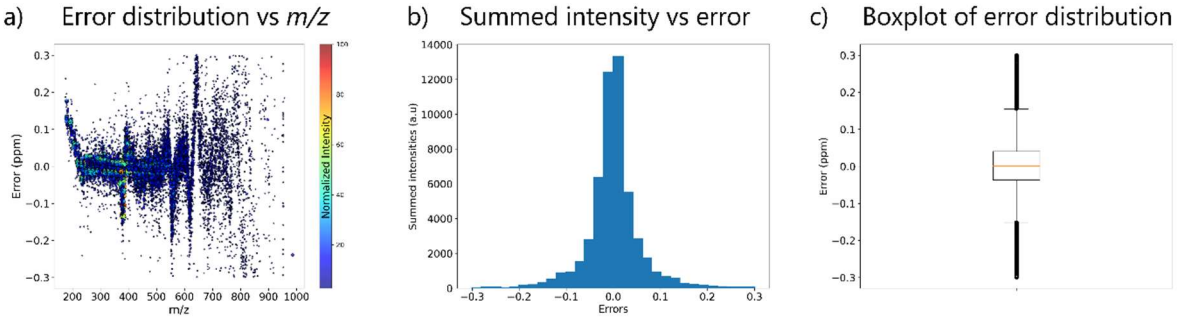


Figure S 2: Three ways to represent the attribution error in PyC2MC: a) error distribution versus m/z ratio with intensity as color code. b) histogram representing the distribution of the intensity versus attribution error. c) Boxplot representation of the error distribution.

This figure shows the different representations of the attribution error available in the software. The above diagrams were generated using the Astrochemistry dataset.

4. Aromaticity index and maximum carbonyl ratio as a color-coding variable for Van Krevelen plots

In addition to the average carbon oxidation state displayed in a Kroll plot, two other variables are commonly used in environmental science to evidence chemical properties of detected species: the maximum carbonyl ratio (MCR) and the aromaticity index (AI). MCR is calculated as follows [28]:

The resulting value gives information on the oxidation of a molecule : [0;0.2] : Very highly oxidized; [0.2;0.5] : highly oxidized; [0.5;0.9] : Intermediately oxidized and [0.9;1] : highly unsaturated. An example of a Van Krevelen diagram of an industrial wildfire sample using MCR as a color-coding variable is displayed in Figure S2.a.

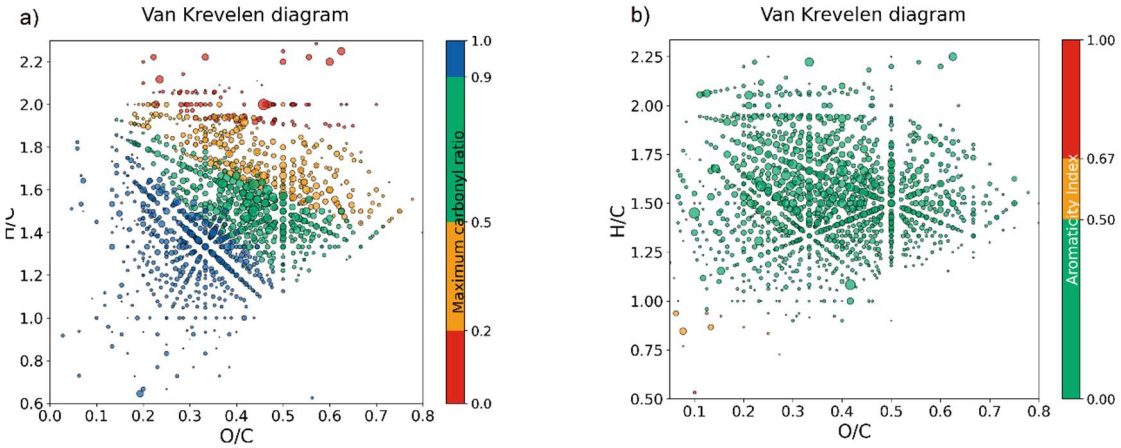


Figure S 3: Van Krevelen diagrams using a) MCR and b)AI as a color-coding variable.

Aromaticity index is a measure of the carbon-carbon double bond density[27]. It is calculated as follows:

$$AI = \frac{DBE_{AI}}{C_{AI}} = \frac{1 + C - O - S - 0.5\,H}{C - O - S - N - P}$$

$$If\ DBE_{AI} \leq 0\ or\ C_{AI} \leq 0, then\ AI = 0$$

AI thus gives information on the saturation and aromaticity of detected species: AI > 0.5 : aromatic species; AI ≥ 0.67 : condensed aromatics. An example of a Van Krevelen diagram using AI as a color-coding variable is displayed in Figure S4.b2. A modified version of the AI is

used to characterize dissolved organic matter (DOM) as in this kind of sample, approximately half of the oxygen is bound using σ-bounds rather than π-bounds:

$$AI_{mod} = \frac{1 + C - 0.5\,O - S - 0.5\,H}{C - 0.5\,O - S - N - P}$$

5. Inter sample comparison

The following figure exemplifies the primary comparative figures of the software. Figure S2.a shows the abundance of the two main compound families in Tholins samples at different ageing steps, exhibiting the oxidation process. However, Figure S2.b shows no variation of the DBE pattern correlated to the ageing process.

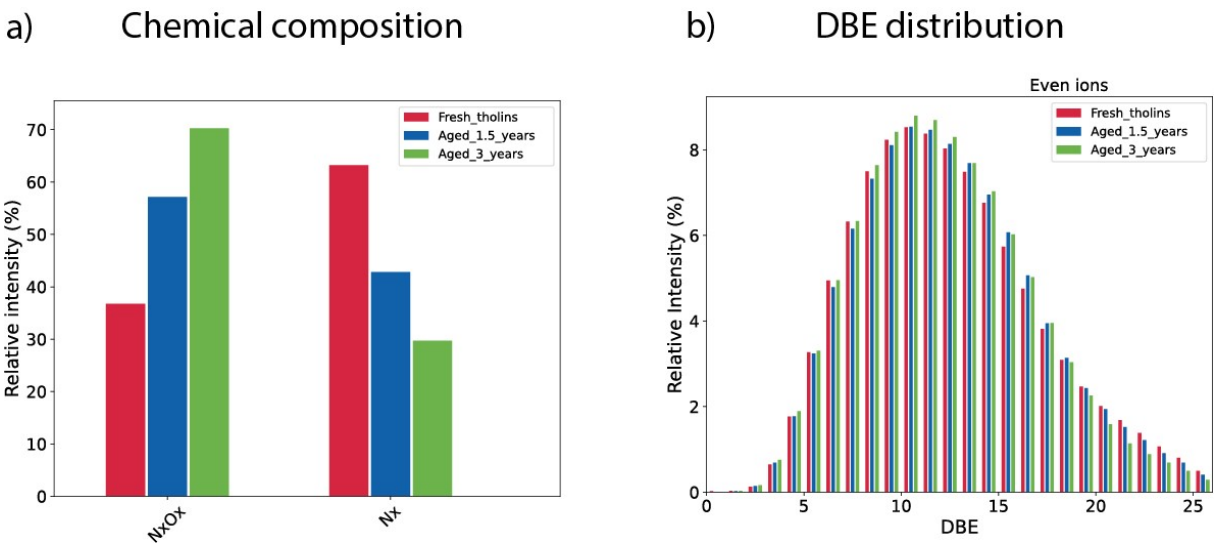


Figure S 4: Demonstration of the inter sample comparison functionalities.

6. Statistical analysis features

In addition to the volcano plot featured in the main body of this article, PyC2MC encompasses other statistical analysis tools such as PCA and HCA (See figure S3). The following diagrams were generated using the astrochemistry dataset and illustrate the separation of the samples according to their age.

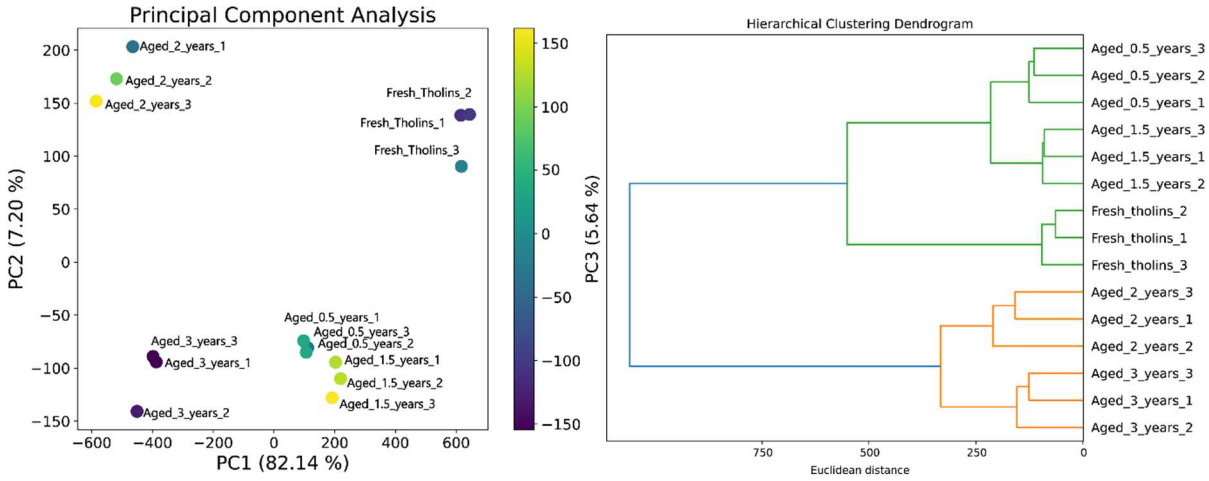


Figure S 5: Statistical analysis features. On the left: Principal Component Analysis (PCA). On the right: Hierarchical Clustering Analysis (HCA).

7. Isotope finder

The additional feature named “Isotope Finder” is provided with its own GUI which is shown in Figure S XX below. On this interface, the user can input the value of the observed split and the tolerance, both in Daltons. Then when the “Find my isotopes” is clicked, a list of the plausible isotope couples and their corresponding error is shown. The solution presenting the smallest error (in absolute value) is highlighted; however, the users should acknowledge

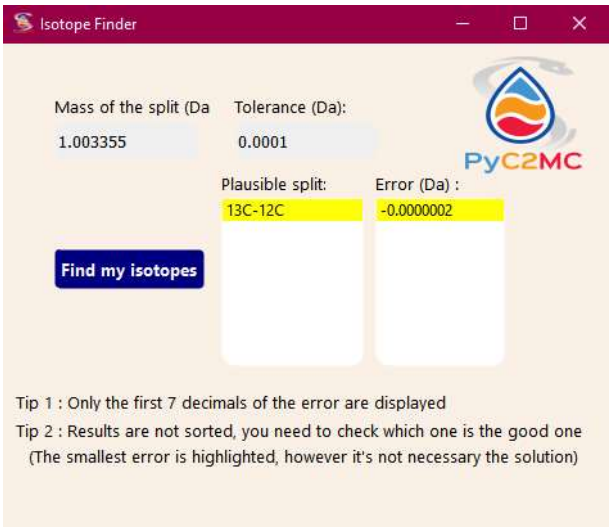


Figure S 6: Isotope finder's GUI developed under PyQT5

that this solution might not be the most plausible one.

3. Conclusion

As of today, the executable version of the software counts above 150 downloads and has been the object of suggestions and revisions that will be applied soon. In parallel, the iC2MC joint lab develops another aspect of PyC2MC focused on data treatment steps such as the *m/z* calibration, the detection of region of interest for hyphenated MS analysis and molecular formula attribution.

The next chapter focuses on the targeted analysis of nickel and vanadium-containing compounds in ship emissions by FTICR MS. This study has been performed by using the preliminary version of PyC2MC among other tools.

4. References

1. Stopka, S.A., et al., *Ambient Metabolic Profiling and Imaging of Biological Samples with Ultrahigh Molecular Resolution Using Laser Ablation Electrospray Ionization 21 Tesla FTICR Mass Spectrometry.* Analytical Chemistry, 2019. **91**(8): p. 5028-5035.

2. Rüger, C.P., et al., *Characterisation of ship diesel primary particulate matter at the molecular level by means of ultra-high-resolution mass spectrometry coupled to laser desorption ionisation--comparison of feed fuel, filter extracts and direct particle measurements.* Anal Bioanal Chem, 2015. **407**(20): p. 5923-37.

3. Marshall, A.G. and R.P. Rodgers, *Petroleomics: Chemistry of the underworld.* 2008. **105**(47): p. 18090-18095.

4. Lacroix-Andrivet, O., et al., *Molecular Characterization of Aged Bitumen with Selective and Nonselective Ionization Methods by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. 2. Statistical Approach on Multiple-Origin Samples.* Energy and Fuels, 2021. **35**(20): p. 16442-16451.

5. Cho, Y., et al., *Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics.* Mass Spectrometry Reviews, 2015. **34**(2): p. 248-263.

6. Rüger, C.P., et al., *Comprehensive chemical comparison of fuel composition and aerosol particles emitted from a ship diesel engine by gas chromatography atmospheric pressure chemical ionisation ultra-high resolution mass spectrometry with improved data processing routines.* Eur J Mass Spectrom (Chichester), 2017. **23**(1): p. 28-39.

7. He, C., et al., *Ionization selectivity of electrospray and atmospheric pressure photoionization FT-ICR MS for petroleum refinery wastewater dissolved organic matter.* Environmental Science: Processes & Impacts, 2021. **23**(10): p. 1466-1475.

8. Bianco, A., et al., *Chemical Characterization of Cloudwater Collected at Puy de Dôme by FT-ICR MS Reveals the Presence of SOA Components.* ACS Earth and Space Chemistry, 2019. **3**(10): p. 2076-2087.

9. Marshall, A.G., *Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Acc. Chem. Res., 1985. **18**: p. 316-322.

10. Marshall, A.G. and T. Chen, *40 years of Fourier transform ion cyclotron resonance mass spectrometry.* International Journal of Mass Spectrometry, 2015. **377**: p. 410-420.

11. Nikolaev, E.N., et al., *Initial Experimental Characterization of a New Ultra-High Resolution FTICR Cell with Dynamic Harmonization.* Journal of The American Society for Mass Spectrometry, 2011. **22**(7): p. 1125-1133.

12. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis.* Anal Chem, 2000. **72**(6): p. 1156-62.

13. Denisov, E., et al., *Orbitrap mass spectrometry with resolving powers above 1,000,000.* International Journal of Mass Spectrometry, 2012. **325-327**: p. 80-85.

14. Corilo, Y., *PetroOrg*. 2012.

15. Orthous-Daunay, F.-R., R. Thissen, and V. Vuitton, *Measured mass to stoichoimetric formula through exhaustive search.* Proceedings of the International Astronomical Union, 2019. **15**(S350): p. 193-199.

16. Röst, H.L., et al., *OpenMS: a flexible open-source software platform for mass spectrometry data analysis.* Nature Methods, 2016. **13**(9): p. 741-748.

17. Niedermeyer, T.H.J. and M. Strohalm, *mMass as a Software Tool for the Annotation of Cyclic Peptide Tandem Mass Spectra.* PLoS ONE, 2012. **7**(9): p. e44913.

18. Merder, J., et al., *ICBM-OCEAN: Processing Ultrahigh-Resolution Mass Spectrometry Data of Complex Molecular Mixtures.* Analytical Chemistry, 2020. **92**(10): p. 6832-6838.

19. Vetter, W., F.W. McLafferty, and F. Turecek, *Interpretation of mass spectra. Fourth edition (1993). University Science Books, Mill Valley, California.* Biological Mass Spectrometry, 1994. **23**(6): p. 379-379.

20. Le Maitre, J., et al., *Structural analysis of heavy oil fractions after hydrodenitrogenation by high-resolution tandem mass spectrometry and ion mobility spectrometry.* Faraday Discuss, 2019. **218**(0): p. 417-430.

21. Chacon-Patino, M.L., S.M. Rowland, and R.P. Rodgers, *Advances in Asphaltene Petroleomics. Part 1: Asphaltenes Are Composed of Abundant Island and Archipelago Structural Motifs.* Energy & Fuels, 2017. **31**(12): p. 13509-13518.

22. Kim, S., R.W. Kramer, and P.G. Hatcher, *Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram.* Analytical Chemistry, 2003. **75**(20): p. 5336-5344.

23. Van Krevelen, D.W., *Graphical-statistical method for the study of structure and reaction processes of coal.* Fuel, 1950. **29**: p. 269-284.

24. Kendrick, E., *A Mass Scale Based on CH2 = 14.0000 for High Resolution Mass Spectrometry of Organic Compounds.* Analytical Chemistry, 1963. **35**(13): p. 2146-2154.

25. Hughey, C.A., et al., *Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra.* Analytical Chemistry, 2001. **73**(19): p. 4676-4681.

26. Kroll, J.H., et al., *Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol.* Nature Chemistry, 2011. **3**(2): p. 133-139.

27. Koch, B.P. and T. Dittmar, *From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter.* Rapid Communications in Mass Spectrometry, 2006. **20**(5): p. 926-932.

28. Zhang, Y., et al., *The maximum carbonyl ratio (MCR) as a new index for the structural classification of secondary organic aerosol components.* Rapid Communications in Mass Spectrometry, 2021. **35**(14).

29. Colby, S.M., et al., *DEIMoS: An Open-Source Tool for Processing High-Dimensional Mass Spectrometry Data.* Analytical Chemistry, 2022. **94**(16): p. 6130-6138.

30. Letourneau, D.R. and D.A. Volmer, *Constellation: An Open-Source Web Application for Unsupervised Systematic Trend Detection in High-Resolution Mass Spectrometry Data.* J Am Soc Mass Spectrom, 2022.

31. Sueur, M., et al., *Selective characterization of petroporphyrins in shipping fuels and their corresponding emissions using electron-transfer matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry.* Fuel, 2023. **332**.

32. Kitson, E., et al., *PyKrev: A Python Library for the Analysis of Complex Mixture FT-MS Data.* Journal of the American Society for Mass Spectrometry, 2021. **32**(5): p. 1263-1267.

33. Gavard, R., et al., *KairosMS: A New Solution for the Processing of Hyphenated Ultrahigh Resolution Mass Spectrometry Data.* Analytical Chemistry, 2020. **92**(5): p. 3775-3786.

34. Heberle, H., et al., *InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams.* BMC Bioinformatics, 2015. **16**(1).

35. McKinney, W. *Data Structures for Statistical Computing in Python*. SciPy.

36. Harris, C.R., et al., *Array programming with NumPy.* Nature, 2020. **585**(7825): p. 357-362.

37. Pedregosa, F., et al., *Scikit-learn: Machine Learning in Python.* 2011.

38. Hunter, J.D., *Matplotlib: A 2D Graphics Environment.* Computing in Science & Engineering, 2007. **9**(3): p. 90-95.

39. Boyer, G. *Chemparse*. 2020 01/01/2022]; Available from: https://pypi.org/project/chemparse/.

40. Raybaut, P., *Spyder-documentation.* Available Online at: Pythonhosted. Org, 2009.

41. Cortesi, D., *PyInstaller Manual — PyInstaller 3.2.1 documentation.* Available at: <https://pyinstaller.readthedocs.io/>, 2017.

42. Schneider, E., et al., *Molecular Characterization of Water-Soluble Aerosol Particle Extracts by Ultrahigh-Resolution Mass Spectrometry: Observation of Industrial Emissions and an Atmospherically Aged Wildfire Plume at Lake Baikal.* ACS Earth and Space Chemistry, 2022. **6**(4): p. 1095-1107.

43. Castilla, C., et al., *Direct Inlet Probe Atmospheric Pressure Photo and Chemical Ionization Coupled to Ultrahigh Resolution Mass Spectrometry for the Description of Lignocellulosic Biomass.* Journal of the American Society for Mass Spectrometry, 2020. **31**(4): p. 822-831.

44. Mase, C., et al., *Molecular Characterization of a Mixed Plastic Pyrolysis Oil from Municipal Wastes by Direct Infusion Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Energy & Fuels, 2021. **35**(18): p. 14828-14837.

45. Maillard, J., et al., *Comparison of soluble and insoluble organic matter in analogues of Titan's aerosols.* Earth and Planetary Science Letters, 2018. **495**: p. 185-191.

46. Guillemant, J., et al., *Evaluating the Benefits of Data Fusion and PARAFAC for the Chemometric Analysis of FT-ICR MS Data Sets from Gas Oil Samples.* Energy & Fuels, 2020. **34**(7): p. 8195-8205.

47. Cho, Y., Y.H. Kim, and S. Kim, *Planar Limit-Assisted Structural Interpretation of Saturates/Aromatics/Resins/Asphaltenes Fractionated Crude Oil Compounds Observed by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Analytical Chemistry, 2011. **83**(15): p. 6068-6073.

48. Lacroix-Andrivet, O., et al., *Molecular Characterization of Formulated Lubricants and Additive Packages Using Kendrick Mass Defect Determined by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Journal of the American Society for Mass Spectrometry, 2022. **33**(7): p. 1194-1203.

49. Fouquet, T.N.J., *The Kendrick analysis for polymer mass spectrometry.* Journal of Mass Spectrometry, 2019. **54**(12): p. 933-947.

50. Caumette, G., et al., *Element speciation analysis of petroleum and related materials.* Journal of Analytical Atomic Spectrometry, 2009. **24**(3): p. 263.

51. Hur, M., S. Kim, and C.S. Hsu, *Petroinformatics*, in *Springer Handbook of Petroleum Technology*, C.S. Hsu and P.R. Robinson, Editors. 2017, Springer International Publishing: Cham. p. 173-198.

52. Jeong, S., et al., *Aerosol emissions from a marine diesel engine running on different fuels and effects of exhaust gas cleaning measures.* Environ Pollut, 2022. **316**(Pt 1): p. 120526.

53. Hur, M., et al., *Statistically Significant Differences in Composition of Petroleum Crude Oils Revealed by Volcano Plots Generated from Ultrahigh Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectra.* Energy & Fuels, 2018. **32**(2): p. 1206-1212.

54. Coursey, J.S., et al., *Atomic Weights and Isotopic Compositions (version 4.1)*. 2015.

Chapter IV. Investigation of Ship Emissions Molecular Composition

1. Introduction and Summary of "Selective characterization of petroporphyrins in shipping fuels and their corresponding emissions using electron-transfer matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry"

The elemental analysis of ship fuels exhibits the presence of heavy metals such as vanadium or nickel. These metals are known to be present in fossil fuels as a degradation product of chlorophylls: the petroporphyrins. The petroporphyrins consist of a tetrapyrrole-based core complexing a metal ion such as vanadium or nickel. In fact, the presence of these metals, at the atomic level, has also been evidenced by elemental composition analyses of ship emissions. The goal of this study is to determine the fate of the petroporphyrins during the combustion process in a ship engine. For this purpose, we used the realistic ship engine described in Chapter II and commercial ship fuels with different V and Ni measured contents and different combustion conditions to emulate the running of a freight ship. The used fuels and corresponding particulate matter (PM) emissions were analyzed by electron-transfer matrix-assisted laser desorption/ionization (ET-MALDI) FTICR MS. This analytical technique allowed for enhancing the ionization of the petroporphyrins in both ship fuels and emissions. First, a preliminary study of the petroporphyrins was performed in ship fuels, allowing to exhibit the predominance of vanadyl petroporphyrins over nickel porphyrins as suggested by the elemental composition analysis. This preliminary characterization also permitted to draw out three main double bond equivalent (DBE) values for the detected porphyrins. As the DBE is a structural descriptor, we were able to propose potential structures for the porphyrins we detected linked with the results of previous research. Then, we focused on the fate of the porphyrins through the combustion process by analyzing the PM and using the fold change as an indicator of the consumption or production of each porphyrin during the combustion process by comparing their relative intensity evolution between the fuels samples and the PM emissions. We were able to highlight that the petroporphyrins are homogeneously consumed during the combustion and that no specific structure, suggested by the DBE, is predominantly degraded. Furthermore, we studied the influence of the engine load on the emission

composition and were able to determine that the porphyrins are more likely to be released under a low engine load rather than under a high load probably due to an incomplete combustion allowing the emission of petroporphyrins as a component of the unburnt fuel droplets. This study was published, and its full-text is available hereafter.

2. Full text article: "Selective characterization of petroporphyrins in shipping fuels and their corresponding emissions using electron-transfer matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry"

Maxime SUEUR[1,2], Christopher P. RÜGER*[2,3,4], Julien F. MAILLARD[1,2], Hélène LAVANANT[1], Ralf ZIMMERMANN[3,4], Carlos AFONSO[1,2]

[1]Normandie Université, COBRA, UMR 6014 et FR 3038, Université de Rouen, INSA de Rouen-Normandie, CNRS, IRCOF, Mont Saint Aignan Cedex

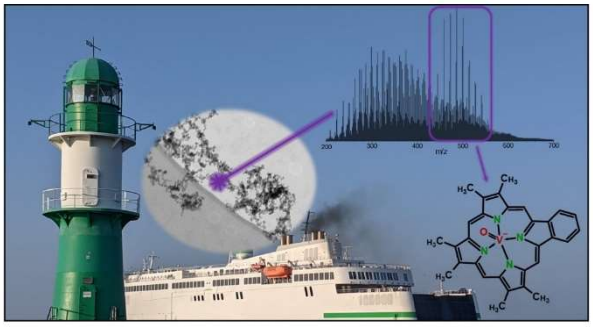
[2]International Joint Laboratory - iC2MC: Complex Matrices Molecular Characterization, TRTG, BP 27, 76700 Harfleur, France.

[3]Joint Mass Spectrometry Centre/Chair of Analytical Chemistry, University of Rostock, 18059 Rostock, Germany.

[4]Department Life, Light & Matter (LLM), University of Rostock, 18051 Rostock, Germany

*** corresponding author: christopher.rueger@uni-rostock.de**

Graphical Abstract



Abstract

In the context of the global transportation of goods, shipping emissions account for a significant proportion of air pollution. Indeed, a focus has been made over recent years on this primary emission source, leading to several regulations with respect to the chemical composition of shipping fuels. However, these regulations mainly concern the fuel sulfur content (FSC) and do not consider other compound classes such as polycyclic aromatic hydrocarbons (PAHs) or metal-containing aromatics, i.e., petroporphyrins, known to be present in bunker fuels. Petroporphyrins are tetrapyrrole-based metal complexes derived from the transformation of chlorophylls through geological time scales. In contrast to PAHs, their fate in the combustion process and effects on environmental health are widely unknown. In this study, we present electron-transfer ionization in matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry (ET-MALDI FTICR MS) for the characterization of vanadyl and nickel porphyrins in shipping feed fuels and primary particulate matter emissions. For the first time, these petroporphyrins could successfully be described in the heavy fuel oil (HFO) feeds but also in the particles emitted by the combustion of the respective fuel on a molecular level. Three main alkylated series of porphyrins were observed; these series can be qualified by their double bond equivalent, i.e. 17, 18, and 20, and correspond to various core structures. Our results highlight the molecular fate of the petroporphyrins through combustion and show that a significant amount of petroporphyrins is released unburned or partially dealkylated to the atmosphere. Furthermore, our results suggest that a higher amount of petroporphyrins might be released in harbors than in open sea, due to a less efficient combustion at lower engine load. This last observation motivates a future specific study on porphyrins' health and environmental effects.

1.Introduction.

Among the various sources of anthropogenic emissions, shipping is one of the largest contributors to the total emissions of the transportation sector. In 2007, Corbett et al. estimated that shipping emissions accounted globally for 60,000 cardiopulmonary and lung cancer deaths annually under the regulation at that time. Most of those deaths can be related to major trade routes such as European and southwest Asian coastlines.[1] Furthermore, other health effects can be associated to shipping emissions via oxidative stress and inflammation.[2] Aside from the environmental health aspects, shipping emissions have a direct impact on climate, depending on their intrinsic radiative forcing. For example, black carbon (BC)[3], CO_2 and O_3 have a positive radiative forcing (RF) as opposed to sulfate particles which have a negative RF. Organic aerosols also induce indirect effects on climate by modifying the optical properties of clouds, causing a negative RF. This indirect effect is estimated to outweigh the greenhouse gas warming effect, leading to a negative net RF from the shipping emissions.[4] Shipping emissions exhibit a unique molecular profile, different from other sources, such as biomass burning or road engine emissions.[5] In 2020, a novel regulation limiting the sulfur content in shipping fuels from 3.5% (m/m) to 0.5% (m/m) by International Marine Organization (IMO) became effective[6]. This regulation is even more restrictive in Sulfur Emission Control Areas (SECAs), where fuel-sulfur content (FSC) has been limited to 0.1% (m/m) since 2015[6]. The aim of this restriction was to decrease the sulfur content of the shipping emissions either by causing a switch from classical bunker or heavy fuel oils (HFO) towards cleaner feed fuel alternatives such as marine gas oil (MGO) or by using exhaust gas after-treatment (scrubber technology) to wash the emissions.[7] Compared to MGO, HFO emissions exhibit a particular toxic mixture with higher concentrations of particulate matter (PM), polycyclic aromatic hydrocarbons (PAH), heavy alkanes and metals, such as nickel and vanadium.[8-13] However, combustion of the cleaner distillate fuel MGO emits higher levels of soot, which is a known carcinogen, and induces strong toxicological effects.[14] Furthermore, most organic species in the aerosol are transferred directly from the fuel as unburnt material, especially at lower power operations.[15] Vanadium and nickel in HFO are mostly present as oxides. However, approximately 10% are present as organometallic compounds such as petroporphyrins.[16, 17] Petroporphyrins are tetrapyrrole-based metal complexes.[18] They are the results of chlorophylls' degradation over geological time through diagenesis and catagenesis

processes[19]. The molecular fate of these porphyrins after combustion in a ship engine is largely unknown. Furthermore, even though, health and environmental effects of heavy metals such as nickel and vanadium are studied[20, 21], the ecological impact of organic species containing such metals is largely unknown.

Bunker fuels and their corresponding combustion engine emissions are highly complex mixtures composed of numerous organic and inorganic species. This chemical complexity requires state-of-the-art instrumentation for molecular-level analysis. Examples of such instrumentation, specifically deployed towards shipping emissions, involve chromatographic separations like gas chromatography selective ion monitoring mass spectrometry (GC-SIM-MS)[22], or , high resolution time of flight mass spectrometry (HR-ToF-MS) coupled to two-dimensional gas chromatography (GC×GC). Other use direct introduction or other means of sample introduction and ionization like thermal-desorption of needle trap devices coupled to photoionization time-of-flight mass spectrometry (NTD-TD-REMPI-ToF-MS)[23], direct inlet probe (DIP) or thermogravimetric analysis (TGA)[24], proton transfer reaction mass spectrometer (PTR-MS)[25]. Finally, high resolution-time of flight-aerosol mass spectrometry (HR-ToF-AMS)[26], consists in HR-ToF-MS coupled to a sampling, separating and ionizing unit allowing real time analysis of aerosols[27]. Nonetheless, despite the existence of established technologies in aerosol science, none of these techniques was able to attribute metalorganic residues in carbonaceous aerosols on a molecular level. Another technique that has been used for the in-depth characterization of shipping fuels and their emissions is the laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry (LDI-FTICR-MS)[28], which promotes the ionization of aromatic and polycyclic aromatic compounds. In this respect, targeting petroporphyrins, with their aromatic functionalities, in particulate matter by LDI might be suitable[29, 30]. LDI time-of-flight mass spectrometry has been used in literature for the characterization of petroporphyrins from various petroleum matrices[29, 31]. Particularly, the Kilpatrick group contributed methods for petroporphyrin purification and molecular analysis[30, 32].

However, direct LDI may cause fragmentation of labile analytes such as the petroporphyrins, potentially cleaving off alkylation side-chains or altering the aromatic backbone.[33] Electron-transfer matrix-assisted laser desorption/ionization (ET-MALDI) has

been shown to circumvent this drawback for the specific compound class of petroporphyrins [31]. In brief, ET matrices allow the ionization of analytes based on their ionization energy (IE): if the matrix' IE is higher than the analyte's IE, its ionization can occur through charge exchange processes,[34] and vanadium and nickel porphyrins exhibit a distinctly low IE (6.5-7.5 eV)[33]. Derivated cyano-phenylenevinylenes (CNPV) has been reported as an efficient matrix for selective ionization of petroporphyrins in crude oil, and especially the α-cyanophenylenevinylene derivative (α-CNPV-CH_3, IE=8.42 eV) recently developed and introduced by Ramirez-Pradilla et al.[33-35] Other ionization techniques are able to detect petroporphyrins, such as atmospheric pressure photoionization (APPI)[36, 37] and positive-ion electrospray ionization-(ESI+) using sodium attachment instead of proton or electron transfer as an ionization channel[38]. Bunker fuel emissions are highly complex, thus high mass accuracy and ultrahigh resolution of the FTICR MS are required to distinguish the targeted petroporphyrins from other molecular classes.

The scope of this study is the characterization of the petroporphyrins in different bunker fuels and the respective emissions obtained after their combustion in a ship diesel engine by ET-MALDI FTICR MS. Indeed, we chose ET-MALDI to address these metalorganic constituents not only in petrochemical feed fuels but also in the linked combustion emissions. This work focuses on tracing the fate of petroporphyrins in the combustion process by comparing the fuels and their corresponding emissions. We aimed at tracing dealkylation processes, oxidation, or other reactions that occur prior to the thermal decomposition of these metalorganic targets in the engine into smaller products, ultimately to metal oxides, NO_x, CO_2, and water. For this purpose, feed and emissions of HFOs with different FSC were studied, and of a porphyrin-free MGO, which was used as a control. To the best of the authors' knowledge, this is the first method aiming at the analysis of petroporphyrins on the molecular level in an environmental aerosol sample. This molecular knowledge should help to understand shipping emissions more in detail and especially the, for now, somewhat unknown and poorly described petroporphyrins fraction.

2.Experimental section

2.1.Fuel samples characteristics and preparation. Four fuels used for shipping with different chemical properties were investigated (Table 1). The heavy fuel oils (HFO) differ in their sulfur contents. Among the four fuels, only the marine gas oil (MGO) and the heavy fuel oil A (HFO A), have a sulfur content less than or equal to 0.5% (m/m), meeting the sulfur requirement from the International Marine Organization[6]. 1 mg of each fuel was diluted in 2 mL of HPLC grade (99.9%) acetonitrile (ACN, CAS: 75-05-8) and put in an ultrasonic bath for 10 min. The ACN was then completely evaporated using a N_2 stream and the samples were diluted in 2 mL of extra pure (99.9%) tetrahydrofuran (THF, CAS: 109-99-9). Furthermore, we also made a blank sample without using fuel to have a control sample.

Table IV-1: Sulfur, vanadium and nickel mass fraction of the fuels.

Fuel	Sulfur content (%)	Vanadium content (ppm)	Nickel content (ppm)
MGO	0.0015	<1	<1
HFO A	0.53	26	20
HFO B	1.02	57	<1
HFO C	2.16	237	23

2.2.Aerosol samples characteristics and preparation. PM was obtained through the combustion of the above-mentioned fuels in a one-stroke 80 kW research ship diesel engine operating at four different power loads during a measurement campaign at the University of Rostock (Germany) in 2020 (Table 2). See Table S 1 for engine details. This engine ran at the four power loads following the ISO 8178-4 E2 norm, i.e. 24 min at 100%, 60 min at 75% and 18 min at 50 and 25% as described by Streibel et al. for a previous campaign[13, 15]. The particulate matter was sampled on pre-baked and cleaned quartz fiber filters (QFF) at specific operating points and subjected to an elemental analysis by an in-situ inductively coupled plasma (ICP). For the analysis referred herein, 1/8 of each QFF was extracted in 1.5 mL of ACN and mixed in an ultrasonic bath for 10 minutes. The extract was then filtered on 0.2 µm PTFE membrane filter using a syringe filter holder and dried under a N_2 flow until the ACN evaporated. Finally, the dry extracts were dissolved in 50µL of ACN. Furthermore, the extraction step was also performed on a blank filter to obtain a control sample.

Table IV-2: Mass of vanadium and nickel quantified on the filter depending on the fuel and power load of the engine.

Sample name	Fuel used	Load (kW)	Vanadium amount on the filter (µg)	Nickel amount on the filter (µg)
PM A	HFO A	20 (25%)	1.37	0.70
PM B	HFO B	20 (25%)	0.73	0.54
PM C	HFO C	20 (25%)	7.39	1.95
PM C'	HFO C	60 (75%)	9.89	2.43

The α-cyanophenylenevynylene ET-matrix (α-CNPV-CH_3) was synthesized as previously reported [33] and dissolved in THF to reach a concentration of 1mg/mL.

2.3.Mass spectrometry. Mass spectra were acquired using a Bruker solarix XR 12T (Paracell™) instrument, equipped with a MALDI source (3.5 mbar) using a Nd:YAG×3 solid-state laser (355nm) in positive mode. The laser was used under the following conditions: focus 40.4%; frequency 2 kHz; number of shots per scan: 100. Different laser power settings were tested within a range of 14 to 30% and found to be optimal at 20% for LDI experiments on fuel extracts, and 15 % or 17 % for ET-MALDI experiments on fuel and particles extracts, respectively (Figure S 1 in the supplementary information) [39]. 200 transients were acquired and added up. Each transient was digitalized using 4M data points and a mass range ranging from *m/z* 110 to 1100, resulting in a transient duration of 1.26 s. The time of flight for trapping ions in the ICR cell was set to 1 ms. For LDI-FTICR experiments, 7 µL of each fuel extract were deposited on the MALDI target. For the MALDI FTICR experiments, 2 µL of matrix solution were mixed with 5 µL of fuel extract and 5 µL of matrix were mixed with 2 µL of particle extract prior to deposition on the MALDI target. More matrix solution had to be used for the study of particle extract because, as shown later, porphyrin signals relative intensity is much lower in the particles than in the fuel. Therefore it required a higher quantity of matrix to enhance porphyrins ionization and to be able to identify the porphyrins among the other signals of particle extracts. For each analysis, triplicate measurements were collected on a same sample.

2.4.Porphyrins attributions. Raw spectra were treated with Data Analysis 5.0 (Bruker Daltonics, Bremen, Germany). Internal calibration was performed using already known $C_xH_yN_1$ and C_xH_y molecular series and $C_xH_yN_4VO$ species identified in HFO C (termed N_1, CH, and N_4VO). The internal calibration list is given in supporting information (See Table S 2). A preliminary porphyrins attribution was performed using the SmartFormula tool from DataAnalysis. To focus on metal-containing ions, the molecular attribution boundaries were set as follow: $C_xH_yN_4VO_{1\text{-}3}S_{0\text{-}1}$ for vanadyl porphyrins and $C_xH_yN_4Ni$ for nickel porphyrins; with a minimal double bond equivalent[40] (DBE) of 17 , a minimum signal-to-noise ratio of 4 and a maximum mass error of 0.5 ppm. More flexible limits were also used to detect any chemical transformation through combustion of the porphyrins, i.e. lower DBE boundary, higher oxygen number (up to 7) and lower number of nitrogen (down to 2). Then, the peak list was also subjected to a Kendrick mass defect analysis[41, 42] using a self-written python 3.7.11 routine to extract each of alkylated porphyrin series , i.e. porphyrins with the same DBE, directly from the mass list. Details on this procedure are given in the supporting information and illustrated on Fig S 2. Within a series, any missing porphyrin were searched manually with DataAnalysis and their isotopic pattern was checked. Several molecular attributions of porphyrins therefore slightly exceeded the maximum error or had a low relative intensity and signal-to-noise ratio. Self-written python routines were used for plotting and data visualization.

3.Results and discussion

3.1.Effect of the ET-matrix on fuel and particle samples. Prior to the in-depth characterization of the porphyrins, the effects and suitability of the ET-matrix was investigated. The main interest of this type of matrix is to hinder the ionization of species having a higher ionization energy (above 8.42 eV) in favor of species having a lower ionization energy, such as PAHs with large conjugated π electron systems or petroporphyrins with their electro-excessive pyrrole cycles. Indeed, the ionization energy (IE) of the α-CNPV-CH3 matrix is higher than the one of the petroporphyrins; this difference of IE, superior to a 0.5 eV gap[43], allows for electron transfer reactions to occur as described in equation (1) and (2) where m is the matrix and A the analyte.

$m + hv \rightarrow m^{+\bullet} + e^{-}$ (1)

$m^{+\bullet} + A \rightarrow m + A^{+\bullet}$ (2)

Figure 1 shows the LDI mass spectra acquired on fuel and PM deposits without matrix addition (Figure 1a and 1c) compared to mass spectra obtained with the ET matrix (Figure 1b and 1d). Mass spectrometric enlargements to the characteristic *m/z* region for petroporphyrins found in crude oils is shown (i.e., between *m/z* 440 and 560). The molecular complexity of the primary ship emissions and corresponding feed fuels expectedly resulted in a high number and density of peaks covering a broad mass range from *m/z* 200 up to 700, in agreement to previous investigations by LDI FTICR MS[28]. Characteristic series of peaks with a delta *m/z* of 14, corresponding to CH_2 alkylation pattern, are depicted with colored dots. To encompass the isobaric complexity, one may observe that the ET-MALDI FTICR mass spectrum of HFO C (Figure 1b) included a total of 26 peaks above a signal-to-noise level of 4 in a range of only 300 mDa (Figure S 3).

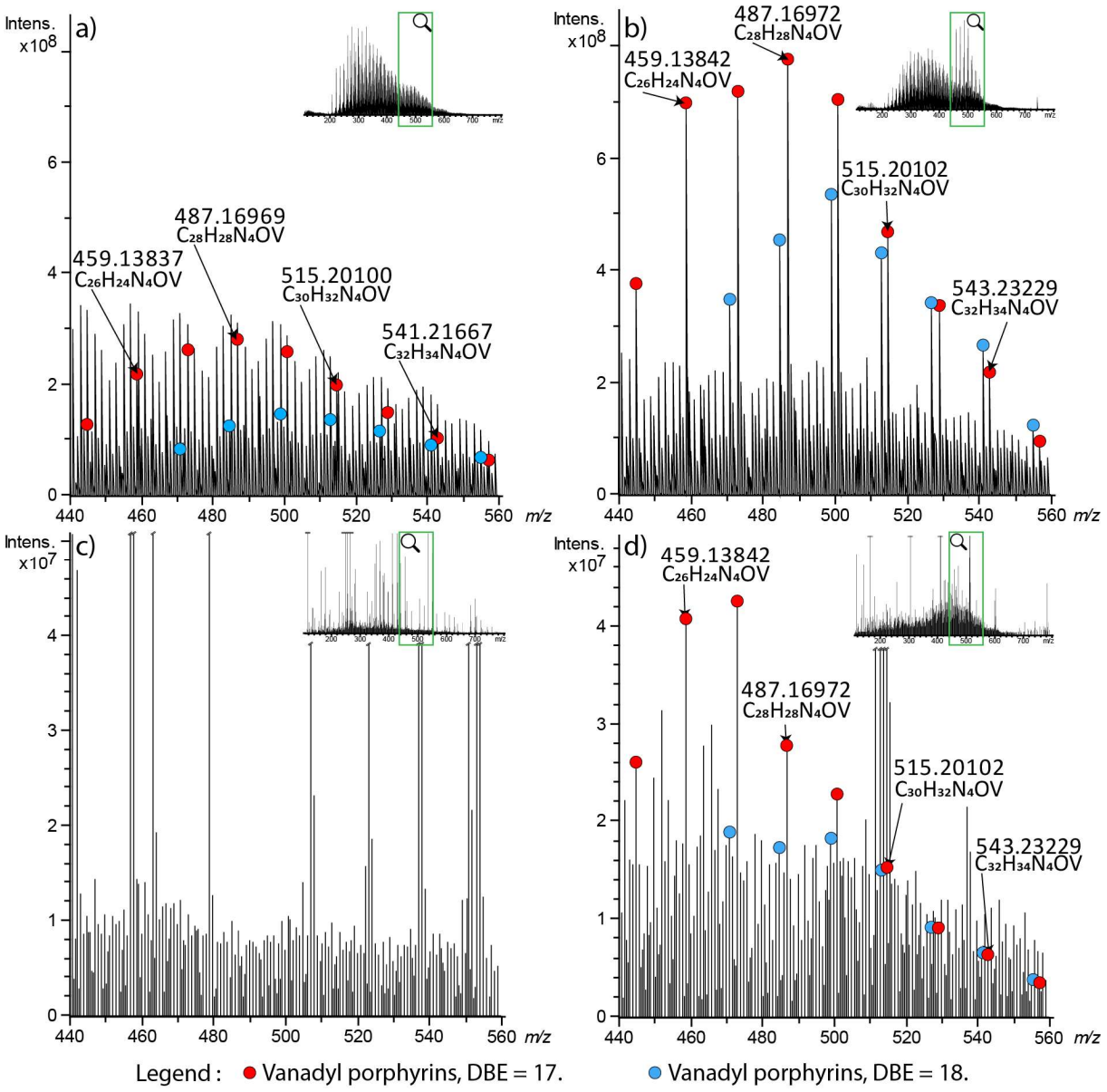


Figure IV-1: Broadband and zoomed mass spectra presenting the influence of the ET-Matrix in the petroporphyrins range of mass to charge ratio. The occurrence of selected petroporphyrin series is given as red (vanadyl, DBE 17) and blue (vanadyl, DBE 18) marks. a): HFO C; b): HFO C with matrix; c)PM C; d): PM C with matrix. No porphyrin species were detected in the PM sample without ET matrix.

The isobaric complexity confirmed the need to use a FTICR MS platform to be able to resolve the peaks of interest among the multitude of signals originating from the sample and matrix signals. In fact, the mass difference between a vanadyl petroporphyrins such as $C_{28}H_{28}N_4OV$ and the $C_{36}H_{23}O_2$ from the $C_xH_yO_2$ oxygenated series present within both fuel and PM mass spectra is below 0.5 mDa. On the LDI mass spectrum of HFO C (Figure 1a), the main distribution centered around *m/z* 350. Classically, it corresponds to the apolar species ionized

by laser desorption, i.e., homologue alkylated series of the CH, S_1 S_2, N_1 and N_2 classes. Another distribution around *m/z* 500 corresponded to the porphyrins' region. In each enlargement, attributed vanadyl porphyrins with a DBE of 17 are marked with a red dot and those with DBE 18 with a blue dot. In the shipping feed fuel sample, the effect of the matrix addition is already easily noticeable in the broadband visualization (insets in figure 1) as the addition of the matrix caused the distribution around *m/z* 500 to appear. The signals corresponding to the vanadyl porphyrins series were significantly more intense in the ET-MALDI mass spectrum of the fuel sample (Figure 1b) than in the LDI mass spectrum of the same fuel sample (Figure 1a), to the point of becoming the most intense peaks of the mass spectrum. In the LDI mass spectrum of the PM sample (Figures 1c), a low-intensity distribution centered around *m/z* 350 could be detected but no porphyrins were evidenced. However, on the ET-MALDI mass spectrum of the PM sample (Figures 1d), the distribution around *m/z* 500 could clearly be distinguished and several signals could be attributed to porphyrins. In fact, the chosen α-CNPV-CH3 matrix was able to significantly enhance the detection of petroporphyrins in both fuel and PM samples. Such enhancement of detection of petroporphyrins, thanks to the selective ionization of low IE molecules, first evidenced in heavy residues, crude oils and asphaltenes[33, 35] from fossil origin could therefore be successfully transferred here to the analysis of environmental PM samples. In the case of the primary ship emission PM, the drastic increase in ionization response by addition of the ET matrix allowed to overcome the low concentrations and made a detection feasible. All following experiments were thereafter performed with the ET matrix to ensure an efficient characterization of petroporphyrins and to compare the petroporphyrin content between feed fuel and PM emission. In addition to the previously characterized fuel samples, marine gas oil (MGO), a light fuel used in freight transport, was analyzed but no petroporphyrins were found, as seen in Figure S 4.

3.2.Preliminary characterization of petroporphyrins. The tetrapyrrole core of the petroporphyrins can chelate various metallic elements such as vanadium, nickel, iron, manganese, or copper[18]. Vanadyl and nickel porphyrins are the only types of petroporphyrins detected in the studied samples. However, even with the ET matrix, no petroporphyrins were detected in the PM A, which resulted from the combustion of the HFO A. Moreover, no nickel porphyrins were evidenced in the Particles B sample.

As a preliminary rough comparison between the fuels and their corresponding emissions, the intensity of vanadyl and nickel porphyrins signals relative to the summed intensity of attributed species in the mass spectra are displayed in Figure 2. We found the relative intensity of petroporphyrins signals were significantly higher in the HFO C than in the other fuels. A student's t test[44] at a confidence level of 95%, performed thanks to the triplicates measurements, confirmed this observation, as shown on Figure 2 and found in Table S 3. The same observations were made with PM C with respect to PM B and PM A. This was consistent with the high vanadium and nickel contents in the corresponding feed fuel C (Table 1). Elemental analysis was used to determine vanadium and nickel contents, but such analyses does not provide information on the chemical speciation of the elements, i.e., either mineral inorganic state or complexed metalorganic condition in porphyrin cores. In fact, in fossil material, a high fraction of vanadium and nickel content are included in porphyrin cores[17, 45]. Notably, we observed a decrease of the relative intensity of porphyrin signals between the HFO A and HFO B sample, which is not the tendency described by the above-mentioned elemental analysis (Table 1). In fact, the vanadium content in HFO B was twice that of HFO A, suggesting that an important fraction of vanadium and nickel in HFO B actually correspond to a mineral inorganic state, which couldn't be detected using ET-MALDI.

Finally, the relative intensity of porphyrins is lower in the particles samples than in the feed fuels, hinting at their consumption through combustion. Due to their higher intensity vanadyl porphyrins are more easily detectable and thus characterizable than nickel porphyrins in the studied samples.

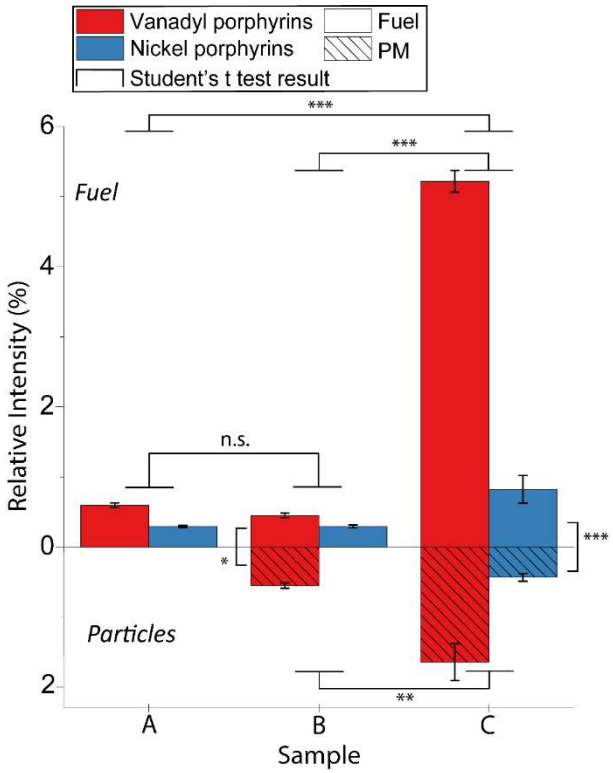


*Figure IV-2: Comparison between vanadyl and nickel petroporphyrins intensities. Results of Student's t test (CL = 95%) showing if a difference is statistically significant or not[44] : n.s.: non-significant with a confidence level of 95%, */**/***: significant with p-value respectively below 0.05, 0.01 and 0.001.*

Figure 3 presents the distribution of the relative intensity of petroporphyrins according to their double bond equivalent (DBE). Interestingly, all samples exhibited the same DBE pattern: etio petroporphyrins were the most intense, followed by rhodo and rhodo-etio porphyrins[19, 46] (Figure 4). On the one hand, this finding lets us hypothesize that these three specific petroporphyrins are the most abundant in the different heavy fuel oils, regardless of their sulfur, vanadium or nickel content. Similar DBE distributions were obtained before and after combustion. A Pearson correlation of the DBE distribution shape between fuel samples and their corresponding emissions confirmed this similarity. The correlation coefficient between HFO C/Particles C and HFO B/Particles B were above 0.95 (0.9771 and 0. 9568, respectively), leading to α-values under the maximum acceptable value (α=0.05) for a confidence level of 95%. The results of these correlations and details on the applied procedure can be found in supporting information (Table S 4). As the combustion process did not affect the DBE pattern, we inferred the combustion did not single out any specific structure. Indeed,

the DBE is a relevant indirect measure to the common petroporphyrin's structure is it allows to discriminate plausible core structures. According to the literature, the most plausible core structures for DBE 17, 18 and 20 petroporphyrins are etio-, DPEP- (deoxophylloerythroetio porphyrin) and rhodo-etio- cores, respectively[47, 48] (Figure 4).

The global maximum DBE observed was 27 in the HFO C; even though only an almost negligible number of signals corresponds to this DBE value, this limit was consistent with the maximum DBE observed by Ramirez-Pradilla et al.[35] in 2019 in crude oil samples using the same matrix. Furthermore, we did not observe any petroporphyrins under the DBE 17, neither in the fuel nor in the PM, which would have hinted at ring opening through the combustion process.

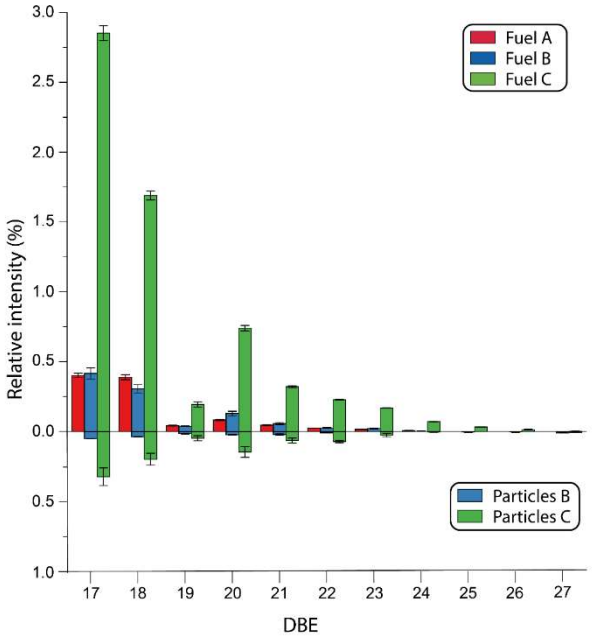


Figure IV-3: Relative intensity of the vanadyl petroporphyrin series depending on their double bond equivalent (DBE) in the studied feed fuel and PM samples (no petroporphyrins detected in Particles A).

3.3.Molecular maps as a tool for molecular characterization. Taking advantage of the molecular attributions, DBE versus carbon number diagrams (DBE vs #C) were plotted and used for fingerprint representation. In general, these molecular maps allow for the visualization of alkylated series (horizontal lines) and an easier in-depth comparison between samples. Figure 4 shows the DBE vs #C maps for vanadyl ($C_xH_yN_4VO$) and nickel ($C_xH_yN_4Ni$) porphyrins in both fuels and PM. In the most complex sample, i.e., HFO C, we were able to identify 290 different

vanadyl porphyrins and 168 nickel porphyrins. In the most complex PM, we attributed 182 unique vanadyl porphyrins and 97 nickel porphyrins. To the best of our knowledge, it is the first time that these many porphyrins were found in this kind of ultra-complex aerosol matrix.

Vanadyl porphyrins in the fuels sample involved abundant alkylated series at DBE 17, 18 and 20, following the DBE pattern found in Figure 3. For HFO A and B, the maximum value of DBE was lower than for HFO C (24 versus 27), which might be caused by the overall lower relative abundance. Even though nickel porphyrins were detected with significantly lower abundance, they followed the same DBE pattern as the vanadyl analogous. Also, the highest DBE value for Ni porphyrin (max DBE: 26) was just under the maximum value of VO porphyrins (max DBE: 27). The most abundant species of these series were found between carbon number 25 and 30, i.e., close to the PAH planar limit defined by Cho et al[49]. This localization in the DBE versus #C maps implies moderatly alkylated porphyrins.

Porphyrins being among the PAH family, one may expect they would be affected by the combustion by similar transformations. A possible chemical transformation of PAH. during combustion is the dealkylation[13, 28, 50, 51]. Indeed, in the combustion process, side chains can be cleaved off prior to further chemical transformation like, for example, ring opening or oxidation. Another chemical transformation involves formation of PAH from ethyne at high temperature, i.e. by pyrogenesis. Porphyrins, however cannot be generated from smaller building blocks.[28] The comparison between the DBE vs #C of the fuel and particles samples

aims to determine if dealkylation or other reaction pathways, such as ring-opening or partial oxidation, take place.

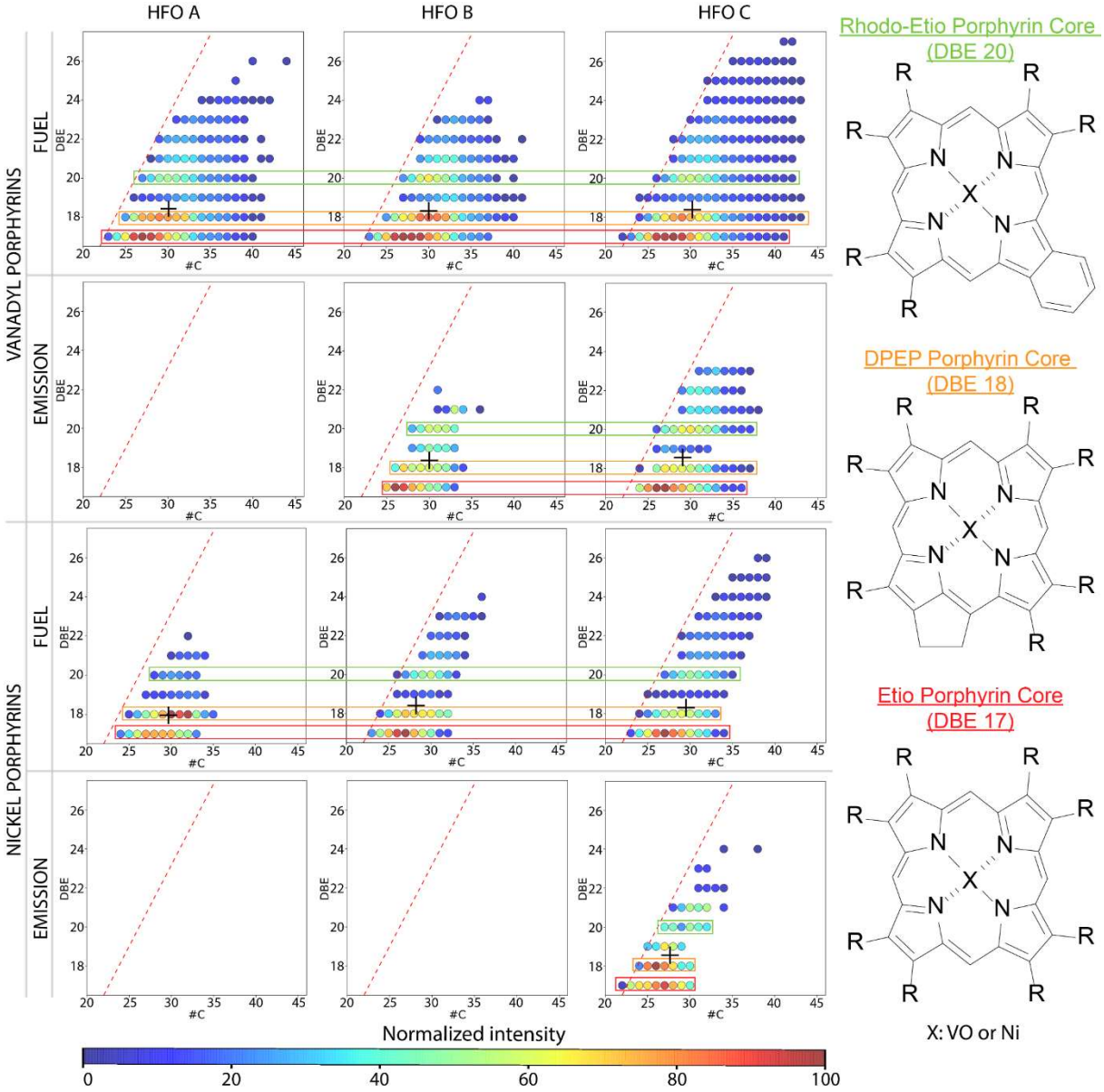


Figure IV-4: DBE vs carbon number diagrams for vanadyl and nickel porphyrins found in the feed fuels and corresponding primary ship diesel particulate matter samples. The proposed core structures for the main alkylated series are given on the right side and based on fossil petrochemical literature [47]. The planar aromatic limit [49]is given as dashed red line. No vanadyl porphyrins were found for the particulate matter of the HFO A feed and no nickel porphyrins for the particulate matter of the HFO A and B. The center of gravity of the DBE vs #C maps, represented by a black cross, corresponds to an intensity-weighted mean value of DBE and #C.

In the primary emission PM samples, a lower number of species could be detected because of the lower total intensity of the porphyrins fraction, *i.e.* the summed intensity of every porphyrin signals. Indeed, many alkylated species in the series were not detected in the PM, both for vanadyl and nickel porphyrins. We calculated a center of gravity for each DBE vs #C distribution using intensity-weighted values (represented by black crosses in Figure 4). We did not observe any shift of this center before and after combustion, which suggests that the petroporphyrins were not specifically de-alkylated during the combustion process but uniformly degraded. No DBE homologue rows and, thus, no structures appeared particurlaly targeted by the combustion process. In addition, the same DBE pattern in the feed fuel and PM was observed, supporting the fact that no specific consumption based on the core structure occured during the initial steps of the combustion process and that porphyrins in the particulate matter are essentially a fraction of unburnt fuel. No porphyrins were formed from higher DBE species (by, for example, the ring-opening of an attached benzene group). A putative structure for each most intense value of DBE are given in the Figure 4. These structures are based on previous work[47]. More hypothetical structures are represented in Figure S 5.

The absence of nickel porphyrins in PM B, was most likely due to the low S/N ratio that did not allow for the detection of species with a very low intensity. Indeed, in PM B, the vanadyl porphyrins fraction only accounted for a rather low fraction of 0.55% of the summed intensity of all attributed species, or total intensity. As vanadyl prorphyrins were usually more abundant than the nickel porphyrins, their fraction of the total intensity was expected to be lower, thereby underneath the detection limit. For PM A, the fact that no porphyrins were detected could also be caused by other combustion characteristics, e.g., more efficient combustion with lower emission of unburned species of HFO A , consequently emitting less vanadium and nickel on the particulate matter, as shown in Table 1. Assuming in a first approximation a uniform combustion behavior, this would lead to drastically lower petroporphyrin fractions in the particulate matter. Indeed, the PM emission of the ship diesel engine fed with HFO A is lower in terms of mass concentration as well as in number concentration of particles in the exhaust, independantly of the used engine load.

3.4.Porphyrins fate through the combustion process. Aiming to analyze the combustion pattern, we used the respective fold change normalized to the overall signal yield (summed intensity of all attributed species) to compare HFO C and PM C. For all porphyrins detected in both feed fuel and PM, the fold change value was calculated from peak intensities as follow:

$$FC = \frac{Peak\ intensity\ in\ PM}{Peak\ intensity\ in\ feed\ fuel}$$

In agreement with petroinformatics largely described by Hur et al.[52], the binary logarithm of the FC (log_2[FC]) was then used for color-coding the DBE vs #C maps (Figure 5). A diagram representing the data treatment strategy for Figure 5 is found in the supporting information (Figure S6). All detected porphyrins are represented: fuel-exclusive and particles-exclusive petroporphyrins are represented by fixed blue and pink colors, respectively. With this approach, the higher the modulus of a log_2[FC] value, the higher the difference between feed fuel and primary emission. Positive values are linked to an increased abundance in the particles, which can only be caused by pyrogenesis or transformation of larger species in smaller constituents, e.g., due to dealkylation or ring-opening reactions. Negative values are caused by a larger abundance in the feed fuel, i.e., consumption and decomposition of the respective petroporphyrin during the combustion process. Additionally, if a log_2[FC] value is close to zero, it means that the represented species is as intense in the fuel as in the particles and no considerable transformation can be assumed or consumption and transformation processes compensate for each other. Figure 5a and 5c present the results for the vanadyl porphyrins; numerous highly-alkylated and -aromatic vanadyl porphyrins were not detected in the particulate matter sample and, thus, are represented as blue dots. Nevertheless, the homogeneity of the FC values demonstrates that no specific combustion process was observed at this molecular level, in agreement with the more general observations for the grouped DBE classes and overall abundance of vanadyl petroporphyrins. Indeed, the fold change values being uniformly distributed under 0 without particles-exclusive porphyrins hints a general consumption with no apparent structural dependence. However, the Figure 5b and 5d representing the nickel porphyrins show that the fold change values vary from negative to positive at DBE 17 to 19 as the carbon number decreases. This observation could be correlated

to a dealkylation process but as the overall intensity of nickel porphyrins is low (under 1% of the total intensity of attributed species in the richest sample), we are unable to affirm this behavior yet. Figure 5a and 5b show the results with the intensities relative to the total intensity of identified species, as opposed to Figure 5c and 5d that use absolute intensities. Each method has its advantages and drawbacks: normalized intensities allow for a homogeneous order of magnitude for the compared values; but, if a species is heavily consumed through combustion, it will increase the relative abundance of all other species in the particles samples, possibly leading to "false" positive fold change values. On the other hand, using absolute intensities palliates this issue but is sensitive to the difference in order of magnitude between the intensities before and after combustion. Hence this second method is limited by the facts that the MALDI method used is here purely qualitative and that it is not possible to correct the difference using the particulate matter mass on the filter and the amount of fuel used in the extraction process.

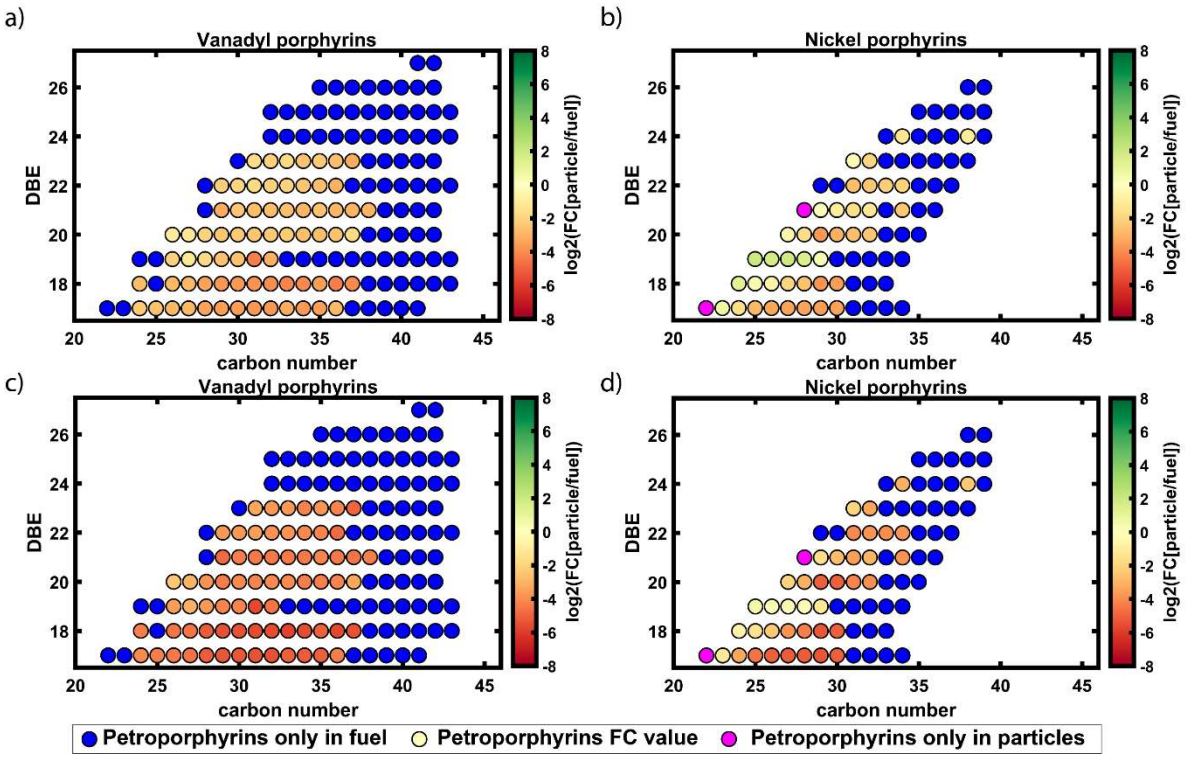


Figure IV-5: DBE vs #C plot using the log_2[FC] values as color code, highlighting the chemical fate of petroporphyrins through the combustion process in the ship diesel engine. a) Vanadyl porphyrins, intensity relative to the total intensity of identified species, b) Nickel porphyrins, intensity relative to the total intensity of identified species, c) Vanadyl porphyrins, absolute intensity, d) Nickell porphyrins, absolute intensity.

3.5.Effect of the engine's power load on porphyrins in PM. Particulate matter was obtained by burning HFO C with a motor load of 75 %, reflecting the more optimized condition used in open-sea long-range transport (i.e. cruise state): PM C'. No petroporphyrins were found in this sample (Figure 6). In previous data, where petroporphyrins could be detected well above the S/N-limit, PM samples were generated with 25% load, corresponding to the power load while maneuvering actions or idling in the harbor, The presence of abundant porphyrins can largely be related to the overall lower combustion efficiency with a 25% power load, as the engine is optimized for the 75% load conditions. Even though the PM emissions with 75% load condition account for the longest distances in open-sea transport, the PM that result from the 25% load is emitted closer to the coastal line or directly in the harbor area and are thus expected to have a stronger influence on local and semi-local environmental health. Figure 6 shows the mass spectra of PM C and PM C' at *m/z* 487 where the most intense vanadyl porphyrin ($C_{28}H_{28}N_4OV$, *m/z* 487.16969) is expected. Interestingly, even though this usually abundant petroporphyrin appears on the PM C mass spectra, it did not appear on PM C' spectra. The PM C spectra exhibits an important variation in the absolute intensity of the peak of interest (*m/z* 487.16969); nevertheless, the use of the of relative intensities, as partly described in Figure S6, we were able to reduce this variation (See Table S5). The absence of peak above the S/N signal in this region for the mass spectra of PM C' sample suggests that the petroporphyrins present in the fuel are consumed at this load to a point not detectable with the presented ET-MALDI approach; nevertheless, the limited size of the sample set

doesn't allow us to confirm it. The main ion (*m/z* 487.23082) shown in PM C' mass spectra are also found in blank spectrum, so they most likely come from the filter and the matrix used.

4.Conclusion

For the first time, this study allowed to evidence petroporphyrins in primary ship diesel emission at the molecular level by the combination of ultra-high resolution mass spectrometry

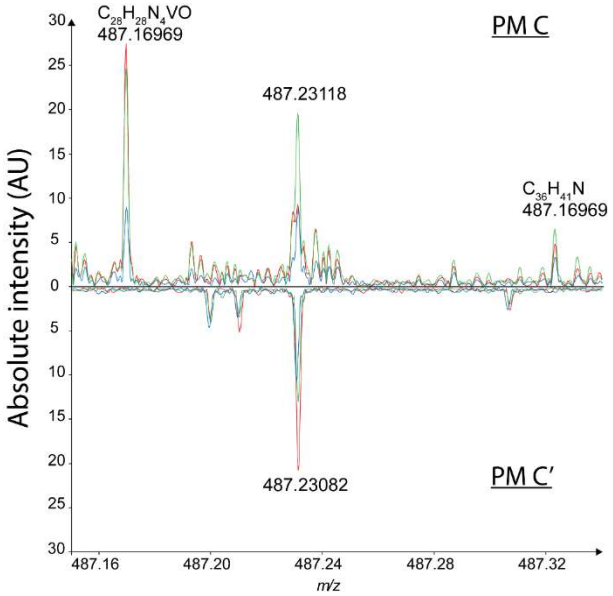


Figure IV-6: ET-MALDI FT-ICR mass spectra at m/z 487 (350 mDa excerpt) of PM samples obtained from the HFO C feed fuel at different power loads: Up) 25% load (20 kW) harbor maneuvering, , and Down) 75% load (60 kW) cruise long-range. Each color corresponds to an analytical replicate.

(FTICR MS) and a selective electron-transfer matrix in ET-MALDI. The α-CNPV-CH3 matrix allowed the characterization of the targeted petroporphyrins because of their highly aromatic character. The evidenced porphyrins consisted of vanadium oxide or nickel complexed inside a tetrapyrrole core, well-known in the fossil feed fuels. These porphyrins were detected, not only in the feed heavy fuel oils, but also in the particles emitted by their combustion under realistic conditions in a one-stroke ship diesel engine.

A general uniform consumption and degradation of the petroporphyrins through the engine combustion was observed with no particular preference for any structural motif, as differentiated by the DBE. Despite the reported effects of a dealkylation on pure PAHs and N/S-PAHs[53], no significant dealkylation effect could be found. Visualized by DBE versus #C maps,

an overall reduction of the chemical space after combustion, most likely due to the lower overall abundance, was found. Moreover, the porphyrins in the particulate matter were chemically very close to those of the feed fuel but found in significantly lower abundance. No transformed classes (oxidized, ring-opening) indicated any further combustion fate of the petroporphyrins, known to be very stable. Nonetheless, to confirm this finding, the occurrence of these products will be further studied in future work utilizing novel sampled particulate matter with higher concentration, other combustion conditions, and a bigger sample set. Here also complementary analytical tools, such as APPI and/or positive-ion electrospray ionization-(ESI+) using sodium attachment, can be deployed.

Concerning environmental health, it has been concluded that higher amounts of petroporphyrins might be released by the combustion of heavy fuel oils in harbor areas or close to the coastal line rather than in open sea, based on different engine loading conditions and combustion efficiency. In this respect, the here molecularly proven emission of metal organics is an exciting factor and motivates related health studies. The molecularly resolved and detected emitted petroporphyrin pattern thus add another group of compounds largely unknown to respiratory toxicity. Furthermore, the presence and fate of petroporphyrins in the primary exhaust particulate matter of a ship engine equipped with exhaust gas treatment technology, i.e., scrubber technology and particle filter, motivates future studies. The fate of the highly aromatic and polar petroporphyrins during these washing and removal steps will be of high interest, as trace metals may potentially be introduced directly into the aqueous phase.

Acknowledgments

This work has been partially supported by University of Rouen Normandy, INSA Rouen Normandy, the Centre National de la Recherche Scientifique (CNRS), European Regional Development Fund (ERDF), Labex SynOrg (ANR-11-LABX-0029), Carnot Institut I2C, the graduate school for research Xl-Chem (ANR-18-EURE-0020 XL CHEM), the European Union's Horizon 2020 Research Infrastructures program (Grant Agreement 731077) and by Region

Normandie. We thank the DFG (ZI 764/28-1) and ANR (ANR-20-CE92-0036) for funding of the research project TIMSAC.

Supporting information

Supporting information contains laser power optimization curves, Kendrick mass defect based extraction, spectrum highlighting the complexity of studied sample, Student's t test and Pearson' correlation results, hypothetical porphyrins core structures and mass spectra suggesting the absence of porphyrins in MGO.

Laser power optimization

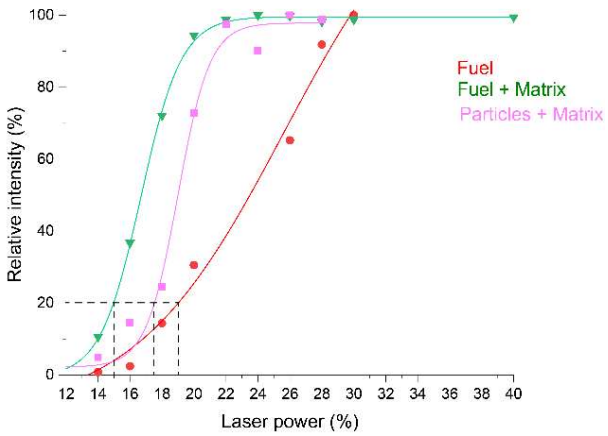


Figure S 7: Laser power optimization curves.

For the three samples displayed in Figure S 1, spectra were recorded at different laser powers (LP) and the total intensity observed was determined, i.e. peak picking was performed with a signal over noise ratio of 4 on the whole spectra and the intensity of every peak was summed. The total intensities for a same sample type were then normalized and plotted against laser power. We chose the laser power corresponding to 20% of the maximum total intensity obtained for each sample type. Laser power had to be optimized in order to obtain a qualitative signal while avoiding the formation of undesired species such as fullerenes.

Engine detail

Engine model	1 VDS 18/15
Method of operation	Four stroke diesel, direct injected, compressor charged
Amount of cylinders	1
Valves	4
Stroke	180 mm
Bore	150 mm
Length of connecting rod	332 mm
Nominal speed	1,500 min^{-1}
Compression ratio	13
Maximum power	80 kW
Nominal power	60 kW

Table S 3: Engine's specifications

This table lists the specifications of the engine used during the measurement campaign which produced our samples. This engine and the running conditions are furthermore detailed by Streiblel et al.[13]

Calibration list

Molecular formula	*m/z* ratio	Molecular formula	*m/z* ratio	Molecular formula	*m/z* ratio
$C_{26}H_{20}$	332.155952	$C_{47}H_{67}N$	645.526803	$C_{26}H_{21}N_2$	361.169925
$C_{27}H_{22}$	346.171602	$C_{48}H_{63}N$	653.495502	$C_{25}H_{20}N_4O_1V_1$	443.107120
$C_{29}H_{24}$	372.187252	$C_{49}H_{65}N$	667.511153	$C_{25}H_{22}N_4O_1V_1$	445.122770
$C_{31}H_{24}$	396.187252	$C_{50}H_{65}N$	679.511153	$C_{26}H_{22}N_4O_1V_1$	457.122770
$C_{33}H_{26}$	422.202902	$C_{50}H_{69}N$	683.542453	$C_{26}H_{24}N_4O_1V_1$	459.138420
$C_{36}H_{30}$	462.234202	$C_{51}H_{63}N$	689.495502	$C_{27}H_{24}N_4O_1V_1$	471.138420
$C_{21}H_{19}N$	285.151201	$C_{50}H_{75}N$	689.589403	$C_{27}H_{26}N_4O_1V_1$	473.154070
$C_{22}H_{17}N$	295.135551	$C_{52}H_{67}N$	705.526803	$C_{28}H_{22}N_4O_1V_1$	481.122770
$C_{23}H_{19}N$	309.151201	$C_{52}H_{73}N$	711.573753	$C_{28}H_{26}N_4O_1V_1$	485.154070
$C_{24}H_{16}N$	318.127726	$C_{52}H_{77}N$	715.605053	$C_{28}H_{28}N_4O_1V_1$	487.169720
$C_{24}H_{25}N$	327.198151	$C_{54}H_{67}N$	729.526803	$C_{29}H_{24}N_4O_1V_1$	495.138420
$C_{26}H_{21}N$	347.166851	$C_{54}H_{77}N$	739.605053	$C_{29}H_{28}N_4O_1V_1$	499.169720
$C_{27}H_{23}N$	361.182501	$C_{55}H_{73}N$	747.573753	$C_{29}H_{30}N_4O_1V_1$	501.185370
$C_{28}H_{25}N$	375.198151	$C_{56}H_{65}N$	751.511153	$C_{30}H_{26}N_4O_1V_1$	509.154070
$C_{29}H_{25}N$	387.198151	$C_{55}H_{79}N$	753.620703	$C_{30}H_{30}N_4O_1V_1$	513.185370
$C_{31}H_{25}N$	411.198151	$C_{57}H_{67}N$	765.526803	$C_{30}H_{32}N_4O_1V_1$	515.201020
$C_{32}H_{27}N$	425.213801	$C_{57}H_{71}N$	769.558103	$C_{31}H_{32}N_4O_1V_1$	527.201020
$C_{40}H_{49}N$	543.385952	$C_{57}H_{79}N$	777.620703	$C_{31}H_{34}N_4O_1V_1$	529.216670
$C_{41}H_{55}N$	561.432902	$C_{57}H_{81}N$	779.636353	$C_{32}H_{34}N_4O_1V_1$	541.216670
$C_{43}H_{51}N$	581.401602	$C_{57}H_{83}N$	781.652003	$C_{33}H_{36}N_4O_1V_1$	555.232320
$C_{44}H_{57}N$	599.448552	$C_{58}H_{79}N$	789.620703	$C_{35}H_{34}N_4O_1V_1$	577.216670
$C_{45}H_{57}N$	611.448552	$C_{59}H_{73}N$	795.573753	$C_{36}H_{32}N_4O_1V_1$	587.201020
$C_{46}H_{57}N$	623.448552	$C_{26}H_{20}N_2$	360.162100	$C_{37}H_{38}N_4O_1V_1$	605.247970
				$C_{39}H_{38}N_4O_1V_1$	629.247970

Table S 4: internal calibration list

This table contains all the ions used to perform the internal calibrations during this study. It is composed of species detected in HFO C such as already known $C_xH_yN_1$ and C_xH_y molecular series and $C_xH_yN_4VO$ species newly identified . Furthermore, two characteristic ions of the matrix were also used: $C_{26}H_{20}N_2$ and $C_{26}H_{21}N_2$.

Kendrick mass defect based extraction

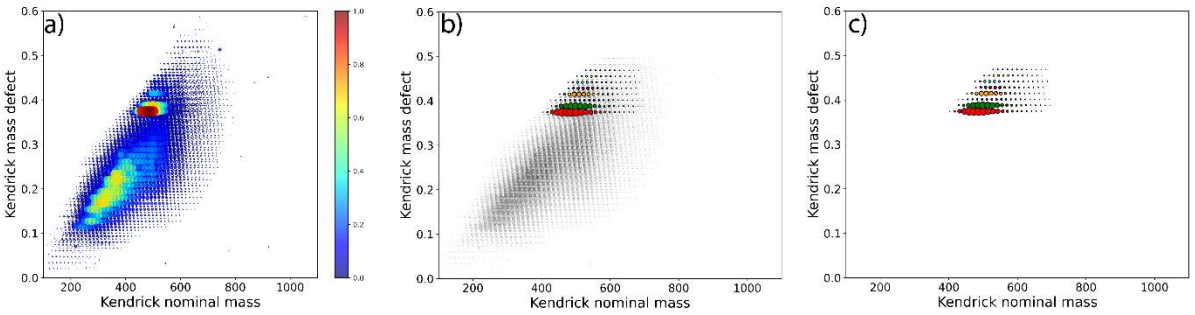


Figure S 8: a) Kendrick's diagram with all species, b) Kendrick's diagram with extracted series stacked on the previous plot, c) Kendrick's diagram with only the extracted series.

The peak picking feature available in Data Analysis 5.0 allowed the Kendrick Mass Defect (KMD) diagram to be plotted (a). Our self-written python routine enables the extraction, both visual and numerical, of selected homologue series. For example, if we enter the *m/z* ratio of $C_{28}H_{28}N_4VO^+$, i.e. *m/z* 487.16972, the program will display a series containing signals with the same KMD, either over the faded global KMD diagram (b) or on a blank diagram (c). The data concerning this specific series, i.e. m/z, intensity, error, etc., can then be extracted to a .csv (comma separated values) file. As seen in Figure S 2) b and c, it is possible to select several series to extract. Then, using this information, we attributed the vanadyl and nickel porphyrins using the following molecular boundaries: $C_xH_yN_4VO_{1\text{-}3}S_{0\text{-}1}$ and $C_xH_yN_4Ni$.

Sample's molecular complexity

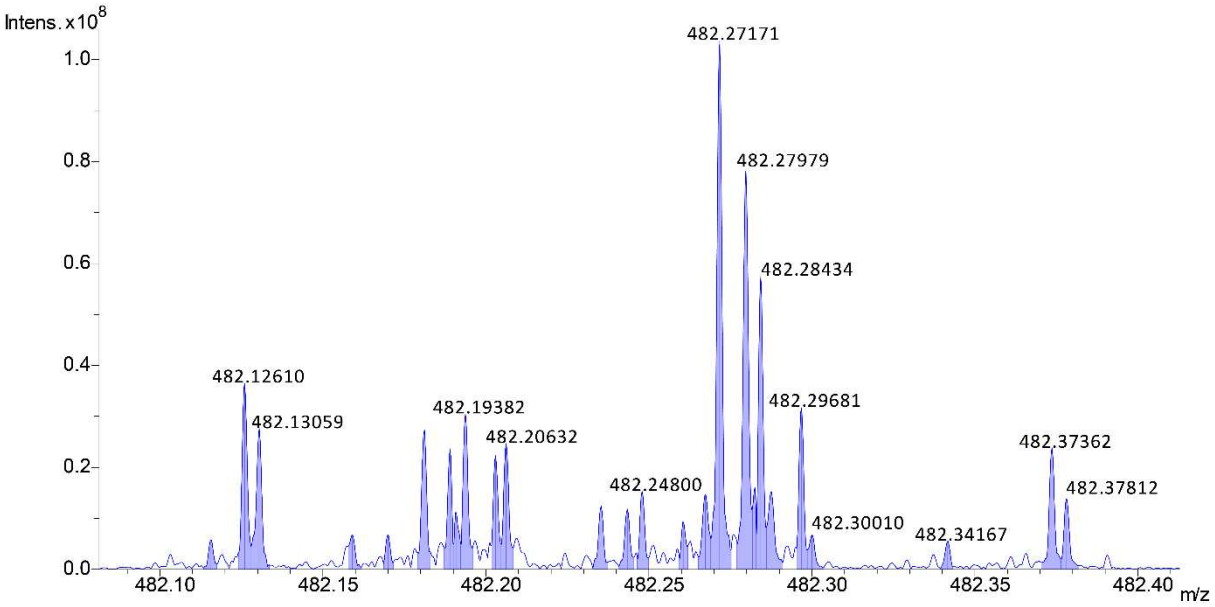


Figure S 9: Sample's molecular complexity

Figure S 3 highlights the complexity of the studied samples. This mass spectrum corresponds to the HFO C sample with the α-CNPV-CH3 matrix and exhibits a total of 26 peaks above the selected S/N ratio (S/N = 4) within the range of 300mDa at *m/z* 482.

Searching for porphyrins in other samples

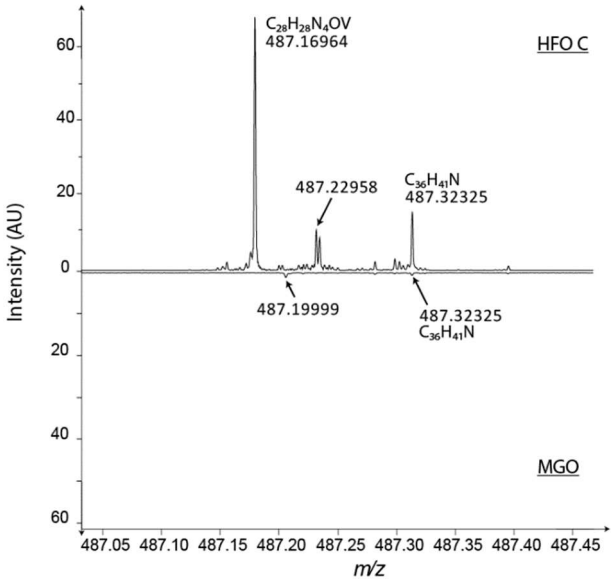


Figure S 10: Two mass spectra zoomed at m/z 487. The upper spectrum corresponds to the HFO C sample and the lower spectrum concerns the MGO sample.

In the mass range displayed in Figure S4, the most intense petroporphyrin found in our main samples, i.e. HFO A, B, C, and Particles A, B, and C, is supposed to be found. However, if we can clearly see it on the HFO C spectrum, this figure demonstrates that this porphyrin isn't found in the MGO sample.

Student's t-test

With the aim of determining if the differences in porphyrins' relative intensity between the samples were significant or not, we realized a Student's t test. We summed the relative intensity of nickel and vanadyl porphyrins for each replicate (3 per sample) and performed the tests on these values. The null hypothesis for these tests is: '*The mean relative intensities of petroporphyrins in the compared samples are equal*' and the chosen alternative hypothesis is '*The mean relative intensities of petroporphyrins in the compared samples are not equal*'. Hence, we realized a two-tailed statistical test as opposed to a one-tailed test whose alternative hypothesis would be: '*The mean relative intensity of petroporphyrins in the studied sample is greater than in the reference sample*'. The t test returns a p value representing the probability that the null hypothesis is true, thus, to be able to affirm that we observe a significant difference between the compared samples with a confidence level of 95%, p must be under 0.05. The p values obtained are shown in table S 1.

Samples compared	p-value
HFO A / HFO B	0.073143
HFO B / HFO C	0.000018
HFO A / HFO C	0.000004
Part. B / Part. C	0.004745
HFO B / Part. B	0.022522
HFO C / Part. C	0.000231

Table S 5: Results of Student's t test

Except between HFO A and B, the observed differences are all significant with a confidence level of 95%.

Pearson's correlation coefficients

The following table contains the Pearson's correlation coefficient obtained by comparing the DBE distribution pattern of fuel samples with the pattern of their corresponding emissions. This comparison was realized under Python using the *scipy[54]* package's "*pearsonr"* function on the mean intensity of each DBE value of the compared samples.

Samples	HFO B/Particles B	HFO C/Particles C
Correlation coefficient	0.9568	0.9771
Significance level α (%)	3.99e-4	2.37e-5

Table S 6: Comparison of DBE distribution patterns using the Pearson correlation coefficient

A correlation coefficient close to 1 exhibits a perfect correlation between the compared datasets, in our case, it would mean that the compared samples have the exact same DBE distribution pattern. However, this only coefficient is not enough to prove similarity. Indeed, we need to be cautious of the significance level (α) which is the probability that the correlation is in fact false. Hence, an α value close to zero is what you look for, even if the commonly accepted maximum α value is 5%.

Porphyrins core

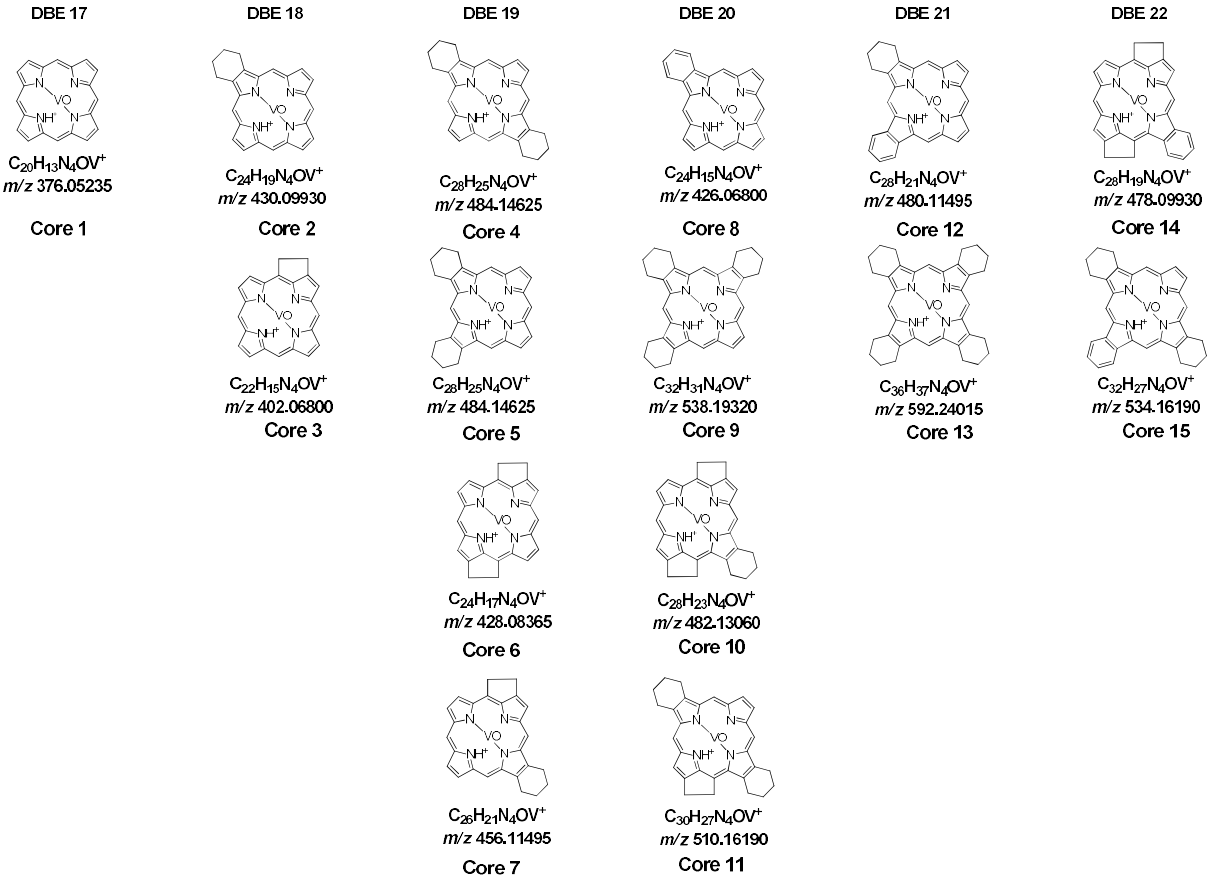


Figure S 11: Possible porphyrin core structures depending on their DBE (Reproduced with the author's agreement[47])

These possible core structures for petroporphyrins were found in previous work by Maillard et al.[47] in which they used trapped ion mobility coupled with Fourier transform ion cyclotron resonance mass spectrometry (TIMS-FTICR MS)

Data treatment workflow

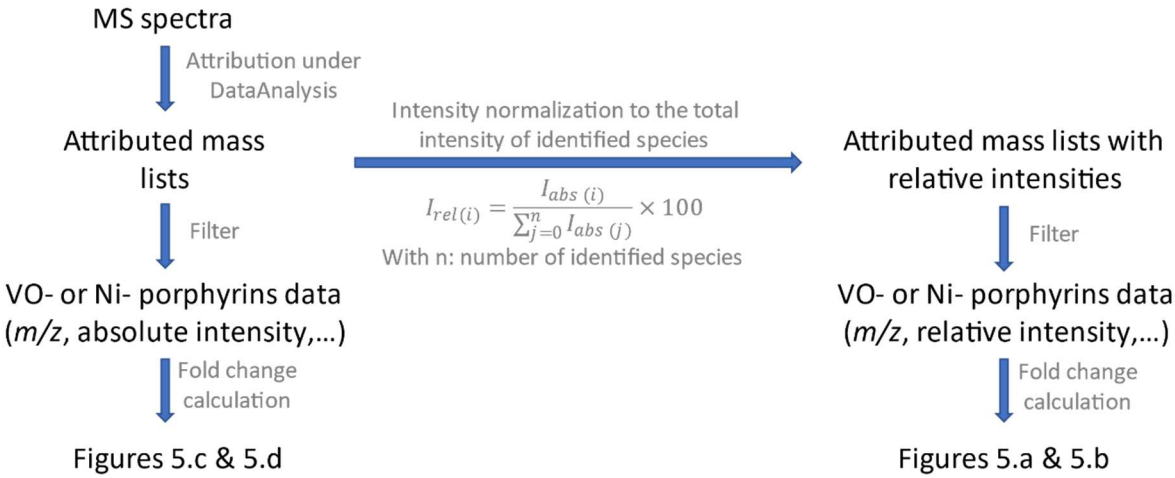


Figure S 12: Data treatment strategy for the establishment of Figure 5

This Figure shows the data treatment strategy used to elaborate Figure 5 and exhibits the difference between Figure 5.a & 5.c and Figure 5.b & 5.d respectively. This difference is the type of intensity chosen to calculate the fold change values. Indeed, the two options have their own advantages and drawbacks, so they both needed to be shown.

Calculation of coefficient of variation

We calculated the coefficient of variation (CV) for both absolute and relative intensities of the peak corresponding to $C_{28}H_{28}N_4OV$ **(***m/z* 487.16969) across the replicates as follows:

$$CV_{abs} = \frac{Standard\ deviation\ of\ abs.intensity}{Mean\ abs.\ intensity} \times 100$$

$$CV_{rel} = \frac{Standard\ deviation\ of\ rel.intensity}{Mean\ rel.\ intensity} \times 100$$

	Absolute intensity	Relative intensity
Coefficient of variation (%)	37.68	20.11

Table S 7: Coefficient of variation values demonstrating the advantage to use relative intensities

3. Conclusion

In this study, we observed that vanadium and nickel petroporphyrins are more likely to be released by a ship when its engine load is low. However, a low engine load is used when the ship is maneuvering in the harbor or near the coastline, *i.e.* near inhabited area. This last observation is of interest from the point of view of environmental health as the toxicity of petroporphyrins is largely unknown. This study was made possible because of the ultrahigh resolving power and excellent mass accuracy of the FTICR MS, allowing to differentiate the petroporphyrins from other compound classes such as the $C_xH_yO_2$ that only presented an *m/z* difference of around 0.5 mDa. However, FTICR MS doesn't provide any information on the isomeric complexity of a sample, contrary to IMS-MS. The last chapter of this thesis focuses on our work on the structural analysis on primary and secondary ship emissions by TWIMS-TOF-MS.

4. References

1. Corbett, J.J., et al., *Mortality from Ship Emissions: A Global Assessment.* Environmental Science & Technology, 2007. **41**(24): p. 8512-8518.

2. Kanashova, T., et al., *Differential proteomic analysis of mouse macrophages exposed to adsorbate-loaded heavy fuel oil derived combustion particles using an automated sample-preparation workflow.* Analytical and Bioanalytical Chemistry, 2015. **407**(20): p. 5965-5976.

3. Corbin, J.C., et al., *Brown and Black Carbon Emitted by a Marine Engine Operated on Heavy Fuel Oil and Distillate Fuels: Optical Properties, Size Distributions, and Emission Factors.* Journal of Geophysical Research: Atmospheres, 2018. **123**(11): p. 6175-6195.

4. Eyring, V., et al., *Transport impacts on atmosphere and climate: Shipping.* Atmospheric Environment, 2010. **44**(37): p. 4735-4771.

5. Czech, H., et al., *A chemometric investigation of aromatic emission profiles from a marine engine in comparison with residential wood combustion and road traffic: Implications for source apportionment inside and outside sulphur emission control areas.* Atmospheric Environment, 2017. **167**: p. 212-222.

6. Agrawal, H., et al., *In-use gaseous and particulate matter emissions from a modern ocean going container vessel.* Atmospheric Environment, 2008. **42**(21): p. 5504-5510.

7. Abadie, L.M., N. Goicoechea, and I. Galarraga, *Adapting the shipping sector to stricter emissions regulations: Fuel switching or installing a scrubber?* Transportation Research Part D: Transport and Environment, 2017. **57**: p. 237-250.

8. Streibel, T., et al., *Aerosol emissions of a ship diesel engine operated with diesel fuel or heavy fuel oil.* Environmental Science and Pollution Research, 2017. **24**(12): p. 10976-10991.

9. Winnes, H. and E. Fridell, *Particle Emissions from Ships: Dependence on Fuel Type.* Journal of the Air & Waste Management Association, 2009. **59**(12): p. 1391-1398.

10. Agrawal, H., et al., *Emission measurements from a crude oil tanker at sea.* Environ Sci Technol, 2008. **42**(19): p. 7098-103.

11. Moldanová, J., et al., *Physical and chemical characterisation of PM emissions from two ships operating in European Emission Control Areas.* Atmospheric Measurement Techniques, 2013. **6**(12): p. 3577-3596.

12. Czech, H., et al., *New directions: Beyond sulphur, vanadium and nickel – About source apportionment of ship emissions in emission control areas.* Atmospheric Environment, 2017. **163**: p. 190-191.

13. Streibel, T., et al., *Aerosol emissions of a ship diesel engine operated with diesel fuel or heavy fuel oil.* Environ Sci Pollut Res Int, 2017. **24**(12): p. 10976-10991.

14. Oeder, S., et al., *Particulate Matter from Both Heavy Fuel Oil and Diesel Fuel Shipping Emissions Show Strong Biological Effects on Human Lung Cells at Realistic and Comparable In Vitro Exposure Conditions.* PLOS ONE, 2015. **10**(6): p. e0126536.

15. Sippula, O., et al., *Particle emissions from a marine engine: chemical composition and aromatic emission profiles under various operating conditions.* Environ Sci Technol, 2014. **48**(19): p. 11721-9.

16. Corbin, J.C., et al., *Trace Metals in Soot and PM2.5 from Heavy-Fuel-Oil Combustion in a Marine Engine.* Environmental Science & Technology, 2018. **52**(11): p. 6714-6722.

17. Caumette, G., et al., *Element speciation analysis of petroleum and related materials.* Journal of Analytical Atomic Spectrometry, 2009. **24**(3): p. 263.

18. Ali, M.F. and S. Abbas, *A review of methods for the demetallization of residual fuel oils.* Fuel Processing Technology, 2006. **87**(7): p. 573-584.

19. Filby Royston H. and V.B.G. J., *Geochemistry of Metal Complexes in Petroleum, Source Rocks, and Coals: An Overview*, in *ACS Symposium Series*. 1987. p. 2-39.

20. Xu, C., et al., *Overlooked Significant Impact of Trace Metals on the Bacterial Community of PM 2.5 in High-Time Resolution.* Journal of Geophysical Research: Atmospheres, 2021. **126**(23).

21. Wang, Y.N., et al., *Temporal dynamics of heavy metal distribution and associated microbial community in ambient aerosols from vanadium smelter.* Science of The Total Environment, 2020. **735**: p. 139360.

22. Reda, A.A., et al., *Gas phase carbonyl compounds in ship emissions: Differences between diesel fuel and heavy fuel oil operation.* Atmospheric Environment, 2015. **112**: p. 370-380.

23. Kleeblatt, J., et al., *Needle trap sampling thermal-desorption resonance enhanced multiphoton ionization time-of-flight mass spectrometry for analysis of marine diesel engine exhaust.* Analytical Methods, 2015. **7**(8): p. 3608-3617.

24. Käfer, U., et al., *Detailed Chemical Characterization of Bunker Fuels by High-Resolution Time-of-Flight Mass Spectrometry Hyphenated to GC × GC and Thermal Analysis.* Energy & Fuels, 2019. **33**(11): p. 10745-10755.

25. Radischat, C., et al., *Real-time analysis of organic compounds in ship engine aerosol emissions using resonance-enhanced multiphoton ionisation and proton transfer mass spectrometry.* Analytical and Bioanalytical Chemistry, 2015. **407**(20): p. 5939-5951.

26. Mueller, L., et al., *Characteristics and temporal evolution of particulate emissions from a ship diesel engine.* Applied Energy, 2015. **155**: p. 204-217.

27. Drewnick, F., et al., *A New Time-of-Flight Aerosol Mass Spectrometer (TOF-AMS)—Instrument Description and First Field Deployment.* Aerosol Science and Technology, 2005. **39**(7): p. 637-658.

28. Rüger, C.P., et al., *Characterisation of ship diesel primary particulate matter at the molecular level by means of ultra-high-resolution mass spectrometry coupled to laser desorption ionisation--comparison of feed fuel, filter extracts and direct particle measurements.* Anal Bioanal Chem, 2015. **407**(20): p. 5923-37.

29. Xu, H., et al., *Characterization of Petroporphyrins Using Ultraviolet−Visible Spectroscopy and Laser Desorption Ionization Time-of-Flight Mass Spectrometry.* Energy & Fuels, 2005. **19**(2): p. 517-524.

30. Rytting, B.M., et al., *Ultrahigh-Purity Vanadyl Petroporphyrins.* Energy & Fuels, 2018. **32**(5): p. 5711-5724.

31. Ryan, D.J. and K. Qian, *Laser-Based Ionization: A Review on the Use of Matrix-Assisted Laser Desorption/Ionization and Laser Desorption/Ionization Mass Spectrometry in Petroleum Research.* Energy & Fuels, 2020. **34**(10): p. 11887-11896.

32. Zhang, Y., et al., *Elucidating the Geometric Substitution of Petroporphyrins by Spectroscopic Analysis and Atomic Force Microscopy Molecular Imaging.* Energy & Fuels, 2019. **33**(7): p. 6088-6097.

33. Giraldo-Dávila, D., et al., *Selective ionization by electron-transfer MALDI-MS of vanadyl porphyrins from crude oils.* Fuel, 2018. **226**: p. 103-111.

34. Ramírez-Pradilla, J.S., C. Blanco-Tirado, and M.Y. Combariza, *Electron-Transfer Ionization of Nanoparticles, Polymers, Porphyrins, and Fullerenes Using Synthetically Tunable α-Cyanophenylenevinylenes as UV MALDI-MS Matrices.* ACS Applied Materials & Interfaces, 2019. **11**(11): p. 10975-10987.

35. Ramírez-Pradilla, J.S., et al., *Comprehensive Petroporphyrin Identification in Crude Oils Using Highly Selective Electron Transfer Reactions in MALDI-FTICR-MS.* Energy & Fuels, 2019. **33**(5): p. 3899-3907.

36. Ramírez-Pradilla, J.S., et al., *Effect of the Ionization Source on the Targeted Analysis of Nickel and Vanadyl Porphyrins in Crude Oil.* Energy & Fuels, 2021. **35**(18): p. 14542-14552.

37. Cho, Y., et al., *Evaluation of Laser Desorption Ionization Coupled to Fourier Transform Ion Cyclotron Resonance Mass Spectrometry To Study Metalloporphyrin Complexes.* Energy & Fuels, 2014. **28**(11): p. 6699-6706.

38. Nyadong, L. and J. Mendez-Arroyo, *Sodium Cationization Electrospray Ionization Orbitrap Mass Spectrometry for Selective Determination of Crude Oil Porphyrins.* Energy & Fuels, 2021.

39. Maillard, J., et al., *Comparison of soluble and insoluble organic matter in analogues of Titan's aerosols.* Earth and Planetary Science Letters, 2018. **495**: p. 185-191.

40. Vetter, W., F.W. McLafferty, and F. Turecek, *Interpretation of mass spectra. Fourth edition (1993). University Science Books, Mill Valley, California.* Biological Mass Spectrometry, 1994. **23**(6): p. 379-379.

41. Müller, W.H., et al., *Dual-polarity SALDI FT-ICR MS imaging and Kendrick mass defect data filtering for lipid analysis.* Analytical and Bioanalytical Chemistry, 2021. **413**(10): p. 2821-2830.

42. McCann, A., et al., *Mass shift in mass spectrometry imaging: comprehensive analysis and practical corrective workflow.* Analytical and Bioanalytical Chemistry, 2021. **413**(10): p. 2831-2844.

43. Knochenmuss, R., et al., *Secondary ion-molecule reactions in matrix-assisted laser desorption/ionization.* Journal of Mass Spectrometry, 2000. **35**(11): p. 1237-1245.

44. Student, *The Probable Error of a Mean.* Biometrika, 1908. **6**(1): p. 1.

45. Desprez, A., et al., *Study of the Size Distribution of Sulfur, Vanadium, and Nickel Compounds in Four Crude Oils and Their Distillation Cuts by Gel Permeation Chromatography Inductively Coupled Plasma High-Resolution Mass Spectrometry.* Energy & Fuels, 2014. **28**(6): p. 3730-3737.

46. McKenna, A.M., et al., *Advances and Challenges in the Molecular Characterization of Petroporphyrins.* Energy & Fuels, 2021.

47. Maillard, J.F., et al., *Structural analysis of petroporphyrins from asphaltene by trapped ion mobility coupled with Fourier transform ion cyclotron resonance mass spectrometry.* Analyst, 2021. **146**(13): p. 4161-4171.

48. Zhao, X., et al., *New Vanadium Compounds in Venezuela Heavy Crude Oil Detected by Positive-ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Scientific Reports, 2015. **4**(1).

49. Cho, Y., Y.H. Kim, and S. Kim, *Planar Limit-Assisted Structural Interpretation of Saturates/Aromatics/Resins/Asphaltenes Fractionated Crude Oil Compounds Observed by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Analytical Chemistry, 2011. **83**(15): p. 6068-6073.

50. Saha, M., et al., *Sources of sedimentary PAHs in tropical Asian waters: Differentiation between pyrogenic and petrogenic sources by alkyl homolog abundance.* Marine Pollution Bulletin, 2009. **58**(2): p. 189-200.

51. Yunker, M.B., et al., *PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH source and composition.* Organic Geochemistry, 2002. **33**(4): p. 489-515.

52. Hur, M., S. Kim, and C.S. Hsu, *Petroinformatics*, in *Springer Handbook of Petroleum Technology*, C.S. Hsu and P.R. Robinson, Editors. 2017, Springer International Publishing: Cham. p. 173-198.

53. Rüger, C.P., et al., *Comprehensive chemical comparison of fuel composition and aerosol particles emitted from a ship diesel engine by gas chromatography atmospheric pressure chemical ionisation ultra-high resolution mass spectrometry with improved data processing routines.* Eur J Mass Spectrom (Chichester), 2017. **23**(1): p. 28-39.

54. Virtanen, P., et al., *SciPy 1.0: fundamental algorithms for scientific computing in Python.* Nat Methods, 2020. **17**(3): p. 261-272.

Chapter V. Structural Study of Aged Aerosols Using Ion Mobility Spectrometry

1. Introduction

This final chapter focuses on our last, but still ongoing, study on the structural elucidation of aged aerosols using ion mobility mass spectrometry.

The study of the aging process of organic aerosols (OA) presents a major interest as the physical and chemical properties of the aerosol change over time and space, transforming primary organic aerosols (POA) into secondary organic aerosols (SOA). As mentioned in I.1.4, secondary organic aerosols tend to be more functionalized and polar than primary organic aerosols, thus increasing their efficiency as cloud condensation nuclei [1]. Hence, the aging of biogenic or anthropogenic OA can eventually lead to climate perturbations as clouds directly interact with both solar and terrestrial radiations [2]. Even if OA aging can lead to macroscopic phenomena, changes with aging are expected to occur at the molecular level, resulting in the obtention of new molecules and the consumption of others. These changes in OA composition have been studied by mass spectrometry [3, 4] and aerosol mass spectrometry [5-7] (AMS). AMS is a technique that allows to obtain the particle size distribution of a PM sample. Although, the use of high-resolution mass spectrometry allows tracking changes in the molecular composition with the determination of molecular formulas, mass spectrometry (MS) alone does not provide information on the isomeric diversity of the sample.

For this purpose, ion mobility spectrometry (IMS) can be coupled to MS for the structural study of OA as shown by Krechmer *et al.* [8] which carried out the first application of IMS-MS for the characterization of both organic aerosol's PM and gas phase species. Indeed, they performed the analysis of filter samples using ESI-IMS-MS and the analysis of gas-phase species with a custom-built nitrate chemical ionization source. For the characterization of the particulate matter, pre-baked quartz fiber filters were used to collect $PM_{2.5}$ samples during the 2013 Southern Oxidant and Aerosol Study (SOAS) [9]. Thanks to ion mobility spectrometry, they were able to separate structural isomers that couldn't be separated using other traditional techniques such as liquid chromatography (LC) or gas chromatography (GC). The

community has shown interest in the use of IMS-MS for the structural characterization of SOA produced from biogenic precursors such as α-pinene or limonene [10, 11]. However, IMS-MS is not limited to biogenic OA and can be used for the analysis of anthropogenic samples such as chemical warfare agents [12] or engine exhaust. Indeed, in their study, C. Rüger *et al.* [13] performed the analysis of both biogenic and anthropogenic PM. In particular, they mimicked atmospheric aging conditions using a high-volume photochemical emission aging flow tube reactor (PEAR) [14] and compared fresh and aged emissions, thus showing the capability of IMS-MS to differentiate between fresh and aged aerosol samples.

The goal of the study presented in this chapter is to use the ability of IMS-MS to provide structural information to investigate the effects of the aging process on ship emissions. The results presented hereafter are the initial results of an ongoing study.

2. Samples and Experimental Setup

For this study, the samples were collected at the University of Rostock during a measurement campaign entitled “Ultrafine particles from transportation – health assessment of sources” (ULTRHAS) . During this campaign, we used a ship diesel engine and a sampling setup close to the one schematized in Figure II-1 except for the scrubber unit. In addition, a PEAR reactor was used online to simulate the photochemical aging of the engine exhaust. The experimental setup used is schematized in Figure V-1.

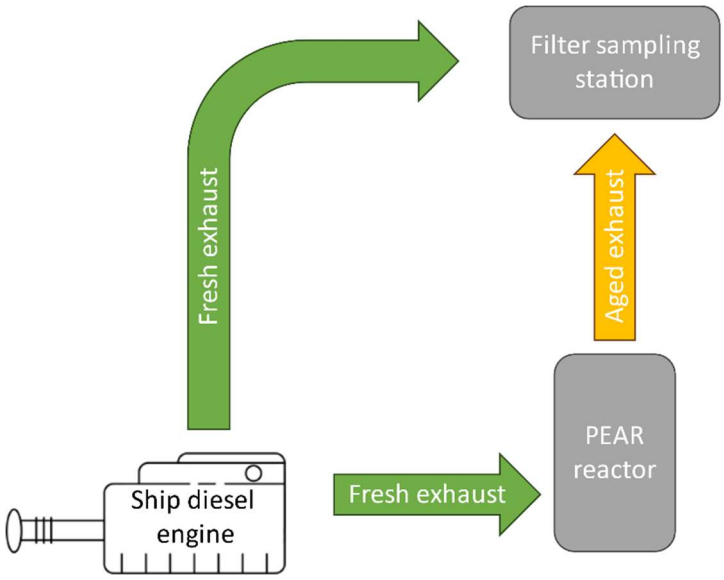


Figure V-1 Scheme of the experimental setup used for the structural elucidation study.

During the campaign, two types of fuel, marine gas oil (MGO) and heavy fuel oil (HFO) were burnt following an engine load cycle that mimicked real-life conditions. The resulting exhaust aerosols were partly directed toward the PEAR reactor in which they underwent artificial aging using a UV lamp and an ozone stream. After the PEAR reactor, the residual ozone was removed using catalytic denuders to stop the oxidation reactions. Finally, both fresh and aged emissions were collected on pre-baked quartz fiber filters (47 mm diameter, Whatman, USA) for 4 hours at a volumetric flow rate of 10 L/min. The collected filters were then stored at −20°C. The final sample set was composed of three filters per aging condition (fresh or aged) and fuel (MGO or HFO) plus three blank filters leading to a total of fifteen filters.

Each filter was split in eight parts. One eighth of each filter was submitted to an extraction in water and another one eighth to an extraction in a dichloromethane (DCM)/methanol (MeOH) mix (1:1 v/v). For both extractions, the filter piece was immersed into 2.5 mL of solvent and put under sonication for 30 min before a filtration step through a polytetrafluoroethylene (PTFE) membrane with a pore size of 0.2 µm. Then the aqueous extract was submitted to solid-phase extraction (SPE) to remove salts while the organic extract was not further prepared before analysis. The SPE protocol is the following: the cartridge (OASIS 6 mL, Waters) is activated with 5 mL of MeOH then loaded with 2 mL of sample, washed with 10 mL of ultrapure water, dried under a N_2 flow for 10 min, and finally eluted with 2 mL of MeOH. For the rest of this chapter, the DCM/MeOH extracts will be identified as “organic extracts”, and the H_2O extracts after the SPE protocol will be referred to as “aqueous extracts”.

The samples were submitted to TWIMS-TOF-MS analysis using the APCI source for the organic extracts and the ESI(+/−) source for the aqueous extract. The parameters for each source are given in table V-1. The instrument was mass calibrated using a sodium formate solution and in ion mobility with the TuneMix (Agilent)

Table V-1: Experimental parameters for APCI, ESI(+) and ESI(–)-TWIMS-TOF MS experiments

	APCI	ESI(+)	ESI(–)
Flow rate (µL/h)	600	300	300
Run duration (min)	15	20	20
Scan rate (Hz)	1	1	1
Corona voltage (V)	4000		
Corona current (µA)	30		
Capillary voltage (V)		2500	2200
Sampling cone	30	30	30
Extraction cone	5	5	5
Cone gas flow (L/h)	30	30	30
Desolvation Temperature (°C)	600	300	300
Desolvation Gas Flow (L/h)	800	800	800
IMS gas flow (mL/min)	70	70	70
IMS wave velocity (m/s)	1000	700	700
IMS wave height (V)	20	20	20

In this study, we performed FTICR MS analysis to facilitate the molecular formula attribution on the TWIMS-TOF MS experiments described hereafter and to control the absence of isobaric interferences. For both FTICR MS and TWIMS MS experiments, we used APCI for the organic extracts and ESI for the aqueous extracts operated in both positive and negative ion mode.

Concerning FTICR MS, for both APCI and ESI experiments, the spectra were acquired over a mass range of *m/z* 90-1200. 100 scans were averaged with a transient length of 1.12 s for ESI experiments while 150 scans were averaged with a transient length of 1.05 s for APCI experiments. Table V-2 presents the parameters used for the APCI and ESI sources. A solution of sodium trifluoroacetate was used for the *m/z* calibration. The molecular formulae were attributed with the DataAnalysis (Version 6.0, Bruker Daltonics, Germany) software with a maximum error of 0.5 ppm, a minimum signal/noise ratio of 7, and the following elemental boundaries: $N_{0-2}O_{0-15}Na_{0-1}$ for ESI(+), $N_{0-2}O_{0-15}$ for ESI(−) and $N_{0-2}O_{0-10}$ for APCI.

Table V-2: Experimental parameters for APCI, ESI(+) and ESI(–)-FTICR MS experiments

	APCI	ESI(+)	ESI(–)
Flow rate (µL/h)	600	300	300
Accumulation time (s)	0.150	0.075	0.100
Capillary voltage (V)	4000	3500	3500
End-plate offset (V)	-500	-500	-1000
Drying gas flow rate (L/min)	2	4	4
Drying gas temperature (°C)	220	200	200
Nebulizing gas pressure	1	1.4	1.4
TOF to detector (ms)	0.7	0.7	0.7
Corona discharge (nA)	4000		
Source temperature (°C)	350		

3. Results and Discussion

As mentioned above, molecular formula attribution shown hereafter was performed with the help of FTICR MS data and verified by using the fine isotopic structure. Unless stated otherwise, the results presented and discussed hereafter concern the samples obtained by the combustion of HFO.

3.1 Organic Extracts Analysis

With the APCI source, we expected to ionize molecules in a wide range of polarity which should be adapted to the mainly apolar species extracted by using a mix of organic solvents. We also wanted to detect the oxygenated species that could be formed during the aging process in the PEAR reactor. Due to a low quantity of samples, we were only able to perform two replicates per aging condition and per fuel used for the combustion.

Figures V-2 show the maps of the drift time *vs* the *m/z* ratio for a fresh sample, or POA (Figure V-2a), and an aged sample, or SOA (Figure V-2b). The maps present a “wrap-around” phenomenon in both sample types. This phenomenon is caused because the time required for all the ions of a single pulse to reach the end of the cell is greater than the IMS scan duration. The largest ions are thus observed on the next scan at very low drift time values, as highlighted

by the red ellipses in Figure V-2. This phenomenon can be avoided by changing the IMS parameters; nevertheless, we decided to let it happen to increase the separation efficiency in the region of interest which is in the range of *m/z* 100-350. Indeed, on the instrument we used, the number of points acquired in IMS is fixed; hence, a way to enhance the IMS resolution over a specific *m/z* region is to increase the number of points acquired in the selected region. This allowed more accurate drift time measurements, for several low intensity APCI analyses. The maps in Figure V-2 show no significant difference at this scale.

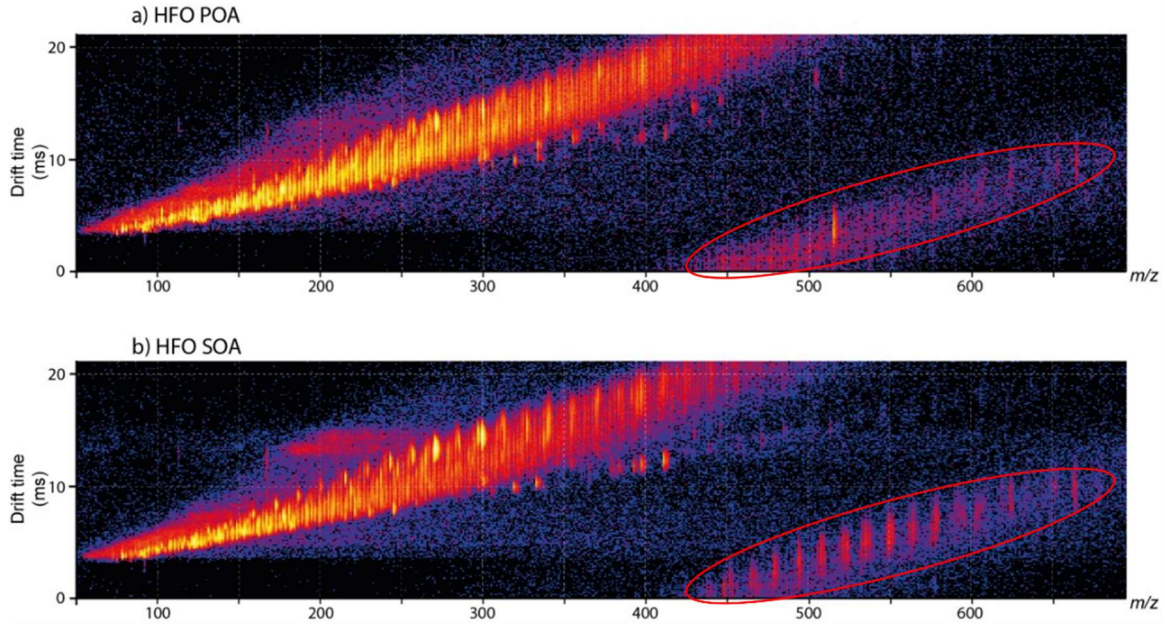


Figure V-2: Drift time vs m/z 2D maps comparing the fresh organic extract (a) and the aged organic extract (b) from APCI(+)-IMS-MS experiments. The wrap-around phenomenon is highlighted by the red ellipses.

We performed a peak-picking using the Driftscope software with an absolute intensity threshold of 50 counts on these 2D maps to search for differences in drift time and FWHM values distribution between the fresh and aged samples. Then, we used the signals between *m/z* 150 and 400 with a relative intensity above 1 %, with respect to the highest peak intensity, to calculate an average drift time and FWHM observed per bin of 10 Da. The results of this operation, displayed on Figures V-3a and b, hint that the species detected in the aged samples are, on average, more compact and isomerically diverse than in the fresh one as they exhibit a lower drift time and a higher FWHM, in average.

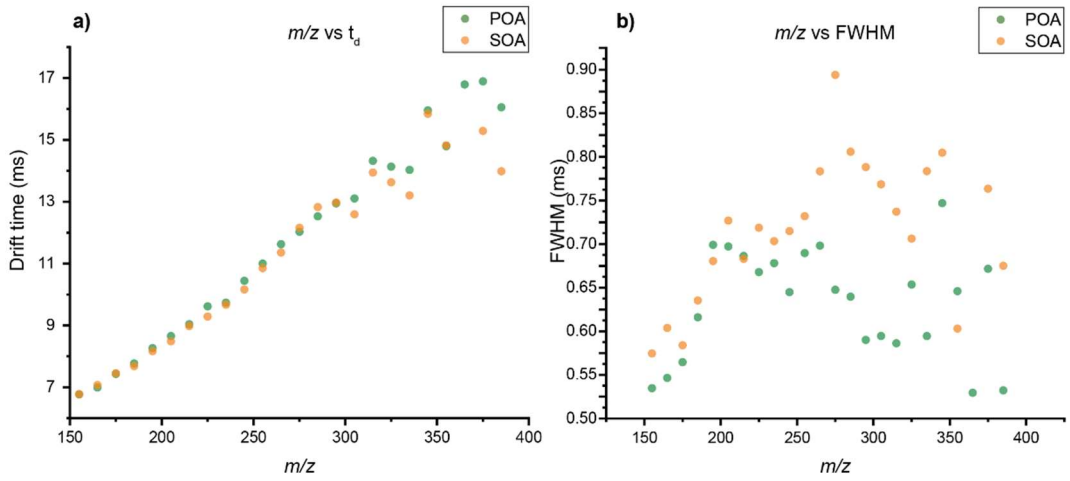


Figure V-3:Drift time (a) and FWHM (b) values averaged over bins of 10 Da as observed for POA (green) and SOA (orange) with APCI-TWIMS-TOF after an automatic peak-picking.

Focusing on a nominal mass exhibiting peaks with high relative intensities in both POA and SOA samples, we searched for more detailed differences. Figure V-4a and Figure V-4b show the enlarged mass spectra at *m/z* 205 of a POA and a SOA sample. On each spectrum, the intensity is normalized to the highest peak displayed; however, the absolute maximum intensity is higher for the POA sample (1.10^4 a.u.) than for the SOA sample (5.10^3 a.u.) (data not shown). In this same *m/z* range, the FTICR MS data allowed us to attribute unambiguously the molecular formula of the different signals as shown on Figure V-4. All these molecules are isobars with carbon numbers ranging from 11 to 15 carbon atoms. All molecular formulas differ by exchanges of CH_4 groups by O atoms from $C_{15}H_{25}^+$ to $C_{11}H_9O^+$. We can see that the species containing more oxygen atoms, at lower mass defects (or m/z values), are relatively more intense in the aged SOA sample than in the fresh POA. That is classically a result of the oxidation that occurs during the artificial aging. The increased intensity of molecules with several oxygen atoms is confirmed by the diagram in Figure V-4c which displays the relative intensities of the species identified at the nominal *m/z* 205 normalized to the highest intensity observed in both samples. This shift in the intensity profile towards species with a higher oxygen number is an expected result of the aging process imposed on our samples and is an indicator of its efficiency. This trend has also been observed at several other *m/z* differing by 14 Da, *i.e.* a $-CH_2$ unit.

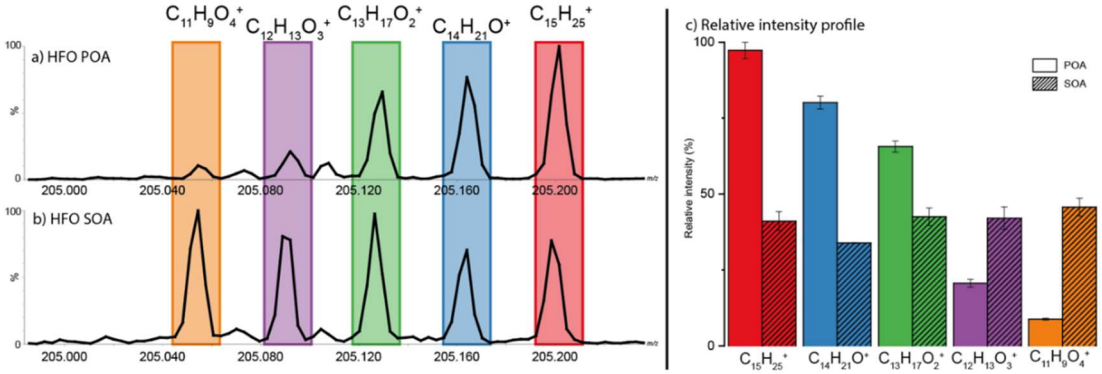


Figure V-4: Comparison of fresh and aged samples in APCI(+)-IMS-MS. Left mass spectra showing nominal mass 205, and right, bar plot of different classes.

From the extracted ion mobility spectra (EIM), we determined the drift times (t_D) and the full-width half maximum (FWHM) which is an indicator of the isomeric diversity (Table V-3). Indeed, the broader the ion mobility peak, the more diverse isomers are likely to contribute to the peak. We compared these values between the POA and SOA to find significative differences hinting at the consumption or production of species based on their structure. To ensure the significance of our result we used a Student's test (or T-test) to evaluate the differences in t_D and FWHM for each ion between the POA and the SOA samples. The values of drift times and FWHM obtained by the Gaussian fitting as well as the p-values resulting from the T-tests are displayed in Table V-3. For comparison, the drift time and FWHM values measured for the TuneMix ions in ESI(+) with the same IMS parameters are displayed in Table V-4.

Table V-3: Drift time and FWHM averaged values and T-tests results for organic extracts analyzed by APCI(+)-IMS-MS (two replicates)

			Drift time (ms)			FWHM (ms)		
Ion	*m/z*	Aerosol type	Average	Std. Dev.	p-value	Average	Std. Dev.	p-value
$C_{11}H_9O_4^+$	205.05	POA	7.42	0.02	0.14	0.666	0.066	0.54
		SOA	7.36	0.01		0.750	0.029	
$C_{12}H_{13}O_3^+$	205.09	POA	7.79	0.04	0.13	0.701	0.038	0.76
		SOA	7.77	0.02		0.724	0.021	
$C_{13}H_{17}O_2^+$	205.12	POA	8.18	0.02	0.13	0.886	0.018	0.21
		SOA	8.25	0.03		1.141	0.068	
$C_{14}H_{21}O^+$	205.16	POA	8.43	0.01	0.34	1.029	0.010	0.68
		SOA	8.48	0.04		1.116	0.148	
$C_{15}H_{25}^+$	205.19	POA	8.87	0.04	0.12	0.959	0.003	0.20
		SOA	8.94	0.06		1.001	0.011	

Table V-4: Tune mix ions drift time and FWHM values as measured with ESI(+)-TWIMS-TOF using the same IMS parameters as the APCI(+)-TWIMS-TOF experiment

m/z	t_d (ms)	FWHM (ms)
118	5.06	0.329
322	9.52	0.614
622	17.37	1.259

The FWHM at *m/z* 205 between 0.67 and 1.14 ms are all larger than the FWHM of the EIM signal of the ion *m/z* 322 of the tune mix, indicating the presence of unresolved isomers in both POA and SOA. None of the p-values in Table V-3 are below 0.05, so we cannot conclude that there are significant differences in t_D or FWHM between the POA and SOA samples for the selected ions. However, these relatively high p-values can also be caused by the lack of analytical replicates due to the low quantity of samples available. Nevertheless, we observed a systematic increase in drift times and FWHM, going from POA to SOA, with only the exception of $C_{11}H_9O_4^+$ which has the highest p-value and lowest intensity in the POA. This would point towards a higher structural diversity in SOA, both in ions with and without oxygen atoms.

We studied, in both POA and SOA, the evolution of the drift time caused by either the addition or loss of 16 u to each of the five ions studied above, either as O atoms or CH_4, thus observing an additional twelve ions with nominal *m/z* ranging from 173 to 237. The results, plotted in Figure V-5, confirm that the drift time, and thus the CCS, of an ion increases with the *m/z* ratio. The increase is roughly linear and the slope is different when adding O atoms or CH_4. Indeed, the addition of CH_4 in the molecular formula leads to a greater increase in the drift time compared to the addition of an oxygen atom. While CH_4 and O are isobars, they don't have the same size or impact on molecular structure. Focusing on ions of the same nominal *m/z* (like *m/z* 205), we observed that the replacement in the molecular formula of a CH_4 group by an oxygen atom systematically decreased the ion's drift time, and consequently collision cross sections. This decrease in drift time for isobars with higher numbers of oxygen atoms is consistent with previous results on polyolefin pyrolysis products [15]. It indicates that the average decrease in drift time and increase in FWHM of the SOA compared to the POA observed in Figures V-3 could largely be caused by the increase in oxygenated species, accompanied by an increase in isomeric diversity.

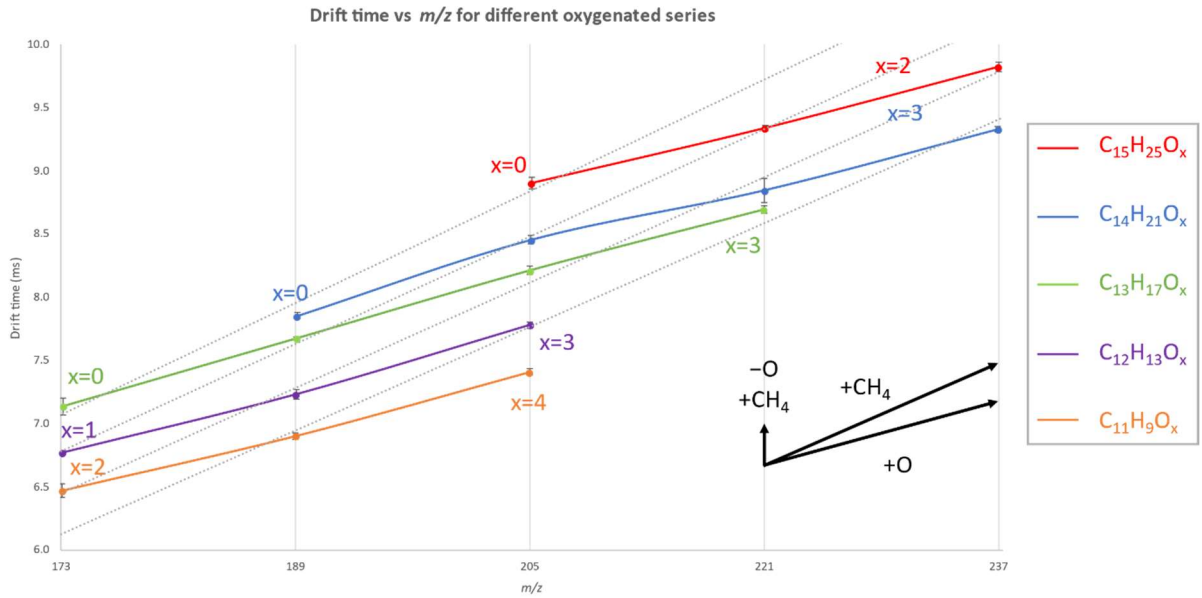


Figure V-5: Diagram plotting the drift time vs m/z *ratio for several oxygenated series with expected evolution of drift time and m/z ratio with the addition and/or removal of a CH_4 group or an oxygen atom. Data from APCI-IMS-MS experiments.*

3.2 Aqueous Extracts Analysis

We used the ESI source, operated in both positive and negative ion modes, for the characterization of the aqueous extracts as we were expecting to find mainly oxygenated polar species due to the extraction protocol. Three analytical replicates were made.

The maps plotting the drift time against the *m/z* ratio of our samples obtained in the positive ion mode are shown in Figures V-6a and V-6b. We observed several distributions of drift time values against *m/z*. However, the secondary distribution, highlighted by the green ellipses, corresponded to a polyethylene (PE) ion series caused by contamination of the instrument source. Hence, for the rest of this part, we will only focus on the main distribution of drift time *versus m/z*. Compared to the APCI experiments, the ESI experiment did not require the use of a higher value of IMS wave velocity, as the higher signal intensity in ESI yielded better-resolved peaks in IMS. Thus, we do not observe a wrap-around phenomenon as with the APCI experiments though it can still be observed slightly from around *m/z* 700. It should be pointed out that the color plot scale is on a log scale. As a consequence, the species displayed in purple correspond to very low-intensity ions barely observable in the mass spectra.

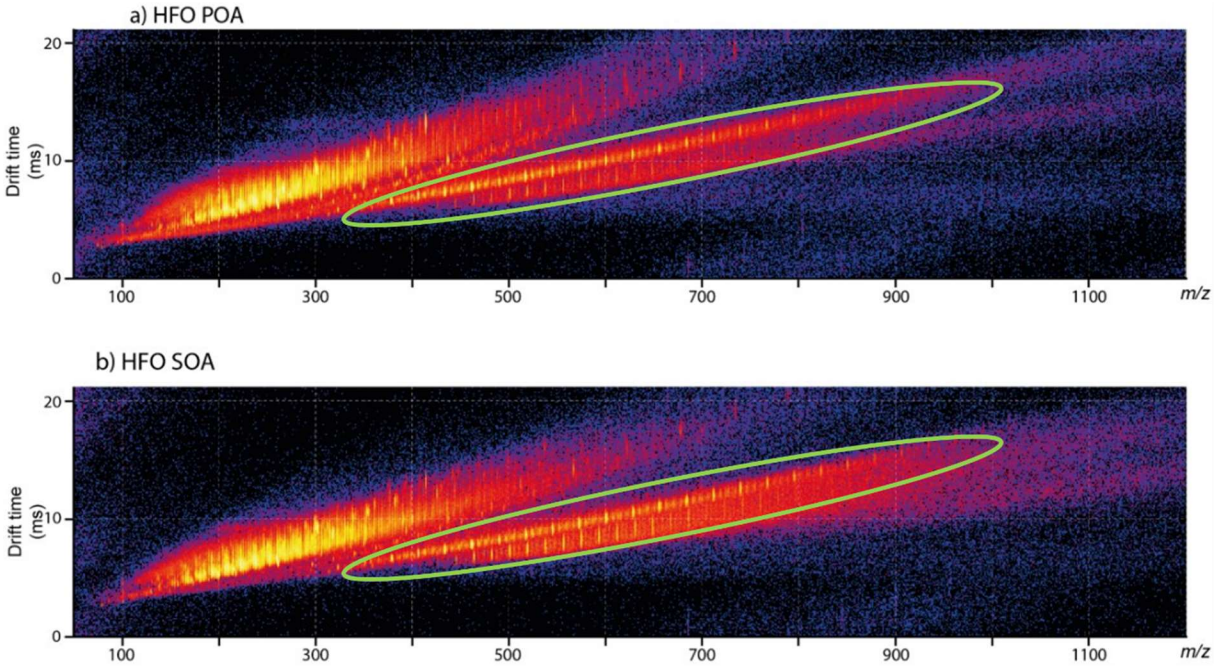


Figure V-6: Drift time vs m/z 2D maps comparing the fresh aqueous extract (a) and the aged aqueous extract (b) from ESI(+)-IMS-MS experiments. The PE distributions are highlighted by the green ellipses.

Similarly to the APCI experiment, we performed a peak-picking on the 2D maps above to search for differences in drift time and FWHM values distribution between the fresh and aged samples. Then we used the species comprised between *m/z* 150 and 400 with a relative intensity above 1 %, with respect to the highest peak intensity , to calculate the average drift time and FWHM observed per bin of 10 Da. The results are shown on Figures V-7 a and b. Figure V-7 a shows that, as for the APCI experiment, the species detected in the aged samples tend to be more compact than those in the fresh samples. However, Figure V-7 b suggests that, as opposed to the previous experiments, the aged samples present less isomeric diversity than the fresh ones.

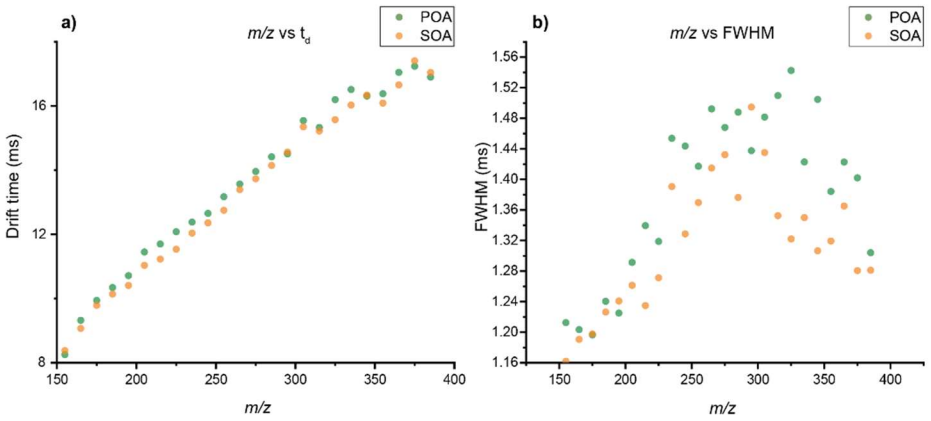


Figure V-7: Drift times (a) and FWHM (b) values, averaged over bins of 10 Da, observed for POA (green) and SOA (orange) with ESI(+)-TWIMS-TOF after an automatic peak-picking.

Focusing on nominal *m/z* 221, we observed four ions corresponding to sodium adducts of compounds with different numbers of oxygen atom , present in both the POA and the SOA. The molecular formulas were confirmed by FTICR MS experiments. The average intensity of molecules with several oxygen atoms was again significantly higher in the SOA compared to the POA samples, as shown in Figures V-8a and V-8b. On each spectrum, the intensity is normalized to the most intense peak displayed; however, the maximum intensity is higher for the POA sample ($7.6.10^4$ a.u.) than for the SOA sample ($2.5.10^4$ a.u.) (data not shown). We also reported the relative intensities of each ion, normalized to the intensity of the most abundant ion in all samples combined, on the diagram shown in Figure V-8c. These results show a shift in the intensity profile towards the species with a higher oxygen number due to the artificial aging, consistently with the observation made for the organic extracts analysis.

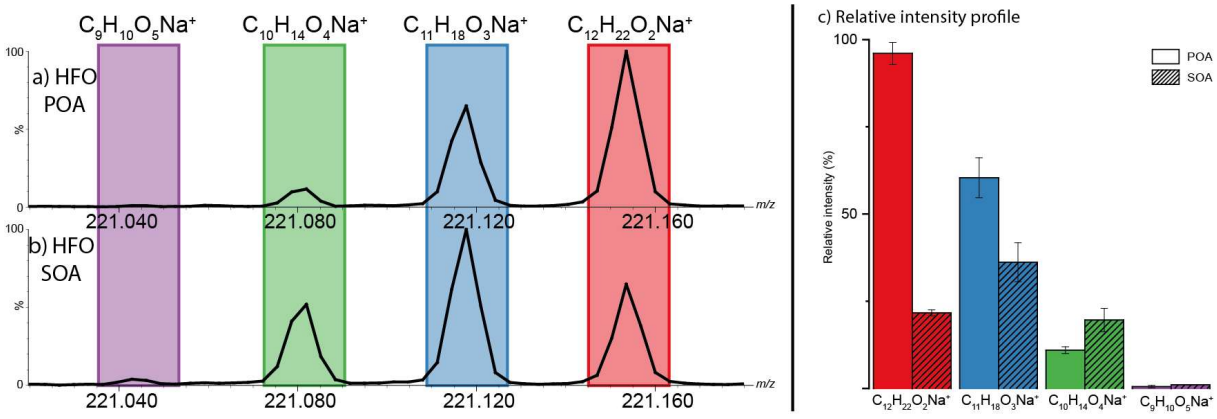


Figure V-8: Comparison of the relative intensity profile between fresh and aged samples analyzed by ESI(+)-IMS-MS.

The drift time and FWHM values for the TuneMix ions are displayed in Table V-5. The drift times (t_D) and the full-width half maximum (FWHM) values for the POA and SOA along with a Student's test (or T-test) to evaluate the differences with p-values are displayed in Table V-6. Here again, the FWHM of all EIM signals for *m/z* 221 were larger than the 0.36 ms observed for the ion *m/z* 322 of the tune mix.

Table V-5: Tune mix ions drift time and FWHM values measured in ESI(+)-TWIMS-TOF

m/z	t_d (ms)	FWHM (ms)
118	3.51	0.225
322	6.61	0.357
622	11.99	0.696
922	17.40	1.089

Table V-6: Drift time and FWHM averaged values and T-tests results for aqueous extracts in ESI(+) (three replicates).

			Drift time (ms)			FWHM (ms)		
Ion	*m/z*	Aerosol type	Average	Std. Dev.	p-value	Average	Std. Dev.	p-value
$C_9H_{10}NaO_5^+$	221.05	POA	5.91	0.01	0.55	0.771	0.048	0.67
		SOA	5.98	0.04		0.787	0.073	
$C_{10}H_{14}NaO_4^+$	221.08	POA	5.9	0.01	0.05	0.774	0.032	0.05
		POA	5.81	0.01		0.653	0.005	
$C_{11}H_{18}NaO_3^+$	221.11	POA	6.03	0.01	0.1	0.724	0.012	0.06
		SOA	6.06	0.01		0.666	0.008	
$C_{12}H_{22}NaO_2^+$	221.15	POA	6.4	0.01	0.01	0.663	0.019	0.01
		SOA	6.53	0.01		0.977	0.043	

A significant change of drift time and FWHM was observed between the POA samples and the SOA samples for $C_{12}H_{22}NaO_2^+$ at *m/z* 221.15. Indeed, the p-values, which characterize the significance of the difference in the mean t_D and FWHM between POA and SOA for this ion are 0.01, strictly lower than the threshold of 0.05. The EIM of this ion in the POA and in the SOA samples are shown in Figures V-9a and V-9b. The lower absolute intensity in the SOA sample indicated this compound was probably consumed and transformed during the aging process. When the intensities are normalized in each sample as shown in Figure V-9b, it was still possible to observe a difference in the peak shape between the POA and the SOA. In fact, the IMS signal appeared to have a bimodal distribution suggesting the presence of two ion

isomer populations. As a consequence, the EIMs in both samples could not be fitted correctly by a single Gaussian fitting. These EIMs hint at the fact that there were at least two different families of structural isomers corresponding to the attributed molecular formula. This hypothesis was verified by performing a deconvolution of the EIMs considering the presence of two Gaussian distributions. Indeed, the output of the double-fitting yielded a cumulative fit statistically matching the raw data ($R^2 > 0.99$ for each replicate) for both POA and SOA samples, as shown in Figures V-9c and V-9d.

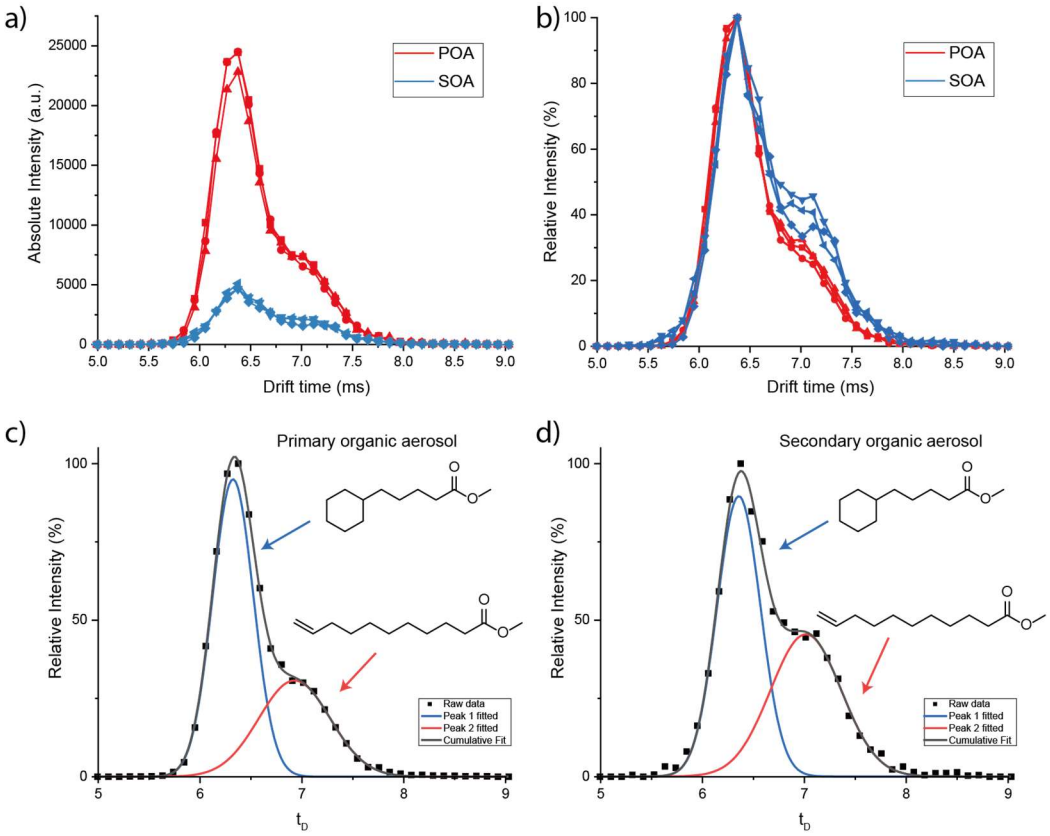


Figure V-9: a) EIM at m/z 221.1535 ($C_{12}H_{22}NaO_2^+$) in absolute intensity. b) EIM at m/z *221.15 in relative intensity. c) Double Gaussian fitting within the POA sample and putative structures. d) Double Gaussian fitting within the SOA sample and putative structures. Data from ESI(+)-IMS-MS experiments.*

The mean results of this fitting are displayed in Table V-7. The relative intensity of the Gaussian peak fit at lower drift time (peak 1) decreased from POA to SOA, while the relative intensity of the Gaussian peak fitted at higher drift times (peak 2) increased. There still are uncertainties concerning the fitting in the SOA samples, as shown by the lower average R^2 coefficient and the higher coefficient of variation values compared to those obtained for the POA sample. This either suggests that the measurements for the SOA samples do not exhibit

enough repeatability and/or that a more complex fitting is necessary to model the isomeric diversity in the SOA samples. Nevertheless, the variation of the intensity ratio between the first and second peaks between POA and SOA hints at a selective consumption of the ions with lower drift times and/or production of ions with higher drift times. This observation still needs to be confirmed by further experiments. The more compact species (lower t_D and CCS) could be more sensitive to the aging process than the larger ones.

Each IMS signal on such complex organic mixtures is expected to involve many isomers. The observation of a distribution with comparatively larger drift times are likely to be associated to a large change in the structure, such as a ring opening or branching. It is known that aromatic structures are typically more compact than species with saturated rings [16]. On the other hand, a change in CCS may be related to the existence of branched species that are more compact than non-branched. The molecular formula $C_{12}H_{22}O_2$ corresponds to a compound with two unsaturations so it cannot be an aromatic compound. We suggest two tentative structures displayed on Figure V-9c and d, that would involve a ring opening to account for the difference in drift times between the two distributions.

Table V-7: Results of the double Gaussian fitting of the EIM at m/z *221.1535 ($C_{12}H_{22}NaO_2^+$). (CV: Coefficient of variation)*

		Peak 1 : lower drift time			Peak 2 : higher drift time		
	Parameter	Mean values	Std. Dev.	CV	Mean values	Std. Dev.	CV
POA	dt (ms)	6.324	0.002	0.03%	6.893	0.028	0.40%
	FWHM (ms)	0.457	0.011	2.38%	0.883	0.028	3.17%
	Rel. Int. (%)	91.86	2.05	2.23%	31.39	1.45	4.60%
	R^2	0.999					
SOA	dt (ms)	6.347	0.004	0.07%	6.934	0.092	1.33%
	FWHM (ms)	0.464	0.040	8.63%	1.013	0.143	14.09%
	Rel. Int. (%)	77.85	9.34	11.99%	42.95	2.75	6.40%
	R^2	0.995					

For the negative ion mode, the maps of the drift time against the m/z ratio of our samples obtained in the negative ion mode are shown in Figures V-10a and b and presents again a PE distribution caused by the source contamination in addition to the main distribution of the sample.

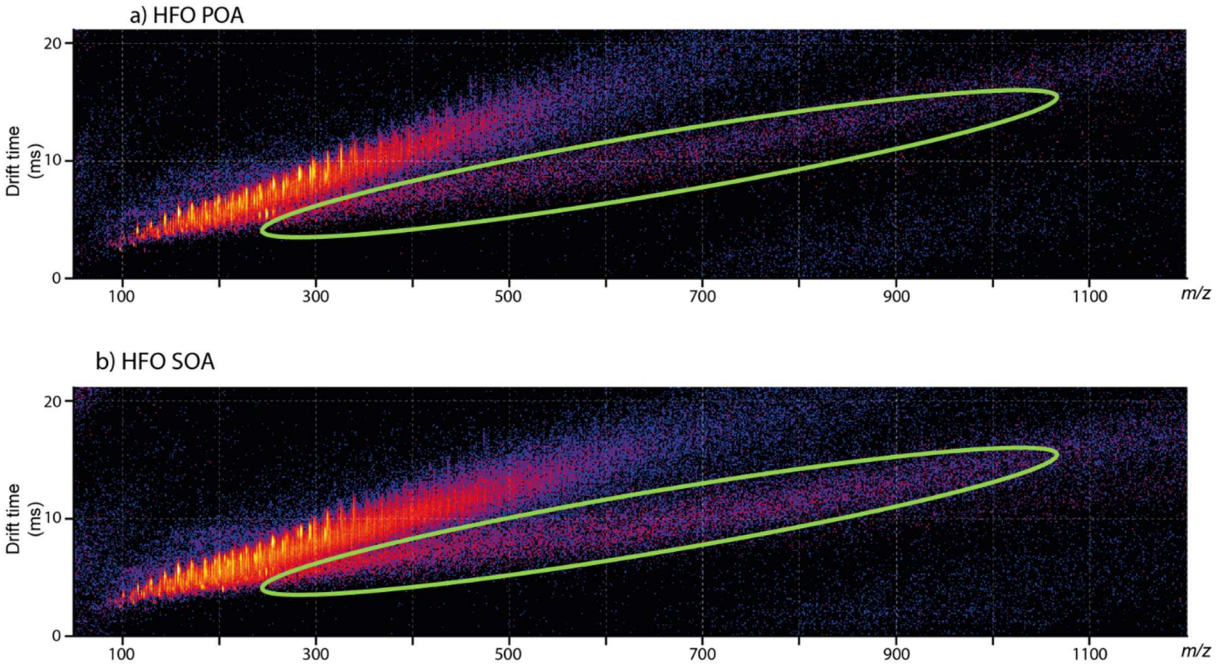


Figure V-10: Drift time vs m/z 2D maps comparing the fresh aqueous extract (a) and the aged aqueous extract (b) recorded by ESI(−)-IMS-MS. PE contaminant distributions are highlighted by the green ellipses.

As for the preceding experiments, we averaged the drift time and FWHM values per bin of 10 Da between *m/z* 150 and 400 after performing a peak-picking on the 2D maps above. The results of this treatment are shown on Figures V-11a and b. For these experiments, it appears that the compounds in the aged samples are more compact than those in the fresh ones, on average. Also, the isomeric diversity seems to be slightly more important in the aged samples than in the fresh, although no clear discrepancy is found.

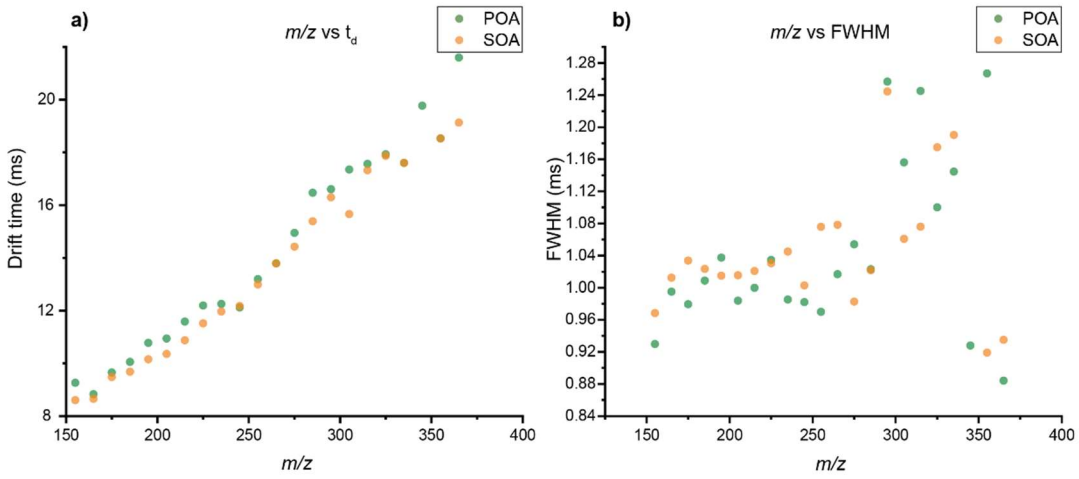


Figure V-11: Drift times (a) and FWHM (b) values averaged over bins of 10 Da observed for POA (green) and SOA (orange) with ESI(−)-TWIMS-TOF after an automatic peak-picking.

We focused on a *m/z* range at which intense compounds were found both in POA and SOA to verify previous observations concerning the shift of the intensity profile towards the species with a higher oxygen number due to the aging process. In the same way as for the APCI and ESI(+) experiments, we focused on the nominal *m/z* 199 based on FTICR experiments and the TOF intensity profile.

Indeed, at *m/z* 199, we were able to confidently attribute four molecular formulas with carbon numbers ranging from 9 to 12 with two to five oxygen atoms (Figure V-12). In the negative ion mode, the ions with higher oxygen numbers had higher relative intensity in the SOA compared to the POA, as shown in Figure V-12a and V-12b. On each spectrum, the intensity is normalized to the most intense peak displayed. The maximum intensity is higher for the POA sample ($3.4.10^4$ a.u.) than for the SOA sample ($2.1.10^4$ a.u.) (data not shown). We also reported the relative intensities of each ion, normalized to the intensity of the most abundant ion in all samples combined, on the diagram shown in Figure V-12c. A shift in the intensity profile towards the species with a higher oxygen number due to the aging was observed agreeing with the previous observations.

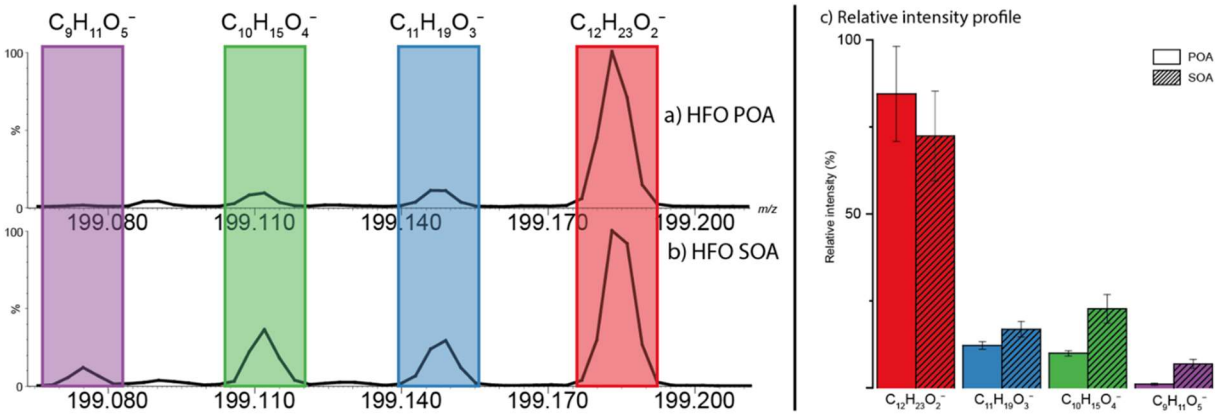


Figure V-12: Comparison of the relative intensity profile between fresh and aged samples in ESI(−).

Here again, the FWHM of all EIM signals at m/z 199 (Table V-8) were larger than the 0.32 ms observed for the ion *m/z* 322 of the tune mix (Table V-9). However, the drift times t_D and FWHM of each ion in POA and SOA and the t-test showed no significant variations, neither for t_D nor for FWHM, as no p-values are below the 0.05 threshold. No significant structural change could therefore be evidenced for any of these ions at *m/z* 199.

Table V-8: Drift time and FWHM averaged values and T-tests results for aqueous extracts in ESI(−).

			Drift time (ms)			FWHM (ms)		
Ion	*m/z*	Aerosol type	Average	Std. Dev.	p-value	Average	Std. Dev.	p-value
$C_{10}H_{15}O_4^-$	199.11	POA	6.09	0.32	0.87	0.509	0.123	0.53
		SOA	6.08	0.35		0.532	0.164	
$C_{11}H_{19}O_3^-$	199.15	POA	6.09	0.34	0.42	0.534	0.106	0.42
		SOA	6.09	0.34		0.534	0.106	
$C_{12}H_{23}O_2^-$	199.18	POA	6.07	0.33	0.74	0.510	0.088	0.79
		SOA	6.08	0.34		0.506	0.085	

Table V-9 Tune mix ions drift time and FWHM values measured in ESI(−)-TWIMS-TOF

m/z	t_d (ms)	FWHM (ms)
113	2.54	0.183
302	5.43	0.317
602	9.43	0.531
1034	18.97	1.232

We noted that the FWHM values of the signals in the EIM of *m/z* 199 were in the range of 0.5 ms, much lower than the values of up to 1 ms for positive APCI on the organic extract or positive ESI on the aqueous extract. The molecules ionized in ESI–, which are usually carboxylic acids, had therefore less isobaric and isomeric diversity than the molecules ionized in the positive ion modes.

4. Conclusion and Perspectives

The preliminary results of this ongoing study showed evidence for structural differences due to the atmospheric aging of ship emissions, especially in the aqueous extracts analyzed with positive ESI. The use of other ion sources, such as atmospheric pressure photoionization could help gather information on other chemical classes of compounds, such as aromatics. Also, further analysis on a more resolutive IMS instrument could be performed to confirm changes in ion mobility spectra as evidence of more changes in structure. Such instruments could be the trapped IMS (TIMS) or the cyclic IMS (cIMS). The use of molecular modeling could be used to propose putative structures with CCS consistent with those determined experimentally and, in this way, obtain information on the structures affected by the aging process. Finally, our hypothesized ring opening proposed to account for the bimodal

distribution in EIM observed in the ESI(+) results at *m/z* 221 needs to be confirmed or rejected by other experiments such as tandem MS experiments isolating and fragmenting the ions at *m/z* 221 to determine the potential structures of the isomers.

5. References

1. Trivitayanurak, W. and P.J. Adams, *Does the POA–SOA split matter for global CCN formation?* Atmospheric Chemistry and Physics, 2014. **14**(2): p. 995-1010.

2. Houghton, J.E.T., et al., *Climate Change 2001: The Scientific Basis*. 2001. p. 881.

3. Nguyen, T.B., et al., *High-resolution mass spectrometry analysis of secondary organic aerosol generated by ozonolysis of isoprene.* Atmospheric Environment, 2010. **44**(8): p. 1032-1042.

4. Reinhardt, A., et al., *Ultrahigh Mass Resolution and Accurate Mass Measurements as a Tool To Characterize Oligomers in Secondary Organic Aerosols.* Analytical Chemistry, 2007. **79**(11): p. 4074-4082.

5. Aiken, A.C., et al., *O/C and OM/OC Ratios of Primary, Secondary, and Ambient Organic Aerosols with High-Resolution Time-of-Flight Aerosol Mass Spectrometry.* Environmental Science & Technology, 2008. **42**(12): p. 4478-4485.

6. Bahreini, R., et al., *Measurements of Secondary Organic Aerosol from Oxidation of Cycloalkenes, Terpenes, and m-Xylene Using an Aerodyne Aerosol Mass Spectrometer.* Environmental Science & Technology, 2005. **39**(15): p. 5674-5688.

7. Ge, X., et al., *Primary and secondary organic aerosols in Fresno, California during wintertime: Results from high resolution aerosol mass spectrometry.* 2012. **117**(D19).

8. Krechmer, J.E., et al., *Ion mobility spectrometry–mass spectrometry (IMS–MS) for on- and offline analysis of atmospheric gas and aerosol species.* Atmos. Meas. Tech., 2016. **9**(7): p. 3245-3262.

9. *Southern Oxydant and Aerosol Study 2013*. [cited 2023; Available from: https://soas2013.rutgers.edu/.

10. Iinuma, Y., et al., *Structural Characterisation of Dimeric Esters in α-Pinene Secondary Organic Aerosol Using N2 and CO2 Ion Mobility Mass Spectrometry.* 2021. **12**(1): p. 17.

11. West, C.P., et al., *Molecular and Structural Characterization of Isomeric Compounds in Atmospheric Organic Aerosol Using Ion Mobility-Mass Spectrometry.* The Journal of Physical Chemistry A, 2023. **127**(7): p. 1656-1674.

12. Steiner, W.E., et al., *Detection of a Chemical Warfare Agent Simulant in Various Aerosol Matrixes by Ion Mobility Time-of-Flight Mass Spectrometry.* Analytical Chemistry, 2005. **77**(15): p. 4792-4799.

13. Rüger, C.P., et al., *Cyclic Ion Mobility Spectrometry Coupled to High-Resolution Time-of-Flight Mass Spectrometry Equipped with Atmospheric Solid Analysis Probe for the Molecular Characterization of Combustion Particulate Matter.* Journal of the American Society for Mass Spectrometry, 2021. **32**(1): p. 206-217.

14. Ihalainen, M., et al., *A novel high-volume Photochemical Emission Aging flow tube Reactor (PEAR).* Aerosol Science and Technology, 2019. **53**(3): p. 276-294.

15. Farenc, M., et al., *Characterization of Polyolefin Pyrolysis Species Produced Under Ambient Conditions by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry and Ion Mobility-Mass Spectrometry.* Journal of the American Society for Mass Spectrometry, 2017. **28**(3): p. 507-514.

16. Lalli, P.M., et al., *Isomeric Separation and Structural Characterization of Acids in Petroleum by Ion Mobility Mass Spectrometry.* Energy & Fuels, 2015. **29**(6): p. 3626-3633.

General Conclusion and Perspectives

During this work, we were able to develop a software for UHRMS data visualization called “Python tools for complex matrices molecular characterization – Viewer edition” or PyC2MC Viewer. This software can process data extracted from mass spectra peak lists to make the commonly used data visualization plots such as DBE vs carbon number maps, van Krevelen diagrams, or Kendrick mass defect diagrams, among others. It also allows basic sample comparison and statistical analyses. PyC2MC Viewer was designed to be a user-friendly tool with an intuitive graphical user interface and a simple installation protocol. In addition, this software was also created to encourage modifications and improvements by its users. This is why it was published as an open-source code on a GitHub public repository. As of now, PyC2MC counts more than 150 downloads of its executable version and has been subjected to several improvements since its first publication.

We evidenced for the first time the presence of vanadyl and nickel petroporphyrins in the exhaust of a ship engine, consistently with the known emission of vanadium and nickel into the atmosphere. We were also able to conclude that the poorly efficient combustion occurring at low engine load caused the release of a greater quantity of petroporphyrins. As this low engine load is mainly used when ships travel close to the coastlines or when they are at berth in ports, it means that petroporphyrins are mainly released close to populated areas which is problematic as their toxicity is yet unknown.

In the last chapter of this dissertation, we presented the initial results of our structural study of aged ship emissions by IMS-MS. These results show the capacity of ion mobility spectrometry to approach the isomeric diversity in organic aerosols. These results show the differences between fresh and photochemically aged particulate matter obtained from heavy fuel oil combustion in a ship engine. A shift of the average oxygen number towards higher values was evidenced in the aged samples, which is caused by the oxidation taking place during the artificial aging. Furthermore, we also observed a bimodal distribution in the ion profile for one of the ions studied, indicating the existence of two isomeric populations. These populations also seem to be differently affected by the aging process as the intensity ratio between the two populations is significantly different before and after the artificial aging.

As perspectives of this work, PyC2MC will be continuously developed and improved, as it was intended. In the future, PyC2MC, will integrate raw data processing, from FTMS signal processing to molecular formula attribution. The data treatment of hyphenated MS techniques (LC-MS, GC-MS, IMS-MS) will be also included. This would eventually lead to an open-source software able to efficiently perform every step of FTMS data treatment and visualization. Then, the study on the aged aerosol should be continued. Indeed, further experiments, such as MS/MS, could help to understand the difference in consumption observed in the experiments on aqueous extract with ESI(+). Furthermore, it would also be interesting to use other ion sources, such as atmospheric pressure photoionization (APPI), to benefit from the complementarity of ion sources. Also, it would be possible to use the collision cross-section prediction to determine the most plausible structures of the studied ions.

Résumé en français

Introduction

Le terme "aérosol" désigne une suspension de particules, liquides ou solides, dans un gaz. Plus précisément, en sciences de l'atmosphère, le terme désigne la phase particulaire également appelée matière particulaire (PM). Ces particules ont un diamètre aérodynamique (D_{ae}) compris entre 10^{-9} et 10^{-4} m [1]. La limite inférieure est déterminée par la taille des molécules, tandis que la limite supérieure correspond à la taille à laquelle les particules et les gouttelettes ont tendance à sédimenter.

Les sources primaires d'émissions d'aérosols peuvent être regroupées en deux types : les sources naturelles et les sources anthropiques. Les sources naturelles se composent principalement d'embruns formés par l'éclatement de bulles à l'interface mer-air [2] et de l'érosion des sols. Néanmoins, des phénomènes tels que les feux de forêts et les éruptions volcaniques sont également des sources naturelles de matière particulaire primaire. Les sources anthropiques sont elles-mêmes divisées en sources mobiles et fixes. Les sources mobiles comprennent le trafic maritime, aérien, ferroviaire et terrestre. Alors que les sources stationnaires désignent la production d'énergie domestique et à grande échelle, qu'elle soit thermique ou mécanique, mais aussi la cuisson et les opérations industrielles [3]. La durée de vie d'un aérosol dans l'atmosphère varie entre quelques heures et quelques semaines et dépends de propriétés telles que sa composition chimique, la taille et la structure de ses particules [4, 5]. Les particules en suspension dans l'air peuvent en être retirées suivant deux mécanismes de dépôt : le dépôt humide qui fait intervenir la formation de nuages et de précipitations : et le dépôt sec qui consiste soit en la sédimentation des particules par gravité soit en l'impact des celles-ci sur une surface.

La composition d'un aérosol organique (OA) dépend grandement de sa source, mais aussi de son état de vieillissement. En effet, la composition d'un aérosol n'est pas fixe dans le temps et est notamment affectée par des réactions photochimiques en présence d'agents oxydants atmosphériques tels que l'ozone. Ainsi, les aérosols organiques vieillis (ou aérosols organiques secondaires, SOA) ont tendance à être plus fonctionnalisés et plus polaires que les

aérosols organiques primaires (POA). De ce fait, les SOA ont également tendance à être des noyaux de condensation de nuages plus efficaces [6].

L'un des principaux moteurs de la recherche sur les aérosols organiques est l'étude de leur impact sur notre environnement, en particulier la manière dont ils contribuent au changement climatique ou provoquent des problèmes de santé. L'impact climatique des constituants des OA peut se traduire aussi bien par un réchauffement de l'atmosphère, causé par des particules absorbants les rayonnements terrestres, que par un refroidissement de l'atmosphère causée par la réflexion ou la dispersion des rayonnements solaire dans l'atmosphère. De nombreuses études montrent que les particules en suspension dans l'air et d'autres formes de pollution atmosphérique sont liées à de graves effets sur la santé, notamment une augmentation de la mortalité et des maladies cardiovasculaires, respiratoires et allergiques [7-9].

Parmi les nombreuses sources d'émissions d'aérosols organiques, nous nous sommes concentrés sur les gaz d'échappement des navires qui, en termes de volume émis par an, sont aussi importants que les émissions des voitures. En outre, les émissions des navires présentent un intérêt particulier non seulement en raison de la quantité émise, mais aussi de leur composition, dont il a été démontré qu'elle avait un effet de refroidissement du climat. En effet, les NO_x, formés par la combustion de diazote à haute température dans les moteurs de navires, sont des précurseurs de l'ozone, dont la formation perturbe la concentration d'agents oxydants dans l'atmosphère, conduisant à un vieillissement accéléré des aérosols et gaz à proximité. Notamment, le méthane voit sa durée de vie dans l'atmosphère diminuer à cause de l'émission de précurseurs d'ozone par les navires, menant ainsi à la réduction de sa contribution à l'effet de serre et donc à une diminution de la température globale [10].

De plus, une proportion importante de ces émissions se produit lorsque les navires sont en transit ou à quai dans les zones portuaires. L'impact sur la santé des gaz d'échappement des navires est donc également non négligeable. En effet, il a été estimé en 2007 que les gaz d'échappement des navires étaient responsables de 60 000 décès par an [11]. Plus récemment, Barregard *et al* [12] ont estimé qu'en 2014, environ 1 500 cas de cardiopathie ischémique étaient dus aux émissions de PM provenant de la navigation dans la région de la

mer Baltique. Ils ont fait à peu près la même estimation pour les cas d'accidents vasculaires cérébraux au cours de la même période et dans la même région.

Afin de réduire l'impact du transport maritime sur la santé et le climat, l'Organisation maritime internationale (OMI) a mis en place des réglementations dans le cadre de la convention sur la pollution marine (MARPOL). L'OMI a également défini des régions sensibles dans lesquelles ses réglementations sont encore plus strictes afin de réduire davantage les émissions de polluants cibles : SOx, NOx et particules. Elles sont appelées zone de contrôle des émissions d'azote (NECA), zone de contrôle des émissions de soufre (SECA) ou zone de contrôle des émissions (ECA), selon qu'elles visent à réduire les émissions d'azote, les émissions de soufre ou les émissions de particules, respectivement.

Il existe deux approches principales pour se conformer aux exigences susmentionnées. Tout d'abord, il y a les stratégies "primaires" qui concernent le moteur lui-même et qui englobent le choix du carburant utilisé, l'optimisation des conditions de combustion, et la recirculation des gaz d'échappement (EGR) [13]. D'autre part, il y a les méthodes dîtes "secondaires" qui visent à améliorer un moteur déjà installé, comme les épurateurs et la réduction catalytique sélective (SCR) [14].

La forte proportion de sulfates est une spécificité des émissions du transport maritime. En effet, les carburants utilisés pour le transport maritime sont principalement des fiouls lourds (HFO) à forte teneur en soufre (FSC). L'utilisation de ce type de carburant remonte à la crise pétrolière des années 1970 et à l'augmentation du prix des distillats moyens de pétrole brut, comme le diesel marin (MDO) ou le gazole marin (MGO), qui en a résulté. Cela a incité les entreprises de transport à utiliser une coupe moins raffinée mais moins chère du pétrole brut : les fiouls lourds (ou fiouls résiduels). Une stratégie « primaire » pour réduire les émissions de sulfates est l'utilisation de MGO ou de MDO à la place du HFO [15] ou encore l'utilisation de carburants alternatifs, tels que le biodiesel [16] ou le gaz naturel liquéfié [17]. Une stratégie « secondaire » de réduction des émissions de sulfate est l'installation d'une unité d'épuration des fumées d'échappement (Scrubber unit en anglais). Dans ce type d'unité, les gaz d'échappement sont pulvérisés avec de l'eau pour éliminer les polluants, en particulier

les sulfates de la phase gazeuse. La phase liquide obtenue en sortie d'épurateur est chargée en polluant extrait des fumées et doit donc être traitée avant d'être rejetée.

Les réglementations en matière de pollution de l'air pour les émissions des navires imposent un suivi régulier afin de s'assurer que les compagnies maritimes s'y conforment. Cependant, l'étude en profondeur des conséquences de la pollution de l'air par les navires nécessite un examen plus poussé. Une des approches possibles est celle utilisée dans ce travail qui consiste à caractériser la composition des émissions des navires, en particulier des particules, au niveau moléculaire afin d'identifier les composés qui pourraient faire l'objet de réglementations futures. Pour ce faire, on utilise la spectrométrie de masse à haute résolution qui permet d'accéder dans la composition moléculaire de l'échantillon étudié. Cette technique peut être couplée à une autre appelée spectrométrie de mobilité ionique pour fournir des informations sur la structure des molécules identifiées par spectrométrie de masse et ainsi fournir des informations sur la complexité isomérique.

Matériel et méthodes

Durand cette thèse, des échantillons réalistes de particules émises par les navires ont été étudiés. Ces échantillons de particules ont été obtenus au cours de plusieurs campagnes de mesure qui ont eu lieu à l'université de Rostock, en Allemagne. Pour ces campagnes, un moteur diesel expérimental a été utilisé pour brûler différents carburants marins dans des conditions de fonctionnement variables afin de reproduire les conditions réelles d'utilisation d'un moteur de bateau. Les deux types de carburant utilisés dans les campagnes sont le HFO et le MGO, c'est-à-dire un carburant lourd à haute teneur en soufre et un carburant plus propre conforme aux réglementations 2020 de l'OMI sur la teneur en soufre. La charge du moteur, c'est-à-dire la puissance de sortie, était un paramètre important car elle permet de simuler différents états de fonctionnement du navire qui dépendent généralement de sa distance par rapport à la côte : une faible charge du moteur est utilisée pour les opérations de manœuvre dans les ports, tandis qu'une charge élevée est utilisée lorsque le navire navigue en haute mer. Le dispositif utilisé pour l'échantillonnage des particules se compose de quatre porte-filtres

reliés à une pompe qui permet à un flux continu et stable de gaz d'échappement de passer à travers les filtres. Les particules ont été recueillies sur des filtres en fibre de quartz précuits et stockées à -20°C jusqu'à l'analyse. Les particules ont ensuite été étudiées par spectrométrie de masse, soit telles quelles sous forme de solide, soit après une étape d'extraction solide-liquide permettant de cibler une catégorie spécifique de composés, tels que les espèces polaires ou apolaires.

Deux techniques on été utilisées pour l'analyse de ces échantillons : la spectrométrie de masse à ultra-haute résolution (FTICR MS pour spectromètre de masse à résonnance cyclotronique ionique par transformée de Fourier) et la spectrométrie de mobilité ionique couplée à la spectrométrie de masse haute résolution (IMS-MS). Les expériences de spectrométrie de masse à ultra-haute résolution ont été menées sur un spectromètre FTICR équipé d'un aimant supraconducteur de 12 T tandis que les expériences de mobilité ionique ont été conduites sur un appareil combinant un analyseur de masse à temps de vol (TOF) et une cellule de mobilité ionique à ondes progressives (TWIMS). Ces deux techniques nécessitent que l'échantillon soit ionisé. Ainsi, nous avons utilisé différentes sources d'ionisation afin de bénéficier de leur complémentarité en termes de sélectivité d'ionisation.

Résultats

Les résultats obtenus pendant cette thèse sont présentés sous forme de trois chapitres. Le chapitre trois porte sur le développement de notre logiciel de traitement et visualisation de données acquises par spectrométrie de masse à ultra-haute résolution (UHRMS). Le chapitre quatre a pour sujet la caractérisation ciblée des pétroporphyrines dans les carburants et émission de navire par UHRMS. Enfin le chapitre cinq se concentre sur les premiers résultats obtenus dans l'étude structurelle des émissions des navires âgés par IMS-MS.

Traitement et visualisation des données de UHRMS pour l'analyse de mélanges complexes

Les mélanges complexes englobent une grande variété d'échantillons provenant de différents domaines de recherche tels que la petroléomique, la protéomique ou les sciences de l'environnement. Ainsi, les mélanges complexes peuvent être, par exemple, du sérum sanguin, du carburant, des eaux usées ou des aérosols organiques. Ce type d'échantillon contient une énorme quantité d'informations au niveau moléculaire qui peuvent être utiles pour mieux comprendre les processus biologiques, industriels ou environnementaux. Cependant, l'accès à ces informations moléculaires nécessite des techniques de pointe telles que les spectromètres de masse à transformée de Fourier (FTMS). En effet, le FTMS, sous la forme d'un FTICR MS ou d'un Orbitrap MS, présente une puissance de résolution et une précision suffisantes pour plonger dans la composition moléculaire de ces mélanges organiques complexes. Néanmoins, comme nous l'avons déjà mentionné, ces techniques haut de gamme génèrent de grandes quantités de données qui doivent être traitées et visualisées efficacement pour en extraire des informations utiles. À cet égard, nous avons développé le logiciel PyC2MC pour "Python tool for complex matrices molecular characterization" au sein du laboratoire international pour la caractérisation moléculaire des matrices complexes (iC2MC). Cet outil se veut facile d'utilisation et permet d'utiliser les représentations les plus courantes en spectrométrie de masse à haute résolution, telles que les diagrammes d'équivalent en double liaisons *vs* nombre de carbones ou les diagrammes de van Krevelen, sans être limité à une marque spécifique d'instruments. Il fournit également des outils d'analyse statistique pour faciliter la comparaison entre les échantillons, avec ou sans réplicats , afin de démontrer la reproductibilité d'une analyse et/ou de mettre en évidence les différences au niveau moléculaire. Ce logiciel a fait l'objet de l'article "PyC2MC : an open-source software solution for visualization and treatment of high-resolution mass spectrometry data" que nous avons publié et dont le texte intégral est disponible dans le chapitre trois.

Un autre objectif de cet outil est d'être utilisé, révisé et amélioré par la communauté des utilisateurs de FTMS. Pour cela, nous l'avons mis en open source et open-acces en même

temps que la publication d'un article décrivant le logiciel. PyC2MC fait l'objet d'un développement et d'une amélioration continus grâce aux suggestions des utilisateurs du FTMS, qu'ils fassent partie ou non de l'équipe du projet original.

Caractérisation ciblée des pétroporphyrines dans les carburants et émission de navire par spectrométrie de masse par résonance cyclotronique ionique à transformée de Fourier avec désorption/ionisation laser assistée par matrice à transfert d'électrons (ET-MALDI FTICR MS)

L'analyse élémentaire des combustibles des navires révèle la présence de métaux lourds tels que le vanadium ou le nickel. On sait que ces métaux sont présents dans les combustibles fossiles sous la forme d'un produit de dégradation des chlorophylles : les pétroporphyrines. Les pétroporphyrines sont constituées d'un noyau à base de tétra pyrrole complexant un ion métallique tel que le vanadium ou le nickel. De plus, la présence de ces métaux, au niveau atomique, a également été mise en évidence par des analyses de la composition élémentaire des émissions des navires. L'objectif de cette étude est de déterminer le devenir des pétroporphyrines au cours du processus de combustion dans un moteur de bateau. Pour ce faire, nous avons utilisé le moteur de bateau réaliste décrit précédemment et des carburants de bateau commerciaux avec différentes teneurs mesurées en V et Ni et différentes conditions de combustion pour émuler le fonctionnement d'un bateau de marchandises. Les carburants utilisés et les émissions de particules correspondantes ont été analysés par désorption/ionisation laser assistée par matrice à transfert d'électrons (ET-MALDI) FTICR MS. Cette technique analytique a permis d'améliorer l'ionisation des pétroporphyrines dans les carburants et les émissions des navires. Tout d'abord, une étude préliminaire des pétroporphyrines a été réalisée dans les combustibles des navires, ce qui a permis de mettre en évidence la prédominance des pétroporphyrines de vanadyle sur les porphyrines de nickel, comme l'a suggéré l'analyse de la composition élémentaire. Cette caractérisation préliminaire a également permis de dégager trois valeurs principales d'équivalent double liaison (DBE) pour les porphyrines détectées. Le DBE étant un descripteur

structurel, nous avons pu proposer des structures potentielles pour les porphyrines que nous avons détectées, en lien avec les résultats de recherches antérieures. Ensuite, nous nous sommes concentrés sur le devenir des porphyrines au cours du processus de combustion en analysant les particules et en utilisant le *fold change* comme indicateur de la consommation ou de la production de chaque porphyrine au cours du processus de combustion en comparant l'évolution de leur intensité relative entre les échantillons de combustibles et les émissions de particules. Nous avons pu mettre en évidence que les pétroporphyrines sont consommées de manière homogène pendant la combustion et qu'aucune structure spécifique, suggérée par la DBE, n'est dégradée de manière prédominante. En outre, nous avons étudié l'influence de la charge du moteur sur la composition des émissions et nous avons pu déterminer que les porphyrines sont plus susceptibles d'être libérées sous une faible charge du moteur que sous une charge élevée, probablement en raison d'une combustion incomplète permettant l'émission de pétroporphyrines en tant que composant des gouttelettes de carburant non brûlées. Cette étude a été publiée et son texte intégral est disponible dans le chapitre quatre.

Étude structurelle d'aérosols organiques artificiellement vieillis par spectrométrie de mobilité ionique

L'étude du processus de vieillissement des aérosols organiques (OA) présente un intérêt majeur car le vieillissement modifie les propriétés physiques et chimiques de l'aérosol dans le temps et dans l'espace. En effet, les aérosols organiques secondaires (SOA) ont tendance à être plus fonctionnalisés et polaires que les aérosols organiques primaires (POA), augmentant ainsi leur efficacité en tant que noyaux de condensation des nuages. Par conséquent, le vieillissement des OA biogéniques ou anthropogéniques peut éventuellement entraîner des perturbations climatiques, car les nuages interagissent directement avec les radiations solaires et terrestres. Même si le vieillissement des OA peut conduire à de tels phénomènes macroscopiques, on s'attend à ce que des changements se produisent au niveau moléculaire avec le vieillissement, entraînant l'obtention de nouvelles molécules et la

consommation d'autres. L'utilisation de la spectrométrie de masse à haute résolution permet de suivre les changements dans la composition moléculaire avec la détermination des formules moléculaires. Néanmoins, la spectrométrie de masse seule ne fournit pas d'informations sur les changements structurels qui se produisent au cours du processus de vieillissement ni sur la diversité isomérique de l'échantillon; c'est pourquoi nous avons décidé d'utiliser la spectrométrie de mobilité ionique à ondes progressives (TWIMS).

Pour cette étude, les échantillons ont été prélevés à l'université de Rostock lors de la campagne de mesure ULTRHAS (*ultrafine particles from transportation - health assessment of sources*). Au cours de cette campagne, nous avons utilisé le moteur expérimental et le dispositif d'échantillonnage mentionnés ci-dessus, ainsi que du MGO et du HFO comme carburant. En outre, un réacteur PEAR (*Photochemical emission aging flowtube reactor*) a été utilisé en ligne pour simuler le vieillissement photochimique des gaz d'échappement du moteur. Les émissions fraîches et vieillies ont été collectées sur des filtres en fibre de quartz précuits qui ont ensuite été coupés en deux. Chaque moitié de filtre a été soumise à une étape d'extraction utilisant soit un solvant aqueux, soit un mélange de solvants organiques.

Dans le chapitre cinq, nous présentons les premiers résultats obtenus avec les extraits organiques et aqueux des échantillons de particules collectés lors de la combustion de HFO dans le moteur. Nous avons utilisé la source d'ionisation chimique à pression atmosphérique (APCI) en mode positif pour étudier les extraits organiques et la source d'ionisation par électrospray (ESI) en modes positif et négatif pour étudier les extraits aqueux.

Les résultats préliminaires de cette étude en cours ont mis en évidence des différences structurelles dues au vieillissement atmosphérique des émissions des navires, en particulier dans les extraits aqueux analysés avec la source d'ionisation positive ESI. L'utilisation d'autres sources d'ions, telles que la photo ionisation à pression atmosphérique, pourrait permettre de recueillir des informations sur d'autres classes chimiques de composés, tels que les aromatiques. En outre, une analyse plus poussée sur un instrument IMS plus résolutif pourrait être effectuée pour confirmer que les changements dans les spectres de mobilité ionique sont la preuve de changements importants dans la structure. Ces instruments pourraient être l'IMS piégée (TIMS) ou l'IMS cyclique (cIMS). La modélisation moléculaire pourrait être utilisée pour

proposer des structures putatives afin d'obtenir des informations sur les structures affectées par le processus de vieillissement.

Conclusions et perspectives

Les principaux objectifs de ce travail de doctorat étaient le développement d'un logiciel visant à faciliter la visualisation des données UHRMS, la caractérisation des carburants et des émissions des navires par l'UHRMS, et l'étude structurelle des émissions des navires âgés par l'IMS-MS.

Au cours de ce doctorat, nous avons pu concevoir et publier notre logiciel de visualisation de données UHRMS appelé "Python tools for complex matrices molecular characterization - Viewer edition" ou PyC2MC Viewer. Ce logiciel peut traiter des données extraites de logiciels constructeur ainsi que des données personnalisées par l'utilisateur pour créer des graphiques de visualisation de données couramment utilisés tels que les cartes DBE en fonction du nombre de carbones, les diagrammes de Van Krevelen ou les diagrammes de défauts de masse de Kendrick, entre autres. Il est également doté de fonctions de comparaison d'échantillons de base et de fonctions d'analyse statistique. PyC2MC Viewer a été conçu pour être un outil facile d'utilisation, comme en témoignent son interface graphique simple et intuitive et son protocole d'installation simple. Néanmoins, notre logiciel a également été créé pour encourager les modifications et les améliorations par ses utilisateurs. C'est pourquoi il a été publié en tant que code brut open-source sur GitHub. A ce jour, PyC2MC compte plus de 150 téléchargements de sa version exécutable et a fait l'objet de plusieurs améliorations depuis sa première publication.

En ce qui concerne la caractérisation des pétroporphyrines de vanadyle et de nickel dans les carburants des navires et leurs émissions correspondantes par UHRMS, nous avons observé des pétroporphyrines dans les gaz d'échappement d'un moteur de navire réaliste, ce qui confirme leur rejet dans l'atmosphère. Nous avons également pu conclure que la combustion peu efficace à faible charge du moteur entraînait la libération d'une plus grande

quantité de pétroporphyrines. Comme cette faible charge du moteur est principalement utilisée lorsque les navires voyagent près des côtes ou lorsqu'ils sont à quai dans les ports, cela signifie que les pétroporphyrines sont principalement libérées près des zones peuplées, ce qui est problématique étant donné que leur toxicité est encore largement inconnue.

De plus, dans le dernier chapitre de cette thèse, nous avons présenté les premiers résultats de notre étude structurelle des émissions de navires par IMS-MS. Ces résultats montrent la capacité de la spectrométrie de mobilité ionique à résoudre la diversité isomérique même dans un mélange complexe. Ces résultats montrent les différences entre les particules fraîches et les particules vieillies par photochimie obtenues à partir de la combustion de fioul lourd dans un moteur de navire réaliste. Ces différences consistent principalement en un déplacement du nombre moyen d'oxygène vers des valeurs plus élevées dans les échantillons vieillis, ce qui correspond à l'oxydation qui a lieu pendant le vieillissement artificiel. En outre, nous avons également observé une distribution bimodale du temps de dérive pour l'un des ions étudiés, ce qui laisse supposer l'existence de deux populations isomères. Ces deux populations d'isomères semblent également être affectées différemment par le processus de vieillissement car le rapport d'intensité entre les deux populations est significativement différent avant et après le vieillissement artificiel.

Enfin, ce travail s'achève sur deux perspectives principales. Tout d'abord, PyC2MC Viewer sera continuellement développé et amélioré, comme prévu. Il fusionnera également avec son autre partie: PyC2MC, qui se concentre sur le traitement des données brutes, le traitement des signaux FTMS, l'attribution de formules moléculaires et le traitement des données des techniques MS hybrides (LC-MS, GC-MS, IMS-MS). Ce projet devrait aboutir à un logiciel libre capable de réaliser efficacement chaque étape du traitement et de la visualisation des données FTMS. Ensuite, l'étude sur l'aérosol vieilli devrait être poursuivie comme mentionné à la fin du cinquième chapitre. En effet, d'autres expériences, telles que la MS/MS, pourraient aider à comprendre la différence de consommation observée dans les expériences sur l'extrait aqueux avec ESI(+). De plus, il serait également intéressant d'utiliser d'autres sources d'ions, comme la photo ionisation à pression atmosphérique (APPI), afin de bénéficier de la complémentarité des sources. Il serait également possible d'utiliser la prédiction de la section efficace de collision pour déterminer les structures les plus plausibles des ions étudiés.

Références

1. Kulkarni, P., P.A. Baron, and K. Willeke, eds. *Aerosol Measurement - Principles, Techniques, and Applications, Third Edition*. 2011.

2. Lewis, E.R. and S.E. Schwartz, *Sea Salt Aerosol Production: Mechanisms, Methods, Measurements and Models—A Critical Review*. Geophysical Monograph Series. 2004: American Geophysical Union.

3. Fuzzi, S., et al., *Particulate matter, air quality and climate: lessons learned and future needs.* Atmos. Chem. Phys., 2015. **15**(14): p. 8217-8299.

4. Seinfeld, J.H. and S.N. Pandis, *Atmospheric chemistry and physics: from air pollution to climate change*. 2016: John Wiley & Sons.

5. Williams, J., et al., *Application of the variability-size relationship to atmospheric aerosol studies: estimating aerosol lifetimes and ages.* Atmospheric Chemistry and Physics, 2002. **2**(2): p. 133-145.

6. Trivitayanurak, W. and P.J. Adams, *Does the POA–SOA split matter for global CCN formation?* Atmospheric Chemistry and Physics, 2014. **14**(2): p. 995-1010.

7. Pope, C.A., et al., *Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution.* Circulation, 2004. **109**(1): p. 71-77.

8. Gauderman, W.J., et al., *The Effect of Air Pollution on Lung Development from 10 to 18 Years of Age.* New England Journal of Medicine, 2004. **351**(11): p. 1057-1067.

9. Bernstein, J.A., et al., *Health effects of air pollution.* Journal of Allergy and Clinical Immunology, 2004. **114**(5): p. 1116-1123.

10. Myhre, G., et al., *Radiative forcing due to changes in ozone and methane caused by the transport sector.* Atmospheric Environment, 2011. **45**(2): p. 387-394.

11. Corbett, J.J., et al., *Mortality from Ship Emissions: A Global Assessment.* Environmental Science & Technology, 2007. **41**(24): p. 8512-8518.

12. Barregard, L., et al., *Impact on Population Health of Baltic Shipping Emissions.* International Journal of Environmental Research and Public Health, 2019. **16**(11): p. 1954.

13. Lamas, M. and C. Rodriguez, *Emissions from marine engines and NOx reduction methods.* Journal of maritime research, 2012. **9**(1): p. 77-81.

14. Lion, S., I. Vlaskos, and R. Taccani, *A review of emissions reduction technologies for low and medium speed marine Diesel engines and their potential for waste heat recovery.* Energy Conversion and Management, 2020. **207**: p. 112553.

15. Winnes, H. and E. Fridell, *Particle Emissions from Ships: Dependence on Fuel Type.* Journal of the Air & Waste Management Association, 2009. **59**(12): p. 1391-1398.

16. Mohd Noor, C.W., M.M. Noor, and R. Mamat, *Biodiesel as alternative fuel for marine diesel engine applications: A review.* Renewable and Sustainable Energy Reviews, 2018. **94**: p. 127-142.

17. Peng, W., et al., *Comprehensive analysis of the air quality impacts of switching a marine vessel from diesel fuel to natural gas.* Environmental Pollution, 2020. **266**: p. 115404.

Appendix 1 : Lists of publications, communications and trainings

Publications

-Selective characterization of petroporphyrins in shipping fuels and their corresponding emissions using electron-transfer matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry; **Sueur, M.**; Rüger, C. P.; Maillard, J. F.; Lavanant, H.; Zimmermann, R.; Afonso, C.; Fuel, 2023.

-PyC2MC: An Open-Source Software Solution for Visualization and Treatment of High-Resolution Mass Spectrometry Data, **Sueur, M.**; Maillard, J. F.; Lacroix-Andrivet, O.; Rüger, C. P.; Giusti, P.; Lavanant, H.; Afonso, C.; JASMS, 2023.

-Wildfire plume ageing in the Photochemical Large Aerosol Chamber (PHOTO-LAC), Czech, H.; Popovicheva, O.; Chernov, D. G.; Kozlov, A.; Schneider, E.; Shmargunov, V. P.; **Sueur, M.**; Rüger, C. P.; Afonso, C.; Uzhegov, V.; et al. Environ. Sci.: Processes Impacts, 2023.

Oral communications

-Rencontres du Club Jeune de la SFSM from 07/03/2022 to 11/03/2022 in Lyon, France: "*Selective characterization of petroporphyrins in shipping fuels and their corresponding emissions using electron-transfer matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry*" (20 min)

-EU FTICR MS short course from 11/07/2022 to 14/07/2022 in Lisboa, Portugal: "*PyC2MC: An Open-Source Software Solution for Visualization and Treatment of High-Resolution Mass Spectrometry Data*" (20 min)

-EU FTICR MS End User School from 12/12/2022 to 16/12/2022 in Lille, France: ''*PyC2MC: An Open-Source Software Solution for Visualization and Treatment of High-Resolution Mass Spectrometry Data*" (40 min)

-Rencontres du Club Jeune de la SFSM from 13/03/2023 to 17/03/2023 in Marseille, France: "*PyC2MC: An Open-Source Software Solution for Visualization and Treatment of High-Resolution Mass Spectrometry Data*" (20 min)

-JFSM 2023 from 05/09/2023 to 08/09/2023 in Marseille, France: "*FTICR MS and advanced data visualization for the targeted and untargeted analysis of environmentally impactful complex mixtures: ship emissions*" (20 min)

Posters

-EU FTICR MS advanced school, from 26/09/2021 to 30/09/2021 in Prague, Czech Republic: "Characterization of organic aerosol using DIP-APCI FTICR MS"

-Analytics, from 05/09/2022 to 08/09/2022 in Nantes, France: "*PyC2MC: An Open-Source Software Solution for Visualization and Treatment of High-Resolution Mass Spectrometry Data*"

-ASMS annual conference, from 04/06/2023 to 08/06/2023 in Houston, USA: "*PyC2MC: An Open-Source Software Solution for Visualization and Treatment of High-Resolution Mass Spectrometry Data*"

Trainings

Scientific (70 h):

-Journées du club jeune de la société française de spectrométrie de masse (2021) : 14 h

-Journées de la société française de spectrométrie de masse (2021) : 28 h

-Journées du club jeune de la société française de spectrométrie de masse (2022) : 14 h

-Journées du club jeune de la société française de spectrométrie de masse (2023) : 14 h

Transversal (65 h)

-Formation NEO CNRS : 1 h

-Research integrity in scientific professions: 16 h

-Formations pour mission d'enseignement: 28 h

-Rédiger et publier un article scientifique : 20 h

Appendix 2 : Full-text article: “An overview of the complementarity of ionization sources for the characterization of complex organic mixtures in Fourier transform mass spectrometry”

A review of the complementarity of the main ion sources for the characterization of complex organic mixtures using Fourier Transform Mass Spectrometry

Charlotte Mase[1,2,3,†], Maxime Sueur[1,2,†], Hélène Lavanant[1,2], Christopher Paul Rüger[2,4], Pierre Giusti[1,2,3], and Carlos Afonso[1,2*]

[1] Univ Rouen Normandie, INSA Rouen Normandie, CNRS, Normandie Univ, COBRA UMR 6014, INC3M FR 3038, F-76000 Rouen, France

[2] International Joint Laboratory – iC2MC: Complex Matrices Molecular Characterization, TRTG, BP 27, 76700 Harfleur, France

[3] TotalEnergies OneTech, Total Research & Technology Gonfreville, BP 27, 76700 Harfleur, France

[4] Joint Mass Spectrometry Centre / Chair of Analytical chemistry, University of Rostock, Dr.-Lorenz-Weg 1, 18059 Rostock, Germany

[†] These authors contributed equally

Corresponding author (*): Carlos Afonso, carlos.afonso@univ-rouen.fr

ORCID

Charlotte Mase: 0000-0001-8820-2595

Maxime Sueur: 0000-0002-5014-0218

Hélène Lavanant: 0000-0002-2963-0327

Christopher P. Rüger: 0000-0001-9634-9239

Pierre Giusti: 0000-0002-9569-3158

Carlos Afonso: 0000-0002-2406-5664

Abstract

Complex organic mixtures are found in many areas of research, such as energy, the environment, health, and even astrochemistry, to name but a few. Despite their inherent diversity, these molecular systems share a common characteristic: their highly complex chemical composition, which holds an extensive potential of information at the molecular level. While numerous analytical techniques can provide access to this information, this review article focuses on Fourier transform mass spectrometry (FTMS). This technique is characterized by unparalleled resolution and mass accuracy, allowing the unravelling of both isobaric and elemental complexity. We present an overview of the main ionization sources employed to characterize such complex organic mixtures, namely electrospray (ESI), atmospheric pressure photoionization (APPI), atmospheric pressure chemical ionization (APCI), atmospheric pressure laser ionization (APLI), and (matrix-assisted) laser desorption ionization ((MA)LDI). We, then, examine the complementarity of these ion sources when used in direct infusion in key research areas. Finally, we discuss the synergistic integration of ultrahigh-resolution mass spectrometry (UHRMS) and these same ionization sources with chromatographic techniques such as gas chromatography, liquid chromatography, and supercritical fluid chromatography.

Keywords: ionization sources, ultrahigh-resolution, complex organic mixtures, hyphenation

1. Introduction

Soils, petroleum, sediments, carbonaceous chondrites, soot, and tholins are examples of highly complex organic mixtures involving numerous chemical functionalities and wide coverage of different chemical spaces. In most cases, complex organic mixtures contain compounds within a relatively low mass range of a hundred to about a thousand Daltons [1-3]. The definition used here excludes natural and anthropogenic high-molecular weight molecules, such as proteins or polymers, and complex mixtures of peptides. Complex organic

mixtures are found in many research fields such as energy, environment, health, planetology, and cultural heritage (Figure 1).

Historically, fossil-derived mixtures, such as crude oil, coal and bitumen, have for decades challenged analytical instrumentation due to their tremendous molecular complexity at the isobaric and isomeric levels [4]. The molecular description of these highly complex mixtures is often referred to as petroleomics, which also includes the study of derivatives, such as refined fuels [5], bitumen [6], lubricants [7], or other petrochemicals. Later, the term metallopetroleomics was introduced to describe metals in fossils. Petroleum and its derivatives have been used for decades and have attracted the interest of the analytical chemistry community due to their challenging complex chemical composition, which has been the driving force behind the development of state-of-the-art analytical instrumentation developments. Nowadays, due to environmental concerns, other complex mixtures that are potential sources of energy and chemicals are being investigated, such as biomass and its derivatives such as bio-oils [8].

The study of carbonaceous aerosols has gained considerable interest in the context of environmental health and climate change. These complex mixtures, either originating from either biogenic or anthropogenic sources, can represent a significant toxicological burden. Another complex matrix is ambient air pollution in the form of either biogenic [9] or anthropogenic [10] organic aerosols, which may pose a toxicological and/or environmental threat. In addition, in the field of environmental sciences, the study of dissolved organic matter (DOM) is essential as it can, for example, act as a chelating agent for metals, affecting their solubility, transport, and toxicity [11]. Note that, the term ‘dissolved’ can be confusing as DOM cannot be considered a chemical solution. By definition, and in contrast to particulate organic matter (POM) [12], DOM refers to all matter and particles that are able to pass through a 0.45 µm filter [13], resulting in a wide range of complex species that are dissolved or in the solid state. Another field is the study of natural organic matter (NOM) [11], which resulted from the decay of plant and animal tissues and is often found in other compartments, such as the pedosphere [14].

In medical and pharmacological research, work in the 'omics' sciences such as metabolomics [15], lipidomics [16], and proteomics [17] are complementary and together are helping to decipher biological processes that underlie the development of diseases such as cancer [18]. However, biological samples can also be considered as complex organic mixtures. They are often subjected to sample preparation procedures that reduce their chemical complexity prior to analysis. We have chosen to include only lipidomics and metabolomics in this review, in accordance with the definition of complex organic mixture mentioned in the introduction.

A less common field, in which complex organic mixtures are studied is astrochemistry, using samples such as aerosols that mimic Titan's atmosphere [19] or meteorites [20]. These synthetic mixtures aim to reproduce extraterrestrial phenomena and exhibit chemical and isotopic compositions that are not naturally found on our planet. However, due to their nature, the number, and quantity of samples available can be dramatically low, especially for meteorites and other extraterrestrial materials.

Cultural heritage is another field where the number and quantity of available samples is low. This subject involves the analysis of any sample that has a historical or artistic significance. From paintings and how they react to their environment [21] to older samples such as oil lamps dating from between the 5^{th} century BC to the 4^{th} century AD [22] or embalming material from the Greco-Roman period [23]. The main aims of this research are to shed light on past events and to conserve works of art.

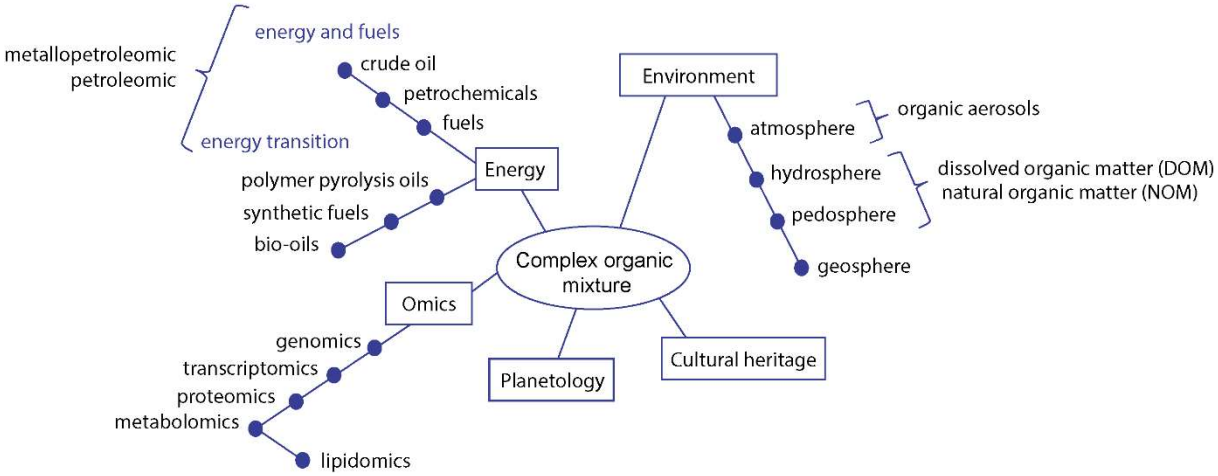


Figure 1. Research field involving complex organic mixtures.

Although the above examples exhibit a wide applicative and chemical diversity, they share a common characteristic: the richness of their chemical composition, either from isobaric, isomeric, or elemental complexity (Figure 2a). Components of isobaric complexity have the same nominal mass and therefore appear at the same *m/z* in mass spectrometry for single-charged species. In the example given in Figure 2a, more than ten peaks are resolved at the nominal mass *m/z* 263. On the other hand, elemental complexity refers to the diversity and number of hetero elements that may be present in the sample. This difference must be taken into account when characterizing complex mixtures, as the analytical methods used to study a sample with isobaric complexity [19] may be different from those used to study a sample with elemental complexity [7]. In any case, this diversity makes complex organic mixtures an important topic. Knowledge of the molecular composition is a prerequisite for understanding the fate and processes in chemical, biological and physical phenomena involving complex mixtures. Therefore, unravelling this vast amount of information at the molecular level is of great importance.

In the last decades, ultra-high resolution mass spectrometry (UHRMS), mainly achieved by Fourier transform mass spectrometry (FTMS), has been used as the main analytical tool for the untargeted characterization of complex organic mixtures. It includes Orbitrap and Fourier transform ion cyclotron resonance (FTICR) analyzers. Their principles have been widely

reported in the literature [24, 25]. Both offer high performance in terms of mass resolution, mass accuracy, and dynamic range compared to other MS analyzers such as time of flight (TOF) (Figure 2b) [3].

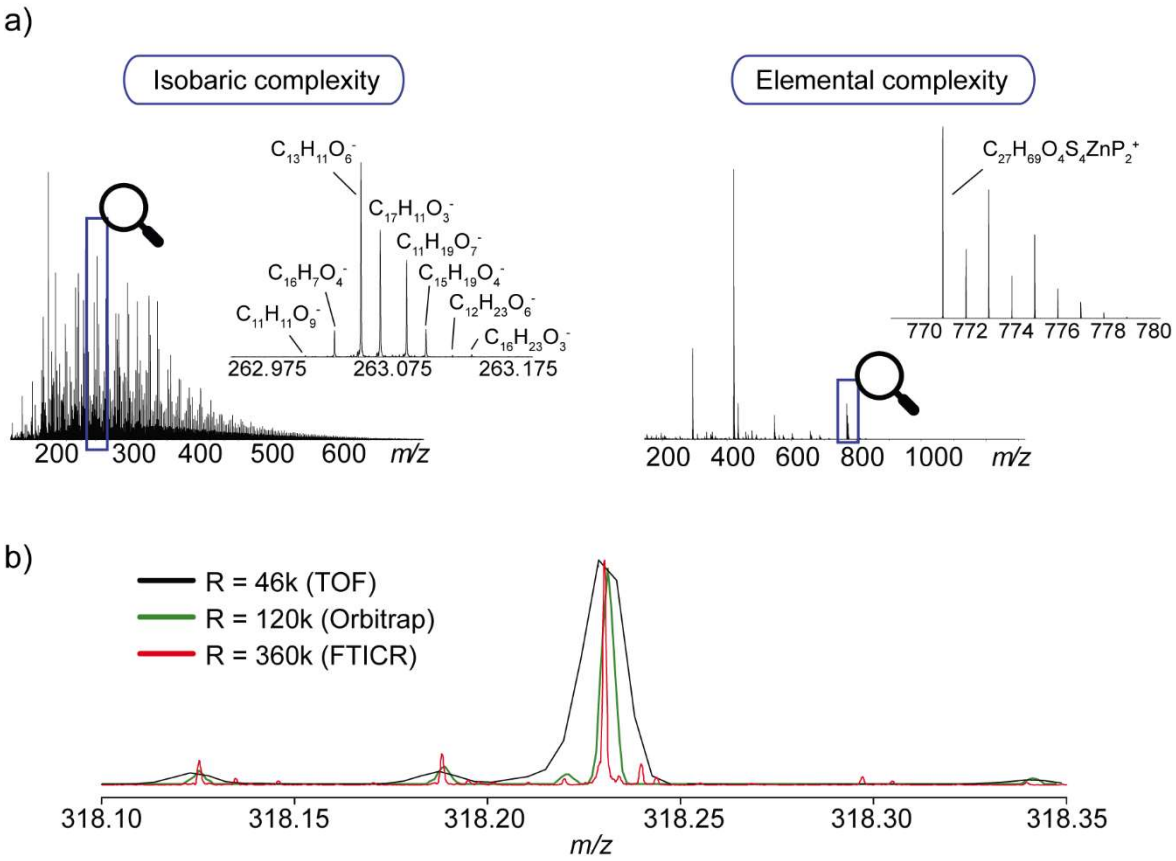


Figure 2. (a) Comparison between isobaric and elemental complexity found in complex organic mixtures. (b) Comparison of performance between the time of flight (TOF), Orbitrap, and FTICR with the zoom of a nominal mass justifying the use of FT-MS analyzer for the molecular characterization of complex organic mixtures. Note that results showed correspond to internal results.

The analytical challenge of mass spectrometry and not the least is the choice of ionization technique, since only molecules that are ionized are detected and no ionization

technique is exhaustive. There is a wide variety of ionization techniques; each ionization scheme has its own characteristics and thus, ionizes its own range family of chemical compounds. Therefore, complementary ionization sources can be used for a broader coverage of the chemical space and, thus, comprehensive analysis of complex organic mixtures. In this review, after an introduction on the classification of the ionization source, we will briefly describe the functionality of five main ionization techniques, namely electrospray (ESI), atmospheric pressure chemical ionization (APCI), atmospheric pressure photoionization (APPI), atmospheric pressure laser ionization (APLI), and laser desorption ionization (LDI), which correspond to the main ionization sources found for the analysis of complex organic mixtures. In a second part, the complementarity of these ionization sources is highlighted, with examples from the fields of energy, environment, omics and pharmaceuticals, planetology, and cultural heritage analysis. Finally, the complementarity of ionization sources used in hyphenated systems for isomer separation are described.

2. Classification of ionization sources

2.1. Introduction

In mass spectrometry, ionization is a key because only those molecules that are efficiently ionized with a particular ion source are detected [26]. There are many ionization techniques, and the physicochemical properties of the sample, such as state of matter, volatility, polarity, molecular mass, chemical, and thermal stability, determine their choice. The choice is also determined by the desired analytical approach: direct infusion or coupling with chromatography. The ionization technique used determines the ionized chemical space and therefore dramatically influences the obtained mass spectra. There is no technique that universally ionizes all the compounds in a complex organic mixture, simultaneously. Each ionization technique has shown its advantages for specific chemical structures or functionalities [27]. The selectivity of an ionization technique may be important when a particular group of compounds is targeted. However, when a broad overview of the existing compounds is needed, such selectivity can also be problematic. Therefore, it is very important

to understand the nature of the various ionization techniques and to be able to apply each one appropriately to study complex and heterogeneous organic mixtures. The combination of different principles and modifications, to improve sensitivity, selectivity, robustness, or other parameters has led to a whole new “zoo” [28, 29] of ionization sources.

Ionization sources can be classified into several categories [30]. The most common classification is based on their mode of operation, specifically whether the ionization takes place in a vacuum or at atmospheric pressure. Historically, the first ionization methods developed were vacuum ionization sources such as electron ionization (EI) and chemical ionization (CI) [31, 32]. However, these ion sources are limited to volatile molecules and EI is considered as a harsh ionization source that induces a lot of fragmentation. Such harsh ionization sources are suitable for the characterization of complex organic mixtures only when they are used after a chromatographic separation to associate molecular ions and fragments based on retention time alignment. LDI is generally performed under reduced pressure conditions and can produce few fragments for small molecules whereas it can produce large intact ions using an organic matrix with MALDI and is therefore consequently considered a soft source. However, atmospheric pressure ionization sources are practically all considered soft ionization methods, primarily due to the collisional cooling, and are therefore suitable for complex organic mixtures even in the absence of chromatographic separation. Ionization sources can also be classified according to the physical state of the samples they can analyze, such as solid samples or solutions. Sample introduction of solid samples can be done by evolved gas approaches such as atmospheric solid analysis probe (ASAP), or direct introduction probe (DIP) or laser desorption. With DIP, ionization can be performed by any preferred technique, such as chemical ionization (APCI) or photoionization (APPI). Samples in solution can be introduced directly into the ionization sources. Finally, they can be classified according to their specificities and limitations in terms of ionization efficiency as a function of polarity and mass. By viewing a sample through different "analytical glasses", it is possible to obtain a complete description of the sample. The graph showing the selectivity of ionization sources as a function of the polarity of the analyte and its molecular weight is widely used in the literature [33-35] and is shown in Figure 3. These two properties are the main parameters that determine the ionization source adapted to a sample. ESI, APPI, APCI, APLI, and LDI are the main ionization sources found in these graphs, mainly because they represent the most common ionization

sources for complex organic mixtures characterization. As can be seen in Figure 4, each of these ionization methods covers a different chemical space. In addition to ESI, which is the most widely used atmospheric pressure source, APCI, and APPI sources continue to be widely used because they provide access to a wider chemical space in terms of polarity and molecular weight. Typically, the APCI range proposed in the literature stops at weakly polar compounds (dotted line in Figure 3). However, as we will see in the next subsection that, under certain conditions, APCI can cover a wider range than APPI in terms of polarity and ionize non-polar compounds (dashed area in Figure 3).

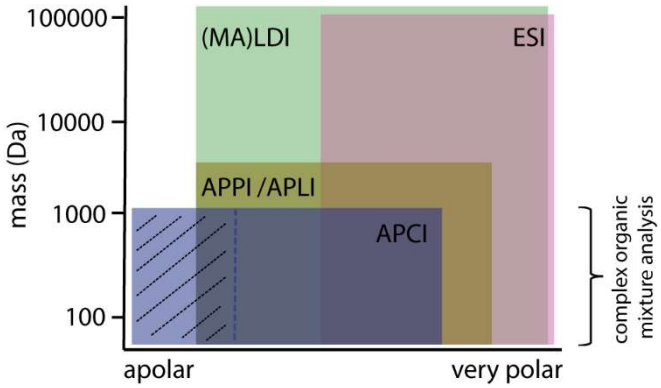


Figure 3. Selectivity of the four main ionization sources used for the chemical description of the complex organic mixture in the function of the polarity and the molecular weight of analytes. APCI: atmospheric pressure chemical ionization, APPI: atmospheric pressure photoionization, APLI: atmospheric pressure laser ionization, ESI: electrospray ionization, and (MA)LDI: (matrix-assisted) laser desorption ionization. The dotted line corresponds to the boundary between the area covered by conventional APCI and the area covered by APCI when hydrocarbons are used as solvents (indicated by the striped area).

2.2. Functional description of ionization sources

The main ionization sources used for the molecular characterization of complex organic mixtures are atmospheric pressure ionization sources such as ESI, APCI, APPI, and APLI as they are commercially available from all instrument vendors. The other main ionization source used for the molecular characterization of complex organic mixtures is the LDI source [36]. Considered as soft ionization sources, they differ in their mode of operation. The functional description of each has been briefly presented in this section including the advantage and the limitation of each for the analysis of complex organic mixtures.

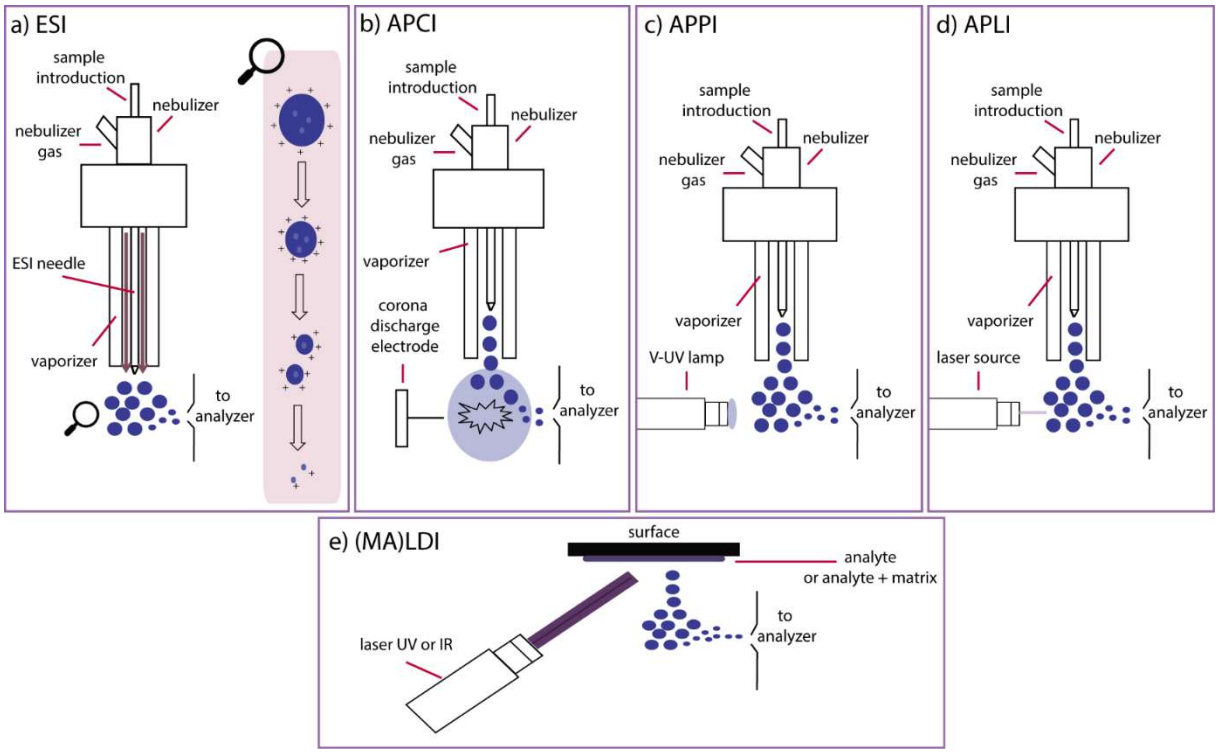


Figure 4. Schematic description of the four main atmospheric pressure ionization sources used for the molecular characterization of complex organic mixtures. (a) Electrospray ionization, ESI, (b) atmospheric pressure chemical ionization, APCI, (c) atmospheric pressure chemical ionization, APPI, (d) atmospheric pressure laser ionization, APLI, and (e) laser desorption ionization, LDI.

2.2.1. Electrospray ionization

ESI is an ionization source that is highly selective for polar species. The ionization step, which takes place in solution, can involve ionization competition and interference between analytes [37]. Ionization is governed by the solution phase properties of the molecules and in particular their acidity and basicity. The pH and the pKa are therefore important factors [38]. Negative mode ESI will favor the ionization of acidic molecules while positive mode will favor the ionization of basic molecules. The use of a modifier such as formic acid or ammonium hydroxide can also shift the equilibrium and thus promote ionization [38]. In addition, ESI has the advantage of avoiding analyte dissociation during ionization and allows the production of multi-charged ions allowing the analysis of high molecular weight compounds such as proteins. The ESI source also allows the ionization of small polar molecules. Ions with an even number of electrons such as protonated $[M+H]^+$ or multi-protonated $[M+nH]^{n+}$ species in the case of a positive mode ionization and deprotonated molecules $[M-H]^-$ in the case of a negative mode ionization are thus mainly produced. The addition of a cation or an anion can also be involved to produce adduct ions such as $[M+Na]^+$ or $[M+Cl]^-$. The formation of adducts can be promoted by the addition of salts such as sodium acetate or ammonium acetate. Multimeric forms such as $[2M+H]^+$ may also be observed.

In electrospray ionization, the analytes are introduced into the solution through a nebulizing needle (glass needle or steel capillary tube) at a low flow rate and then transferred to the gas phase according to a desolvation-ionization phenomenon (Figure 4a). In the first step, a spray of highly charged microdroplets is formed by breaking the Taylor cone at the end of the nebulization needle, under the action of an electric field and a nebulization gas (usually N_2). Then a hot drying gas (N_2), arriving in counter-current, assists the evaporation of the solvent from the microdroplets whose size decreases until reaching the Rayleigh limit where the coulombic repulsions are such that there is fission of microdroplets. Several evaporation-fission cycles are necessary to dissolve the ions from the last droplets produced (nanodroplets).

The ESI source, described as a soft ionization source, has a wide ionization range in terms of mass, allowing the ionization of intact small and large molecules. It remains

nevertheless dedicated only to the analysis of soluble molecules and more particularly polar to ionic compounds, which makes it a very selective source [39]. In the case of complex organic mixtures, the use of a selective source can be essential for the detection of polar compounds present in low proportions. The detection of specific compounds has proven to be essential for understanding reactivity. For example, the unique signatures of DOM in Antarctic Lakes were obtained using ESI in negative ion mode [40]. The identification of each DOM helps to explain reactivity in different aquatic processes such as disinfection by-product formation and photochemistry. ESI in both positive and negative ion modes have also allowed selective detection of basic and neutral nitrogen-containing compounds respectively in a vacuum gas oil, which represented less than 100 ppm of nitrogen. By comparison, refractory compounds were quickly identified in feed and effluent. This information can used to adjust processes [41]. In addition to its ionization selectivity, the ESI source has also become a necessary tool for the characterization of large biomolecules [42-44]. It has been shown that in this type of complex organic mixture, it is essential to select appropriate pH and solvents [45, 46].

2.2.2. Plasma ionization and direct introduction probes

2.2.2.1. APCI

The atmospheric pressure chemical ionization source (APCI) is a less selective ionization source. The ionization takes place in the gas phase and is governed by the properties of the gas phase such as ionization energy and proton affinity [47, 48]. It is described as a robust, reliable, and very sensitive source due to an efficient ionization less sensitive to matrix effects including ion suppression. Therefore, it can be considered for a wide range of applications. It is commonly used for the analysis of small polar and non-polar compounds with low ESI response. However, operation at high temperatures requires thermally stable compounds, or induces some thermal degradation of labile species.

Ionization is initiated by a corona discharge generating nitrogen plasma [49]. The analytes are introduced through a tube heated at high temperatures (400-550°C) for

volatilization (Figure 4b). The use of a nebulizing gas facilitates the formation of the aerosol at atmospheric pressure. A high potential applied to a needle located just after this tube creates a corona discharge that ionizes the nebulizing gas and then the solvent molecules and analytes in the aerosol via primary $N_2^{+\bullet}$ ions [48, 50]. Since the first step is the formation of a nitrogen plasma, the driving force for ionization is therefore the ionization energy of nitrogen which is 15.6 eV. This explains why APCI can in principle ionized a very wide range of molecules in terms of polarity. In the positive ionization mode, protonated molecules $[M+H]^+$ or radical ions $M^{+\bullet}$ can be formed. The predominance of one or the other process depends on the solvent used and the physicochemical properties of the analyte such as the proton affinity and ionization energy. In the presence of an aprotic solvent, the charge exchange process is favored, whereas conversely, the process of proton transfer is favored in the presence of a protic solvent or traces of moisture [51]. Furthermore, the APCI source has shown some promise in the analysis of saturated hydrocarbons when small hydrocarbons are used as solvents [52, 53].

Young Hwan *et al.* showed the evolution and abundance sensitivity of molecular classes in a petroleum sample by changing the solvent composition from toluene and methanol or acetonitrile to a single component system consisting only of toluene [54]. For example, they observed a greater increase in the relative abundance of molecular ions for sulfur-containing compounds. Similarly, Sanguineti *et al.* used APCI with heptane as a solvent to characterize bio-oils from microalgae [55]. They highlighted the presence of diverse classes of bio-oil components such as sulfur-, phosphorus-, and chlorine-containing compounds, which are unfavorable compounds for refining processes, in addition to the expected nitrogen-containing compounds and hydrocarbon compounds. Hourani *et al.* used APCI with isooctane to detect polycyclic aromatic sulfur heterocycles in vacuum gas oil [56]. In their study, they demonstrated the usefulness of the APCI source for characterizing of these species without any derivatization and with limited fragmentation.

Furthermore, direct introduction probes can be used with APCI for the analysis of solids, such as the atmospheric solids analysis probe (ASAP) and the direct insertion probe (DIP). The sample deposited on a glass capillary is desorbed thanks by a heated desolvatation gas and then ionized by the corona discharge. These sources allow for the direct analysis of solid and liquid samples with little or no sample preparation. They are simple and fast methods

that require no sample preparation and are suitable for liquid or solid samples if they are vaporizable. These ionization methods are therefore suitable for many samples and fields of applications. In their article, McEwen *et al. r*eported examples of applications where the ASAP is used [57]. These include drug analysis, tissue analysis, polymers and polymer additives, and petroleum. For their part, DIP-APCI FT-MS has been used in particular for the characterization of additives [58], and copolymers [59].

2.2.3. Photoionization

2.2.3.1 Atmospheric pressure photoionization (APPI)

As APCI, APPI is a less selective ionization source compared to ESI. APPI allows the ionization of a wide range of compound classes including non-polar compounds that are difficult to detect by ESI [60]. It has a low susceptibility to ion suppression and matrix effects, wide linear dynamic range, and very low chemical noise since most impurities and solvent molecules are not ionized [61]. As with the APCI source, the ionization occurs in the gas phase. Radical cations $M^{+\bullet}$ and to a lesser extent protonated/deprotonated species are formed by the absorption of V-UV photons and a subsequent complex reaction scheme (Figure 4c) [62],[63]. V-UV photons can be generated by using plasma discharge lamps filled with inert gas, such as krypton (emitting at 10.6/10.0 eV), argon (11.7 eV), or xenon (9.6/8.4 eV) [64-68]. The mechanism of photoionization with the APPI source also implies that the molecules are in the gas phase through an initial step of sample nebulization and evaporation. In the most simplistic gas phase reaction, the molecule, presenting a UV chromophore, will absorb a photon and form a radical ion with an odd number of electrons. Furthermore, detailed work, which would exceed the scope of this review, is presented by Kauppila *et al*. In practice, this direct APPI ionization is often not very efficient [69].

To improve the ionization efficiency in APPI, a dopant can be added to the sample [47, 70]. Especially in the presence of protic solvents protonated molecules $[M+H]^+$ can be produced [71]. The dopant can also act as an intermediate in the formation of protonated

molecules. Bruins *et al.* showed that the use of anisole as a dopant allowed for a 100-fold increase in the analytical sensitivity of analytes with low proton affinities in acetonitrile [72]. The same observation was reported by Kauppila *et al.,* who showed that the ionization efficiency was 1-2 orders of magnitude higher with dopant than without in the study of seven polycyclic aromatic hydrocarbons (PAHs) [73]. The same group also proved that the charge exchange is favored for solvents with low proton affinity and that proton transfer is enhanced with the addition of protic solvent. This approach can be advantageous for the identification of specific compounds in various organic mixtures. It should be pointed out that APPI without dopant is often not very efficient and when it is used with toluene (IE 8.8 eV) as a dopant it favors the ionization of molecules with low ionization energy such as aromatic molecules.

Direct introduction of solids can also be used with atmospheric pressure photoionization. In fact, Castilla *et al.* used DIP-APPI FT-MS to characterize lignocellulosic biomass samples [74]. It was compared to DIP-APCI. Using the same direct introduction probe for ground wood samples, the authors proved that APPI was more specific for lignin degradation compounds, while APCI covered a greater variety of oxygenated compounds. Podgorski *et al.* also used a custom APPI source, called desorption APPI, DAPPI), with a similar principle to DIP allowing for the in-source desorption of intact bio-char after combustion (at 250°C) and after pyrolysis (at 400°C) [75].

The simplicity, robustness, broad addressable chemical space, and reduced matrix effects make APPI to be very popular in many applications, especially for complex mixture analysis at high-resolution mass spectrometry platforms. In the field of complex energy transition matrices, both APPI and ESI sources are the main atmospheric pressure ionization concepts used. For example, APPI FT-MS has been used by Chiaberge *et al.* [76] for the untargeted characterization of crude oil samples allowing for the separation of 14 samples according to their geographical origin by using statistical analysis of the obtained FT-MS data. Also, Vetere *et al.* [77] used this ion source for the qualitative evaluation of sulfur-containing species, which are difficult to detect, in heavy crude oil samples after chromatographic separation. APPI is also widely used in the environmental field. For example, Headley *et al.* [78] used APPI FT-MS to evaluate the adsorption of components of the naphthenic acid fraction of oil sands by carbohydrate-based materials.

2.2.3.2. Atmospheric pressure laser ionization (APLI)

The mechanism of atmospheric pressure laser ionization (APLI) differs from conventional lamp-based APPI in the way that instead of a single-step V-UV photoionization, APLI is based on a stepwise two-photon ionization scheme (Figure 4d). The APLI mechanism is based on the classical resonantly enhanced multi-photon ionization performed under elevated (atmospheric) pressure conditions (REMPI). Due to its mechanism and similarity to APPI, APLI is particularly efficient for non-polar aromatic species such as polycyclic aromatic hydrocarbons (PAHs). However, due to the first step of REMPI, APLI targets and enhances species whose first excited states have long lifetimes [79]. Another specificity of the APLI is the need for a high photon density for the second absorption. This requires the use of a laser, because a high photon density also allows for a very efficient ionization [80]. Furthermore, the UV wavelength range used in APLI allows for a better ionization efficiency and selectivity than in APPI; in fact, in this range, most solvents and mobile phase molecules are transparent (*e.g.,* H_2O, CH_3CN, and MeOH) [81]. Unlike APPI, APLI is less influenced by the use of a dopant [82].

APLI can be used for the detection of low concentration compounds in the gas phase as shown by Schmidt *et al*. [83]. APLI is often found in coupling with separation techniques or compared to other ion sources as the study by Panda *et al.* [84]. Here they use APLI for the characterization of sulfur-containing crude oil fraction and investigate the influence of the nebulization temperature on the resulting mass spectra. They also evidence the formation of mainly radical ions when using APLI. APLI is shown to efficiently ionize non-polar compounds containing 1 or 2 sulfur atoms. Furthermore, nitrogen-rich crude oil shows a wide range of ionized species, but with an emphasis on non-polar aromatic hydrocarbons. In addition to aromatic hydrocarbons, aromatic sulfur compounds were also ionized by ASPLI [85].

2.2.4. Laser desorption ionization (LDI)

LDI allows the production of ions from a solid or liquid sample. The sample, in solution or powder form, is deposited on a conductive metal plate. Once the solvent is evaporated (for liquid deposition), the plate is introduced into the source, generally under vacuum. By laser irradiation, the photons tear off the molecules from the plate surface in the form of aggregates which will lead, in the case of ionization in positive mode, to radical ions $M^{+\bullet}$, or protonated molecules $[M+H]^+$ (Figure 4e). Adducts such as $[M+Na]^+$ can also be observed. Ionization mechanisms have been described in detail by Knochenmuss [86]. Compounds that are ionized show very little decomposition. Therefore, like the ESI source, it is considered a soft ionization source and will work well with complex mixtures containing compounds with chromophores. LDI can be therefore successfully applied to the study of aromatic rich samples for a variety of complex organic mixtures [87-89].

In some cases, the use of a matrix is essential. In fact, there are situations where the laser is unable to directly desorb and ionize the intact compounds. This is called matrix-assisted laser desorption ionization (MALDI). MALDI is a process that involves the photo volatilization of a co-crystallized sample with a small molecule, called matrix, in excess proportion [90]. The matrix is, in most cases, a small organic molecule with a high absorption coefficient at the laser wavelength. The analyte is incorporated in low proportion into a solid or liquid organic mixture. Compared to LDI, it is the matrix that absorbs the energy of the photons emitted by the laser. The matrix would consequently reach in an excited state. The main process reported is the desorption of the ions formed by proton transfer between the photon excited of the matrix and the analyte. Transfers between two matrix molecules can be observed by the same proton transfer process [91]. Radical species also coexist through the charge transfer process.

MALDI allows the study of molecules over a wide range of mass (> 100,000 Da) and polarity. In addition, and compared to other techniques, it offers ease of use, high robustness, and high sensitivity. Its softness makes it a source of choice for the analysis of fragile species found in the biology/health research field such as neuropeptides [92], oligonucleotides [93], or amino acids [94]. Furthermore, specific matrices can also be used to selectively enhance the ionization of a specific compound family present at a trace level, *e.g.* vanadyl and nickel petroporphyrins in crude oil and its derivatives, as shown by Giraldo-Dávila *et al*. [95] and Sueur et al. [10]. In their respective studies, they showed that the use of an electron-transfer

(ET) matrix, instead of the commonly-used proton-transfer based matrix, greatly increased the selectivity of species with lower ionization energy (IE) than the ET matrix. Indeed, the ET matrix hinders the ionization of higher IE by being easily ionized by MALDI and transferring their charge to the low IE analyte by an exchange of electrons. On the other hand, Mase et al. compared 12 matrices including proton-transfer and electron-transfer matrices for the characterization of wood pyrolysis oils [96]. They showed that some matrices enabled the ionization of a large number of species such as graphene oxide or dithranol, while others were specific to certain classes such as ferulic acid (FA) or 1,5-naphtalenediamine (1,5-DAN).

3. Ion source complementarity

The range of desolvation/desorption and ionization mechanisms makes each ion source more or less selective, either in terms of mass range or polarity. This selectivity, which at first sight might be seen as a drawback, is in fact a strength, since the combination of data obtained with different ionization sources can be used to achieve more comprehensive characterization of complex organic mixtures. In this review, we have selected several studies and reviews in the research fields of energy, environment, health, planetology, and cultural heritage that show how different ionization sources are used to provide complementary information on complex organic mixtures. Note that this list of articles is non-exhaustive published mainly within the last two decades.

3.1. Field of energy

One of the key issues in the energy sector is crude oil and its derivatives. These samples have been extensively characterized by FT-MS using various ion sources: ESI [97], APPI [98], APCI [99], APLI [100], and LDI [101]. However, the use of several sources is often required to achieve the desired level of molecular information for both comprehensive and targeted descriptions of petroleum-like mixtures.

Gaspar *et al.* [102] investigated the complementarity of five AP ion sources: ESI, heat-assisted ESI (H-ESI), APPI, APCI, and APLI, for the untargeted characterization of the asphaltene fraction of crude oil using a 12 T LTQ-FTICR. The asphaltene fraction was obtained by performing a SARA fractionation [103]. This study shows that APCI, APPI, and APLI allow the ionization of more compounds than ESI and its heated variant (Figure 5). In fact, ESI does not detect the non-polar compounds that make up a large part of the asphaltene fraction, while nitrogen-containing species dominate because they are selectively ionized by ESI. Nevertheless, ESI and H-ESI are shown to be particularly efficient for nitrogen containing species. On the other hand, APPI, APCI and APLI show a similar distribution of compound classes, i.e. non-polar species without heteroatoms. These three sources allowed the detection of compounds with lower double bond equivalent (DBE) values and were also able to ionize a wider range of analytes, such as sulphur-containing species, which were completely absent in the ESI and H-ESI results. It is also noteworthy that APPI allowed the identification of the highest number of unique species.

With the aim of characterizing "supercomplex" crude oil mixtures, Panda *et al.* [104] performed the analysis of a vacuum gas oil (VGO) fraction by FTICR MS by using ESI, MALDI, APPI, APLI, and APCI as ion sources. Their study focuses on the ionization of polycyclic aromatic sulphur heterocycles (PASH) after a derivatization step that adds a methyl group to the PASHs. It is shown that only ESI and MALDI allow to observe the derivatized form of the PASHs. Furthermore, the compound distributions obtained by ESI and MALDI are very similar. Conversely, APPI, APLI, and APCI only give the nonderivatized forms, because of the high-temperature nebulizer, which allows for the thermal removal of the methyl group. The results for all ionization sources highlight the ability of APCI to ionize high molecular weight species not detected by the other four sources. However, both APPI and APLI were able to ionize a wider range of compounds.

In their study to characterize of aged bitumen samples, Lacroix-Andrivet *et al.* [6] used negative mode ESI to selectively ionize polar and acidic heteroatom-containing species, which include markers of aging. The results of this selective ionization were used to predict the aging state of bitumen samples. On the other hand, APPI was used in positive mode as a non-selective ion source to screen for low-polarity compounds that could be used to determine the

origin of the sample. Kondily *et al.* [105] evaluated the complementarity and combination of different AP ion sources, ESI, APPI, and APCI for the characterization of light crude oil. This study shows the complementarity of the three sources as demonstrated in previous examples. The authors also perform dual-source combinations to try to broaden the range of ionized molecules. Although some species can only be ionized by combining of two sources, these dual sources are more susceptible to matrix effects than the standard ion sources.

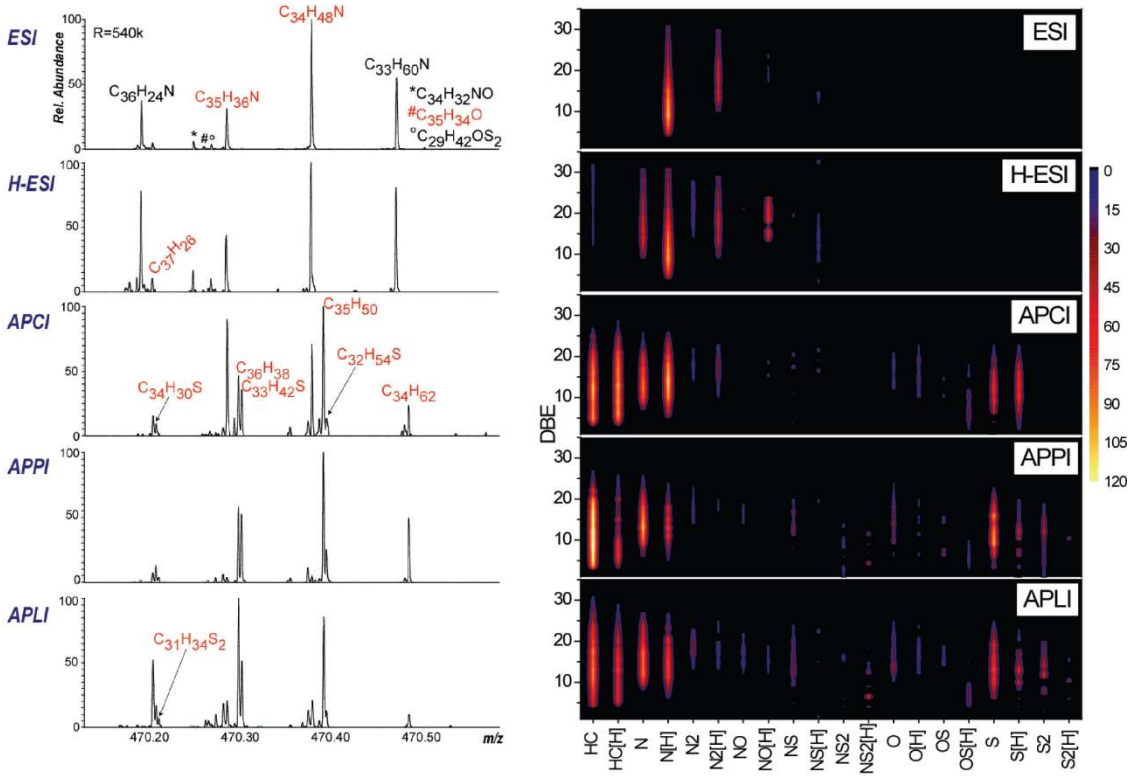


Figure 5. Enlargement at *m/z* 470 of mass spectra obtained with the different sources (left) and the DBE distribution of the individual compounds class (right) using ESI, H-ESI, APCI, APPI, and APLI for the characterization of asphaltenes [102].

Compared to petroleum-based materials and differences in molecular composition, the main source used for the characterization of bio-oils, whether from lignocellulosic biomass or algae, is ESI. In fact, ESI is particularly indicated for the analysis of polar to midpolar species, with a selective ionization of acidic and basic compounds in negative and positive ion modes, respectively. Since these types of species are common in bio-oils, ESI is adapted and therefore

more widely used. However, several authors have shown the complementarity of ESI with other ionization sources to obtain a complete molecular description.

Hertzog *et al.* have used three ionization sources coupled with FTICR to characterize a bio-oil obtained from fast pyrolysis of lignocellulosic biomass [8]. ESI and LDI were used in both positive and negative ion modes and APPI was used in positive ion mode only. As expected, the same molecular family $C_xH_yO_z$ was observed with each ionization source. However, major differences were however observed in the DBE *versus* carbon number plots showing the complementarity between these ionization sources. ESI allowed the ionization of compounds with higher oxygen atom content and lower DBE values whereas APPI allowed the ionization of less polar compounds with high DBE values and low oxygen atom content. LDI was considered to be intermediate in terms of DBE and the number of oxygen atoms between both ESI and APPI sources. However, a significant part of the features detected by LDI was specifically detected by this ionization method, providing important additional molecular information.

Mase *et al.* have also investigated the contribution of ionization sources in combination with FTICR for the characterization of bio-oils [106]. APCI was compared with ESI and APPI, in the positive ion mode. Different dopants in ESI and dilution solvents in APCI were also compared. The main observation was that the atmospheric pressure ionization sources differed significantly in their ionization selectivity. The APCI source proved to be particularly attractive because it ionized aliphatic molecules in addition to the other bio-oil components also ionized by APPI and ESI.

3.2. Environment

Motivated by environmental and climatic changes, the in-depth analysis of environmental samples has gained more and more interest. Environmental complex organic mixtures encompass a wide variety of samples, from seawater to soil and ambient air and thus require specific analytical methods [107].

Among the complex organic mixtures in the environment, dissolved organic matter (DOM) plays an important role in the global carbon cycle and the resulting physical or chemical processes [40]. Kurek *et al*. attempted at deciphering the complexity of DOM using ESI (−) and APPI-FTICR MS [108]. They analyzed water samples from three sites in Florida: Suwanee River, Kissimmee River, and Wakulla Springs. Here, the dopant-dependent selectivity of APPI is demonstrated using two dopants, tetrahydrofuran (THF) and toluene. Indeed, THF is shown to be capable of ionizing not only N-containing species, similar to toluene, but also for the ionization of S-containing species. However, THF and toluene seem to be complementary as they both allow the ionization of different species. Multi-doped APPI and ESI (−) thus allow the detection of a wide range of compounds.

In addition to the biogenic DOM, anthropogenic DOM is of great research interest and can be found in wastewater from a city or a factory. As an example, He *et al.* [109] studied wastewater samples from a petroleum refinery using ESI (+/−) and APPI-FTICR MS. In this study, the use of three different ionization techniques enabled the identification of compounds specific to anthropogenic DOM when compared to biogenic DOM from the Suwannee River standard sample. When examining a specific molecular class (N_1, O_1 or HC) in Figure 6, each ionization source allowed the ionization of different compounds that provided a more in-depth description of the sample, demonstrating the complementarity of the three ionization sources. In fact, as shown in the Figure 6, ESI (−) favored the ionization of neutral nitrogen-containing species such as carbazoles or benzocarbazoles, which are easily deprotonated in solution, while ESI (+) favored the ionization of basic nitrogen-containing species such as those with quinoline or pyridine moieties. Regarding the oxygen-containing species, ESI (−) proved to be very efficient for the ionization of phenolic compounds on DBE 4 and fatty acids. On the other hand, APPI allowed the ionization of species with higher aromaticity in each compound class.

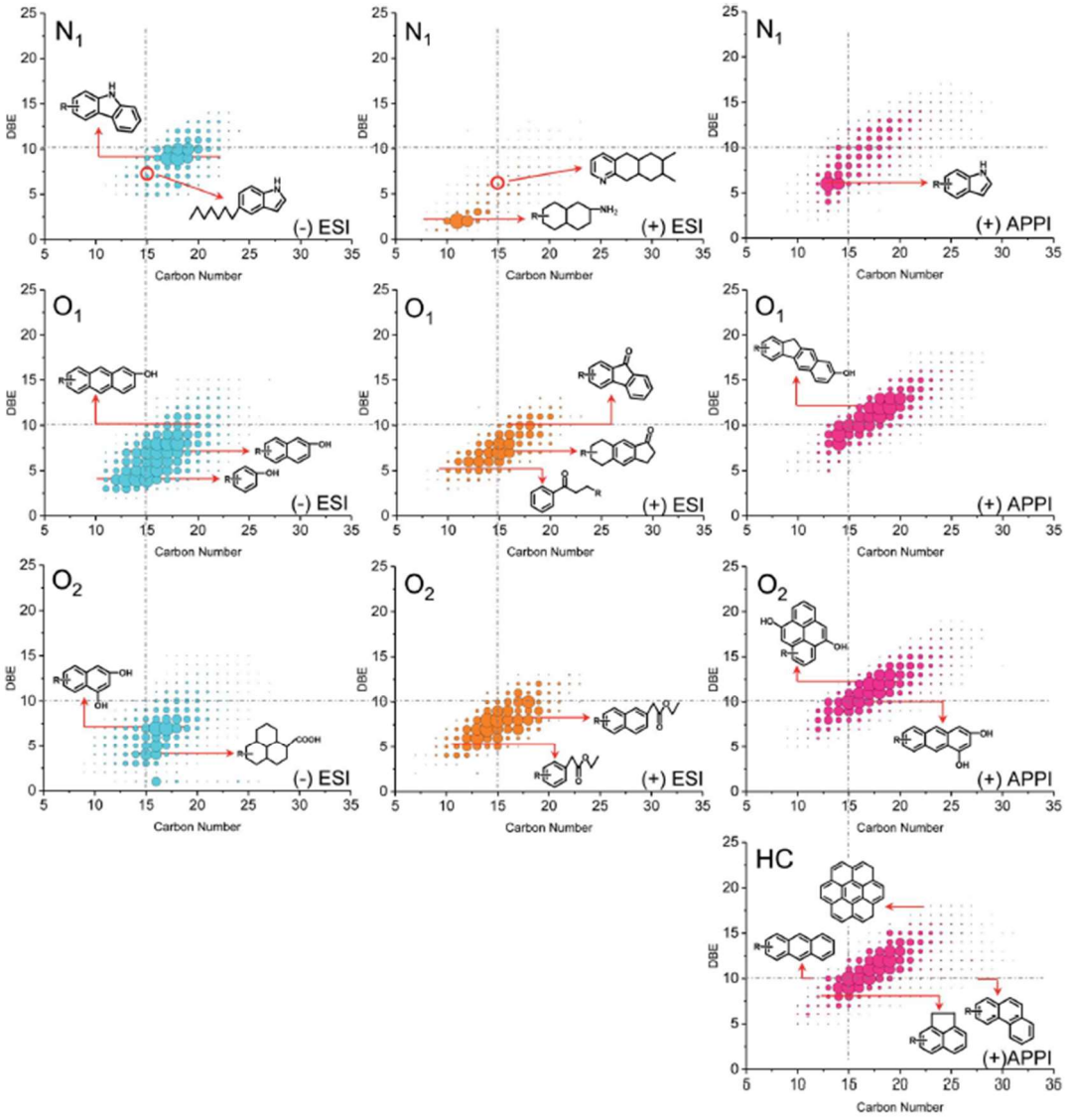


Figure 6. DBE versus carbon number plots of N_1, O_1, O_2, and CH class detected by ESI (+/−) and APPI (+) FTICR MS for a refinery wastewater DOM showing the complementarity of the ionization sources.*[109]*

Natural organic matter (NOM), or humus, is produced by entropy-driven decay of plant and animal tissues. Its complexity has been shown by Vinci *et al.* [110] in their study by using both ESI in negative mode and APPI in positive mode with an LTQ-Orbitrap. They analyzed 6 fractions of a soil sample, 3 organo-soluble and 3 water-soluble. They showed the selectivity of APPI towards the aromatic and/or nitrogen-containing species, while ESI (−) favored compounds with high oxygen/carbon ratio, which were not detected by APPI. The use of two ion sources allowed for a more detailed molecular characterization of the 6 fractions. In

addition, Luo *et al.* also used ESI (+/−) and APPI FT-MS to characterize contaminated soil samples [111].

Another category of environmentally relevant complex mixtures is organic aerosols (OA). OA can originate from either biogenic or anthropogenic sources and elevated concentrations of OA in inhabited areas can pose significant health risks [112]. In addition, their impact on the environment is not negligible and has been studied since the 1950s [113]. Traditionally, OA has been characterized by a combination of sophisticated extraction and separation techniques, and targeted analyses [114]. However, in the last decades, untargeted analyses of OA have emerged in order to identify as many compounds as possible [115]. This in an approach for which FT-MS is well suited. Kuang *et al.* characterized anthropogenic fine particles from Hong Kong ambient air by ESI and APPI-FTICR MS in positive ion mode [116]. APPI allowed the detection of highly aromatic PAH while ESI allowed the detection of more aliphatic species, especially in the case of nitro-PAHs. This study identified many more non-functionalized PAHs were identified than the usually targeted 16 high-priority PAHs and the 74 PAHs described in the cancer risk assessment. OA can also be divided into two fractions: water-soluble organic aerosol (WSOA) and water-insoluble organic aerosol (WIOA). Both fractions can be studied as shown by Choi *et al.* who compared WSAO and WIAO using ESI and APPI-FTICR MS operated in positive mode [117]. It is also possible to consider only one of the fractions, as done by Schneider *et al.* in their comparison of biogenic and anthropogenic OA using electrospray ionization in both negative and positive modes [9].

3.3. Omics and pharmaceutical analysis

Living organisms are governed by a plethora of biological complex processes that take place from the molecular to the cell scale. The omics sciences aim to understand and characterize these processes. Mass spectrometry is used in several fields of research: genomics, transcriptomics, metabolomics, lipidomics, and proteomics. In this context, FTMS and the use of several ion sources are of particular interest. However, each field has its own specificities in terms of sample preparation, MS analysis, and data treatment/interpretation.

Only lipidomics and metabolomics will be considered in this review according to the complex organic mixture definition given in the introduction.

Metabolomics is a recent science that studies all the metabolites in a cell, an organ, or an organism. Advanced analytical methods and especially mass spectrometry combined with statistical analysis have allowed the understanding of some metabolic mechanisms [118]. The chemical complexity of the mixture requires the use of several ion sources. Gray *et al.* used the complementarity of ESI and APCI sources coupled to an FTICR mass spectrometer to determine the cold acclimation mechanism in *Arabidopsis* by metabolic fingerprinting [119]. The data generated by both sources were aligned and processed to obtain a list of all unique masses. Then, statistical analysis using PCA was used to conclude the acclimation of the plant. Similarly, Ahoroni *et al*. used ESI and APCI in a complementarity manner for the non-targeted metabolomic analyses, focusing on strawberry [120]. By combining both sources using both polarities, they identified about 5500 different mass peaks. They nevertheless underlined the fact that mass spectrometry alone did not allow the identification of isomers and that therefore potentially a much higher number of species existed. Despite this drawback, the study has revealed the process of accumulation of compounds that serve as raw materials and building blocks for the production of ripening-associated metabolites. Still in the field of metabolomics, Calabrese et *al*. used MALDI with different matrices in addition to the classically used ESI for the analysis of plant root exudate metabolites [121]. Plant root exudates are complex mixtures with a high salt concentration which induces matrix effects in ESI, while MALDI is more salt tolerant. Here, the complementarity of the two sources was demonstrated, as ESI and MALDI showed selectivity towards small molecules with low aromaticity and highly aromatic molecules, respectively.

Lipidomic studies are interested in the specific role of lipid molecular species in health. Lipids consist of a large number of structural and functional molecular species covering a wide range of mass and polarity. Thanks to the development of soft ionization techniques such as ESI, APCI, and MALDI, mass spectrometry has become a widely used technique for the characterization of such matrices allowing a rapid and sensitive analysis. The review by Hu *et al.* provides a good overview of the recent advances in mass spectrometry in lipidomics [122]. The identification of all the lipids in a mixture is complicated in a single analysis, so the use of

different ionization sources is recommended. The ESI source in positive and negative ion mode is the most used ion source. It is mainly used for identification with tandem mass spectrometry MS/MS or coupled with liquid chromatography for quantification. In ESI, signal suppression can be observed and is a risk. For this reason, the MALDI source is often used as a complement because of its lower sensitivity to matrix effects. However, MALDI cannot be used in on-line coupling with liquid chromatography and has limitations for quantification. Nevertheless, MALDI has seen in great development because it allows an imaging approach [123].

In the field of pharmaceutical analysis, Deschamps *et al.* have reviewed the principles of the two mass spectrometers Orbitrap and FTICR, and highlight their application, development, and future perspectives [124]. They showed the main use of the ESI source in this field. However, due to the important matrix effect found in these samples, they insist on the quality of the sample preparation. Other ionization sources, such as MALDI, APPI, and APCI, are used in cases where sample preparation is not possible but are less frequently found. Still related to matrix effects, coupling with liquid chromatography is widely used for the characterization of these types of samples. In this case, it is also the ESI source that will be most used. Finally, there is an interest in mass imaging using LDI and MALDI sources, which allow a great understanding of the bio distribution, metabolism, and accumulation of drugs in the human body.

3.4. Planetology and astrochemistry

Mass spectrometers are regularly carried on spacecraft exploring the solar system to understand the origin, distribution, and evolution of organic matter. Innovative space instruments are therefore needed. In their study, Selliez *et al.* have compared two instruments, a home-built space-designed Orbitrap, and a laboratory Benchmark FTICR mass spectrometer [125]. The first one was composed of a laser ablation ionization (LAb) source and the second one with an LDI source. Except for the difference observed due to the dynamic range, resolution, and mass range of the analyzers, both ionization sources presented similarities and allowed the ionization of the same nitrogen-containing compounds.

Few comparisons of ionization sources have been performed on planetology samples, mainly due to the scarcity of samples and the small quantities available. In addition, the main works on planetology samples used LDI as the ionization source [19, 20]. Hertzog *et al.* have compared both APPI and ESI in positive and negative ion modes for the characterization of the Murchison meteorite [126]. Among the 16000 unique features detected, only 4 % were common to all analyses, clearly demonstrating the complementarity of the ionization sources. The main difference is in the classes of ionized molecules (Figure 7). ESI (−) allowed the ionization of oxygen and sulfur-containing molecules whereas ESI (+) allowed the ionization of nitrogen-containing molecules. APPI (+) allowed the ionization of less polar molecules such as hydrocarbons and heteroatomic compounds with lower O/C, N/C, and S/C ratios compared to ESI. The same observation was made by Naraoka *et al.*, who also compared ESI and APPI in both ionization modes for the characterization of the carbonaceous asteroid Ryugu [127]. ESI allowed the ionization of molecules with higher O/C ratios, while APPI showed a narrower distribution at lower O/C ratios when ionizing hydrocarbon compounds.

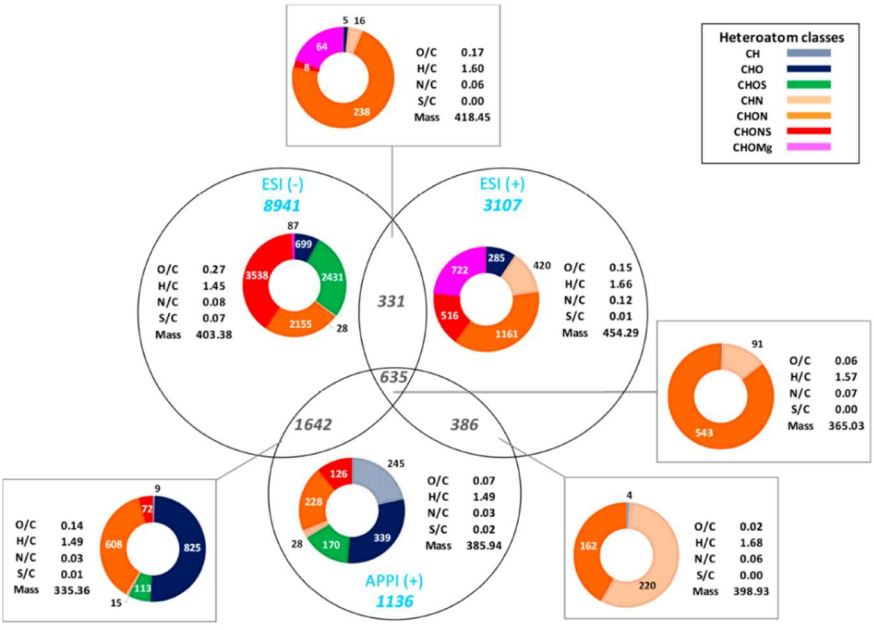


Figure 7. Venn diagram achieved from data obtained in ESI and APPI-FTICR MS in both positive and negative ionization modes. The heteroatom class distributions and the

corresponding weighted average values gathered in the tables are given for specific and common features [126].

3.5. Cultural heritage

Among the characterization techniques used in cultural heritage studies, ultra-high resolution mass spectrometry has shown particular interest, although it is considered a destructive technique and requires a sample collection, which is not always easy to obtain. Therefore, ionization techniques that requires little sample quantity and sample preparation are preferred such as MALDI.

Some work had focused on the complementarity of the ionization sources to determine which would provide better information. Vahur *et al.* have compared the MALDI and the APCI in positive ion mode for the analysis of dammar resin [128], which is a widespread varnish material found in paintings. They used 10 different solvents, because, with the years, the properties of the resin such as solubility changes, and it is therefore important to find the appropriate dilution solvent. They conclude that both MALDI and APCI sources are needed to get a complete picture of the dammar. MALDI allowed the ionization of the more polar components of the resin and APCI allowed the ionization of nonpolar components such as hydrocarbon and slightly oxidized hydrocarbon molecules. Similarly, Peets *et al.* have compared MALDI and ESI in both positive and negative ion modes for the characterization of red pigments from various plants, fungi, and insects [129]. Even though ESI allowed the ionization of the main components of interest and could be considered the final technique, MALDI proved to be interesting in the case of some red dyes by giving characteristic distributions.

Another interesting work compared the four main ionization sources found, APCI, APPI, ESI, and LDI in negative ion mode for the characterization of Scotch whisky [130]. Approximately 700 common molecular formulas were observed but each ionization source allowed the ionization of unique species. Among the four ionization sources, LDI was the most

distinctive in terms of the number of molecular formulas assigned. However, regarding the van Krevelen diagrams (Figure 8), LDI presented a distribution more concentrated with cask extractives, likely lignin-derivatives compounds whereas APCI and ESI allowed the ionization of highly oxygenated compounds such as carbohydrates and more aliphatic molecules such as fatty acids or alcohols. The authors have thus shown each ionization source should be chosen according to the desired information.

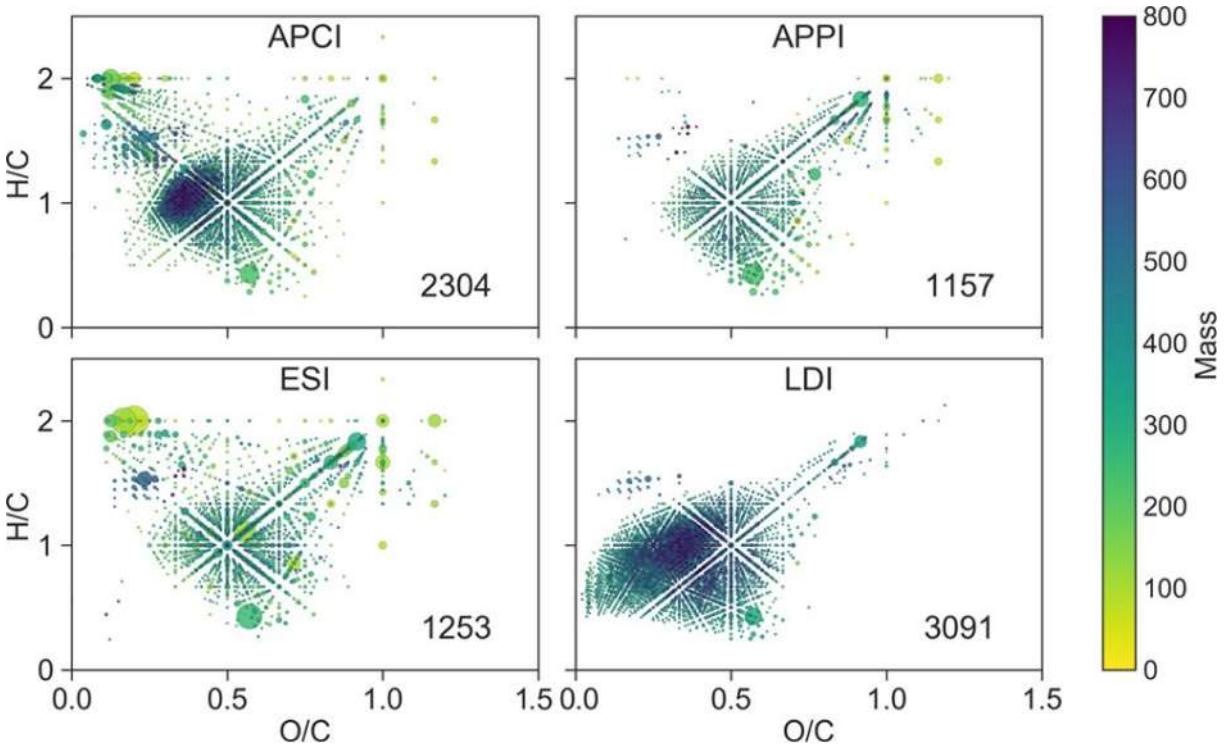


Figure 8. van Krevelen diagrams obtained with APCI, APPI, ESI, and LDI-FTICR-MS in negative ion mode showing the complementarity of the ionization sources for the characterization of a scotch whisky [130].

4. Ionization sources for hyphenated chromatographic methods

4.1. Introduction

Hyphenated chromatography and FT-MS were described as powerful tools for the characterization of complex organic mixtures allowing comprehensive molecular description at both isobaric and isomeric levels. FT-MS allowed the required high measurement accuracy to determine accurate molecular formulas and resolution for isobar distinction and the separation techniques highlight isomers and reduce the ion suppression effects extending the dynamic range. Various chromatographic methods have been coupled to FT-MS [131]. However, coupling to FT-MS involves some compromises and is not look as straightforward as hyphenation with other analyzers such as time of flight (TOF). Table 2 reports an overview of hyphenated chromatography (including liquid chromatography, supercritical fluid chromatography, and gas chromatography) and FT-MS techniques used for the characterization of complex organic mixtures. The ionization source, operating mode, and sample preparation used for each application were specified. We can see that the ionization sources used are not the same depending on the chromatography method used. The APCI source is widely used for gas chromatography, while the ESI source is preferred for liquid chromatography. This is due to the compatibility of the ionization source and the chromatography method, with the APCI source working in the gas phase and the ESI source working in the solution. However, we can also see that different ionization methods can be used with the same chromatographic system and allow giving complementarity results. The last part of this paper will be focused on this aspect.

4.2. Complementarity of ionization sources in hyphenated methods

4.2.1. Gas chromatography

APCI was the most retrieved ionization source interface between GC and FT-MS for the characterization of complex organic mixtures as observed in Table 2. APCI is a common ionization technique used for complex matrices in petroleum samples and pyrolysis oil.

Several studies have reported the complementarity of ionization sources with hyphenated gas chromatography and FT-MS techniques. In 2014, Ortiz *et al.* used gas

chromatography coupled with FTICR (9 T) for the characterization of environmental samples [132]. They used the EI and the CI with different gases. The authors showed that the observed distribution and composition were influenced by the ionization technique and the choice of reagent. EI led to a pronounced hydrocarbon contribution due to fragmentation, CI with ammonia led to more-prominent nitrogen-containing species, and CI with methane produced results between the other two methods. They demonstrated the complementarity between GC-FTICR MS with EI and CI sources and direct infusion FTICR MS with APPI source by comparison with results obtained by direct infusion with APPI source. Kondyli *et al*. compared the EI and the APPI sources for the characterization of gas condensate by GC-Orbitrap MS [133]. Both sources allowed the ionization of different compound classes such as hydrocarbons compounds and heteroatom-containing compounds (oxygen, nitrogen, and sulfur) but with different intensities (Figure 9). Especially, the heteroatom-containing species were well ionized with the APPI. The lower intensity of the sulfur-containing compounds in EI was explained by the fact these species were unstable and fragmented very easily. They showed that the high energy of EI generated a lot of fragmentation and made the complexity even higher and therefore the characterization more difficult. However, they highlighted the use of EI low energy, which allowed the reduction of the complexity and therefore the distinction between fragment and molecular ions. Finally, they demonstrated that APPI is a good complementary source of EI for the characterization of complex organic mixtures allowing the detection of a wide number of compounds covering a large polarity range.

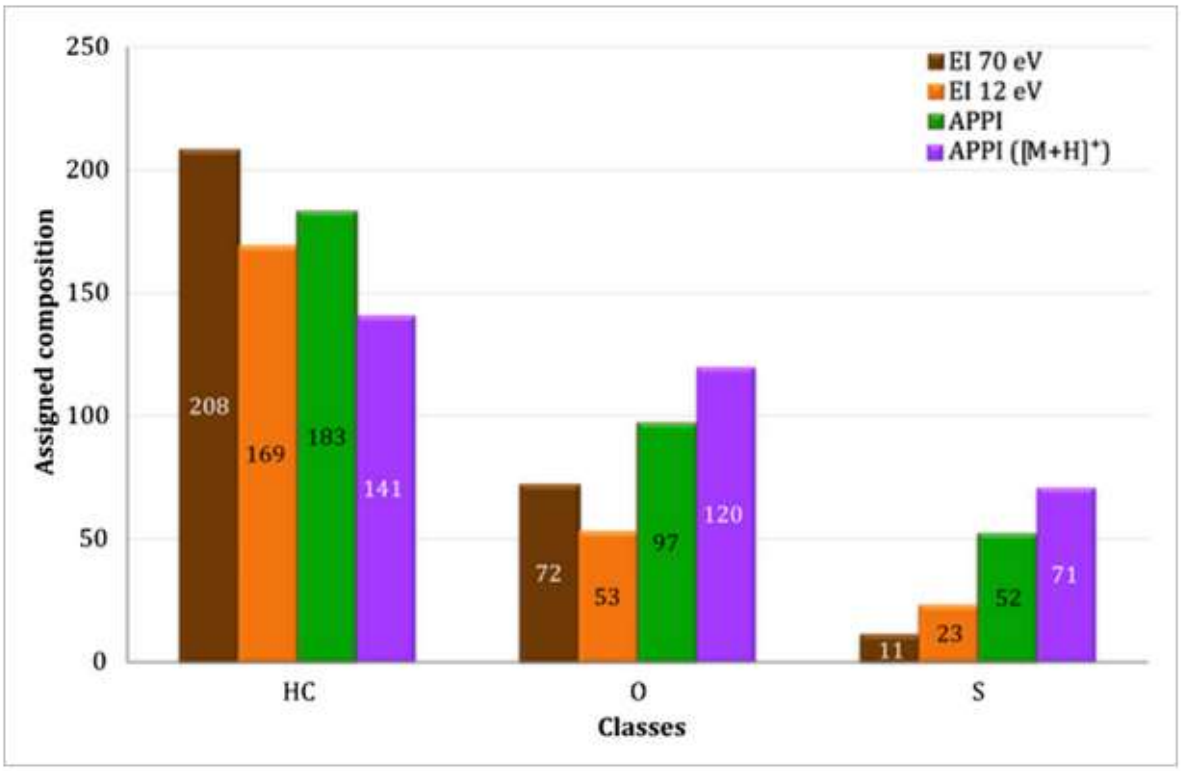


Figure 9. Comparison of class distribution between EI high and low energy and APPI for the characterization of gas condensate by GC-Orbitrap MS [133].

4.2.2. Liquid chromatography

As mentioned above, the coupling of FTMS with chromatography also allows to bypass the ionization competition and signal suppression that can occur with direct infusion. Reymond *et al*. compared the coupling LC-FTICR and the direct infusion (DI) FTICR using an ESI source in negative ion mode for the characterization of lignocellulosic-based bio-oil [134]. A bimodal distribution was obtained in LC-FTICR with a first distribution including oxygenated compounds between 3 and 8 and a second distribution with oxygenated compounds between 12 and 18 whereas only a single distribution was obtained in DI-FTICR. Liquid chromatography provided valuable additional information on heavy polar molecules. To obtain these good results in LC-FTICR, they shown that optimization of the desolvation and ionization steps is necessary to obtain a well-controlled analytical method. In their study, they also compared ESI

and APCI sources. They have shown that ESI in negative ion mode provided a detailed characterization of the bio-oil with the assignment of more than 5000 molecular formulas compared to the APCI. The same conclusion was reached by the same team using ESI and APCI sources in both positive and negative ion modes in the study of liquid-liquid extraction fractions from lignocellulosic biomass [135]. However, the combination of both ionization sources allowed to achieve a comprehensive characterization of the sample by detection of lipid, phenolic, and carbohydrate compounds (Figure 10).

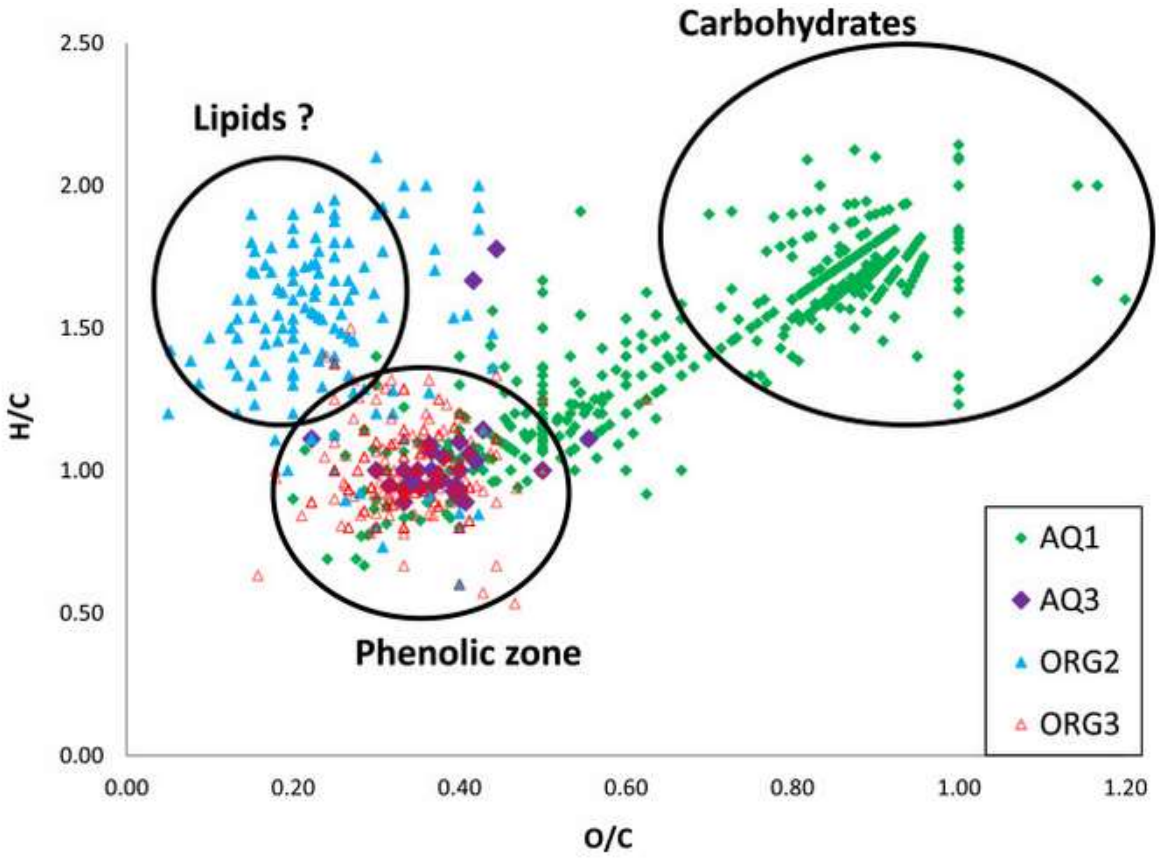


Figure 10. van Krevelen diagram obtained by ESI and APCI in both positive and negative ion modes in LC-FT-MS for the characterization of liquid-liquid fractions (AQ for aqueous and ORG for organic) of lignocellulosic-based biomass.[158]

To continue screening ionization sources for coupling with liquid chromatography, we found the work of Labadidi *et al*. comparing four ionization techniques (ESI, APPI, APCI, and APLI) for the characterization of deasphalted crude oil using high-performance liquid chromatography (HPLC) FTICR MS [136]. Using normal-phase liquid chromatography with n-hexane and isopropyl alcohol mobile phases, they separated the compound classes of the sample into two main peaks (Figure 11). The first peak included hydrocarbons and other non-polar compounds, while the second peak nitrogen-containing species. As expected, ESI allowed the ionization of only nitrogen-containing compounds. APPI offered the widest spectrum of classes allowing the identification of non-polar oxygen-containing compounds that were less observed with both APCI and APLI sources. The results obtained in APCI and APLI are quite similar. Here, the use of the four ionization sources allowed to highlight the selectivity of each one and to show that APPI was the best adapted to give the widest molecular information on this type of sample.

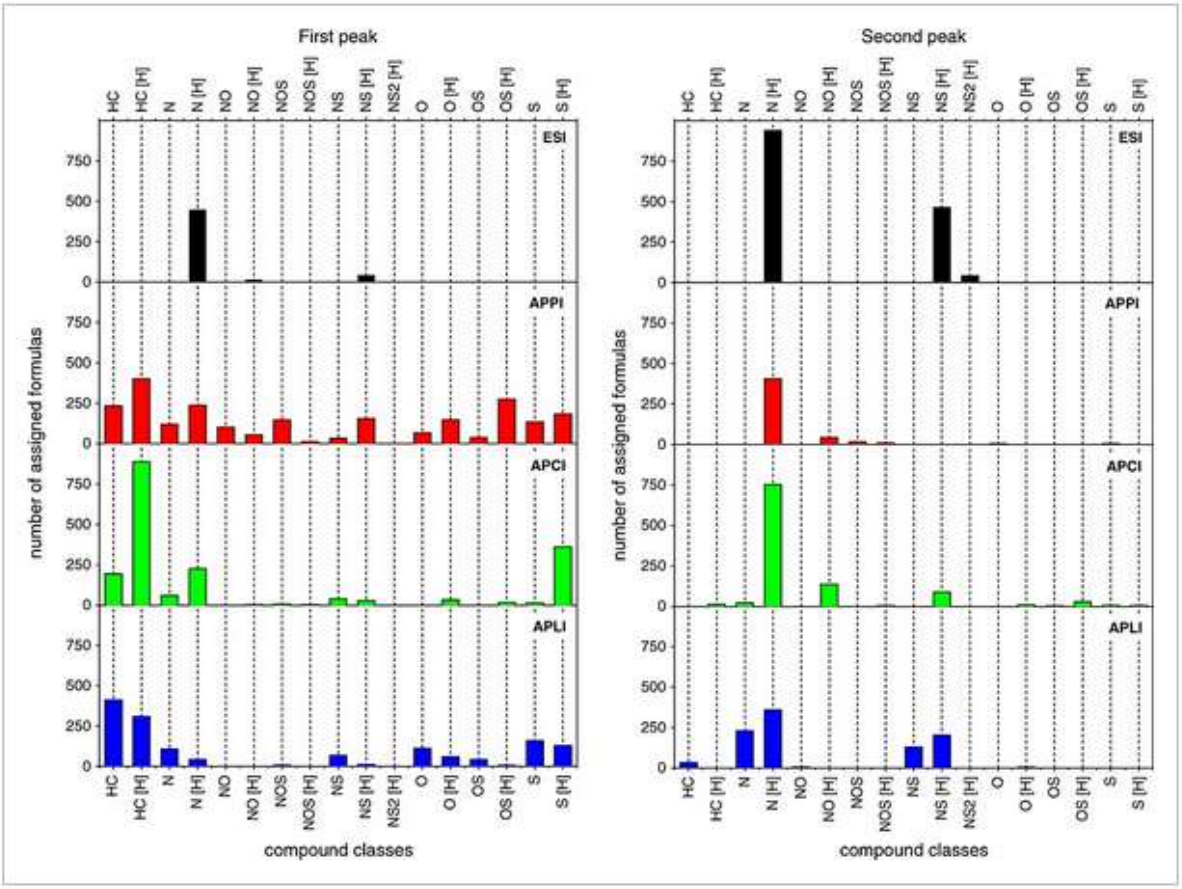


Figure 11. Class distribution based on the number of assigned formulas using ESI, APPI, APCI, and APLI sources obtained for the two peaks separated in LC-FTICR MS on a deasphalted crude oil [136].

Table 2. Overview of some hyphenated chromatographic and FT MS techniques. The ionization source, mode of use, and sample preparation used were specified for each.

Reference	Type of matrix	Hyphenated method	Ionization source	Mode	Sample preparation
Liquid chromatography					
[137]	Microbial and mammalian proteomic samples	UHPLC-Orbitrap MS	ESI	Positive	Cells preparation
[138]	Transgenic M. truncatula Hairy Roots	LC-FTICR MS	ESI	Negative	Extraction
[139]	Plant species	UPLC-FTICR MS	ESI	Negative	Extraction
[140,141]	Oil sands process-affected water from Syncrude Canada	LC-Orbitrap MS	ESI	Positive and negative	Filtration with 0.45 μm filter
					Acidification to pH = 2 and extracted by dichloromethane
					The residue dissolved in acetone
[16]	Complex biological sample	HPLC-Orbitrap MS	ESI	Positive and negative	MTBE-based lipid extraction
[142]	Pyrolysis bio-oils	LC-Orbitrap MS	ESI	Positive and negative	Dilution
[143]	Diesel, vacuum gas oil, and heavy oil	LC-Orbitrap MS	ESI	Positive and negative	Dilution

[144]	Oil sands from Northern Alberta	LC-Orbitrap MS	APCI	Positive and negative	Extraction with dichloromethane
[135]	Water-soluble fraction of biomass	LC-FTICR MS	ESI and APCI	Positive and negative	Liquid-liquid fractionation
					Filtration
[134]	Fast pyrolysis bio-oil from softwood	LC-FTICR MS	ESI and APCI	Positive and negative	Dilution
[145]	Nitrogen-rich crude oil	LC-FTICR MS	Home-built APLI	Positive	Dilution
[136]	Crude oil sample	HPLC-FTICR MS	ESI, APPI, APCI and APLI	Positive	Desaphalted
Supercritical fluid chromatography					
[146,147]	Vacuum gas oil	SFC-FTICR MS	APPI	Positive	Dilution in the mobile phase
[144]	Oil sands from Northern Alberta	SFC-Orbitrap MS	APCI	Positive and negative	Extraction with dichloromethane
[148]	Lipids in animal plasma	SFC-Orbitrap MS	ESI	Positive and negative	Extraction using Bligh and Dyer's method
Gas chromatography					
[149]	Gasoline samples	GC-FTICR MS	EI	Positive	No sample preparation
[150]	Gasoline samples	GC-FTICR MS	CI	Positive	Headspace

[132]	naphthenic acid fraction compounds of oil sands	GC-FTICR MS	EI and CI	Positive	Filtration under vacuum, acidification to pH = 4.5, andextractiond with Strata X-A solid phase extraction sorbent
					Methylation using BF_3/methanol
[151]	Marine sediments	GC-Orbitrap MS	APPI	Positive and negative	Soxhlet extraction
[133]	Gas condensed of Arabic origin	GC-Orbitrap MS	EI and APPI	Positive	Dilution
[152]	Oil sands from the Athabasca River basin	GC-FTICR MS	APCI	Positive	Filtration under vacuum, acidification to pH = 4.5, andextractiond with Strata X-A solid phase extraction sorbent
					Methylation using BF_3/methanol
[153]	Pyrolysis softwood bio-oil and its esterified products	GC-FTICR MS	APCI	Positive	Esterification with n-butanol in the presence of H_2SO_4

[154]	Sediment core	GC-FTICR MS	APCI	Positive	Extraction with dichloromethane
[50]	Particulate matter from diesel fuel and heavy fuel oil combustion	GC-FTICR MS	APCI	Positive	Extraction with methanol and dichloromethane
					Filtration with PTFE membrane filter
[155]	German brown coal pyrolysis	GC-FTICR MS	APCI	Negative	Extraction with benzene
[156]	Petroleum + lignocellulosic biomass	GC-FTICR MS	APCI	Positive	Thermogravimetry
[157]	Food sample	GC-FTICR MS	APCI	Positive	GPC then dilution in toluene
[158]	Diesel sample	GC-FTICR MS	APCI	Positive	Extraction with methanol
[85]	Crude oil sample	GC-FTICR MS	APSPLI	Positive	No sample preparation
[159]	Crude oil sample	GC-FTICR MS	APLI	Positive	No sample preparation
[160]	Gas oil	GC×GC-FTICR MS	APLI	Positive	No sample preparation

5. Conclusion

The interest of FTMS lies in its unsurpassed resolving power and mass accuracy. These features allow in-depth characterization of complex organic mixtures, helping us to better understand our environment, from biological processes at the cellular level to the origin of life by analyzing the organic content of meteorites. However, to make FTMS analysis as

comprehensive as possible, users should look forward to taking advantage of complementary ion sources. In particular, the use of a highly selective ion source such as ESI can be of great interest for the detection of low abundance polar species, but can lead to the loss of the big picture when non-polar molecules are the main components of the mixture.

Indeed, as shown in this review, the use of multiple ionization techniques significantly expands the chemical space accessible to the user in terms of polarity, aromaticity, and mass. This complementarity of sources can also be added to the separative power of techniques compatible with FTMS, such as LC, GC or SFC, to achieve a higher level of complexity in characterization, i.e. isomer separation. However, the more ion sources and hyphenations that are set up, the more complex the resulting data. This, and the increase in measurement time, are the main factors limiting the number of ion sources and separation techniques used for the characterization of complex organic mixtures. Therefore, efforts are constantly being made to shorten and simplify the processing of such complex data, and in some cases, the analysis can be automated to reduce experimental time. In general, it is not always feasible to use multiple ionization methods, but the choice of selective or non-selective ionization method is critical to obtaining the expected information.

Overall, FTMS users must make a compromise between time and depth of characterization. Even though FTMS is one of the techniques that allows one to dive the deepest into a sample.

Author contributions

Charlotte Mase: conceptualization, visualization, writing – original draft preparation. Maxime Sueur: conceptualization, visualization, writing – original draft preparation. Hélène Lavanant: supervision, conceptualization, writing – review and editing. Christopher Paul Rüger: supervision, conceptualization, writing – review and editing. Pierre Giusti: supervision, conceptualization, writing – review and editing. Carlos Afonso: supervision, conceptualization, writing – review and editing.

Acknowledgements

This work has been partially supported by University of Rouen Normandy, the European Regional Development Fund (ERDF, HN0001343), Labex SynOrg (Grant ANR-11-LABX-0029), Carnot Institute I2C, the Graduate School for Research XL-Chem (Grant ANR-18EURE-0020), the European Union's Horizon 2020 Research Infrastructures program (Grant Agreement 731077), région Normandie. Access to the CNRS research infrastructure Infranalytics (FR2054) is gratefully acknowledged. We thank the DFG (ZI 764/28-1) and ANR (ANR-20-CE92-0036) for funding the research project TIMSAC.

References

1. Wilkins, C.L., *Hyphenated techniques for analysis of complex organic mixtures.* Science, 1983. **222**(4621): p. 291-6.

2. Laskin, J., et al., *Chemical analysis of complex organic mixtures using reactive nanospray desorption electrospray ionization mass spectrometry.* Anal Chem, 2012. **84**(16): p. 7179-87.

3. Smith, D.F., et al., *21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures.* Anal Chem, 2018. **90**(3): p. 2041-2047.

4. Fernandez-Lima, F.A., et al., *Petroleum crude oil characterization by IMS-MS and FTICR MS.* Anal Chem, 2009. **81**(24): p. 9941-7.

5. Rüger, C.P., et al., *Comprehensive chemical comparison of fuel composition and aerosol particles emitted from a ship diesel engine by gas chromatography atmospheric pressure chemical ionisation ultra-high resolution mass spectrometry with improved data processing routines.* Eur J Mass Spectrom (Chichester), 2017. **23**(1): p. 28-39.

6. Lacroix-Andrivet, O., et al., *Molecular Characterization of Aged Bitumen with Selective and Nonselective Ionization Methods by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. 2. Statistical Approach on Multiple-Origin Samples.* Energy & Fuels, 2021. **35**(20): p. 16442-16451.

7. Lacroix-Andrivet, O., et al., *Molecular Characterization of Formulated Lubricants and Additive Packages Using Kendrick Mass Defect Determined by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* J Am Soc Mass Spectrom, 2022. **33**(7): p. 1194-1203.

8. Hertzog, J., et al., *Combination of electrospray ionization, atmospheric pressure photoionization and laser desorption ionization Fourier transform ion cyclotronic resonance mass spectrometry for the investigation of complex mixtures - Application to the petroleomic analysis of bio-oils.* Anal Chim Acta, 2017. **969**: p. 26-34.

9. Schneider, E., et al., *Molecular Characterization of Water-Soluble Aerosol Particle Extracts by Ultrahigh-Resolution Mass Spectrometry: Observation of Industrial Emissions and an Atmospherically Aged Wildfire Plume at Lake Baikal.* ACS Earth and Space Chemistry, 2022. **6**(4): p. 1095-1107.

10. Sueur, M., et al., *Selective characterization of petroporphyrins in shipping fuels and their corresponding emissions using electron-transfer matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry.* Fuel, 2023. **332**.

11. Nebbioso, A. and A. Piccolo, *Molecular characterization of dissolved organic matter (DOM): a critical review.* Anal Bioanal Chem, 2013. **405**(1): p. 109-24.

12. Menges, J., et al., *Particulate Organic Matter Mobilization and Transformation Along a Himalayan River Revealed by ESI-FT-ICR-MS.* Journal of Geophysical Research: Biogeosciences, 2022. **127**(12).

13. Thurman, E.M., *Organic geochemistry of natural waters.* 2012.

14. Bahureksa, W., et al., *Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation.* Environ Sci Technol, 2021. **55**(14): p. 9637-9656.

15. Stopka, S.A., et al., *Ambient Metabolic Profiling and Imaging of Biological Samples with Ultrahigh Molecular Resolution Using Laser Ablation Electrospray Ionization 21 Tesla FTICR Mass Spectrometry.* Anal Chem, 2019. **91**(8): p. 5028-5035.

16. Triebl, A., et al., *Lipidomics by ultrahigh performance liquid chromatography-high resolution mass spectrometry and its application to complex biological samples.* J Chromatogr B Analyt Technol Biomed Life Sci, 2017. **1053**: p. 72-80.

17. Han, X., A. Aslanian, and J.R. Yates, 3rd, *Mass spectrometry for proteomics.* Curr Opin Chem Biol, 2008. **12**(5): p. 483-90.

18. Darebna, P., et al., *Changes in the expression of N- and O-glycopeptides in patients with colorectal cancer and hepatocellular carcinoma quantified by full-MS scan FT-ICR and multiple reaction monitoring.* J Proteomics, 2017. **153**: p. 44-52.

19. Maillard, J., et al., *Comparison of soluble and insoluble organic matter in analogues of Titan's aerosols.* Earth and Planetary Science Letters, 2018. **495**: p. 185-191.

20. Danger, G., et al., *Unprecedented Molecular Diversity Revealed in Meteoritic Insoluble Organic Matter: The Paris Meteorite's Case.* The Planetary Science Journal, 2020. **1**(3).

21. van den Brink, O.F., et al., *Matrix-assisted laser desorption/ionization Fourier transform mass spectrometric analysis of oxygenated triglycerides and phosphatidylcholines in egg tempera paint dosimeters used for environmental monitoring of museum display conditions.* J Mass Spectrom, 2001. **36**(5): p. 479-92.

22. Tokarski, N.G.C.R.J.M.H.C., *Analysis of archaeological triacylglycerols by high resolution nanoESI, FT-ICR MS and IRMPD MS/MS: Application to 5th century BC–4th century AD oil lamps from Olbia (Ukraine).* International Journal of Mass Spectrometry, 2009. **284**(1-3): p. 47-56.

23. Hertzog, J., et al., *Unravelling the Egyptian embalming materials by a multi-method approach comprising high-resolution mass spectrometry.* Journal of Archaeological Science: Reports, 2023. **48**.

24. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis.* Anal Chem, 2000. **72**(6): p. 1156-62.

25. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: a primer.* Mass Spectrom Rev, 1998. **17**(1).

26. Awad, H., M.M. Khamis, and A. El-Aneed, *Mass Spectrometry, Review of the Basics: Ionization.* Applied Spectroscopy Reviews, 2014. **50**(2): p. 158-175.

27. Cho, Y., et al., *Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics.* Mass Spectrom Rev, 2015. **34**(2): p. 248-63.

28. Alberici, R.M., et al., *Ambient mass spectrometry: bringing MS into the "real world".* Anal Bioanal Chem, 2010. **398**(1): p. 265-94.

29. Liu, P., et al., *Recent advances of electrochemical mass spectrometry.* Analyst, 2013. **138**(19): p. 5519-39.

30. Gross, J.H., *Mass Spectrometry: a textbook.* Springer Cham, 2017.

31. Dempster, A.J., *A new Method of Positive Ray Analysis.* Physical Review, 1918. **11**(4): p. 316-325.

32. Munson, M.S. and F.-H. Field, *Chemical ionization mass spectrometry. I. General introduction.* Journal of the American Chemical Society, 1966. **88**(12): p. 2621-2630.

33. Wang, Y., et al., *Current state of the art of mass spectrometry-based metabolomics studies – a review focusing on wide coverage, high throughput and easy identification.* RSC Advances, 2015. **5**(96): p. 78728-78737.

34. Li, D.X., et al., *Gas chromatography coupled to atmospheric pressure ionization mass spectrometry (GC-API-MS): review.* Anal Chim Acta, 2015. **891**: p. 43-61.

35. Halket, J.M., et al., *Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS.* J Exp Bot, 2005. **56**(410): p. 219-43.

36. Fenner, N.C. and N.R. Daly, *Laser Used for Mass Analysis.* Review of Scientific Instruments, 1966. **37**(8): p. 1068-1070.

37. King, R., et al., *Mechanistic investigation of ionization suppression in electrospray ionization.* Journal of the American Society for Mass Spectrometry, 2000. **11**(11): p. 942-950.

38. Wilm, M., *Principles of electrospray ionization.* Mol Cell Proteomics, 2011. **10**(7): p. M111 009407.

39. Kruve, A., *Strategies for Drawing Quantitative Conclusions from Nontargeted Liquid Chromatography-High-Resolution Mass Spectrometry Analysis.* Anal Chem, 2020. **92**(7): p. 4691-4699.

40. Niu, X.Z., et al., *Characterisation of dissolved organic matter using Fourier-transform ion cyclotron resonance mass spectrometry: Type-specific unique signatures and implications for reactivity.* Sci Total Environ, 2018. **644**: p. 68-76.

41. Mikhaylova, P., et al., *Molecular analysis of nitrogen-containing compounds in vacuum gas oils hydrodenitrogenation by (ESI+/-)-FTICR-MS.* Fuel, 2022. **323**.

42. Mann, M. and M. Wilm, *Electrospray mass spectrometry for protein characterization.* Trends Biochem Sci, 1995. **20**(6): p. 219-24.

43. Smith, R.D., et al., *New developments in biochemical mass spectrometry: electrospray ionization.* Anal Chem, 1990. **62**(9): p. 882-99.

44. Wunschel, D.S., et al., *Electrospray ionization Fourier transform ion cyclotron resonance analysis of large polymerase chain reaction products.* J Am Soc Mass Spectrom, 2000. **11**(4): p. 333-7.

45. Ryan Bonfiglio, R.C.K., Timithy V olah, Kara Merkle *The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds.* Rapid Commun Mass Spectrom, 1999. **13**(12): p. 1175-1185.

46. Null, A.P., L.T. George, and D.C. Muddiman, *Evaluation of sample preparation techniques for mass measurements of PCR products using ESI-FT-ICR mass spectrometry.* J Am Soc Mass Spectrom, 2002. **13**(4): p. 338-44.

47. Marotta, E., et al., *Atmospheric pressure photoionization mechanisms.* International Journal of Mass Spectrometry, 2003. **228**(2-3): p. 841-849.

48. Andrade, F.J., et al., *Atmospheric pressure chemical ionization source. 1. Ionization of compounds in the gas phase.* Anal Chem, 2008. **80**(8): p. 2646-53.

49. Carroll, D.I., et al., *Atmospheric pressure ionization mass spectrometry. Corona discharge ion source for use in a liquid chromatograph-mass spectrometer-computer analytical system.* Analytical Chemistry, 1975. **47**(14): p. 2369-2373.

50. Schwemer, T., et al., *Gas Chromatography Coupled to Atmospheric Pressure Chemical Ionization FT-ICR Mass Spectrometry for Improvement of Data Reliability.* Anal Chem, 2015. **87**(24): p. 11957-61.

51. Vaikkinen, A., T.J. Kauppila, and R. Kostiainen, *Charge Exchange Reaction in Dopant-Assisted Atmospheric Pressure Chemical Ionization and Atmospheric Pressure Photoionization.* J Am Soc Mass Spectrom, 2016. **27**(8): p. 1291-300.

52. Moulian, R., et al., *Chemical Characterization Using Different Analytical Techniques to Understand Processes: The Case of the Paraffinic Base Oil Production Line.* Processes, 2020. **8**(11).

53. Manheim, J.M., et al., *Fragmentation of Saturated Hydrocarbons upon Atmospheric Pressure Chemical Ionization Is Caused by Proton-Transfer Reactions.* Anal Chem, 2020. **92**(13): p. 8883-8892.

54. Kim, Y.H. and S. Kim, *Improved abundance sensitivity of molecular ions in positive-ion APCI MS analysis of petroleum in toluene.* J Am Soc Mass Spectrom, 2010. **21**(3): p. 386-92.

55. Sanguineti, M.M., et al., *Analysis of impact of temperature and saltwater on Nannochloropsis salina bio-oil production by ultra high resolution APCI FT-ICR MS.* Algal Research, 2015. **9**: p. 227-235.

56. Hourani, N., et al., *Atmospheric pressure chemical ionization Fourier transform ion cyclotron resonance mass spectrometry for complex thiophenic mixture analysis.* Rapid Commun Mass Spectrom, 2013. **27**(21): p. 2432-8.

57. McEwen, C. and S. Gutteridge, *Analysis of the inhibition of the ergosterol pathway in fungi using the atmospheric solids analysis probe (ASAP) method.* J Am Soc Mass Spectrom, 2007. **18**(7): p. 1274-8.

58. Lattimer, R.P. and M.J. Polce, *Direct probe CI-MS and APCI-MS for direct materials analysis.* Journal of Analytical and Applied Pyrolysis, 2011. **92**(2): p. 355-360.

59. Whitson, S.E., et al., *Direct probe-atmospheric pressure chemical ionization mass spectrometry of cross-linked copolymers and copolymer blends.* Anal Chem, 2008. **80**(20): p. 7778-85.

60. Huba, A.K., K. Huba, and P.R. Gardinali, *Understanding the atmospheric pressure ionization of petroleum components: The effects of size, structure, and presence of heteroatoms.* Sci Total Environ, 2016. **568**: p. 1018-1025.

61. Zimmermann, R. and L. Hanley, *Photoionization and Photo-Induced Processes in Mass Spectrometry*. 2021: Wiley. 440.

62. Raffaelli, A. and A. Saba, *Atmospheric pressure photoionization mass spectrometry.* Mass Spectrom Rev, 2003. **22**(5): p. 318-31.

63. Thomas, M.J., et al., *Solvent and Flow Rate Effects on the Observed Compositional Profiles and the Relative Intensities of Radical and Protonated Species in Atmospheric Pressure Photoionization Mass Spectrometry.* Anal Chem, 2022. **94**(12): p. 4954-4960.

64. Neumann, A., et al., *APPI (10.6/10.0 eV), APCI, and APLI (266 nm) for gas chromatography high resolution mass spectrometry of standards and complex mixtures.* Analytical Chemistry, 2023.

65. Short, L.C., S.S. Cai, and J.A. Syage, *APPI-MS: effects of mobile phases and VUV lamps on the detection of PAH compounds.* J Am Soc Mass Spectrom, 2007. **18**(4): p. 589-99.

66. McEwen, C.N., *GC/MS on an LC/MS instrument using atmospheric pressure photoionization.* International Journal of Mass Spectrometry, 2007. **259**(1-3): p. 57-64.

67. Hieta, J.P., et al., *Soft X-ray Atmospheric Pressure Photoionization in Liquid Chromatography-Mass Spectrometry.* Anal Chem, 2021. **93**(27): p. 9309-9313.

68. Neumann, A., et al., *Detailed Comparison of Xenon APPI (9.6/8.4 eV), Krypton APPI (10.6/10.0 eV), APCI, and APLI (266 nm) for Gas Chromatography High Resolution Mass Spectrometry of Standards and Complex Mixtures.* J Am Soc Mass Spectrom, 2023.

69. Tiina J. Kauppila, J.S., *Photoionization at Elevated or Atmospheric Pressure: Applications of APPI and LPPI* Photoionization and Photo-Induced Processes in Mass Spectrometry: Fundamentals and Applications, 2020.

70. Kauppila, T.J., J.A. Syage, and T. Benter, *Recent developments in atmospheric pressure photoionization-mass spectrometry.* Mass Spectrom Rev, 2017. **36**(3): p. 423-449.

71. Syage, J.A., *Mechanism of [M + H]+ formation in photoionization mass spectrometry.* J Am Soc Mass Spectrom, 2004. **15**(11): p. 1521-33.

72. Kauppila, T.J., R. Kostiainen, and A.P. Bruins, *Anisole, a new dopant for atmospheric pressure photoionization mass spectrometry of low proton affinity, low ionization energy compounds.* Rapid Commun Mass Spectrom, 2004. **18**(7): p. 808-15.

73. Kauppila, T.J., et al., *Atmospheric pressure photoionization mass spectrometry. Ionization mechanism and the effect of solvent on the ionization of naphthalenes.* Anal Chem, 2002. **74**(21): p. 5470-9.

74. Castilla, C., et al., *Direct Inlet Probe Atmospheric Pressure Photo and Chemical Ionization Coupled to Ultrahigh Resolution Mass Spectrometry for the Description of Lignocellulosic Biomass.* J Am Soc Mass Spectrom, 2020. **31**(4): p. 822-831.

75. Podgorski, D.C., et al., *Characterization of pyrogenic black carbon by desorption atmospheric pressure photoionization Fourier transform ion cyclotron resonance mass spectrometry.* Anal Chem, 2012. **84**(3): p. 1281-7.

76. Chiaberge, S., et al., *Classification of crude oil samples through statistical analysis of APPI FTICR mass spectra.* Fuel Processing Technology, 2013. **106**: p. 181-185.

77. Vetere, A., D. Pröfrock, and W. Schrader, *Qualitative and Quantitative Evaluation of Sulfur-Containing Compound Types in Heavy Crude Oil and Its Fractions.* Energy & Fuels, 2021. **35**(10): p. 8723-8732.

78. Headley, J.V., et al., *Atmospheric Pressure Photoionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Characterization of Tunable Carbohydrate-Based Materials for Sorption of Oil Sands Naphthenic Acids.* Energy & Fuels, 2013. **28**(3): p. 1611-1616.

79. Droste, S., et al., *A silica-based monolithic column in capillary HPLC and CEC coupled with ESI-MS or electrospray-atmospheric-pressure laser ionization-MS.* Electrophoresis, 2005. **26**(21): p. 4098-103.

80. Streibel, T. and R. Zimmermann, *Resonance-enhanced multiphoton ionization mass spectrometry (REMPI-MS): applications for process analysis.* Annu Rev Anal Chem (Palo Alto Calif), 2014. **7**: p. 361-81.

81. Constapel, M., et al., *Atmospheric-pressure laser ionization: a novel ionization method for liquid chromatography/mass spectrometry.* Rapid Commun Mass Spectrom, 2005. **19**(3): p. 326-36.

82. Kauppila, T.J., H. Kersten, and T. Benter, *The ionization mechanisms in direct and dopant-assisted atmospheric pressure photoionization and atmospheric pressure laser ionization.* J Am Soc Mass Spectrom, 2014. **25**(11): p. 1870-81.

83. Schmidt, S., et al., *Atmospheric pressure laser ionization. An analytical technique for highly selective detection of ultralow concentrations in the gas phase.* Anal Chem, 1999. **71**(17): p. 3721-9.

84. Panda, S.K., et al., *Atmospheric pressure laser ionization (APLI) coupled with Fourier transform ion cyclotron resonance mass spectrometry applied to petroleum samples analysis: comparison with electrospray ionization and atmospheric pressure photoionization methods.* Rapid Commun Mass Spectrom, 2011. **25**(16): p. 2317-26.

85. Ruger, C.P., et al., *Atmospheric Pressure Single Photon Laser Ionization (APSPLI) Mass Spectrometry Using a 157 nm Fluorine Excimer Laser for Sensitive and Selective Detection of Non- to Semipolar Hydrocarbons.* Anal Chem, 2021. **93**(8): p. 3691-3697.

86. Knochenmuss, R., *The Coupled Chemical and Physical Dynamics Model of MALDI.* Annu Rev Anal Chem (Palo Alto Calif), 2016. **9**(1): p. 365-85.

87. Qian, K., et al., *Laser Desorption Ionization of Large Condensed Polycyclic Aromatic Hydrocarbons in Petroleum Pitch and Thermally Treated Products.* Energy & Fuels, 2022. **36**(16): p. 8674-8683.

88. Dhahak, A., et al., *Analysis of Products Obtained from Slow Pyrolysis of Poly(ethylene terephthalate) by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Coupled to Electrospray Ionization (ESI) and Laser Desorption Ionization (LDI).* Industrial & Engineering Chemistry Research, 2020. **59**(4): p. 1495-1504.

89. Blackburn, J.W.T., et al., *Laser Desorption/Ionization Coupled to FTICR Mass Spectrometry for Studies of Natural Organic Matter.* Anal Chem, 2017. **89**(8): p. 4382-4386.

90. Dreisewerd, K., *The desorption process in MALDI.* Chem Rev, 2003. **103**(2): p. 395-426.

91. Liao, P.-C. and J. Allison, *Ionization processes in matrix-assisted laser desorption/ionization mass spectrometry: Matrix-dependent formation of [M + H]+ vs [M + Na]+ ions of small peptides and some mechanistic comments.* Journal of Mass Spectrometry, 1995. **30**(3): p. 408-423.

92. Kutz, K.K., J.J. Schmidt, and L. Li, *In situ tissue analysis of neuropeptides by MALDI FTMS in-cell accumulation.* Anal Chem, 2004. **76**(19): p. 5630-40.

93. Li, Y., et al., *High-resolution MALDI Fourier transform mass spectrometry of oligonucleotides.* Anal Chem, 1996. **68**(13): p. 2090-6.

94. Jing, Z., W. Hao-Yang, and G. Yin-Long, *Amino Acids Analysis by MALDI Mass Spectrometry Using Carbon Nanotube as Matrix.* Chinese Journal of Chemistry, 2005. **23**(2): p. 185-189.

95. Giraldo-Dávila, D., et al., *Selective ionization by electron-transfer MALDI-MS of vanadyl porphyrins from crude oils.* Fuel, 2018. **226**: p. 103-111.

96. Mase, C., et al., *Contribution of LDI and MALDI for the Characterization of a Lignocellulosic-Based Pyrolysis Bio-Oil.* J Am Soc Mass Spectrom, 2023.

97. Mapolelo, M.M., et al., *Characterization of naphthenic acids in crude oils and naphthenates by electrospray ionization FT-ICR mass spectrometry.* International Journal of Mass Spectrometry, 2011. **300**(2-3): p. 149-157.

98. Abdul Jameel, A.G., et al., *Characterization of deasphalted heavy fuel oil using APPI (+) FT-ICR mass spectrometry and NMR spectroscopy.* Fuel, 2019. **253**: p. 950-963.

99. Tose, L.V., et al., *Analyzes of hydrocarbons by atmosphere pressure chemical ionization FT-ICR mass spectrometry using isooctane as ionizing reagent.* Fuel, 2015. **153**: p. 346-354.

100. Gaspar, A., et al., *Characterization of Saturates, Aromatics, Resins, and Asphaltenes Heavy Crude Oil Fractions by Atmospheric Pressure Laser Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.* Energy & Fuels, 2012. **26**(6): p. 3481-3487.

101. Terra, N., et al., *Characterization of Crude Oil Molecules Adsorbed onto Carbonate Rock Surface Using LDI FT-ICR MS.* Energy & Fuels, 2022. **36**(12): p. 6159-6166.

102. Gaspar, A., et al., *Impact of different ionization methods on the molecular assignments of asphaltenes by FT-ICR mass spectrometry.* Anal Chem, 2012. **84**(12): p. 5257-67.

103. Vazquez, D. and G.A. Mansoori, *Identification and measurement of petroleum precipitates.* Journal of Petroleum Science and Engineering, 2000. **26**(1-4): p. 49-55.

104. Panda, S.K., J.T. Andersson, and W. Schrader, *Characterization of Supercomplex Crude Oil Mixtures: What Is Really in There?* Angewandte Chemie, 2009. **121**(10): p. 1820-1823.

105. Kondyli, A. and W. Schrader, *Evaluation of the combination of different atmospheric pressure ionization sources for the analysis of extremely complex mixtures.* Rapid Commun Mass Spectrom, 2020. **34**(8): p. e8676.

106. Mase, C., et al., *Contribution of atmospheric pressure chemical ionization mass spectrometry for the characterization of bio-oils from lignocellulosic biomass: Comparison with electrospray ionization and atmospheric pressure photoionization.* Journal of Analytical and Applied Pyrolysis, 2022. **167**.

107. Kim, S., et al., *Analysis of environmental organic matters by Ultrahigh-Resolution mass spectrometry-A review on the development of analytical methods.* Mass Spectrom Rev, 2022. **41**(2): p. 352-369.

108. Kurek, M.R., et al., *Deciphering Dissolved Organic Matter: Ionization, Dopant, and Fragmentation Insights via Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry.* Environ Sci Technol, 2020. **54**(24): p. 16249-16259.

109. He, C., et al., *Ionization selectivity of electrospray and atmospheric pressure photoionization FT-ICR MS for petroleum refinery wastewater dissolved organic matter.* Environ Sci Process Impacts, 2021. **23**(10): p. 1466-1475.

110. Vinci, G., A. Piccolo, and M. Bridoux, *Complementary ESI and APPI high resolution mass spectrometry unravel the molecular complexity of a soil humeome.* Anal Chim Acta, 2022. **1194**: p. 339398.

111. Luo, R. and W. Schrader, *Getting a better overview of a highly PAH contaminated soil: A non-targeted approach assessing the real environmental contamination.* J Hazard Mater, 2021. **418**: p. 126352.

112. Mauderly, J.L. and J.C. Chow, *Health effects of organic aerosols.* Inhal Toxicol, 2008. **20**(3): p. 257-88.

113. Haagen-Smit, A.J., *Chemistry and Physiology of Los Angeles Smog.* Industrial & Engineering Chemistry, 2002. **44**(6): p. 1342-1346.

114. Johnston, M.V. and D.E. Kerecman, *Molecular Characterization of Atmospheric Organic Aerosol by Mass Spectrometry.* Annu Rev Anal Chem (Palo Alto Calif), 2019. **12**(1): p. 247-274.

115. Nizkorodov, S.A., J. Laskin, and A. Laskin, *Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry.* Phys Chem Chem Phys, 2011. **13**(9): p. 3612-29.

116. Kuang, B.Y., et al., *Aromatic formulas in ambient PM(2.5) samples from Hong Kong determined using FT-ICR ultrahigh-resolution mass spectrometry.* Anal Bioanal Chem, 2018. **410**(24): p. 6289-6304.

117. Choi, J.H., et al., *In-depth compositional analysis of water-soluble and -insoluble organic substances in fine (PM(2.5)) airborne particles using ultra-high-resolution 15T FT-ICR MS and GCxGC-TOFMS.* Environ Pollut, 2017. **225**: p. 329-337.

118. Brown, S.C., G. Kruppa, and J.L. Dasseux, *Metabolomics applications of FT-ICR mass spectrometry.* Mass Spectrom Rev, 2005. **24**(2): p. 223-31.

119. Gray, G.R. and D. Heath, *A global reorganization of the metabolome in Arabidopsis during cold acclimation is revealed by metabolic fingerprinting.* Physiologia Plantarum, 2005. **124**(2): p. 236-248.

120. Aharoni, A., et al., *Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry.* OMICS, 2002. **6**(3): p. 217-34.

121. Calabrese, V., et al., *Direct introduction MALDI FTICR MS based on dried droplet deposition applied to non-targeted metabolomics on Pisum Sativum root exudates.* Talanta, 2023. **253**: p. 123901.

122. Hu, C., et al., *Analytical strategies in lipidomics and applications in disease biomarker discovery.* J Chromatogr B Analyt Technol Biomed Life Sci, 2009. **877**(26): p. 2836-46.

123. Goto-Inoue, N., et al., *Imaging mass spectrometry for lipidomics.* Biochim Biophys Acta, 2011. **1811**(11): p. 961-9.

124. Deschamps, E., et al., *Advances in Ultra-High-Resolution Mass Spectrometry for Pharmaceutical Analysis.* Molecules, 2023. **28**(5).

125. Selliez, L., et al., *High-resolution mass spectrometry for future space missions: Comparative analysis of complex organic matter with LAb-CosmOrbitrap and laser desorption/ionization Fourier transform ion cyclotron resonance.* Rapid Commun Mass Spectrom, 2020. **34**(10): p. e8645.

126. Hertzog, J., H. Naraoka, and P. Schmitt-Kopplin, *Profiling Murchison Soluble Organic Matter for New Organic Compounds with APPI- and ESI-FT-ICR MS.* Life (Basel), 2019. **9**(2).

127. Naraoka, H., et al., *Soluble organic molecules in samples of the carbonaceous asteroid (162173) Ryugu.* Science, 2023. **379**(6634): p. eabn9033.

128. Vahur, S., et al., *Analysis of dammar resin with MALDI-FT-ICR-MS and APCI-FT-ICR-MS.* J Mass Spectrom, 2012. **47**(3): p. 392-409.

129. Peets, P., et al., *Instrumental techniques in the analysis of natural red textile dyes.* Journal of Cultural Heritage, 2020. **42**: p. 19-27.

130. Kew, W., et al., *Complementary Ionization Techniques for the Analysis of Scotch Whisky by High Resolution Mass Spectrometry.* Anal Chem, 2018. **90**(19): p. 11265-11272.

131. Gosset-Erard, C., et al., *Hyphenation of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) with separation methods: The art of compromises and the possible - A review.* Talanta, 2023. **257**: p. 124324.

132. Ortiz, X., et al., *Characterization of naphthenic acids by gas chromatography-Fourier transform ion cyclotron resonance mass spectrometry.* Anal Chem, 2014. **86**(15): p. 7666-73.

133. Kondyli, A. and W. Schrader, *High-resolution GC/MS studies of a light crude oil fraction.* J Mass Spectrom, 2019. **54**(1): p. 47-54.

134. Reymond, C., et al., *A rational strategy based on experimental designs to optimize parameters of a liquid chromatography-mass spectrometry analysis of complex matrices.* Talanta, 2019. **205**: p. 120063.

135. Reymond, C., et al., *Characterization of liquid-liquid extraction fractions from lignocellulosic biomass by high performance liquid chromatography hyphenated to tandem high-resolution mass spectrometry.* J Chromatogr A, 2020. **1610**: p. 460569.

136. Lababidi, S. and W. Schrader, *Online normal-phase high-performance liquid chromatography/Fourier transform ion cyclotron resonance mass spectrometry: effects of different ionization methods on the characterization of highly complex crude oil mixtures.* Rapid Commun Mass Spectrom, 2014. **28**(12): p. 1345-52.

Characterization of organic aerosols from ship emissions by high resolution mass spectrometry:Development of new analytical methods and data visualization software.

Abstract

Organic aerosols (OA), whether of anthropogenic or natural origin, have a major impact on the environment, both in terms of climate changes and health effects. Among the many sources of OA, the maritime transport sector occupies a far from negligible place, on the one hand, because of the annual volume of exhaust emitted and their nature, and on the other because of the proximity of emissions to inhabited areas such as ports or coastlines. Although the composition of these exhausts is monitored by means of routine analyses, it is important to study them in depth in order to fine-tune regulations on marine traffic. However, the analysis of the molecular-scale composition of a mixture as complex as ship emissions requires the use of techniques such as mass spectrometry (MS), and in particular Fourier transform mass spectrometry (FTMS). However, FTMS generates a large amount of data and generally requires the use of processing and visualization software to extract and highlight relevant information. Also, although FTMS provides information on the molecular composition of a sample, this technique does not allow the evaluation of the isomeric diversity, unlike ion mobility spectrometry (IMS). So, with the aim to characterize ship emissions, we adopted three lines of work: the development of open-access software under Python to facilitate the processing and visualization of FTMS data for the characterization of complex mixtures, the targeted characterization of petroporphyrins in naval fuels and their combustion products by Fourier transform ion cyclotron resonance (FTICR) mass spectrometry using electron transfer matrix-assisted laser desorption/ionization (ET-MALDI), and the study of structural modifications caused by photochemical aging on ship emissions using IMS-MS.

Keywords: mass spectrometry, FTICR, ion mobility spectrometry, organic aerosols, particulate matter, software development, Python, ET-MALDI.

Caractérisation d'aérosols organique provenant d'émissions de bateaux par spectrométrie de masse haute résolution: Développement de nouvelles méthodes analytiques et d'un logiciel de visualisation de données.

Résumé

Les aérosols organiques (OA), qu'ils soient d'origine anthropogène ou naturelle, ont un grand impact sur l'environnement, tant d'un point de vue climatique que sanitaire. Parmi les nombreuses sources d'OA, le secteur du transport maritime occupe une place loin d'être négligeable, d'une part à cause du volume annuel de gaz d'échappement et de leur nature ; d'autre part à cause de la proximité des émissions avec des zones habitées telles que les ports ou les côtes. Même si la composition de ces gaz d'échappement est contrôlée au moyen d'analyses de routine, il est important de l'étudier en profondeur afin d'affiner les réglementations. Cependant, une analyse de la composition à l'échelle moléculaire d'un mélange aussi complexe que les émissions de navire requiert l'utilisation de techniques telle que la spectrométrie de masse (MS), et notamment la spectrométrie de masse à transformée de Fourier (FTMS). Cependant, la FTMS génère une grande quantité de données et nécessite généralement l'utilisation de logiciels de traitement et de visualisation afin d'extraire et de mettre en évidence les informations pertinentes. De plus, bien que la FTMS fournisse des informations sur la composition moléculaire d'un échantillon, cette technique ne permet pas d'en évaluer la diversité isomérique, contrairement à la spectrométrie de mobilité ionique (IMS). Ainsi, afin de caractériser les émissions de navire, nous avons adopté trois axes de travail : le développement d'un logiciel open-access sous python pour faciliter le traitement et la visualisation des données FTMS pour la caractérisation de mélanges complexes, la caractérisation ciblée de pétroporphyrines dans les carburants navals et leur produits de combustion par spectrométrie de masse à résonnance cyclotronique des ions par transformée de Fourier (FTICR) grâce à la désorption/ionisation laser assistée par matrice à transfert d'électrons (ET-MALDI), et l'étude des modification structurelles provoquées par le vieillissement photochimique sur les émissions de navires par IMS-MS.

Mots clés : Spectrométrie de masse, FTICR, spectrométrie de mobilité ionique, aérosols organiques, matière particulaire, développement de logiciel, Python, ET-MALDI.