



HAL
open science

Photoinduced degradation of contaminants in environmental matrices : Laboratory modelisation versus direct analysis

Edith Nicol

► **To cite this version:**

Edith Nicol. Photoinduced degradation of contaminants in environmental matrices : Laboratory modelisation versus direct analysis. Analytical chemistry. Institut Polytechnique de Paris, 2021. English. NNT : 2021IPPAX084 . tel-03627887

HAL Id: tel-03627887

<https://theses.hal.science/tel-03627887v1>

Submitted on 1 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Étude de la dégradation photochimique de contaminants en matrices environnementales : modélisation en laboratoire versus analyses directes

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°626 : École Doctorale de l'Institut Polytechnique de Paris
(ED IP Paris)
Spécialité de doctorat : Chimie

Thèse présentée et soutenue à Palaiseau, le 04/11/2021 par

Edith NICOL

Composition du Jury :

Hélène Lavanant	Rapporteur
Maître de conférences, Université Rouen Normandie (Laboratoire de Chimie Organique, Bioorganique, Réactivité et Analyse)	
Frédéric Aubriet	Rapporteur
Professeur, Université de Lorraine (Laboratoire de Chimie et Physique, Approche multi-échelle des Milieux Complexes)	
Corinne Gosmini	Présidente
Directrice de recherche, Ecole polytechnique (Laboratoire de Chimie Moléculaire)	
Olivier Chapleur	Examineur
Chercheur, INRAE (Unité Procédés biotechnologiques au service de l'environnement)	
Guillaume Van der Rest	Examineur
Professeur, Université Paris Sud (Institut de Chimie Physique)	
Stéphane Bouchonnet	Directeur de thèse
Ingénieur de Recherche, Ecole polytechnique (Laboratoire de Chimie Moléculaire)	

REMERCIEMENTS

Je souhaite tout d'abord remercier tous les membres de mon jury. Merci à Madame le Docteur Hélène Lavanant et à Monsieur le Professeur Frédéric Aubriet d'avoir accepté d'être rapporteurs de cette thèse de doctorat. Merci également à Madame le Docteur Corinne Gosmini, Monsieur le Docteur Olivier Chapleur et Monsieur le Professeur Guillaume Van der Rest d'avoir participé à l'évaluation de ce travail.

Je tiens aussi à remercier Monsieur le Docteur Cédric Tard, Directeur du Laboratoire de Chimie Moléculaire de l'Ecole Polytechnique, d'avoir soutenu ce travail de thèse au laboratoire et de m'avoir soutenue lorsque j'en avais besoin.

Je tiens à remercier tout particulièrement Monsieur le Docteur Stéphane Bouchonnet, Directeur de cette thèse. Merci de m'avoir encouragée à faire une thèse, de m'avoir fait confiance, guidée et écoutée. Merci pour ta disponibilité vers tes collègues et tes doctorants malgré les charges de travail que représentent la recherche et l'enseignement. Tu as toujours su trouver un moment pour répondre à mes demandes. Merci aussi de m'aider à prendre confiance en moi en m'impliquant dans tes projets de recherche et d'enseignement.

Un grand merci à mes collègues de l'INRIA et du Centre de Mathématiques appliquées, le Docteur Marc Lavielle et Madame Yao Xu pour leur aide précieuse lors de l'élaboration de SPIX. Nous avons réussi à trouver les bons mots pour nous comprendre, merci de votre patience envers nous et pour toutes vos bonnes idées.

Merci aussi à Madame le Docteur Sophie Bourcier pour m'avoir soutenue et conseillée pendant cette thèse. Tes conseils m'ont donné de nouvelles idées et ont permis d'avancer dans les moments les moins simples.

Merci à toi Zsuzsanna d'avoir partagé le FT avec moi et d'avoir travaillé avec moi autant sur tes problématiques que sur les miennes, notre travail a donné de très bons résultats c'était un plaisir de le faire avec toi !

Merci aux précédents directeurs du laboratoire, Monsieur le Docteur Gilles Ohanessian et Madame le Docteur Corinne Gosmini, de m'avoir soutenue dans ce projet et dans tous les autres projets. Merci Gilles de m'avoir donné l'opportunité d'intégrer le laboratoire (DCMR) à l'époque, tu as toujours tout fait pour tirer le meilleur de chacun de nous et pour nous obtenir le plus possible auprès des tutelles. Merci Corinne de m'avoir fait confiance et de m'avoir soutenue auprès de l'école pour cette thèse, cela n'aurait pas eu lieu sans toi.

Merci à Anne-Flo, Christophe, Cindy, Gilles A., Olivier et Sylvaine pour votre aide précieuse au quotidien. Merci à vous pour votre bonne humeur, même quand les pannes se succèdent, et vos grands éclats de rire.

Merci à mes collègues de laboratoire, Audrey, Bibi, Greg, Duncan et Louis pour leurs moments de discussion aussi bien sur ce travail que sur d'autres sujet tous aussi intéressants.

Je n'oublierai pas ceux qui ont partagé mon bureau pendant cette thèse, Carine et Adrien. Merci à vous pour votre patience, votre bonne humeur et de m'avoir fait rire les jours où le FT ne marchait pas. C'était un plaisir de partager ces moments avec vous !

Merci aux thésards du laboratoire, anciens et présents, de m'avoir donné envie de concrétiser ce travail en vous voyant faire le vôtre.

Merci aussi aux « permanents » qui ne sont pas restés permanents pendant cette thèse, Carine, Christian, Cindy, Françoise, Gilles F, Guillaume, Julia et Thérèse et qui m'ont écoutée parler de tout avec toujours autant d'intérêt.

Un grand merci à Guillaume van der Rest pour m'avoir un jour dit que je devrais faire une thèse. La petite graine a germé et en voilà le résultat.

Merci aussi à mes parents, mes sœurs et beaux-frères, neveux et nièces, qui ont été présents et m'ont questionnée tout au long de ce travail. Vous m'avez permis de sortir la tête de mon travail de temps en temps. Merci à vous d'être là pour toutes mes questions existentielles !

Un immense merci à Alix, pour tout ton soutien, ta bienveillance, ton optimisme et ta patience au quotidien. Merci d'avoir toujours essayé de trouver des solutions à mes problèmes même quand je ne voulais pas les entendre. Merci d'avoir tout pris en main pendant ces années et de m'avoir « supportée » comme tu le dis si joliment. Merci aussi à mes enfants, Maxence et Lucie, d'avoir su me faire revenir à la réalité le soir en sortant du labo et de m'avoir fait oublier l'espace d'un instant le stress qu'engendre une thèse en parallèle d'un travail à temps plein (et quelques confinements). La recherche est peut-être un travail de passionné, mais ma passion je la tiens de vous ! Merci à vous trois pour votre soutien, votre confiance et ces moments de rire et de joie que vous m'apportez chaque jour, je n'aurais pas la même volonté sans vous.

PUBLICATIONS SCIENTIFIQUES

ARTICLES PUBLIES DANS LE CADRE DE CETTE THESE

Cazzaniga, N.; Varga, Z.; Nicol, E.; Bouchonnet, S. UV-visible photodegradation of naproxen in water - Structural elucidation of photoproducts and potential toxicity. *European journal of mass spectrometry* **2020**, 26, 400-408

<https://doi.org/10.1177/1469066720973412>

Nicol, E.; Xu, Y.; Varga, Z.; Kinani, S.; Bouchonnet, S.; Lavielle, M. SPIX: a new software package to reveal chemical reactions at trace amounts in very complex mixtures from high-resolution mass spectra data sets. *Rapid Communications in Mass Spectrometry* **2021**, 35, e9015.

<https://doi.org/10.1002/rcm.9015>

Nicol, E.; Varga, Z.; Vujovic, S.; Bouchonnet, S. Laboratory scale UV–visible degradation of acetamidrid in aqueous marketed mixtures - Structural elucidation of photoproducts and toxicological consequences. *Chemosphere* **2020**, 248, 126040.

<https://doi.org/10.1016/j.chemosphere.2020.126040>

Varga, Z.; Nicol, E.; Bouchonnet, S. Photodegradation of benisothiazolinone: Identification and biological activity of degradation products. *Chemosphere* **2020**, 240, 124862.

<https://doi.org/10.1016/j.chemosphere.2019.124862>

Nicol, E.; Chayata, H.; Genty, C.; Bouchonnet, S.; Bourcier, S. Photodegradation of cyprodinil under UV–visible irradiation – chemical and toxicological approaches. *Rapid Communications in Mass Spectrometry* **2016**, 30, 2201-2211.

<https://doi.org/10.1002/rcm.7685>

ARTICLES PUBLIES EN MARGE DES TRAVAUX DE THESE

Chalvin C, Drevensek S, Gilard F, Mauve C, Chollet C, Morin H, et al. Sclareol and linalyl acetate are produced by glandular trichomes through the MEP pathway. *Horticulture Research* **2021**, 8(1), 206.

<https://doi.org/10.1038/s41438-021-00640-w>

Abdelmouleh, M.; Lalande, M.; Nicol, E.; Frison, G.; van der Rest, G.; Pouilly, J. C. Chemical processes involving 18-crown-6-ether in activated noncovalent complexes with protonated peptides. *Chemphyschem* **2021**, 22, 1243-1250.

<https://doi.org/10.1002/cphc.202100075>

Nagornov, K. O.; Tsybin, O. Y.; Nicol, E.; Kozhinov, A. N.; Tsybin, Y. O. Fourier transform ion cyclotron resonance mass spectrometry at the true cyclotron frequency. *Mass Spectrometry Reviews* **2021**, 1-24.

<https://doi.org/10.1002/mas.21681>

Gonçalves, N. P. F.; Varga, Z.; Nicol, E.; Calza, P.; Bouchonnet, S. Comparison of Advanced Oxidation Processes for the Degradation of Maprotiline in Water—Kinetics, Degradation Products and Potential Ecotoxicity. *Catalysts* **2021**, 11, 240.

<https://doi.org/10.3390/catal11020240>

Nagornov, K. O.; Kozhinov, A. N.; Nicol, E.; Tsybin, O. Y.; Touboul, D.; Brunelle, A.; Tsybin, Y. O. Narrow Aperture Detection Electrodes ICR Cell with Quadrupolar Ion Detection for FT-ICR MS at the Cyclotron Frequency. *Journal of the American Society for Mass Spectrometry* **2020**, 31, 2258-2269.

<https://doi.org/10.1021/jasms.0c00221>

Thaunay, F.; Calvo, F.; Nicol, E.; Ohanessian, G.; Clavaguéra, C. Infrared Spectra of Deprotonated Dicarboxylic Acids: IRMPD Spectroscopy and Empirical Valence-Bond Modeling. *ChemPhysChem* **2019**, 20, 803-814.

<https://doi.org/10.1002/cphc.201800947>

Nicol, E.; Clavaguéra, C.; Ohanessian, G. Vibrational spectroscopy of deprotonated peptides containing an acidic side chain. *International Journal of Mass Spectrometry* **2019**, 435, 42-50.

<https://doi.org/10.1016/j.ijms.2018.10.006>

Douix, S.; Dossmann, H.; Nicol, E.; Dufлот, D.; Giuliani, A. Spectroscopy and Photodissociation of the Perfluorooctanoate Anion. *Chemistry – A European Journal* **2018**, 24, 15572-15576.

<https://doi.org/10.1002/chem.201801997>

Bianco, A.; Deguillaume, L.; Vaïtilingom, M.; Nicol, E.; Baray, J.-L.; Chaumerliac, N.; Bridoux, M. Molecular Characterization of Cloud Water Samples Collected at the Puy de Dôme (France) by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environmental Science & Technology* **2018**, 52, 10275-10285.

<https://doi.org/10.1021/acs.est.8b01964>

De Vaugelade, S.; Nicol, E.; Vujovic, S.; Bourcier, S.; Pirnay, S.; Bouchonnet, S. Ultraviolet–visible phototransformation of dehydroacetic acid – Structural characterization of photoproducts and global ecotoxicity. *Rapid Communications in Mass Spectrometry* **2018**, 32, 862-870.

<https://doi.org/10.1002/rcm.8104>

De Vaugelade, S.; Nicol, E.; Vujovic, S.; Bourcier, S.; Pirnay, S.; Bouchonnet, S. UV–vis degradation of α -tocopherol in a model system and in a cosmetic emulsion—Structural elucidation of photoproducts and toxicological consequences. *Journal of Chromatography A* **2017**, 1517, 126-133.

<https://doi.org/10.1016/j.chroma.2017.08.015>

Katari, M.; Nicol, E.; Steinmetz, V.; van der Rest, G.; Carmichael, D.; Frison, G. Improved Infrared Spectra Prediction by DFT from a New Experimental Database. *Chemistry – A European Journal* **2017**, 23, 8414-8423.

<https://doi.org/10.1002/chem.201700340>

Chayata, H.; Lassalle, Y.; Nicol, É.; Manolikakes, S.; Souissi, Y.; Bourcier, S.; Gosmini, C.; Bouchonnet, S. Characterization of the ultraviolet–visible photoproducts of thiophanate-methyl using high performance liquid chromatography coupled with high resolution tandem mass spectrometry—Detection in grapes and tomatoes. *Journal of Chromatography A* **2016**, 1441, 75-82.

<https://doi.org/10.1016/j.chroma.2016.02.078>

TABLE DES MATIÈRES

Introduction générale	-1-
Chapitre I - Etude bibliographique.....	- 11 -
I.1. Les contaminants d'intérêt émergent.....	- 13 -
I.1.1. Problématiques liées aux CECs	- 13 -
I.1.2. Identification et devenir des CECs dans l'environnement.....	- 14 -
I.1.3. Molécules retenues pour cette étude	- 16 -
I.2. Photochimie des contaminants d'intérêt émergent.....	- 18 -
I.2.1. Principe de la photodégradation	- 19 -
I.2.2. Principaux mécanismes de photochimie	- 23 -
I.3. L'analyse non ciblée de mélanges complexes.....	- 24 -
I.3.1. Représentations graphiques	- 26 -
I.3.2. Analyse statistique des données	- 35 -
Chapitre II - Matériels et méthodes.....	- 51 -
II.1. Photoréacteurs utilisés dans cette étude	- 53 -
II.1.1. Photoréacteur à lampe à vapeur de mercure	- 53 -
II.1.2. Enceinte de vieillissement accéléré à lampe à arc de Xénon	- 55 -
II.2. Préparation des échantillons	- 56 -
II.2.1. Préparation des solutions pour la caractérisation de photoproduits par LC-MS et GC-MS.....	- 57 -
II.2.2. Préparation des échantillons pour l'identification de photoproduits par infusion directe.....	- 58 -
II.3. Tests de toxicité	- 58 -
II.3.1. Estimation <i>in silico</i> de la toxicité potentielle d'un composé	- 58 -
II.3.2. Etude de toxicité <i>in vitro</i>	- 60 -

II.4.	Méthodes physicochimiques d'analyse	- 61 -
II.4.1.	Couplage LC-MS <i>versus</i> introduction directe, atouts et limites	- 61 -
II.4.2.	La spectrométrie de masse	- 62 -
Chapitre III - Identification de produits de photodégradation de CECs par LC-MS. -		77 -
III.1.	Photodégradation UV-Visible du cyprodinil – Approches chimique et toxicologique.....	- 79 -
III.1.1.	Introduction des travaux portant sur la photodégradation du cyprodinil.....	- 79 -
III.1.2.	Synthèse des principaux résultats de cette étude.....	- 113 -
III.2.	Photodégradation de la benzisothiazolinone : identification et activité biologique des produits de dégradation	- 114 -
III.2.1.	Présentation des travaux publiés	- 114 -
III.2.2.	Synthèse des principaux résultats de cette étude.....	- 140 -
III.3.	Dégradation UV-Visible de l'acétamipride dans des matrices aqueuses commerciales – Elucidation structurale des photoproduits et conséquences toxicologiques	- 141 -
III.3.1.	Présentation des travaux relatifs à la dégradation photochimique de l'acétamipride.....	- 141 -
III.3.2.	Synthèse des principaux résultats de cette étude.....	- 173 -
III.4.	Photodégradation UV-Visible du naproxène dans l'eau – Elucidation structurale des photoproduits et détermination de leur toxicité potentielle	- 174 -
III.4.1.	Présentation des travaux publiés	- 174 -
III.4.2.	Synthèse des principaux résultats de cette étude.....	- 196 -
III.5.	Conclusion sur les travaux publiés	- 197 -
Chapitre IV - Développement d'un logiciel d'assistance à l'analyse de populations de spectres de masse – application au suivi réactionnel de mélanges hautement complexes.....		- 211 -

IV.1. Concepts de subjectivité et variabilité en chimie analytique	- 213 -
IV.1.1. Définitions de la subjectivité et de la variabilité	- 213 -
IV.1.2. Comparaison d'images.....	- 214 -
IV.1.3. Comparaison d'échantillons naturels	- 220 -
IV.2. SPIX : un logiciel innovant pour caractériser des réactions chimiques en milieu hautement complexes.....	- 221 -
IV.2.1. Test de SPIX sur un modèle simple : photolyse de l'acétamipride dans de l'eau ultrapure	- 222 -
IV.2.2. Test de SPIX sur des données complexes : photolyse de l'acétamipride à très faible concentration dans une solution aqueuse d'acide fulvique	- 227 -
IV.2.3. SPIX : un nouveau logiciel pour l'identification de réactions chimiques au sein de mélanges complexes, à partir de données de spectrométrie de masse à ultra haute résolution	- 232 -
IV.2.4. Conclusion sur les travaux publiés.....	- 261 -
IV.3. Conclusion sur le développement de l'outil SPIX.....	- 262 -
Conclusion générale	-267-

TABLE DES ABREVIATIONS

ACN	acétonitrile
ACP	Analyse en Composantes Principales
AINS	Anti-Inflammatoire Non Stéroïdien
ANOVA	<i>Analysis of variance</i> , analyse de la variance
CDOM	<i>Chromophoric Dissolved Organic Matter</i> , matière organique dissoute chromophore
CEC	<i>Contaminant of Emerging Concern</i> , contaminant d'intérêt émergent
CID	<i>Collision Induced Dissociation</i> , dissociations induites par collisions
CMAp	Centre de Mathématiques Appliquées
DBE	<i>Double Bond Equivalent</i> , nombre de cycles et/ou doubles liaisons
DCM	dichlorométhane
DDT	dichlorodiphényltrichloroéthane
DIMS	<i>Direct Infusion Mass Spectrometry</i> , spectrométrie de masse avec introduction directe par infusion
DOM	<i>Dissolved Organic Matter</i> , matière organique dissoute
DP	<i>Degradation product</i> , produit de dégradation
EC ₅₀	<i>Half maximal effective concentration</i> , concentration efficace médiane
EPA	Agence de protection environnementale des Etats Unis (<i>U.S. Environmental Protection Agency</i>)
ESI	<i>ElectroSpray Ionization</i> , ionisation par électronébulisation
FA	<i>Formic Acid</i> , Acide formique
FC	Fold-change
FT-ICR	<i>Fourier Transform Ion Cyclotron Resonance</i> , résonance cyclotronique ionique à transformée de Fourier
FT-IR	Infrarouge à Transformée de Fourier
FT-MS	<i>Fourier Transform Mass Spectrometer</i> , spectromètre de masse à transformée de Fourier
GC	<i>Gas Chromatography</i> , chromatographie en phase gazeuse
GC-MS	<i>Gas Chromatography Mass Spectrometry</i> , chromatographie gazeuse couplée à la spectrométrie de masse
HMRS	<i>High Resolution Mass Spectrometry</i> , spectrométrie de masse à haute résolution
HPLC	<i>High Performance Liquid Chromatography</i> , chromatographie en phase liquide à haute performance
INRIA	Institut national de recherche en informatique et en automatique
IGC ₅₀	<i>Half maximal Growth Inhibition Concentration</i> , concentration médiane d'inhibition de la croissance

ISO	<i>International Standards Organization</i> , Organisation internationale de normalisation
IUPAC	<i>International Union of Pure and Applied Chemistry</i> , Union internationale de chimie pure et appliquée
KMD	<i>Kendrick Mass Defect</i> , défaut de masse de Kendrick
KM	<i>Kendrick Mass</i> , masse de Kendrick
KMN	<i>Kendrick Mass Nominal</i> , masse nominale de Kendrick
LC	<i>Liquid Chromatography</i> , chromatographie en phase liquide
LD ₅₀	<i>Half maximal lethal dose</i> , dose létale médiane
MS	<i>Mass Spectrometry</i> , spectrométrie de masse
MS/MS	<i>Tandem Mass Spectrometry</i> , spectrométrie de masse en tandem
<i>m/z</i>	Rapport masse sur charge
ONG	Organisation non gouvernementale
PC	<i>Principal Component</i> , composante principale
PLS	<i>Partial least square</i> , régression par les moindres carrés partiels
PLS-DA	Analyse discriminante PLS
PP	Photoproduit
<i>ppm</i>	<i>part-per million</i>
QSAR	<i>Quantitative Structure Activity Relationship</i> , relation quantitative structure/activité
Q-Tof	<i>Quadrupole-Time of flight</i> , quadripôle couplé à un temps de vol
REACH	<i>Registration, Evaluation and Authorisation of Chemicals</i> , enregistrement, évaluation et autorisation des substances chimiques
RMN	Résonance Magnétique Nucléaire
RT	<i>Retention Time</i> , temps de rétention
SFC	<i>Supercritical Fluid Chromatography</i> , chromatographie en phase supercritique
TP	<i>Transformation Product</i> , produit de transformation
t-test	test de Student
UV	Ultraviolet

LISTE DES FIGURES

Figure I-1 : Devenir des pesticides dans les ressources en eau potable et dans les stations de traitement d'eau potable (d'après Melin et al., 2020) ³⁹	- 15 -
Figure I-2 : Voies de phototransformation des polluants en milieu aquatique (d'après Richard et al., 2007) ⁴⁹	- 19 -
Figure I-3 : Exemples de fonctionnalités chimiques présentes dans les substances humiques (d'après Vione et al., 2014) ⁵²	- 20 -
Figure I-4 : Principaux effets de déplacement de la position et de l'intensité de la bande d'absorption d'un groupement chromophore (d'après Guedira, 2013) ⁵³	- 22 -
Figure I-5 : Chemins réactionnels classiques de photolyse (d'après Burrows et al., 2002) ⁵⁰	- 23 -
Figure I-6 : Représentation des différents systèmes d'analyse chimique adaptés à la gestion de projets d'analyse de mélanges complexes. Idéalement, le but est de limiter les coûts par de faibles temps d'analyse et de traitement tout en maximisant les résolutions moléculaires et analytiques. (d'après Schmitt-Kopplin et al. 2019) ¹	- 25 -
Figure I-7 : Diagramme de Kendrick simulé pour des composés de formule brute C_xH_yO (en rouge) et $C_xH_yO_2$ (en bleu).....	- 28 -
Figure I-8 : Diagrammes de Kendrick basés sur différents défauts de masse. (a) Défaut de masse de Kendrick sur une base $-CH_2-$ (b) KMD sur une base $-H_2-$ (c) KMD sur une base $-H_2O-$ et (d) KMD sur une base $-O-$ (d'après Kim et al., 2003) ⁷⁸	- 30 -
Figure I-9 : Diagramme de Kendrick en 3 dimensions. (a) échantillon de pétrole Nord-Américain, (b) échantillon du gisement de production associé. Les échelles de couleur représentent les abondances relatives des ions pour chacun des spectres (d'après Rodgers et al., 2005) ⁸²	- 31 -
Figure I-10 : Diagrammes de van Krevelen en 3 dimensions d'échantillons de pétroles bruts (a) non dégradés, (b) biodégradés. Ces diagrammes n'incluent que des espèces contenant des atomes d'oxygène. (d'après Rodgers et al., 2005) ⁸²	- 32 -
Figure I-11 : Diagramme de van Krevelen d'un échantillon de matière organique dissoute. Les lignes représentent des réactions chimiques (A) méthylation, déméthylation ou élongation de chaîne ; (B) hydrogénation ou déshydrogénation ; (C) hydratation ou déshydratation ; (D) oxydation ou réduction (d'après Kim et al., 2003) ⁷⁸	- 33 -
Figure I-12 : Représentation en 3 dimensions de diagrammes de van Krevelen réalisés à partir de spectres de masse d'huiles d'olives originaires du Portugal. Les différentes variétés d'olives représentées sont : A : Galega Vulgar, B: Madural, C : Cobrançosa, D : Carrasquenha, E : Blanqueta, F : Picual, G : Arbequina, H : Cordovil de Serpa et I : Verdeal Alentejana (d'après Martins et al., 2020) ⁹⁵	- 34 -
Figure I-13 : Diagrammes de van Krevelen en 3 dimensions des classes N, NO et NO_2 de deux carburants fossiles : le charbon (en bleu) et le pétrole brut (en rouge) (d'après Wu et al., 2004) ⁹⁶	- 35 -

Figure I-14 : Volcano plot correspondant à la combinaison du fold change avec un t-test. Les points de couleur représentent les lipides dont l'intensité varie de façon significative entre les cellules saines et les cellules atteintes de mucoviscidose ($FC > 2$ et $p\text{-value} < 0,05$). Les lipides représentés par des points roses voient leur abondance diminuer dans les cellules atteintes comparées aux cellules saines tandis que les lipides représentés par des points verts voient leur abondance augmenter (d'après Dei Cas et al., 2020). ⁹⁸	- 39 -
Figure I-15 : Résultat de l'analyse ACP de données FT-ICR de vins CN et CB (CB : Côte de Beaune ; CN : Côte de Nuits) pour différents crus (2007, 2008 et 2009) (d'après Roullier-Gall et al., 2014) ¹¹¹	- 42 -
Figure I-16 : Analyse statistique PLS-DA des vins de Côte de Beaune (CB) et Côte de Nuits (CN) après 5, 4 et 3 ans de mise en bouteille pour les crus 2007, 2008 et 2009. (A) graphique des scores de PLS-DA en fonction des années, les deux premières composantes expliquent 40,2% de la variabilité. (B) graphique des scores de PLS-DA en fonction des terroirs, les deux premières composantes expliquent 33,4% de la variabilité totale. (d'après Roullier-Gall et al., 2014). ¹¹¹	- 42 -
Figure II-1 : Prototype de photoréacteur à lampe à vapeur de mercure.	- 54 -
Figure II-2 : Photographie du Q-SUN Xe-1	- 55 -
Figure II-3 : Spectre d'émission solaire à midi en été (violet) et comparaison avec la lampe à arc de Xénon équipée de filtre de simulation de la lumière solaire. ¹³¹	- 56 -
Figure II-4 : Domaines d'application des différentes sources d'ionisation en fonction des caractéristiques des molécules à étudier.....	- 63 -
Figure II-5 : Schéma d'une source d'électronébullisation	- 64 -
Figure II-6 : (A) Schéma d'émission des ions dans la source ESI. (B) Schéma du cône de Taylor avec émission de micro gouttelettes. (d'après Kebarle P., 2000) ¹⁵⁴	- 65 -
Figure II-7 : Photographie du spectromètre de masse FT-ICR SolarixXR 9.4T utilisé pour cette étude	- 67 -
Figure II-8 : Schéma du FT-ICR SolarixXR 9.4T utilisé pour cette étude.....	- 68 -
Figure II-9 : A) photographie de la cellule paracell installée au laboratoire. B) design de la paracell (d'après Boldin, I.A., et al. 2011) ¹⁵⁹	- 68 -
Figure II-10 : Mouvement d'un ion dans un champ électromagnétique. v_t : fréquence axiale, v_m : fréquence magnétron, v_c : fréquence cyclotronique (d'après Marshall, A.G., 1998) ¹⁵⁸	- 70 -
Figure III-1: Chemical structures of cyprodinil (Cypro) and cyprodinil-D ₅ (Cypro-D ₅).....	- 85 -
Figure III-2: LC/MS chromatograms recorded for the cyprodinil reference solution (a) and for the same solution irradiated for 4 h (b).....	- 89 -
Figure III-3: Suggested dissociation pathways for cyprodinil protonated on N1	- 90 -
Figure III-4: Suggested dissociation pathways for cyprodinil protonated on N2	- 90 -
Figure III-5: Suggested dissociation pathways for cyprodinil protonated on N3	- 90 -
Figure III-6: Response-dose curves of reference and 4 h photolyzed solutions of cyprodinil with the <i>Vibrio fischeri</i> bioluminescence inhibition test.	- 102 -
Figure III-7: UV-Vis absorption spectrum of BIT (1.2 mM) in ultrapure water at different irradiation times ranging from 0 to 60 minutes	- 123 -
Figure III-8: Suggested pathways for the photodegradation of BIT.....	- 130 -

Figure III-9: Chemical structures of acetamiprid (left-hand) - with numbering of carbon and nitrogen atoms - and acetamiprid-D ₃ (right-hand)	147 -
Figure III-10: LC-MS chromatograms of the aqueous solution of acetamiprid (10 mg.L ⁻¹) at t ₀ and after 60 min of irradiation.....	150 -
Figure III-11: Mechanisms proposed for the photo-induced formation of A, B, PP2 isomers and PP5....	151 -
Figure III-12: Mechanisms involved in the collision-induced dissociation of [PP2a + H] ⁺ ions.....	154 -
Figure III-13: Photo-induced mechanisms postulated for the formation of PP1 and PP3.....	156 -
Figure III-14: Mechanisms postulated for the formation of PP4a and PP4b isomers	156 -
Figure III-15: Results of the <i>in silico</i> toxicity estimations for acetamiprid and its photoproducts	160 -
Figure III-16: EC ₈₀ values determined for aqueous solutions of acetamiprid, Polysect Ultra SL [®] and Roseclear Ultra [®] at different irradiation times.....	161 -
Figure III-17: Relative abundances of Naproxen and its photoproducts over time: compounds of major abundances (NAP, PP1 and PP6) on left side and compounds of minor abundances (PP2, PP3, PP4, PP5 and PP7) on right side. On the y-axis, abundances are normalized to that of Naproxen at t ₀ . The x-axis corresponds to the irradiation time in minutes.	183 -
Figure III-18: Mechanisms proposed for the photodegradation of Naproxen in water.....	187 -
Figure IV-1 : Exemple d'images projetées lors de l'école d'hiver de spectrométrie de masse. ²⁸⁵ a) comparaison de deux mandalas comprenant 22 différences. b) comparaison de deux spectres simulés comprenant 15 différences.	216 -
Figure IV-2 : Graphiques obtenus par le logiciel SPIX pour les données issues de l'irradiation de l'acétamipride dilué dans l'eau ultrapure à 10 ppm. a) comparaison avant (t ₀) et après 30 minutes d'irradiation (t ₃₀) par un test de Student. b) comparaison sur l'ensemble de la cinétique	224 -
Figure IV-3 : Zoom sur le pic de l'ion de l'acétamipride protoné (m/z 223.0747) a) spectres simulés lors de la comparaison t ₀ / t ₃₀ minutes, avant irradiation en bleu, après 30 minutes d'irradiation en rouge. b) boîte à moustache des intensités moyennes pour les deux conditions étudiées (condition 1 : t ₀ ; condition 2 : t ₃₀). c) spectres simulés pour la totalité de la cinétique. La couleur représente le temps d'irradiation. d) courbe de dégradation déterminée à partir des intensités des points de cinétique pour le même ion.....	225 -
Figure IV-4 : Graphiques obtenus par le logiciel SPIX pour les données issues de l'irradiation de l'acétamipride dilué à 0,4 ppm dans l'acide fulvique à 200 ppm. a) comparaison avant (t ₀) et après 30 minutes d'irradiation (t ₃₀) par un test de Student après soustraction des « blancs » b) comparaison pour les données de la cinétique après soustraction des « blancs »	228 -
Figure IV-5: Chemical structures of maprotiline (left) and acetamiprid (right)	243 -
Figure IV-6: m/z 278.19056 signal (protonated maprotiline) extracted for each irradiation time and fitting model associated with degradation kinetics: $8.19 \times 10^{-5} + 4.37 \times 10^{10} \times \exp(-0.23x)$; $r^2 = 0.991$	245 -
Figure IV-7: m/z 294.18552 signal extracted for each irradiation time and associated kinetic model ...	248 -
Figure IV-8: Mass spectra of acetamiprid in mixture with fulvic acid. (A) Mass spectrum before photolysis; (B) mass spectrum after 30 min photolysis; (C) zoom on the protonated acetamiprid peak (m/z 223.0746) in (A); (D) zoom on the protonated acetamiprid peak in (B)	249 -

Figure IV-9: Visual result provided by the SPIX software after processing of mass spectra series recorded from samples taken before and after 30 min photolysis. The differences in ion intensities are given on the y-axis, positive values corresponding to decreased intensity after irradiation. The associated p value is given by the color scale: the more the color tends toward red, the more statistically significant the difference.....- 251 -

SUPPLEMENTARY INFORMATIONS

Supplementary Information III-1: Suggested dissociation pathways for [TP9 + H] ⁺	- 104 -
Supplementary Information III-2: Suggested dissociation pathways for [TP10 + H] ⁺	- 104 -
Supplementary Information III-3: Suggested dissociation pathways for [TP11 + H] ⁺	- 105 -
Supplementary Information III-4: Suggested dissociation pathways for [TP12 + H] ⁺	- 105 -
Supplementary Information III-5: Suggested dissociation pathways for [TP13 + H] ⁺	- 106 -
Supplementary Information III-6: Suggested dissociation pathways for [TP14 + H] ⁺	- 106 -
Supplementary Information III-7: Suggested dissociation pathways for [TP14 + H] ⁺	- 107 -
Supplementary Information III-8: Suggested dissociation pathways for [TP15 + H] ⁺	- 107 -
Supplementary Information III-9: Suggested dissociation pathways for [TP16 + H] ⁺	- 108 -
Supplementary Information III-10: Suggested dissociation pathways for [TP17 + H] ⁺	- 108 -
Supplementary Information III-11: Suggested dissociation pathways for [TP18 + H] ⁺	- 109 -
Supplementary Information III-12: Suggested dissociation pathways for [TP19 + H] ⁺	- 109 -
Supplementary Information III-13: Suggested dissociation pathways for [TP2 + H] ⁺	- 110 -
Supplementary Information III-14: Suggested phototransformation pathways of cyprodinil.....	- 110 -
Supplementary Information III-15: Suggested phototransformation pathways of cyprodinil.....	- 111 -
Supplementary Information III-16: Suggested phototransformation pathways of cyprodinil.....	- 111 -
Supplementary Information III-17: Suggested phototransformation pathways of cyprodinil.....	- 112 -
Supplementary Information III-18: Protocol for the synthesis of PP1 (Benzothiazolinone).....	- 134 -
Supplementary Information III-19: UV-Vis absorption spectrum of BIT (1.2 mM) in ultrapure water	- 134 -
Supplementary Information III-20: Exact m/z values and corresponding formulae of collision-induced product ions from [M+H] ⁺ and [M-H] ⁻ ions in LC-MS/MS.....	- 135 -
Supplementary Information III-21: Consecutive CO eliminations from hydroxylated isomers of BIT and PP1.. ..	- 136 -
Supplementary Information III-22: Dissociation pathways [PP5 - H] ⁻ under collisional activation.....	- 137 -
Supplementary Information III-23: Dissociation pathways of PP6 isomers	- 138 -
Supplementary Information III-24: Relative amount of BIT as a function of irradiation time for concentrations of BIT ranging from 5 to 100 ppm in ACN.....	- 139 -
Supplementary Information III-25: Relative amount of BIT as a function of irradiation time for concentrations of BIT ranging from 12.5 to 50 ppm in water	- 139 -
Supplementary Information III-26: Emission spectrum of the mercury lamp (HPL-N125W/542 E27 SC (Philips, Ivry-sur-Seine, France)	- 163 -
Supplementary Information III-27: Relative amount of acetamiprid (estimated by integrating the area of the corresponding chromatographic peak on the [M+H] ⁺ ionic current in positive ESI) as a function of the irradiation time in pure water, Polysect® and Roseclear® solutions	- 163 -

<i>Supplementary Information III-28: Relative abundances of photoproducts of acetamiprid as a function of irradiation time (normalized data)</i>	- 164 -
<i>Supplementary Information III-29: CID data for [PP2+H]⁺ isomers and suggested dissociation pathways for PP2b and PP2c (for PP2a see Figure III-12 in the main document)</i>	- 165 -
<i>Supplementary Information III-30: CID data for [PP1+H]⁺ and suggested dissociation pathways</i>	- 166 -
<i>Supplementary Information III-31: CID data for [PP3+H]⁺ and suggested dissociation pathways</i>	- 167 -
<i>Supplementary Information III-32: CID data for [PP4+H]⁺ isomers and suggested dissociation pathways</i>	- 168 -
<i>Supplementary Information III-33: CID data for [PP5+H]⁺ and suggested dissociation pathways</i>	- 170 -
<i>Supplementary Information III-34: Possible mechanisms involving the formation of a 5- or 6-membered ring from A to provide a photoproduct carrying 3 Deuterium atoms</i>	- 171 -
<i>Supplementary Information III-35: LC-MS chromatograms of the Polysect Ultra SL[®] and Roseclear Ultra[®] mixtures diluted in water at t₀ and after irradiation for 240 min. Peaks marked with a star correspond to isomers of triticonazole.</i>	- 172 -
<i>Supplementary Information III-36: Detailed results of in silico tests</i>	- 172 -
<i>Supplementary Information III-37: In silico toxicity estimations using the T.E.S.T. software</i>	- 192 -
<i>Supplementary Information III-38: Dissociations of molecular ions of the photoproducts detected by GC-ESI-MS</i>	- 193 -
<i>Supplementary Information III-39: Collision-induced ESI⁺ mass spectra of naproxen and its photoproducts</i>	- 194 -
<i>Supplementary Information III-40: Dissociations of the pseudomolecular ions (MH⁺) of the photoproducts detected by LC-ESI-MS</i>	- 195 -
<i>Supplementary Information IV-1: State of the art of modern approaches in managing high-resolution mass spectrometry data</i>	- 254 -
<i>Supplementary Information IV-2: Current kinetic models in SPIX</i>	- 257 -
<i>Supplementary Information IV-3: Chemicals, reagents and sample preparation</i>	- 258 -
<i>Supplementary Information IV-4: Exported data from the SPIX software after assignation of a kinetic model</i>	- 260 -

LISTE DES TABLEAUX

Tableau I-1 : Présentation des molécules retenues pour cette étude.....	- 18 -
Tableau II-1 : Effet de la résolution sur l'identification de contaminants dans des matrices de complexité moyenne (miels) à élevées (nourriture animale) (d'après Kellmann et al., 2009) ¹⁴⁷	- 72 -
Table III-1: Elemental composition, double-bond equivalency (DBE), molecular weight, retention times (RT) and main CID mass spectra obtained in LC/ESI-MS for cyprodinil and its phototransformation products-	91 -
Table III-2: Retention times and main CID mass spectra of [TP + H] ⁺ obtained in LC/ESI-MS for C ₁₄ H ₁₆ N ₃ and its isomers.....	- 92 -
Table III-3: Retention times and main CID mass spectra of [TP + H] ⁺ obtained in LC/ESI-MS for C ₁₄ H ₁₅ N ₃ O isomers	- 92 -
Table III-4: Retention times and main CID mass spectra of [TP + H] ⁺ obtained in LC/ESI-MS for C ₁₄ H ₁₅ N ₃ O ₂ isomers	- 92 -
Table III-5: Proposed structures and associated in silico predicted toxicity values for oral rat LD ₅₀ for cyprodinil and its photoproducts TP1, TP2 and TP14.....	- 99 -
Table III-6: Proposed structures and associated in silico predicted toxicity values for oral rat LD ₅₀ for cyprodinil and isomers of TP3	- 99 -
Table III-7: Proposed structures and associated in silico predicted toxicity values for oral rat LD ₅₀ for cyprodinil and its photoproducts TP9 to TP13.....	- 99 -
Table III-8: Proposed structures and associated in silico predicted toxicity values for oral rat LD ₅₀ for cyprodinil and its photoproducts TP15 to TP19.....	- 100 -
Table III-9: In silico predicted toxicity values and results for developmental toxicity and mutagenicity tests for cyprodinil and its photoproducts	- 101 -
Table III-10: Photoproducts of BIT detected in LC-MS and GC-MS couplings	- 125 -
Table III-11: Toxicity values estimated by the T.E.S.T. software for BIT and its photoproducts except PP6b.....	- 132 -
Table III-12: Retention times, main ions, and elucidated structures of the photoproducts PP1 to PP3 detected in GC-MS/MS (electron ionization).	- 184 -
Table III-13: Retention times, main ions, and elucidated structures of the photoproducts PP4 to PP7 detected by LC-ESI-MS.	- 185 -
Table III-14: Toxicity values estimated by the T.E.S.T. software for Naproxen and its photoproducts....	- 188 -
Tableau III-15 : Synthèse des résultats publiés.....	- 198 -
Tableau IV-1 : Différences observées pour la première étude (x : variabilité de l'intensité du pic entre les deux spectres comparés).....	- 217 -
Tableau IV-2 : Tableau récapitulatif des différences observées pour la première étude (x : variabilité de l'intensité du pic entre les deux spectres comparés)	- 218 -
Tableau IV-3 : Comparaison des résultats obtenus lors du test visuel.....	- 219 -

Tableau IV-4 : Extraction par le logiciel SPIX des rapports m/z ayant une p-value < 0,05 pour la cinétique et pour la comparaison t0 – t30. Identification des formules brutes par le logiciel DataAnalysis, ions associés à ces formules, regroupement des massifs isotopiques et suppression des harmoniques - 226 -

Tableau IV-5 : Extraction par le logiciel SPIX des rapports m/z auxquels est associée une p-value < 0.05 pour la cinétique, et la comparaison t0 – t30. Identification des formules brutes par le logiciel DataAnalysis et regroupement des massifs isotopiques - 229 -

Table IV-6: Ions extracted and associated kinetic models related to the photodegradation of Maprotiline in wastewater with $r^2 > 0.9$ (data ordered by decreasing intensity) - 246 -

Table IV-7: m/z values for which the intensity significantly varied between series of spectra recorded before and after 30 min of irradiation (n = 6). Ions are listed by increasing p value - 251 -

INTRODUCTION GENERALE

Dans le domaine de l'environnement, la caractérisation de contaminants émergents nécessite d'être capable de détecter des traces de composés, souvent inconnus, au sein de mélanges complexes. La notion de complexité varie selon le domaine auquel on se réfère ; elle est discutée en chimie analytique par Schmitt-Kopplin et al., dans un article publié en 2019 en introduction de la revue *Faraday Discussion* intitulé « *Systems chemical analytics: introduction to the challenges of chemical complexity analysis* ». ¹ La conception scientifique de la complexité s'applique par exemple aux milieux naturels en raison des interactions entre les organismes et leur écosystème, mais aussi entre les organismes et micro-organismes constituant le microbiote. La difficulté des analyses de mélanges complexes réside non seulement dans la quantité d'information présente dans le milieu, mais aussi dans sa constante évolution temporelle. Dans le cas de l'analyse de substrats naturels, à la fois complexes et variables, les auteurs placent la spectrométrie de masse et surtout la spectrométrie de masse à haute résolution parmi les outils les plus adaptés pour déterminer des molécules contenant des hétéroatomes, directement à partir de mélanges complexes. ¹

Plusieurs techniques sont utilisées pour l'analyse de mélanges complexes, notamment en métabolomique ; il s'agit par exemple de la résonance magnétique nucléaire (RMN), de l'analyse infrarouge à transformée de Fourier (FT-IR), de la spectroscopie Raman et de la spectrométrie de masse (MS). ²⁻⁴ Cette dernière est aujourd'hui la technique la plus sensible, la plus précise en terme de quantification, et celle offrant la meilleure gamme dynamique. ^{5,6} Concernant l'analyse de mélanges de produits inconnus, l'utilisation du mode balayage en MS (enregistrement de tous les signaux sur une gamme de rapports m/z donnée) permet d'obtenir des données exploitables aussi bien pour des analyses ciblées que non ciblées, il permet également de réaliser des analyses rétrospectives (concept de *frozen data*).

Les analyses non ciblées permettent la mise en évidence de produits de transformation ou de métabolites de micropolluants. La diversité des sources d'ions aujourd'hui disponibles en spectrométrie de masse en font une technique quasiment universelle. ⁷⁻¹⁰ Le couplage de cette dernière avec des techniques chromatographiques telles que la chromatographie liquide (LC) ou gazeuse (GC) ou encore avec la mobilité ionique permet

d'élargir l'éventail des molécules étudiées en séparant les molécules de différentes tailles et polarités ainsi que les isomères non différenciables par spectrométrie de masse.^{7,11}

Les échantillons issus de l'environnement présentent un double défi : ils sont à la fois complexes au sens où les substrats sont constitués d'une très grande diversité de molécules et ils contiennent des analytes d'intérêt à l'état de trace ; leur analyse est ainsi souvent décrite comme « chercher une aiguille dans une botte de foin ». Pour ces deux raisons, les spécialistes de l'environnement portent depuis les années 2000 un intérêt croissant à la HRMS couplée à la GC ou à la HPLC, comme le montre l'augmentation quasi exponentielle du nombre d'études consacrées à l'analyse de polluants organiques persistants (pesticides, hydrocarbures aromatiques polycycliques, etc.) par ces techniques.^{7,12-19}

L'étude des micropolluants seule n'est pas suffisante en matière de risque environnemental car les produits de transformation et/ou les métabolites de ces contaminants sont également potentiellement toxiques pour l'environnement et les écosystèmes. Peu de produits de transformation (TP) sont identifiés dans les échantillons environnementaux. Lors des analyses dites ciblées, où le spectromètre ne détecte que des molécules préalablement sélectionnées, il est nécessaire de connaître la masse de chaque analyte. Les TP ou métabolites inconnus non recherchés lors de l'analyse ne sont donc pas détectés ; ils ne peuvent pas non plus être retrouvés ultérieurement dans l'échantillon. Procéder à l'analyse rétrospective d'un échantillon ancien nécessite d'avoir accès à un maximum d'informations, ce que peut offrir une analyse non ciblée réalisée en mode dit « de balayage ».²⁰ Ces informations sont utiles pour établir la présence d'un contaminant d'intérêt émergent *a posteriori* dans un échantillon environnemental.

Les produits de photodégradation par irradiation UV-Visible de polluants organiques peuvent résulter de photolyse directe ou indirecte. La photolyse directe peut s'opérer en surface des sols, des eaux, des fruits et feuilles des fruits et légumes, alors que la photolyse indirecte est de plus en plus utilisée comme traitement de désinfection/décontamination dans les stations d'épuration *via* l'utilisation de catalyseurs chimiques. Les deux procédés sont susceptibles de conduire à la formation de TP toxiques.²¹⁻³⁰

Le sujet de cette thèse est double : il porte d'une part sur la caractérisation de contaminants d'intérêt émergent formés par photolyse directe en milieux aqueux et, d'autre part, sur la comparaison de deux approches - modélisation en laboratoire *versus* approche « directe » - pour la détermination de ces composés. L'approche consistant à modéliser la photodégradation en laboratoire pour caractériser les photoproduits et établir la toxicité de ces derniers a été initiée au laboratoire dans les années 2010.³⁰⁻³² Elle présente l'avantage d'utiliser des concentrations en molécules mères supérieures à celles rapportées dans les milieux naturels, permettant ainsi que la détection des TP ne soit pas limitée par les seuils de détection des instruments. De plus, les prélèvements effectués à différents temps de la photolyse permettent d'obtenir une information sur l'ordre d'apparition des TP et de proposer des mécanismes de formation de ces derniers. Cependant, ce type d'approche présente trois limites principales : (1) les choix - inévitablement subjectifs - relatifs à la phase stationnaire, à la méthode chromatographique et au mode d'ionisation ne permettent pas d'être exhaustif quant aux TP identifiés ; (2) il s'agit d'une approche longue et coûteuse en temps machine comme en temps humain ; (3) les résultats obtenus à partir d'expériences en laboratoire ne sont pas totalement extrapolables à des phénomènes se produisant en situation réelle comme par exemple en milieu naturel.

Ayant constaté les limitations susmentionnées, nous avons entrepris, dans une deuxième partie de ce travail, de développer un logiciel de traitement de données complexes obtenues par spectrométrie de masse. Ce logiciel, baptisé « SPIX » a été développé en étroite collaboration avec le Centre de Mathématiques Appliquées de l'Ecole polytechnique et l'INRIA ; il permet de conduire une analyse statistique des résultats basée sur les tests de Student et de Wilcoxon, pour comparer deux populations d'échantillons ou des séries d'échantillons (pour une étude cinétique, par exemple). Ce dernier point est totalement innovant et permet d'identifier facilement et simplement des produits de dégradation et leur évolution au cours du temps, y compris au sein d'une matrice très complexe. Ce logiciel permet de s'affranchir partiellement mais significativement des incertitudes inhérentes à la subjectivité de l'analyste.

Ce manuscrit inclut quatre chapitres.

Le premier présente l'état de l'art sur les contaminants d'intérêt émergent et les approches majoritairement utilisées dans la littérature pour l'analyse de mélanges complexes.

Le second chapitre porte sur la présentation des instruments, techniques et méthodes analytiques utilisés lors de cette étude.

Le troisième chapitre présente – majoritairement sous forme d'articles publiés dans des revues scientifiques – les résultats relatifs à la caractérisation de photoproduits de polluants organiques, résultats obtenus par une approche utilisant la modélisation de la photodégradation en laboratoire.

Le dernier chapitre porte sur le développement du logiciel d'analyses statistiques SPIX et démontre son intérêt pour le suivi de réactions chimiques au sein de mélanges complexes au cours du temps.

Enfin, le bilan de ces résultats et les perspectives ouvertes par ce travail sont présentés dans la conclusion générale.

RÉFÉRENCES BIBLIOGRAPHIQUES

- (1) Schmitt-Kopplin, P.; Hemmler, D.; Moritz, F.; Gougeon, R. D.; Lucio, M.; Meringer, M.; Muller, C.; Harir, M.; Hertkorn, N. Systems chemical analytics: introduction to the challenges of chemical complexity analysis. *Faraday Discuss* **2019**, *218*, 9-28.
- (2) Cherni, A.; Piersanti, E.; Anthoine, S.; Chaux, C.; Shintu, L.; Yemloul, M.; Torresani, B. Challenges in the decomposition of 2D NMR spectra of mixtures of small molecules. *Faraday Discuss* **2019**, *218*, 459-480.
- (3) Goodacre, R. The blind men and the elephant: challenges in the analysis of complex natural mixtures. *Faraday Discuss* **2019**, *218*, 524-539.
- (4) Kuhn, S.; Colreavy-Donnelly, S.; Santana de Souza, J.; Borges, R. M. An integrated approach for mixture analysis using MS and NMR techniques. *Faraday Discuss* **2019**, *218*, 339-353.
- (5) Shulaev, V.; Cortes, D.; Miller, G.; Mittler, R. Metabolomics for plant stress response. *Physiol Plant* **2008**, *132*, 199-208.
- (6) Ghaste, M.; Mistrik, R.; Shulaev, V. Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap Based High Resolution Mass Spectrometry in Metabolomics and Lipidomics. *Int J Mol Sci* **2016**, *17*, 816.
- (7) Hernandez, F.; Sancho, J. V.; Ibanez, M.; Abad, E.; Portoles, T.; Mattioli, L. Current use of high-resolution mass spectrometry in the environmental sciences. *Analytical and bioanalytical chemistry* **2012**, *403*, 1251-1264.
- (8) Lin, L.; Yu, Q.; Yan, X.; Hang, W.; Zheng, J.; Xing, J.; Huang, B. Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabonomics? A serum metabonomic study of kidney cancer. *The Analyst* **2010**, *135*, 2970-2978.
- (9) Alexandre-Gouabau, M. C.; Courant, F.; Moyon, T.; Kuster, A.; Le Gall, G.; Tea, I.; Antignac, J. P.; Darmaun, D. Maternal and cord blood LC-HRMS metabolomics reveal alterations in energy and polyamine metabolism, and oxidative stress in very-low birth weight infants. *J Proteome Res* **2013**, *12*, 2764-2778.
- (10) Arroyo-Manzanares, N.; Di Mavungu, J. D.; Uka, V.; Malysheva, S. V.; Cary, J. W.; Ehrlich, K. C.; Vanhaecke, L.; Bhatnagar, D.; De Saeger, S. Use of UHPLC high-resolution Orbitrap mass spectrometry to investigate the genes involved in the production of secondary metabolites in *Aspergillus flavus*. *Food additives & contaminants. Part A, Chemistry, analysis, control, exposure & risk assessment* **2015**, *32*, 1656-1673.
- (11) Leyva, D.; Tose, L. V.; Porter, J.; Wolff, J.; Jaffe, R.; Fernandez-Lima, F. Understanding the structural complexity of dissolved organic matter: isomeric diversity. *Faraday Discuss* **2019**, *218*, 431-440.
- (12) Ibanez, M.; Guerrero, C.; Sancho, J. V.; Hernandez, F. Screening of antibiotics in surface and wastewater samples by ultra-high-pressure liquid chromatography coupled to hybrid quadrupole time-of-flight mass spectrometry. *Journal of chromatography. A* **2009**, *1216*, 2529-2539.
- (13) Nurmi, J.; Pellinen, J. Multiresidue method for the analysis of emerging contaminants in wastewater by ultra performance liquid chromatography-time-of-flight mass spectrometry. *Journal of chromatography. A* **2011**, *1218*, 6712-6719.

(14) Diaz, R.; Ibanez, M.; Sancho, J. V.; Hernandez, F. Target and non-target screening strategies for organic contaminants, residues and illicit substances in food, environmental and human biological samples by UHPLC-QTOF-MS. *Analytical Methods* **2012**, *4*, 196-209.

(15) Bundy, J. G.; Davey, M. P.; Viant, M. R. Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **2008**, *5*, 3-21.

(16) Lankadurai, B. P.; Nagato, E. G.; Simpson, M. J. Environmental metabolomics: an emerging approach to study organism responses to environmental stressors. *Environmental Reviews* **2013**, *21*, 180-205.

(17) Munoz, K.; Schmidt-Heydt, M.; Stoll, D.; Diehl, D.; Ziegler, J.; Geisen, R.; Schaumann, G. E. Effect of plastic mulching on mycotoxin occurrence and mycobiome abundance in soil samples from asparagus crops. *Mycotoxin research* **2015**, *31*, 191-201.

(18) Sollic, M.; Roy-Lachapelle, A.; Sauve, S. Quantitative performance of liquid chromatography coupled to Q-Exactive high resolution mass spectrometry (HRMS) for the analysis of tetracyclines in a complex matrix. *Anal Chim Acta* **2015**, *853*, 415-424.

(19) Hogenboom, A. C.; Niessen, W. M.; Little, D.; Brinkman, U. A. Accurate mass determinations for the confirmation and identification of organic microcontaminants in surface water using on-line solid-phase extraction liquid chromatography electrospray orthogonal-acceleration time-of-flight mass spectrometry. *Rapid communications in mass spectrometry : RCM* **1999**, *13*, 125-133.

(20) Hernandez, F.; Ibanez, M.; Gracia-Lor, E.; Sancho, J. V. Retrospective LC-QTOF-MS analysis searching for pharmaceutical metabolites in urban wastewater. *J Sep Sci* **2011**, *34*, 3517-3526.

(21) Ibanez, M.; Sancho, J. V.; Pozo, O. J.; Hernandez, F. Use of quadrupole time-of-flight mass spectrometry in environmental analysis: elucidation of transformation products of triazine herbicides in water after UV exposure. *Analytical chemistry* **2004**, *76*, 1328-1335.

(22) Peschka, M.; Petrovic, M.; Knepper, T. P.; Barcelo, D. Determination of two phototransformation products of bentazone using quadrupole time-of-flight mass spectrometry. *Analytical and bioanalytical chemistry* **2007**, *388*, 1227-1234.

(23) Perez, S.; Eichhorn, P.; Barcelo, D. Structural characterization of photodegradation products of enalapril and its metabolite enalaprilat obtained under simulated environmental conditions by hybrid quadrupole-linear ion trap-MS and quadrupole-time-of-flight-MS. *Analytical chemistry* **2007**, *79*, 8293-8300.

(24) Kosjek, T.; Perko, S.; Heath, E.; Kralj, B.; Zigon, D. Application of complementary mass spectrometric techniques to the identification of ketoprofen phototransformation products. *Journal of mass spectrometry : JMS* **2011**, *46*, 391-401.

(25) Postigo, C.; Sirtori, C.; Oller, I.; Malato, S.; Maldonado, M. I.; de Alda, M. L.; Barcelo, D. Solar transformation and photocatalytic treatment of cocaine in water: Kinetics, characterization of major intermediate products and toxicity evaluation. *Appl Catal B-Environ* **2011**, *104*, 37-48.

(26) Postigo, C.; Sirtori, C.; Oller, I.; Malato, S.; Maldonado, M. I.; Lopez de Alda, M.; Barcelo, D. Photolytic and photocatalytic transformation of methadone in aqueous solutions under solar irradiation: kinetics, characterization of major intermediate products and toxicity evaluation. *Water Res* **2011**, *45*, 4815-4826.

(27) Chayata, H.; Lassalle, Y.; Nicol, E.; Manolikakes, S.; Souissi, Y.; Bourcier, S.; Gosmini, C.; Bouchonnet, S. Characterization of the ultraviolet-visible photoproducts of

thiophanate-methyl using high performance liquid chromatography coupled with high resolution tandem mass spectrometry-Detection in grapes and tomatoes. *Journal of chromatography. A* **2016**, *1441*, 75-82.

(28) Nicol, E.; Genty, C.; Bouchonnet, S.; Bourcier, S. Structural elucidation of metolachlor photoproducts by liquid chromatography/high-resolution tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM* **2015**, *29*, 2279-2286.

(29) Lassalle, Y.; Nicol, E.; Genty, C.; Bourcier, S.; Bouchonnet, S. Structural elucidation and estimation of the acute toxicity of the major UV-visible photoproduct of fludioxonil - detection in both skin and flesh samples of grape. *Journal of mass spectrometry : JMS* **2015**, *50*, 864-869.

(30) Lassalle, Y.; Jellouli, H.; Ballerini, L.; Souissi, Y.; Nicol, E.; Bourcier, S.; Bouchonnet, S. Ultraviolet-vis degradation of iprodione and estimation of the acute toxicity of its photodegradation products. *Journal of chromatography. A* **2014**, *1371*, 146-153.

(31) Souissi, Y.; Kinani, S.; Bouchonnet, S.; Bourcier, S.; Malosse, C.; Sablier, M.; Creusot, N.; Mombelli, E.; Ait-Aissa, S. Photolysis of estrone generates estrogenic photoproducts with higher activity than the parent compound. *Environmental science and pollution research international* **2014**, *21*, 7818-7827.

(32) Coffinet, S.; Rifai, A.; Genty, C.; Souissi, Y.; Bourcier, S.; Sablier, M.; Bouchonnet, S. Characterization of the photodegradation products of metolachlor: structural elucidation, potential toxicity and persistence. *Journal of mass spectrometry : JMS* **2012**, *47*, 1582-1593.

Chapitre I - ÉTUDE BIBLIOGRAPHIQUE

I.1. LES CONTAMINANTS D'INTERET EMERGENT

Un contaminant d'intérêt émergent (*CEC: Contaminant of emerging concern*) est un produit chimique introduit dans l'environnement, le plus souvent via le cycle de l'eau, qui peut être potentiellement dangereux pour l'être humain ou l'environnement et qui n'a pas fait l'objet de critères ou de normes à ce jour.³³⁻³⁵ Ces substances ne sont pas nécessairement de nouveaux produits, elles peuvent être présentes dans l'environnement depuis plusieurs années mais leur présence ou les conséquences délétères de celles-ci ne sont connues que depuis peu.

I.1.1. Problématiques liées aux CECs

L'un des premiers CECs à avoir été identifié dans les années 1960 a été le DDT (dichlorodiphényltrichloroéthane). Après des années d'utilisation en tant qu'insecticide pour lutter contre le paludisme, la peste bubonique et bien d'autres maladies, son impact sur l'être humain et l'environnement s'est avéré désastreux. Son utilisation en agriculture a été interdite pour la première fois en 1972 aux Etats Unis.³⁶

Depuis la catastrophe sanitaire causée par le DDT, de nombreuses molécules sont considérées comme polluants émergents et donc classées prioritaires.³⁷ Leur détection dans l'environnement n'est pas facile car ces polluants sont présents en faible concentration (de quelques ng/L à quelques µg/L) et leurs produits de dégradation ne sont pas ou peu connus. Mais la limite de détection dans l'environnement a diminué d'un facteur mille entre 1990 et 2010 et les techniques analytiques modernes permettent de détecter de plus en plus de métabolites.³⁸

Parmi les CECs se trouvent des produits cosmétiques, des agents tensioactifs, des perturbateurs endocriniens (prouvés ou suspectés), des pesticides, des produits pharmaceutiques, des nanoparticules, des plastifiants, des retardateurs de flamme, leurs produits de dégradation, etc. La diversité chimique de ces composés mais aussi la diversité des matrices environnementales au sein desquelles ils se trouvent font qu'il est très difficile de prédire s'ils vont demeurer dans l'eau ou migrer vers des matrices solides (sol ou sédiments). Ces points rendent les techniques de détection et de traitement très

difficiles à mettre en place de façon homogène et harmonieuse. En plus du véritable challenge analytique dû à la faible concentration de ces composés au sein de matrices complexes et variables, il importe aussi de souligner le défi que cela représente pour les études de toxicologie. En effet, même si certaines de ces molécules sont interdites depuis de nombreuses années, elles peuvent s'accumuler et/ou se dégrader dans l'environnement. En effet, ces micropolluants organiques sont, ou ont pu être, rejetés dans l'environnement de façon continue, menant à ce que l'on peut appeler une pseudo persistance. On observe souvent pour ces micropolluants des effets de toxicité chronique et même intergénérationnelle avec des risques non négligeables de bioaccumulation.³⁴ A l'heure actuelle, les conséquences sanitaires et environnementales les plus préoccupantes concernent l'effet de perturbation endocrinienne sur l'être humain et/ou sur les écosystèmes.

Par ailleurs, une fois dans l'environnement, ces composés sont en contact avec d'autres composants et polluants, créant un effet cocktail. Même si la toxicité individuelle de certains de ces composés est parfois connue, ce n'est pas le cas de tous ces composés ; de plus, les effets dus aux interactions potentielles avec la matrice (agonistes ou antagonistes) sont encore inconnus. La détermination d'un seuil maximal réglementaire doit aussi prendre en compte les différentes générations et la place de l'individu dans la chaîne alimentaire avec le risque de biomagnification. Ces différents points rendent longue et difficile la mise en place de normes pour chacun des CECs.

I.1.2. Identification et devenir des CECs dans l'environnement

La présence et le devenir de ces polluants dans l'environnement est une question d'actualité mondiale. Les CECs peuvent provenir de différentes sources telles que l'industrie, l'agriculture mais aussi les eaux urbaines (voir Figure I-1) et finissent par intégrer le cycle de l'eau. Leur dégradation peut provenir de voies très variées, telles que la photolyse, l'hydrolyse, l'ozonolyse, la chloration, la biodégradation, etc.

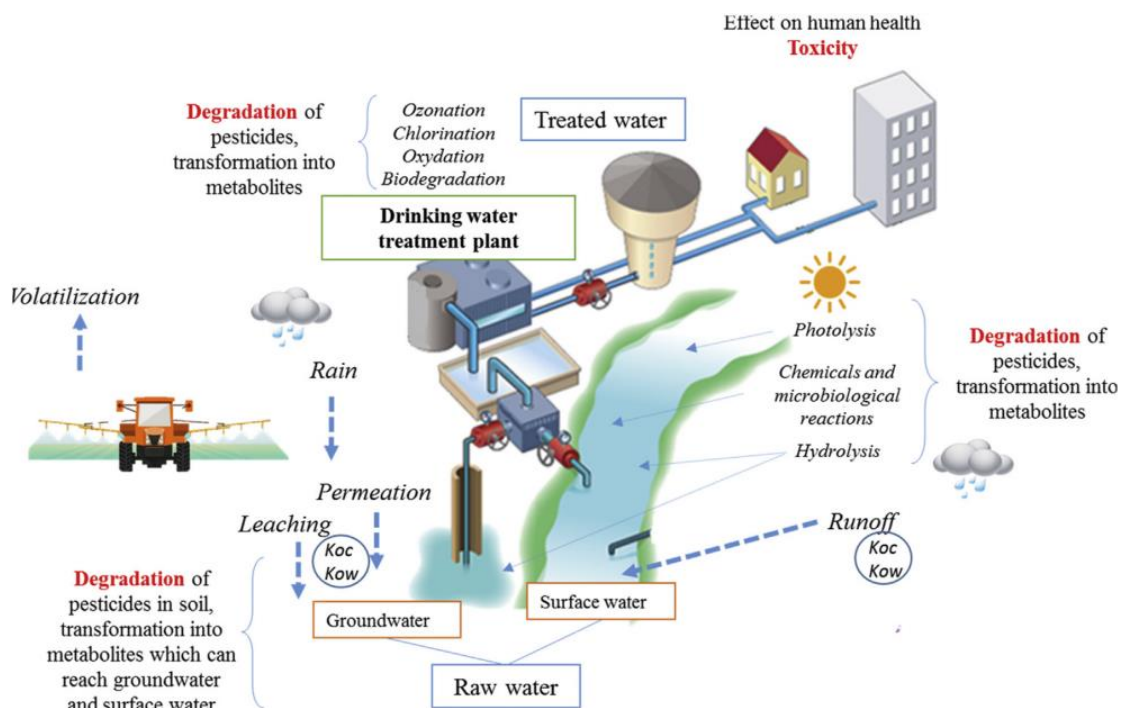


Figure I-1 : Devenir des pesticides dans les ressources en eau potable et dans les stations de traitement d'eau potable (d'après Melin et al., 2020)³⁹

La détection de ces molécules et de leurs produits de dégradation dans l'environnement est un véritable challenge vu la diversité et la complexité des milieux environnementaux dans lesquels ils se trouvent (matrices alimentaires, eaux de surface, eaux souterraines, eaux de station d'épuration, sols, sédiments...) et la faible concentration à laquelle ils se trouvent. Même si les polluants initialement rejetés dans l'environnement peuvent être identifiés et quantifiés, la majorité des produits de dégradation est encore inconnue. Ces produits sont de plus en plus recherchés grâce aux techniques analytiques actuelles et peuvent s'avérer plus toxiques que la molécule mère.⁴⁰⁻⁴⁶ La plupart des analyses permettant de déterminer la concentration en polluants sont des analyses ciblées, c'est-à-dire qu'on ne considère que les molécules recherchées ; les produits de transformation et les nouveaux contaminants ne sont donc pas détectables et une analyse rétrospective n'est pas possible. L'arrivée de nouveaux spectromètres de masse à haute résolution (HRMS) a permis de faciliter les approches non ciblées, l'avantage étant de pouvoir rechercher ou caractériser *a posteriori* de nouveaux contaminants au sein des échantillons. De plus la haute résolution permet une meilleure spécificité et sélectivité des analyses.³⁷

Afin de mettre en œuvre l'identification et la caractérisation d'un plus grand nombre de CECs, la Commission Européenne a financé en 2005 le projet NORMAN.⁴⁷ Ce projet a pour but de promouvoir un réseau permanent de laboratoires de référence, de centres de recherche universitaires ou industriels, des organismes de normalisation et des ONG (organisations non gouvernementales). Ce réseau a pour missions de :

- Faciliter les échanges de données sur la présence et les effets des CECs de façon plus rapide et plus large ;
- Mettre en place des méthodes communes de prélèvement et de mesures (chimiques et biologiques) dans le but d'améliorer la qualité et la comparabilité des données ;
- Transmettre, avec plus de transparence, des informations et des données de suivi relatives aux CECs ;
- Créer un forum indépendant et compétent pour mener des débats scientifiques et techniques portant sur les CECs.

A l'heure actuelle, les pesticides et leurs produits de dégradation connus sont soumis à la réglementation européenne pour l'eau destinée à la consommation humaine. Cette réglementation fixe un seuil de concentration individuelle pour les pesticides de 0.1 µg/L et un seuil de concentration cumulé de 0.5 µg/L.⁴⁸ Mais la plupart des produits de dégradation sont encore inconnus et n'entrent donc pas dans le calcul de cette concentration cumulée.

Les experts du réseau NORMAN révisent régulièrement la liste des substances émergentes et priorisent les études menant à la création de normes pour ces polluants et leurs produits de dégradation.

I.1.3. Molécules retenues pour cette étude

Depuis plusieurs années, le laboratoire étudie la dégradation des polluants dans l'environnement.³⁰⁻³² En effet, même si certains polluants sont connus et leur effets sur l'être humain et l'environnement largement étudiés, l'impact des produits de dégradation

de ces composés ne l'est pas toujours. Pour ce travail de thèse nous avons choisi d'étudier la photodégradation de quatre polluants émergents, des pesticides mais aussi des produits pharmaceutiques, qui se trouvent dans l'environnement et/ou dans nos produits de consommation. L'intérêt était double :

1) dans un premier temps, nous souhaitions élargir nos connaissances sur l'avenir de polluants suite à leur rejet dans l'environnement, afin de permettre le suivi des molécules mères mais aussi des produits de dégradation, tant pour des analyses environnementales qu'alimentaires ;

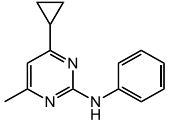
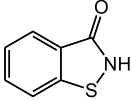
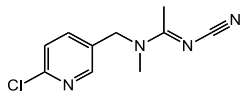
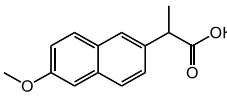
2) dans un second temps, il s'agissait d'approfondir nos connaissances des mécanismes de photolyse de polluants organiques possédant des fonctions chimiques communes (cycles aromatiques, groupes hydroxy, halogènes, etc.) afin de tenter de prédire les mécanismes de dégradation potentielle d'un polluant donné sur la base de sa seule structure chimique.

Ce travail a également porté sur l'étude de la dégradation de certains de ces polluants en milieux complexes telles que des matrices commerciales ou même en présence d'acide fulvique. Cela a permis d'étudier les interactions pouvant avoir lieu lors de la dégradation de ces polluants au sein d'un échantillon réel.

Les polluants choisis pour cette étude sont le cyprodinil, l'acétamipride, la benzisothiazolinone et le naproxène ; ces molécules sont décrites dans le Tableau I-1.

CHAPITRE I : ÉTUDE BIBLIOGRAPHIQUE

Tableau I-1 : Présentation des molécules retenues pour cette étude

Molécule	Cyprodinil	Benzisothiazolinone	Acétamipride	Naproxène
Nom IUPAC	4-cyclopropyl-6-methyl-N-phenylpyrimidin-2-amine	1,2-benzothiazol-3-one	N-[(6-chloropyridin-3-yl)methyl]-N'-cyano-N-methylethanimidamide	(2S)-2-(6-methoxynaphthalen-2-yl)propanoic acid
Structure				
Formule brute	C ₁₄ H ₁₅ N ₃	C ₇ H ₅ NOS	C ₁₀ H ₁₁ ClN ₄	C ₁₄ H ₁₄ O ₃
CAS	121552-61-2	2634-33-5	160430-64-8	22204-53-1
Usage	fongicide	fongicide et biocide	insecticide	médicament
Famille	anilino-pyrimidines	isothiazolinones	néonicotinoïdes	anti-inflammatoires non stéroïdiens
Prescription	Fongicide systémique	Bactéricide et conservateur	Insecticide systémique	Douleurs faibles ou modérées, fièvre ou inflammations
Solubilité dans l'eau	13 mg/L à 25°C	1,1 g/L à 20°C	4,25 g/L à 25°C	15,9 mg/L à 25°C
log Kow	4,00	0,76	0,80	3,18
pKa	4,44	7,30	0,70	4,15
Autorisation de mise sur le marché européen	Oui jusqu'en avril 2021	Proscrit dans les cosmétiques Consommation domestique : autorisé	Banni en France depuis 2018 (dérogations pour certaines cultures) Autorisé en Europe	oui

1.2. PHOTOCHEMIE DES CONTAMINANTS D'INTERET EMERGENT

Parmi tous les processus susceptibles de dégrader les CECs, la photodégradation fait partie de ceux induisant généralement de nombreux photoproduits.⁴⁹ La photodégradation est une réaction chimique de décomposition d'un composé induite par la lumière. Cette dégradation peut avoir lieu à différents moments : dès l'épandage des

pesticides, sur les feuilles soumises à l'irradiation solaire ; elle peut également avoir lieu dans les eaux de surfaces ou à la surface des sols soumis à l'exposition solaire. La dégradation photochimique sous l'action de la lumière UV est aussi un procédé couramment utilisé dans les stations d'épuration comme alternative à la désinfection par le chlore.

A l'heure actuelle, les études de photodégradation de polluants sont nombreuses ; beaucoup sont axées sur les aspects cinétiques mais peu portent sur l'identification et la caractérisation structurale des photoproduits.^{50,51}

I.2.1. Principe de la photodégradation

Il existe deux types de réactions de photodégradation : la photodégradation directe et la photodégradation indirecte, cette dernière étant également appelée photodégradation induite ou sensibilisée (Figure I-2).⁴⁹

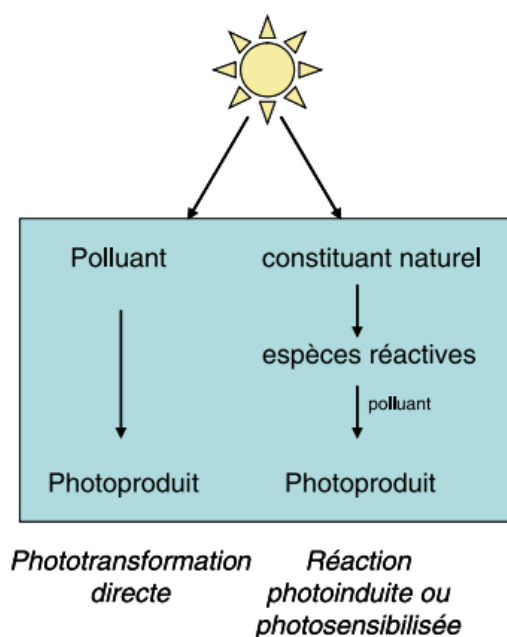


Figure I-2 : Voies de phototransformation des polluants en milieu aquatique (d'après Richard et al., 2007)⁴⁹

Dans le cas de la photolyse directe, l'absorption de photons UV a lieu directement, grâce au(x) groupement(s) chromophore(s) de la molécule tandis que la photolyse indirecte implique un catalyseur, souvent nommé photosensibilisateur, qui absorbe les photons

puis transfère son énergie au composé. Ces photosensibilisateurs peuvent être naturellement présents dans le milieu ou ajoutés par l'homme, notamment dans les stations d'épurations. En eaux de surface, les photosensibilisateurs sont majoritairement les anions nitrate et nitrite ainsi que la fraction de la matière organique dissoute chromophore capable d'absorber la lumière solaire (CDOM – Chromophoric Dissolved Organic Matter).⁵² Cette matière organique dissoute est composée d'une large famille de composés dont les divers groupements fonctionnels peuvent être des dérivés de la lignine, d'acides organiques, d'hydrocarbures, de cycles aromatiques ou même d'hétérocycles (Figure I-3). Certaines de ces espèces peuvent produire - sous l'action de l'irradiation UV - des photosensibilisateurs tels que les radicaux OH \cdot , CO $_3^{\cdot-}$, l'oxygène singulet (1O_2) ou même l'état triplet de la CDOM ($^3CDOM^*$) qui par la suite transfèrent leur énergie aux molécules présentes dans le milieu induisant ainsi leur dégradation.⁵²

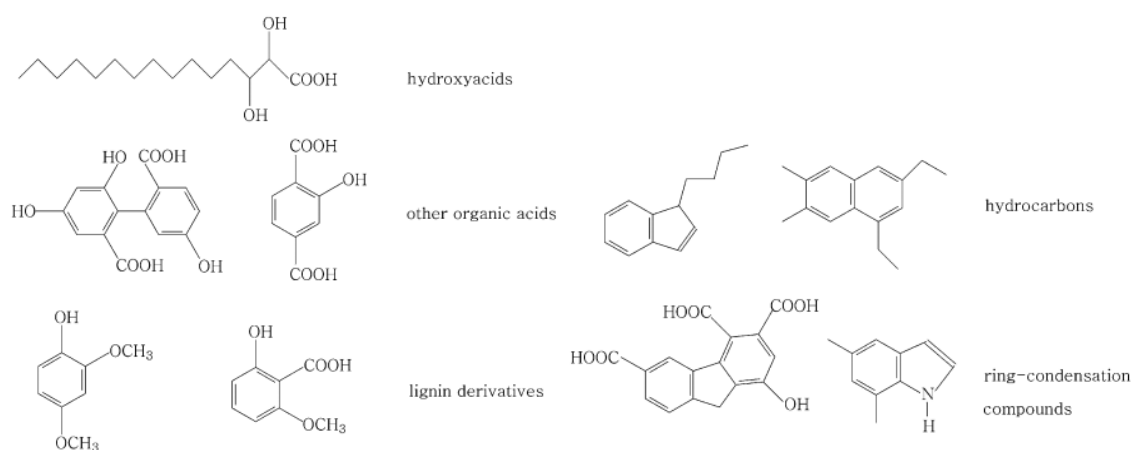


Figure I-3 : Exemples de fonctionnalités chimiques présentes dans les substances humiques (d'après Vione et al., 2014)⁵²

La photochimie directe est principalement régie par deux lois. La première loi, dite loi de Grotthus et Draper, stipule que seule la lumière absorbée par une molécule peut être capable d'induire un changement photochimique de cette molécule. Cette absorption dépend du coefficient d'absorption molaire ϵ de la molécule. Plus ϵ est élevé dans l'UV, plus la molécule subira des photodégradations directes sous l'effet de l'irradiation solaire naturelle.

La deuxième loi de la photochimie, appelée loi de Stark-Einstein, stipule que « l'absorption de la lumière par une molécule est un processus à un seul quantum, de sorte que la somme des rendements quantiques du processus direct doit être unitaire », ce qui signifie que les photoréactions directes sont induites par l'absorption d'un seul photon. L'efficacité de la photolyse est directement liée à l'énergie des photons et donc inversement proportionnelle à leur longueur d'onde. Cette énergie doit correspondre à la différence entre l'état fondamental et l'état excité potentiel de la molécule irradiée.

Pour qu'elle ait lieu, la réaction photochimique dépend donc de deux paramètres : le coefficient d'absorption molaire de la molécule irradiée et l'énergie du photon incident.

L'absorption est favorisée notamment par la présence de groupements chromophores tels que les cycles aromatiques, les systèmes π conjugués mais aussi certains groupements fonctionnels et hétéroatomes. La longueur d'onde d'absorption dépend de la nature des orbitales mises en jeu.⁵³

Transition $\sigma \rightarrow \sigma^*$

Les liaisons σ des composés organiques sont des liaisons très stables. La transition d'un électron d'une orbitale moléculaire liante σ vers une orbitale moléculaire non liante σ^* demande de ce fait beaucoup d'énergie. La bande d'absorption correspondante se trouve donc dans l'UV-lointain.

Transition $n \rightarrow \pi^*$

Le coefficient d'absorption molaire du passage d'une orbitale moléculaire non-liante n à une orbitale moléculaire antiliante π^* est faible. Cette transition se retrouve dans le cas de molécules contenant un hétéroatome porteur de doublets électroniques libres appartenant à un système insaturé. La plus connue de ces bandes d'absorption est celle correspondant au groupement C=O, qui est située entre 270 et 280 nm.

Transition $n \rightarrow \sigma^*$

Pour les alcools, les éthers, les amines et les dérivés halogénés, le transfert d'un électron du doublet n de l'hétéroatome à un niveau σ^* donne une bande d'absorption d'intensité moyenne se situant à l'extrême limite du proche-UV.

Transition $\pi \rightarrow \pi^*$

Cette transition concerne les composés possédant une double liaison isolée. Elle conduit à une forte bande d'absorption vers 165-200 nm.

La position et l'intensité de la bande d'absorption dépendent de l'environnement autour du groupement chromophore. En effet, la conjugaison, les substituants présents ou même les effets de solvants peuvent avoir un effet sur la valeur de la longueur d'onde maximale d'absorption λ_{\max} (effets hypsochrome et bathochrome), mais aussi sur le coefficient d'absorption molaire ϵ (effets hyperchrome et hypochrome), comme indiqué Figure I-4.

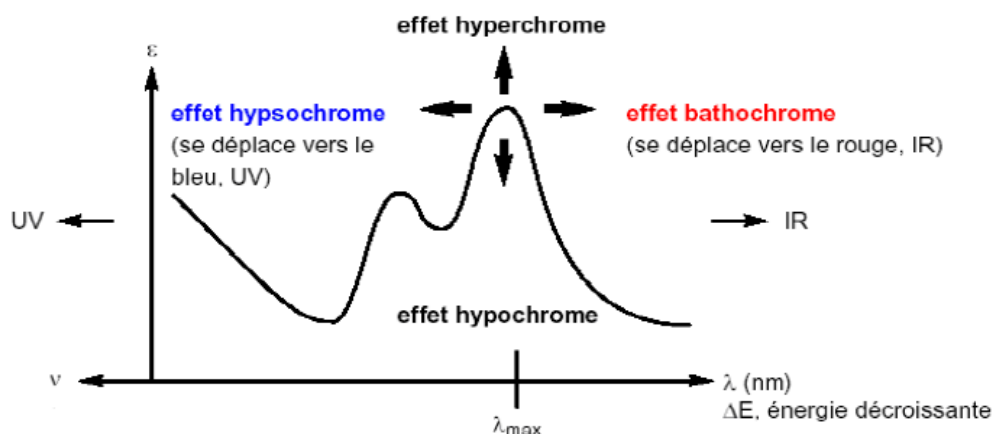


Figure I-4 : Principaux effets de déplacement de la position et de l'intensité de la bande d'absorption d'un groupement chromophore (d'après Guedira, 2013)⁵³

En milieu naturel, la largeur de la région UV du rayonnement solaire qui parvient à la surface de la terre se trouve réduite par l'absorption et le réfléchissement dans la troposphère des longueurs d'ondes inférieures à 290 nm. Seuls les groupements contenant des insaturations ou des hétéroatomes vont donc avoir la possibilité de réagir sous irradiation solaire.⁵⁴ Les transitions observées seront donc principalement les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$.

I.2.2. Principaux mécanismes de photochimie

A l'heure actuelle les réactions photochimiques ayant lieu en mélanges complexes sont loin d'être parfaitement connues. Ceci dit, les réactions photoinduites débutent souvent par la formation d'espèces radicalaires ; il s'en suit des réactions parmi lesquelles on peut citer la déhalogénéation, la substitution nucléophile, l'oxydation, la décarboxylation, l'isomérisation, la cyclisation ou même des réarrangement de type photo-Fries, photo-Claisen ou Norrish.⁵⁵⁻⁵⁷

Les mécanismes induits par irradiation lumineuse relèvent très majoritairement de l'un des processus suivants : l'homolyse, l'hétérolyse ou la photoionisation (voir Figure I-5).⁵⁰

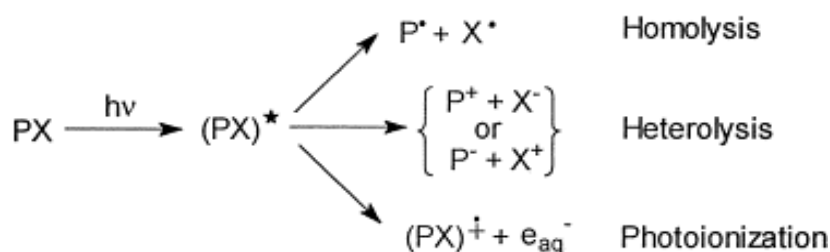


Figure I-5 : Chemins réactionnels classiques de photolyse (d'après Burrows et al., 2002)⁵⁰

En plus des très nombreuses voies possibles de dégradations des molécules, plusieurs points peuvent poser problème pour la caractérisation des produits de photodégradation. La faible concentration des polluants et donc celle de leurs nombreux produits de dégradation dans les matrices environnementales impose une limite de détection de la technique d'analyse très basse. La complexité et la variabilité des matrices environnementales pose des problèmes de performance (spécificité et sélectivité en particulier) et de répétabilité analytique. De nombreuses études visent donc tout d'abord à étudier la dégradation des polluants en laboratoire puis en milieux naturels. Néanmoins, la sensibilité et la résolution de certains spectromètres de masse permettent de tendre vers une analyse non ciblée de mélanges complexes pour en détecter et caractériser les produits de dégradation.

1.3. L'ANALYSE NON CIBLEE DE MELANGES COMPLEXES

L'analyse non ciblée, c'est-à-dire non sélectivement dédiée à l'analyse de composés préalablement définis, constitue, avec l'analyse quantitative et l'élucidation structurale, l'un des principaux défis de la chimie analytique moderne. Elle consiste à tenter d'extraire un maximum d'information à partir de l'analyse d'échantillons plus ou moins complexes, dont la composition est peu ou pas connue. Lorsque la matrice est complexe, ce qui est généralement le cas en chimie de l'environnement (eaux, nuages, sols, matrices animales et végétales, etc.), le défi est de taille : pouvoir détecter sans ambiguïté le maximum de molécules, même à l'état de traces.

Depuis de nombreuses années, la spectrométrie de masse est reconnue comme la technique analytique la plus performante dans ce contexte, parce qu'elle est spécifique, sélective, sensible, et qu'elle peut être aisément couplée à une technique séparative comme la chromatographie. Ainsi, les couplages GC-MS (Chromatographie gazeuse couplée à la spectrométrie de masse) et LC-MS (Chromatographie liquide couplée à la spectrométrie de masse) fournissent des données en 3 dimensions (temps de rétention, m/z et intensité de signal) permettant l'identification de plusieurs dizaines - voire centaines - de molécules. Cependant, les analyses non ciblées utilisant le couplage GC-MS ou LC-MS sont confrontées à deux limites principales. La première est liée au concept de subjectivité qui est largement discuté au chapitre IV : le choix d'une technique chromatographique exclut de fait la détection de certaines molécules. La seconde est imputable au fait qu'un système chromatographique, même très performant, ne peut pas séparer efficacement les molécules d'un échantillon dès lors que le nombre de ces dernières dépassent plusieurs centaines.

La spectrométrie de masse à haute résolution (analyseurs à secteur magnétique, Q-ToF, Orbitrap et FT-ICR) peut permettre de s'affranchir du couplage avec une technique séparative (GC, HPLC). En effet, une mesure de m/z très précise et très résolutive permet de discriminer des ions isobares sans ambiguïté. Très supérieure à celle des autres analyseurs, la résolution exceptionnelle de l'analyseur FT-ICR fait de ce dernier un outil de choix pour l'analyse non ciblée de mélanges complexes.⁵⁸⁻⁶⁰ Une analyse ainsi réalisée, dite analyse DIMS (pour *direct infusion mass spectrometry*) peut enregistrer plusieurs

milliers d'ions en moins de 5 minutes, fournissant ainsi une véritable empreinte de l'échantillon. Par ailleurs, l'absence de séparation chromatographique simplifie le développement analytique et diminue les temps d'analyse. La Figure I-6, montre que les techniques couplées de type LC-MS ou GC-MS - aujourd'hui très utilisées dans le domaine des analyses de mélanges complexes - sont insuffisamment performantes dans ce domaine et que la haute résolution en introduction directe permet d'obtenir le meilleur compromis entre qualité des résultats et coût de l'analyse.

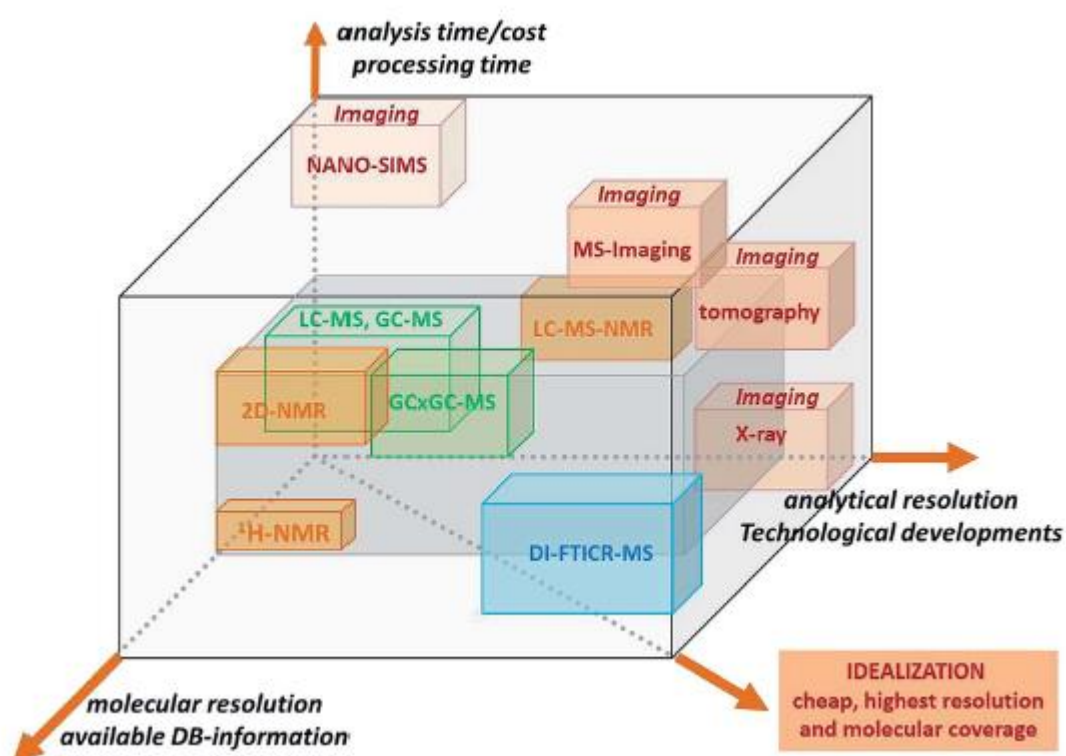


Figure I-6 : Représentation des différents systèmes d'analyse chimique adaptés à la gestion de projets d'analyse de mélanges complexes. Idéalement, le but est de limiter les coûts par de faibles temps d'analyse et de traitement tout en maximisant les résolutions moléculaires et analytiques. (d'après Schmitt-Kopplin et al. 2019)¹

Néanmoins, les analyses DIMS ne permettent pas la distinction des isomères ni l'attribution d'ions fragments ; elles augmentent les phénomènes de suppression ionique avec les sources de type ESI (*electrospray ionization*).

En spectrométrie de masse à très haute résolution (HRMS), la complexité des spectres obtenus est souvent telle qu'il n'est pas possible de comparer visuellement ces derniers sans introduire une importante subjectivité de l'opérateur. Il existe des méthodes de visualisation graphique permettant d'identifier rapidement des différences significatives entre des échantillons de populations différentes (échantillon traité vs non traité ; sain vs malade, etc.). La variabilité intra-individus au sein d'une même population étant non négligeable, il faut souvent avoir recours à des analyses statistiques pour comparer plus finement les résultats relatifs aux populations. Au regard de la complexité inhérente aux échantillons de domaines d'investigation comme la pétrologie, l'environnement, l'alimentaire, la métabolomique ou la santé, il peut s'avérer nécessaire de combiner plusieurs outils, graphiques et/ou statistiques, pour extraire des informations pertinentes des *big data* obtenues.

I.3.1. Représentations graphiques

De nombreuses représentations graphiques sont utilisées afin de pouvoir exploiter les spectres de masses obtenus à partir d'échantillons complexes. Celles-ci sont généralement employées pour des comparaisons d'échantillons deux à deux, le plus souvent pour étudier l'effet d'un traitement sur un échantillon. Ces comparaisons étant visuelles, elles restent limitées à un faible nombre de figures car le cerveau humain est incapable de comparer efficacement des populations de figures. Parmi les nombreuses représentations graphiques existantes, ne sont présentées ici que les plus couramment utilisées : le diagramme de Kendrick et le diagramme de van Krevelen.

I.3.1.1) Le diagramme de Kendrick

En 1963, Kendrick propose une nouvelle façon de représenter des spectres de masses, basée sur une échelle simplifiée pour laquelle la masse de 14,0000 Da est assignée au groupement CH₂ (au lieu de 14,0157 Da).⁶¹ Cette échelle permet de regrouper des molécules par séries. Ces séries sont constituées de molécules de classes (nombre

d'hétéroatomes) homologues et de degrés d'insaturation équivalents mais de niveaux d'alkylation différents. Ainsi, les molécules d'une même série ne différant que d'une unité CH_2 , ont le même défaut de masse de Kendrick (KMD) – nombre entier sur l'échelle du même nom – et sont donc plus facilement identifiables.

Le défaut de masse de Kendrick se calcule à partir de la masse nominale de Kendrick (KMN) et de la masse exacte de Kendrick (KM) définies par les équations 1 et 2.

$$KMD = NKM - KM \quad (\text{eq. 1})$$

$$KM = IUPAC \text{ mass} * \left(\frac{14}{14.01565} \right) \quad (\text{eq. 2})$$

La représentation graphique obtenue en traçant le KMD en fonction du KMN est nommée diagramme de Kendrick ; elle permet d'établir facilement des corrélations entre les différents ions d'un spectre de masse complexe simplement à partir des masses exactes de ces derniers. Le défaut de masse étant unique pour chaque élément, il est possible d'attribuer une formule brute à toutes les molécules d'une même famille (même nombre d'hétéroéléments) à partir de la formule brute d'un seul des ions de cette famille. Ainsi, la Figure I-7 représente un diagramme de Kendrick simulé de molécules ionisées issues des familles $\text{C}_x\text{H}_y\text{O}$ et $\text{C}_x\text{H}_y\text{O}_2$ où les molécules alignées horizontalement possèdent le même nombre d'hétéroatomes et d'insaturations mais diffèrent par le nombre de groupes alkyles (flèche verte Figure I-7). La diminution du nombre d'insaturations se constate par un écart de 0.01340 sur le défaut de masse de Kendrick et par un écart de 2 sur la masse nominale de Kendrick (flèche violette Figure I-7) ; il est également possible de lier deux familles entre elles, en se basant sur des différences de KMD et de KM caractéristiques d'un groupement chimique donné (CH_2O par exemple, écart indiqué par la flèche noire).

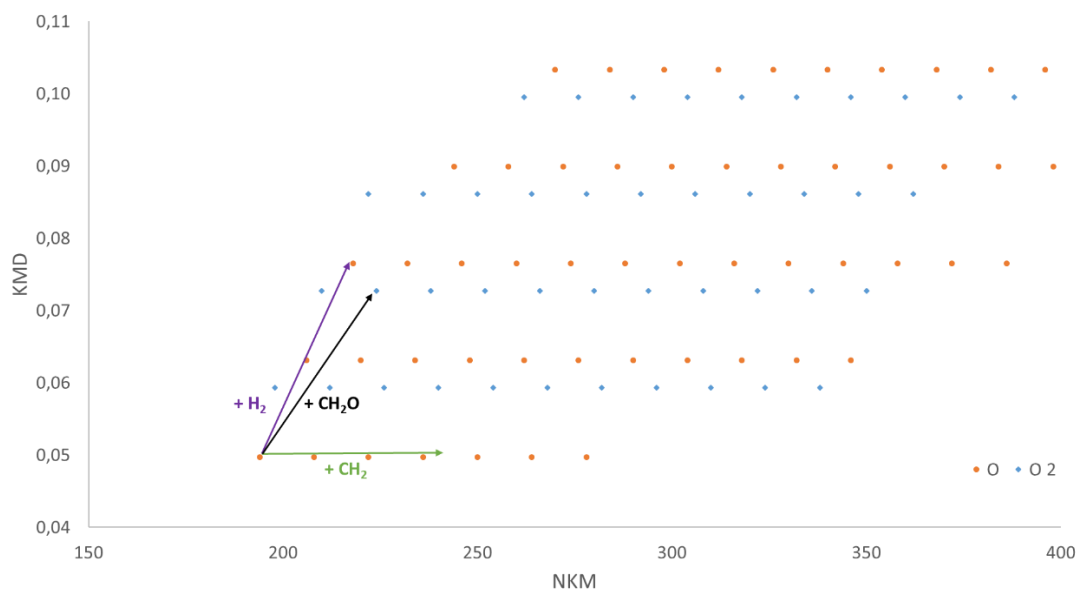


Figure I-7 : Diagramme de Kendrick simulé pour des composés de formule brute C_xH_yO (en rouge) et $C_xH_yO_2$ (en bleu)

Dans un premier temps le diagramme de Kendrick a été utilisé en pétroliométrie, pour mettre en évidence des familles de molécules différant d'un motif CH_2 au sein de mélanges complexes.^{62,63} Il a ensuite été étendu à des analyses d'échantillons environnementaux complexes tels que la matière organique dissoute⁶⁴⁻⁶⁷ et la lignine^{68,69}. Il est aussi utilisé en métabolomique⁷⁰⁻⁷², en lipidomique⁷³ et en protéomique pour l'étude de phosphopeptides.⁷⁴

L'utilisation du diagramme de Kendrick pour des échantillons de plus en plus variés a permis une modification de ce dernier en modifiant la nature du groupement étudié pour le calcul du défaut de masse.⁷⁵ Dans ce cas, le défaut de masse de Kendrick pour le groupement (X) se calcule de la façon suivante (équations 3 et 4) :

$$KMD_{(X)} = NKM_{(X)} - KM_{(X)} \quad (\text{eq. 3})$$

$$KM_{(X)} = \text{Observed mass} * \left(\frac{\text{nominal mass of X}}{\text{exact mass of X}} \right) \quad (\text{eq. 4})$$

Les séries de molécules différant par le nombre de groupements CH_2 ne sont pas les seules à présenter un intérêt. Ainsi, l'approche de Kendrick a été appliquée à la caractérisation de phénomènes de chloration par Jobst *et al.*⁷⁶, d'oxydation par Taguchi *et al.*⁷⁷. Grâce

au diagramme de van Krevelen (dont le principe est expliqué au paragraphe I.3.1.2), Kim *et al.* ont pu mettre en évidence la présence probable de transitions chimiques entre molécules lors de l'analyse de matière organique naturelle ; il s'agissait de différences de longueur de chaînes alkyles, de méthylations/déméthylations, d'hydrogénations/déshydrogénations, d'hydratations/déshydratations et d'oxydations/réductions.⁷⁸ Ils ont donc tracé les différents diagrammes de Kendrick représentés Figure I-8, sur la base des KMD des groupements, molécules et atomes suivants : -CH₂-, -H₂O-, -H₂- et -O-. La Figure I-8 montre que le diagramme tracé sur la base du KMD de -H₂- permet d'identifier des séries de molécules dont les structures diffèrent d'un motif C₂H₄O. De la même façon, le diagramme tracé sur la base du KMD de -H₂O- révèle des séries de molécules dont les structure diffèrent d'un motif CH₂O, celui utilisant le KMD de l'atome d'oxygène révèle des structures différant de C₂H₂ et celui utilisant le KMD de -CH₂- révèle des séries de molécules pouvant être corrélées par le remplacement de 2 atomes d'hydrogène par un atome d'oxygène (oxydation). Au-delà de l'intérêt analytique direct fourni par ces informations, cette approche permet d'utiliser quelques formules brutes identifiées pour assigner une formule brute à presque tous les ions constitutifs du spectre de masse. Cette autre utilisation du diagramme de Kendrick est d'autant plus intéressante que, même en haute résolution avec une précision de moins de 1 ppm, une mesure à *m/z* 500 peut correspondre jusqu'à 21 formules brutes.⁷⁹ Ce type de traitement permet donc de limiter les formules chimiques possibles, en corrélant les ions de hauts poids moléculaires - pour lesquels plusieurs formules brutes sont possibles - avec un ion de faible poids moléculaire pour lequel il est facile d'assigner une seule formule brute.

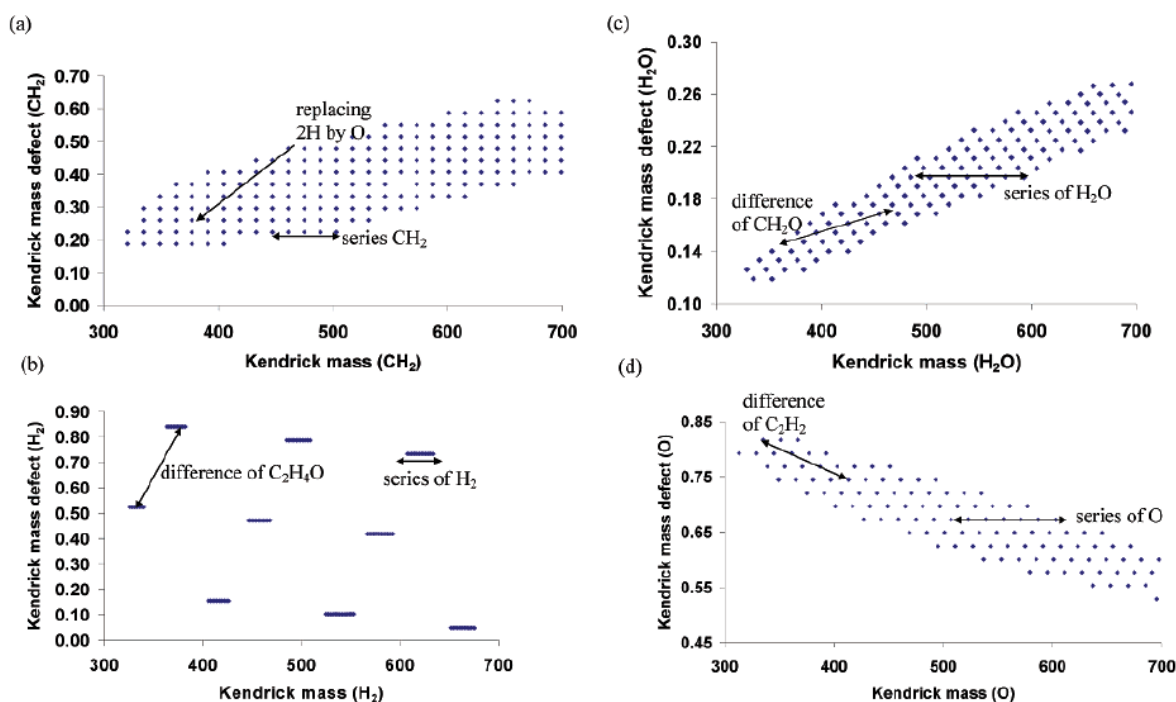


Figure I-8 : Diagrammes de Kendrick basés sur différents défauts de masse. (a) Défaut de masse de Kendrick sur une base $-CH_2$ - (b) KMD sur une base $-H_2$ - (c) KMD sur une base $-H_2O$ - et (d) KMD sur une base $-O$ - (d'après Kim et al., 2003)⁷⁸

Il est également possible d'utiliser les molécules identifiées pour réaliser une calibration interne des échantillons et permettre ainsi d'augmenter la précision sur la mesure de masse. Cette approche permet l'identification de composés n'appartenant pas à une série identifiée grâce au diagramme de Kendrick.^{80,81}

Le diagramme de Kendrick ayant été réalisé pour chacun des échantillons à étudier, l'opérateur les compare visuellement et cherche à repérer des différences traduisant un phénomène chimique, l'augmentation du nombre d'insaturations lors d'un traitement par exemple. La comparaison peut aussi être affinée en ajoutant une troisième dimension au diagramme : l'intensité de chaque ion dans le spectre de masse. Cette troisième dimension peut être visualisée via une échelle de couleur, comme sur la Figure I-9, ou via la taille des points représentés dans le diagramme. La Figure I-9, issue d'un article de Rodgers *et al.*⁸², représente les diagrammes de Kendrick 3D de deux échantillons de pétrole : un extrait brut de gisement et le pétrole obtenu après raffinage du brut. La comparaison des diagrammes montre l'effet du raffinage sur les molécules constituant les échantillons, en termes de nombre d'atomes de carbone (axe x) et de nombre

d'insaturations (axe y). Cette comparaison montre que le gisement contient plus d'espèces ayant un KMD élevé que le pétrole raffiné, ce qui signifie qu'il contient plus d'espèces insaturées.

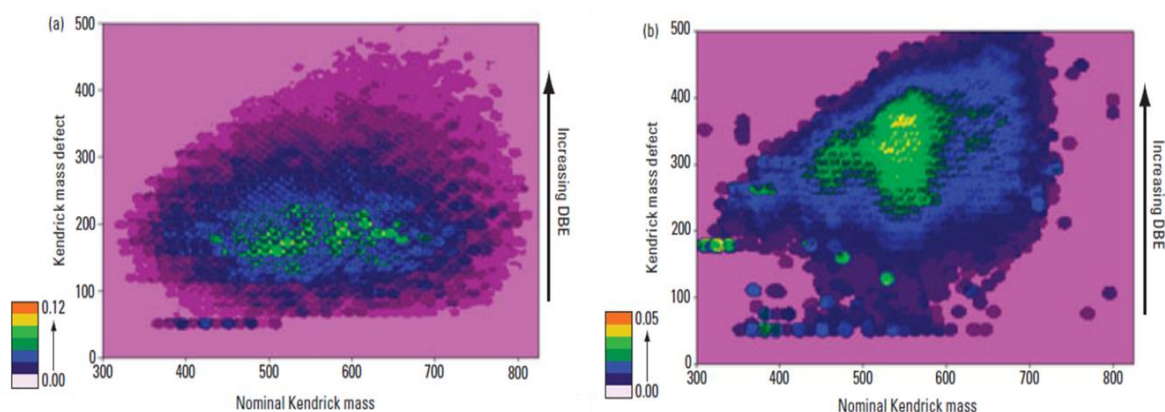


Figure 1-9 : Diagramme de Kendrick en 3 dimensions. (a) échantillon de pétrole Nord-Américain, (b) échantillon du gisement de production associé. Les échelles de couleur représentent les abondances relatives des ions pour chacun des spectres (d'après Rodgers et al., 2005)⁸²

1.3.1.2) Le diagramme de van Krevelen

Le diagramme de van Krevelen a été utilisé pour la première fois pour obtenir une analyse élémentaire visuelle de nombreux échantillons de charbons.⁸³ Le tracé du graphique nécessite d'avoir résolu la composition élémentaire des ions constituant le spectre. L'axe y du diagramme de van Krevelen représente le rapport du nombre d'atomes d'hydrogène sur le nombre d'atomes de carbone (H/C) ; il permet de trier les molécules en fonction de leur degré d'insaturation. L'axe z traduit l'abondance relative des ions qui, en représentation plane, peut être visualisée par une échelle de couleur ou par un diamètre de point. La coordonnée de l'axe x, quant à elle, est choisie sur la base de l'élément chimique X auquel on s'intéresse : nombre d'atome de X / nombre d'atomes de carbone, X pouvant être de l'azote, de l'oxygène ou du soufre. La comparaison de diagrammes de van Krevelen, en mettant en évidence des zones de similarité ou de différence, permet l'observation de tendances quant à l'effet d'un procédé. Ces tendances indiquent des relations structurales entre familles de composés liées par des réactions chimiques (ajout

ou élimination d'un élément).⁸³ La Figure I-10, issue d'un article de Rodgers *et al.*⁸², représente les diagrammes de van Krevelen tracés à partir des spectres de masse d'échantillons de pétroles avant et après dégradation. En classant les composés en fonction du nombre d'atomes d'oxygène qu'ils contiennent, cette représentation illustre le fait que la biodégradation favorise un enrichissement en espèces chimiques incluant 2 atomes d'oxygène (notées « O₂ » sur la Figure I-10).

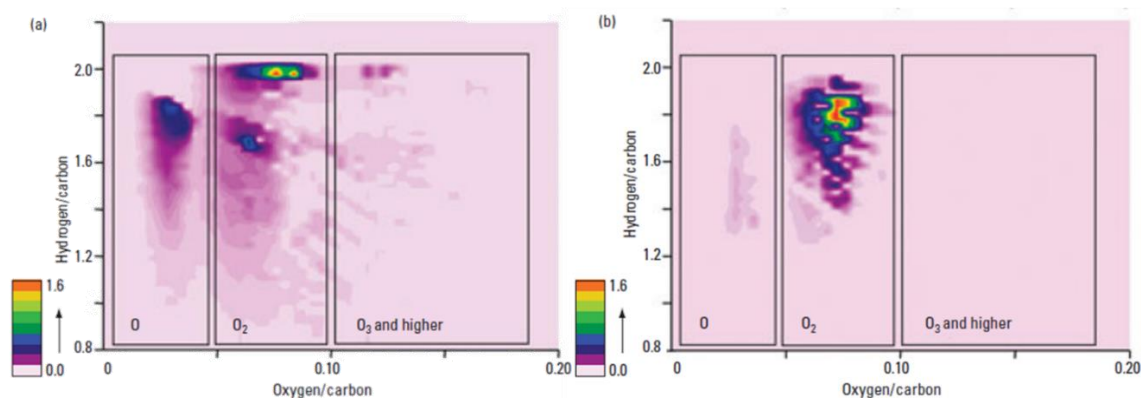


Figure I-10 : Diagrammes de van Krevelen en 3 dimensions d'échantillons de pétroles bruts (a) non dégradés, (b) biodégradés. Ces diagrammes n'incluent que des espèces contenant des atomes d'oxygène. (d'après Rodgers *et al.*, 2005)⁸²

Ces diagrammes ont d'abord été utilisés pour des analyses de pétrologie.⁸⁴⁻⁸⁶ Très rapidement, l'utilisation de ce type de graphique s'est généralisée à d'autres mélanges complexes, par exemple à l'étude d'échantillons environnementaux.⁸⁷⁻⁹² Elle permet également de différencier les origines d'échantillons alimentaires complexes.^{93,94}

La Figure I-11, issue d'un article de Kim *et al.*⁷⁸, représente le diagramme de van Krevelen tracé à partir du spectre de masse d'un échantillon de matière organique dissoute (DOM). Des lignes correspondant à des réactions chimiques se distinguent nettement ; elles permettent d'établir des liens entre les molécules constituant l'échantillon pour mieux le caractériser.

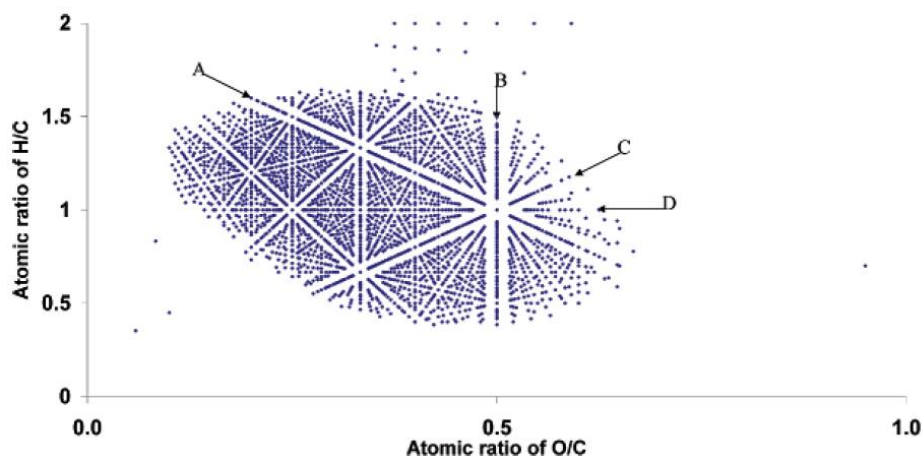


Figure I-11 : Diagramme de van Krevelen d'un échantillon de matière organique dissoute. Les lignes représentent des réactions chimiques (A) méthylation, déméthylation ou élongation de chaîne ; (B) hydrogénation ou déshydrogénation ; (C) hydratation ou déshydratation ; (D) oxydation ou réduction (d'après Kim et al., 2003)⁷⁸

Le diagramme de van Krevelen permet également de différencier des classes de composés présents dans l'échantillon (lipides, protéines, sucres, lignines, tanins...) car chacune de ces familles de molécules va se trouver localisée à un emplacement spécifique du diagramme ; il est ainsi possible d'appréhender la composition globale de l'échantillon. Cette utilisation du diagramme de van Krevelen est encore plus efficace lorsqu'une troisième dimension (intensité des pics) est ajoutée à la représentation, comme illustré Figure I-12 où des huiles d'olive sont comparées en fonction de la variété des olives utilisées.⁹⁵ Les différentes classes de molécules sont aisément reconnaissables et la différenciation entre les variétés d'olives en est grandement facilitée.

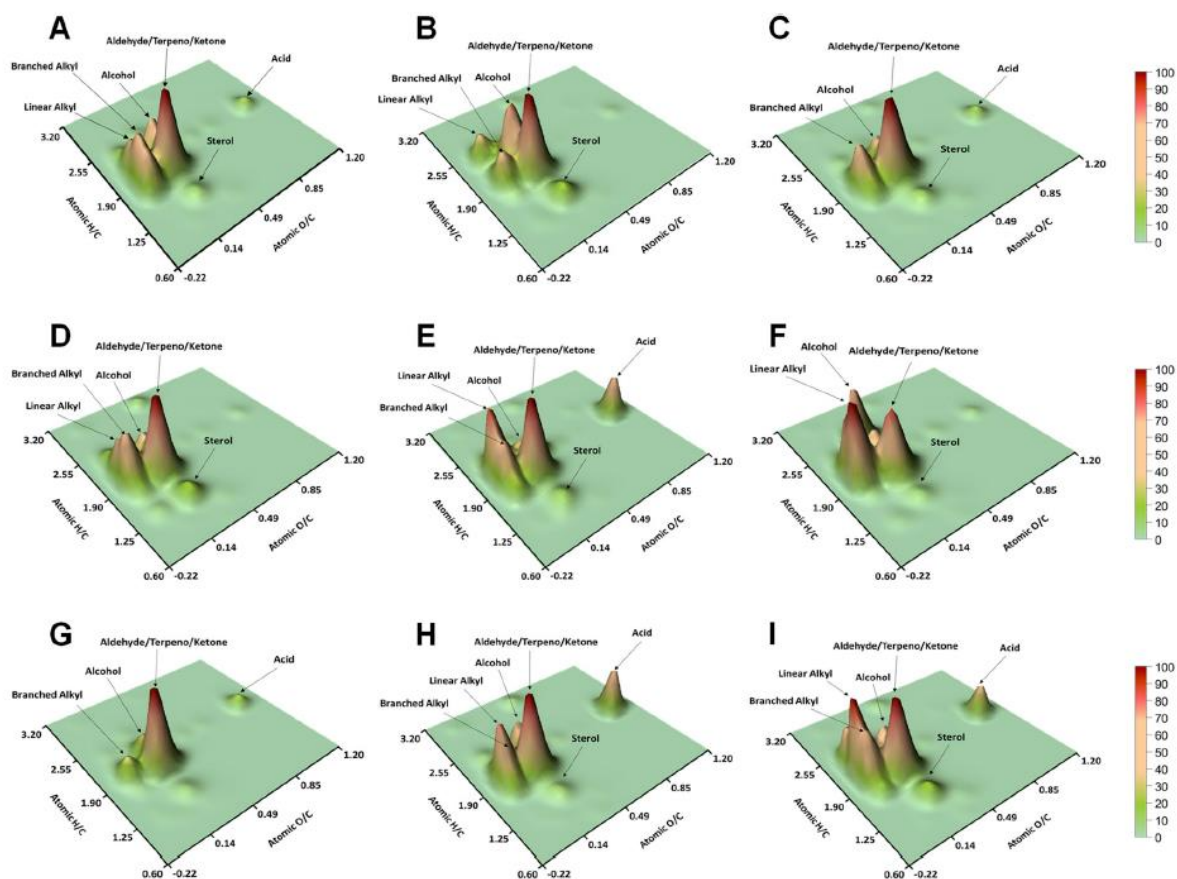


Figure I-12 : Représentation en 3 dimensions de diagrammes de van Krevelen réalisés à partir de spectres de masse d'huiles d'olives originaires du Portugal. Les différentes variétés d'olives représentées sont : A : Galega Vulgar, B: Madural, C : Cobrançosa, D : Carrasquenha, E : Blanqueta, F : Picual, G : Arbequina, H : Cordovil de Serpa et I : Verdeal Alentejana (d'après Martins et al., 2020)⁹⁵

Une autre représentation du diagramme de van Krevelen en 3 dimensions (Figure I-13) a été proposée par Wu *et al.* ; elle consiste à porter les rapports H/C, O/C et N/C sur les axes x, y et z.⁹⁶ Cette représentation fournit une meilleure séparation des différentes classes en termes de compositions élémentaires. La répartition des familles de molécules dans le diagramme est généralement représentée dans différents plans pour en faciliter la visualisation. Les diagrammes de van Krevelen de deux sortes de carburants fossiles, le charbon et l'huile brute, sont comparés Figure I-13 ; la représentation en 3 dimensions montre que l'échantillon de charbon contient plus de molécules aromatiques que l'échantillon d'huile brute.

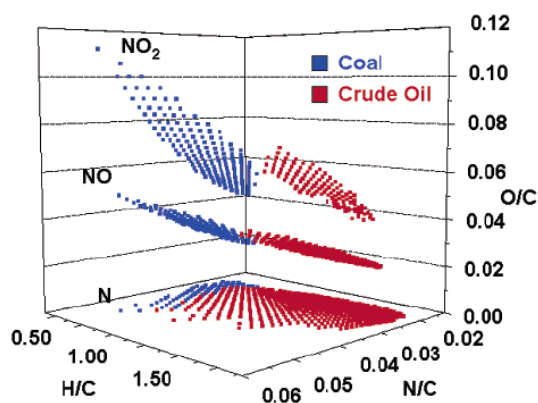


Figure I-13 : Diagrammes de van Krevelen en 3 dimensions des classes N, NO et NO₂ de deux carburants fossiles : le charbon (en bleu) et le pétrole brut (en rouge) (d'après Wu et al., 2004)⁹⁶

Il est généralement pertinent de combiner plusieurs représentations et de les associer à des techniques d'analyses statistiques afin de tirer le maximum d'informations des analyses réalisées sur les échantillons et d'affiner leur différenciation.⁹⁴

I.3.2. Analyse statistique des données

La grande variabilité des individus au sein d'une population rend nécessaire l'analyse d'échantillons de plusieurs individus afin de pouvoir obtenir une conclusion statistique fiable sur ladite population. Les données issues de spectrométrie de masse à haute résolutions contenant souvent plusieurs milliers de pics et les représentations visuelles décrites section I.3.1 ne permettant pas de comparer à l'œil nu un grand nombre d'individus sans introduire de subjectivité opérateur, des outils d'analyse statistique s'avèrent nécessaires. Deux types d'approches sont généralement utilisées pour traiter une telle richesse d'information : les statistiques univariées et multivariées.

Les méthodes d'analyse statistique univariées reposent sur la comparaison de moyennes ou de médianes. Les plus utilisées sont l'analyse de variance (ANOVA pour *analysis of variances*) et le test de Student (t-test). Ces méthodes ne permettent cependant pas de traiter facilement des ensembles de données multidimensionnelles et il peut être utile de les coupler à des analyses statistiques multivariées avant d'envisager l'interprétation

biologique ou chimique des données. Les analyses multivariées les plus utilisées en spectrométrie de masse à haute résolution sont l'analyse en composante principale (ACP) et l'analyse discriminante par régression des moindres carrés (PLS-DA). Les approches univariées et multivariées sont décrites ci-après.

1.3.2.1) Analyses statistiques univariées

Les analyses statistiques univariées sont basées sur la comparaison des moyennes et/ou des médianes de résultats ; elles permettent de mettre en évidence une relation entre deux échantillons pour une variable donnée. Par exemple, une analyse univariée permet d'établir si la présence d'une usine au bord d'une rivière augmente significativement la concentration d'un polluant donné dans l'eau de cette rivière en comparant des populations d'échantillons d'eau prélevés en aval et en amont de l'usine. Les analyses univariées plus communément utilisées en spectrométrie de masse à très haute résolution sont le test de Student (t-test) et l'ANOVA (F-test). Le test de Student est utilisé lorsque la comparaison porte sur les moyennes de deux échantillons d'individus alors que l'approche ANOVA est utilisée lorsque la comparaison porte sur plus de deux échantillons d'individus. Dans le cas où des valeurs aberrantes sont identifiées ou suspectées, le test de Student est remplacé par le test de Wilcoxon qui compare les médianes et non les moyennes.

Ces tests sont basés sur l'hypothèse suivante : si les deux échantillons d'individus proviennent de la même population, les résultats produisent la même moyenne. Si les deux moyennes sont statistiquement différentes, alors les échantillons proviennent de populations différentes. On pose donc l'hypothèse H_0 qui correspond à l'équation 5 pour un test de Student et à l'équation 6 pour un test ANOVA ; la lettre μ désigne la moyenne arithmétique et la lettre (A, B, C...) une population d'échantillons.

$$\mu_A = \mu_B \quad (\text{eq. 5})$$

$$\mu_A = \mu_B = \mu_C \dots \quad (\text{eq. 6})$$

Il existe différents tests de Student, parmi lesquels le t-test non-apparié, utilisé pour comparer deux groupes d'échantillons n'ayant a priori aucun lien, et le t-test apparié, utilisé pour deux groupes d'échantillons présumés dépendants. Dans les deux cas, il s'agit d'un test statistique paramétrique dont la formule dépend de la moyenne et de l'écart-type calculés pour chacun des deux groupes d'échantillons. Pour chaque variable identifiée, la valeur t du test est calculée selon l'équation 7 dans le cas d'échantillons appariés et selon l'équation 8 dans le cas d'échantillons non appariés.

$$t = \frac{m}{s/\sqrt{n}} \quad (\text{eq. 7})$$

avec m : moyenne des différences entre chaque mesure des deux échantillons à comparer

s : écart-type des différences entre chaque mesure des deux échantillons à comparer

n : taille de la série

$$t = \frac{m_A - m_B}{\sqrt{\frac{S^2}{n_A} + \frac{S^2}{n_B}}} \quad (\text{eq. 8})$$

avec m_A et m_B : moyennes des groupes A et B

n_A et n_B : tailles des groupes A et B

S^2 : variance commune aux deux groupes

Le degré de significativité (*p-value*) correspond au risque indiqué par la table de Student pour la valeur absolue de t. Plus la *p-value* (probabilité que l'hypothèse H_0 soit vraie) est faible plus il y a de « chance » que les deux échantillons soient significativement différents. Plus la *p-value* est élevée, plus le risque de commettre une erreur en affirmant que les moyennes des échantillons sont différentes est élevé. Les seuils de *p-value* conventionnels sont généralement fixés à 0,05 ou 0,01 mais il convient d'adapter cette valeur au type de comparaison effectuée : le seuil appliqué à une problématique relative à la santé humaine sera *a priori* inférieur à celui considéré pour l'optimisation d'une synthèse chimique par exemple.

Les résultats d'un t-test peuvent être représentés sous la forme d'un *volcano plot* où est tracée, pour chaque variable, la valeur absolue du logarithme décimal de la *p-value* ($-\log_{10}(p\text{-value})$) en fonction des valeurs du logarithme binaire du *fold change* ($\log_2(\text{fold change})$). Le *fold change* (FC) est le rapport des intensités moyennes des populations comparées : soit A la valeur moyenne d'une variable dans l'échantillon avant traitement et B la valeur moyenne de cette même variable dans l'échantillon après traitement, le *fold change* correspond à B/A. Un graphique de points en forme de volcan apparaît, sur lequel les points les plus élevés sur l'axe des Y sont les plus significatifs et les points les plus éloignés de 0 sur l'axe des X montrent une variabilité plus grande.⁹⁷ Généralement les valeurs intéressantes sont celles ayant une valeur absolue de *fold change* supérieure à 2 (l'intensité moyenne varie au moins d'un facteur 2 entre les deux conditions à comparer) et une *p-value* dont le seuil est déterminé en fonction du taux de faux positifs ou de faux négatifs acceptés. Sur le *volcano plot* représenté Figure I-14, qui représente les résultats d'une comparaison par un test de Student des intensités des lipides observés dans des cellules atteintes de mucoviscidose et dans des cellules saines, les points de couleur rose indiquent les variables dont l'intensité moyenne est réduite pour les cellules atteintes de mucoviscidose comparées aux cellules saines tandis que les points verts représentent les variables dont l'intensité moyenne est supérieure pour les cellules atteintes.⁹⁸ Dans ce cas, la limite de la *p-value* est fixée à 0,05. Ce type de graphique permet de visualiser facilement des tendances lors de la comparaison de deux populations d'échantillons mais il ne permet pas la comparaison de plus de 2 populations.

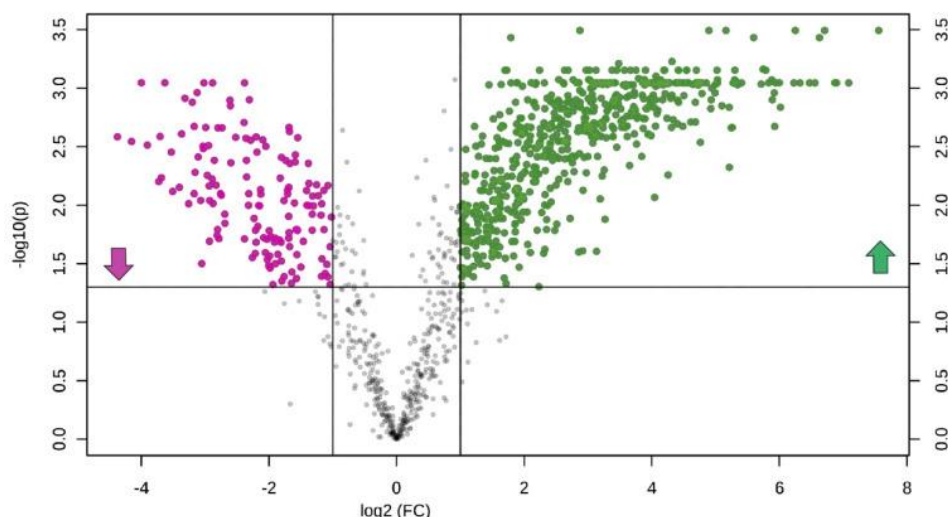


Figure I-14 : Volcano plot correspondant à la combinaison du fold change avec un t-test. Les points de couleur représentent les lipides dont l'intensité varie de façon significative entre les cellules saines et les cellules atteintes de mucoviscidose ($FC > 2$ et $p\text{-value} < 0,05$). Les lipides représentés par des points roses voient leur abondance diminuer dans les cellules atteintes comparées aux cellules saines tandis que les lipides représentés par des points verts voient leur abondance augmenter (d'après Dei Cas et al., 2020).⁹⁸

Les analyses statistiques univariées sont utilisées dans différents domaines tels que la métabolomique^{98,99}, l'analyse de stupéfiants¹⁰⁰ ou les traitements vétérinaires.¹⁰¹ Ce type d'analyse est souvent utilisé pour identifier les variables pour lesquelles il y a une différence significative entre les échantillons. Par la suite, des analyses multivariées y sont associées afin de synthétiser les variables et sélectionner celles susceptibles de devenir des biomarqueurs potentiels. Ces méthodes sont très faciles à appliquer et à interpréter. Cependant, il est fréquent que les variables individuelles ne suffisent plus à caractériser ou interpréter un phénomène. En milieu clinique par exemple, ce sont parfois des études de rapports d'abondance entre deux métabolites qui permettent de diagnostiquer une maladie.¹⁰²

1.3.2.2) Méthodes d'analyse multivariée

Les analyses statistiques multivariées sont un outil de choix pour traiter des jeux de données de grande dimension. De nombreuses méthodes peuvent être utilisées afin d'extraire les informations d'intérêt, réduire un jeu de données, déterminer les sous-classes et comparer des populations d'individus en étudiant les relations entre les

variables. Les analyses multivariées sont séparées en deux groupes : les analyses supervisées et les analyses non supervisées.¹⁰³

Parmi les analyses statistiques supervisées, les analyses PLS (régression par les moindres carrés partiels) et l'analyse discriminante PLS (PLS-DA) sont les plus utilisées dans le cadre d'analyses de mélanges complexes par spectrométrie de masse à haute résolution. Elles sont privilégiées lorsque le nombre d'individus est faible et que le nombre de variables est élevé. Ces méthodes nécessitent une information *à priori*, qui est définie par des propriétés qualitatives spécifiques et préalablement connues. Ces méthodes sont utilisées afin d'affiner la distinction entre les catégories données. La PLS-DA permet de faciliter l'identification des variables responsables de la plus grande discrimination entre les groupes.^{104,105}

Les analyses statistiques non supervisées, quant à elles, ne nécessitent pas d'information préalable sur la nature des échantillons. Ces méthodes sont dites descriptives, elles aident à la visualisation des données. L'analyse en composante principale (ACP) est la plus rencontrée dans le contexte d'analyses de matrices complexes par spectrométrie de masse à haute résolution. Il s'agit d'un algorithme mathématique qui permet de réduire la dimensionnalité des données tout en conservant les informations relatives aux variations les plus importantes. Cette réduction est réalisée en identifiant des directions, appelées composantes principales, suivant lesquelles la variation des données est maximale. A chaque individu est attribué un nombre très restreint de variables, comparé aux centaines ou milliers de variables initiales. Les individus peuvent ainsi être représentés graphiquement, permettant une évaluation visuelle des similitudes et différences entre eux et, potentiellement, leur répartition en catégories.^{106,107} La première composante (PC1) est celle qui explique la plus grande variance parmi l'ensemble des données. La seconde composante (PC2), orthogonale à la première, explique la majeure partie de la variance résiduelle. D'autres composantes (PC3, PC4, etc.) peuvent être considérées mais leur représentation graphique, ajoutée à celles des premières, complique considérablement la visualisation des résultats. C'est pourquoi l'approche ACP est le plus souvent limitée à deux composantes. A partir des résultats de l'ACP, échantillons et variables peuvent être représentés sur une projection bidimensionnelle permettant de regrouper les échantillons par classe, mais également de

déterminer les relations entre échantillons et variables ; un exemple de représentation 2D est donné ci-après. Il importe de considérer que plus le nombre d'individus est grand, plus les erreurs relatives sont faibles ; il est généralement admis qu'un minimum de 10 à 15 individus est requis pour mener une ACP pertinente.

Les analyses statistiques multivariées sont utilisées dans de nombreux domaines tels que l'alimentaire,^{94,108-112} l'analyse d'échantillons environnementaux,¹¹³⁻¹¹⁵ la protéomique,^{116,117} la métabolomique,¹¹⁸⁻¹²⁰ et l'imagerie.^{121,122} Au vu de la grande quantité de variables issues de spectres de masse à haute résolution, il est parfois difficile d'identifier les composantes qui expliquent au mieux les différences entre les populations. Des méthodes de sélection de variables¹²³ ou même des méthodes parcimonieuses¹²⁴ doivent être appliquées afin de supprimer les variables non pertinentes pour la classification ou la prédiction.¹²⁵ Ces méthodes ont prouvé leur efficacité mais doivent être utilisées avec précaution afin d'éviter de perdre des informations intéressantes par sur-ajustement et de traiter correctement les corrélations aléatoires.

En conclusion, les outils d'analyse multivariée permettent d'obtenir une compréhension globale de nombreuses variables en simultané et d'étudier comment elles sont corrélées entre elles. L'ACP est généralement utilisée en premier lieu, pour une première interprétation des données sans *a priori* visant à révéler des groupes. Mener ensuite une approche PLS-DA permet de mieux identifier les variables qui discriminent au mieux ces différents groupes.

Une étude conduite par Roullier-Gall *et al.*, portant sur des analyses par spectrométrie de masse à haute résolution de vins issus de trois crus (2007, 2008 et 2009) et de deux terroirs (Côte de Beaune et Côte de Nuits), illustre parfaitement l'intérêt d'associer les approches ACP et PLS-DA.¹¹¹ Une première analyse en ACP a été effectuée, Figure I-15, qui montre une meilleure séparation des résultats analytiques en fonction des années plutôt que des terroirs. Il apparaît clairement que la première des composantes (axe horizontal) permet la séparation des crus de 2008 et de 2009, tandis que la seconde composante (axe vertical) permet de séparer le cru 2007 des deux autres. A elles deux, ces composantes expliquent 63,9% de la variation globale observée pour l'ensemble des individus.

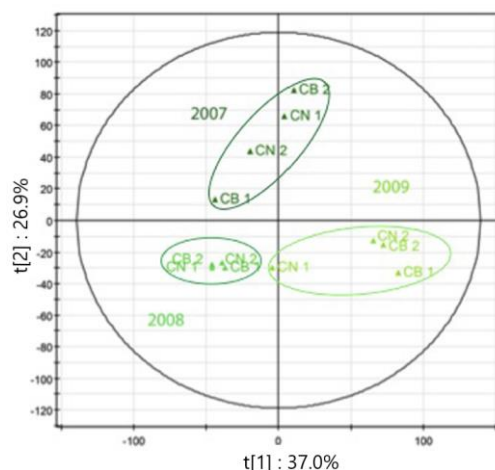


Figure I-15 : Résultat de l'analyse ACP de données FT-ICR de vins CN et CB (CB : Côte de Beaune ; CN : Côte de Nuits) pour différents crus (2007, 2008 et 2009) (d'après Roullier-Gall et al., 2014)¹¹¹

Une analyse statistique supervisée a ensuite été entreprise sur la base de ces résultats : une première PLS-DA pour laquelle les groupes sont les différents crus, une seconde où les groupes sont les terroirs. Les résultats sont présentés Figure I-16. L'image (A) permet de voir que deux composantes suffisent à séparer clairement les trois années. L'image (B) montre que les Côtes de Nuits et les Côtes de Beaune sont séparés seulement par la première composante tandis que la seconde discrimine plutôt l'année.

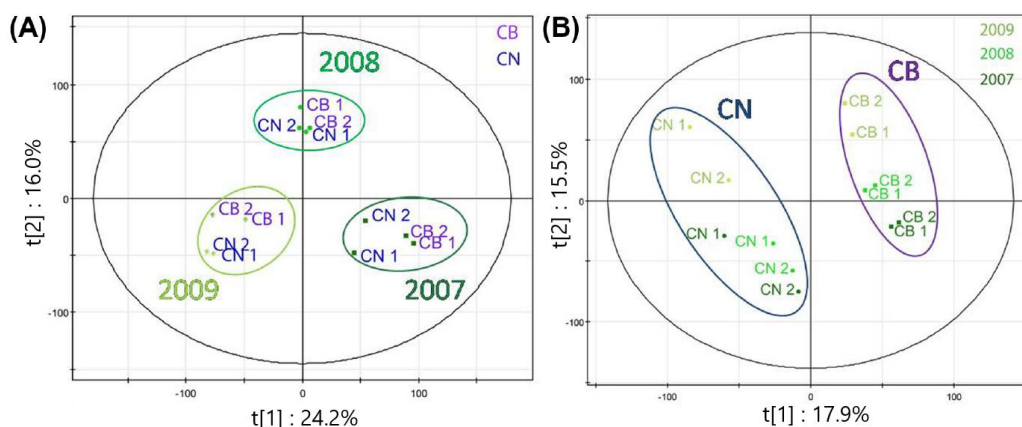


Figure I-16 : Analyse statistique PLS-DA des vins de Côte de Beaune (CB) et Côte de Nuits (CN) après 5, 4 et 3 ans de mise en bouteille pour les crus 2007, 2008 et 2009. (A) graphique des scores de PLS-DA en fonction des années, les deux premières composantes expliquent 40,2% de la variabilité. (B) graphique des scores de PLS-DA en fonction des terroirs, les deux premières composantes expliquent 33,4% de la variabilité totale. (d'après Roullier-Gall et al., 2014).¹¹¹

Ces résultats soulignent le fait que le terroir a un impact certain sur la complexité chimique du vin, mais que le temps, c'est-à-dire le vieillissement en bouteille, est nécessaire pour expliquer la signature chimique des différents vins. Cette étude a été couplée à une projection des rapports m/z ainsi filtrés par PLS-DA sur des diagrammes de van Krevelen, indiquant des signatures chimiques spécifiques des vins de Côte de Beaune et de Côte de Nuits. Dans sa globalité, cette étude montre que, à partir de données issues de HR-MS, la combinaison d'approches statistiques et de représentations graphiques spécifiques permet de discriminer les vins sur la base de critères chimiques liés aux terroirs et millésimes.

De manière générale, la structure des données et la façon dont l'information y est représentée dictent s'il convient d'utiliser des méthodes univariées ou multivariées. Le plus souvent, la façon dont le phénomène étudié s'exprime est inconnu, c'est pourquoi la littérature montre une constante augmentation des recherches combinant l'utilisation de méthodes univariées et multivariées.^{98-101,126} En métabolomique par exemple, les méthodes multivariées sont généralement utilisées pour mettre en évidence des évolutions corrélées d'abondances de différents métabolites alors que les méthodes univariées révèlent des évolutions indépendantes d'abondances de métabolites. Naturellement, la démarche la plus rigoureuse consiste à utiliser la complémentarité des approches ACP et PLS-DA pour l'obtention d'un diagnostic plus global.¹²⁷

RÉFÉRENCES BIBLIOGRAPHIQUES

- (1) Schmitt-Kopplin, P.; Hemmler, D.; Moritz, F.; Gougeon, R. D.; Lucio, M.; Meringer, M.; Muller, C.; Harir, M.; Hertkorn, N. Systems chemical analytics: introduction to the challenges of chemical complexity analysis. *Faraday Discuss* **2019**, *218*, 9-28.
- (30) Lassalle, Y.; Jellouli, H.; Ballerini, L.; Souissi, Y.; Nicol, E.; Bourcier, S.; Bouchonnet, S. Ultraviolet-vis degradation of iprodione and estimation of the acute toxicity of its photodegradation products. *Journal of chromatography. A* **2014**, *1371*, 146-153.
- (31) Souissi, Y.; Kinani, S.; Bouchonnet, S.; Bourcier, S.; Malosse, C.; Sablier, M.; Creusot, N.; Mombelli, E.; Ait-Aissa, S. Photolysis of estrone generates estrogenic photoproducts with higher activity than the parent compound. *Environmental science and pollution research international* **2014**, *21*, 7818-7827.
- (32) Coffinet, S.; Rifai, A.; Genty, C.; Souissi, Y.; Bourcier, S.; Sablier, M.; Bouchonnet, S. Characterization of the photodegradation products of metolachlor: structural elucidation, potential toxicity and persistence. *Journal of mass spectrometry : JMS* **2012**, *47*, 1582-1593.
- (33) Erickson, B. E. Analyzing the ignored environmental contaminants. *Environ Sci Technol* **2002**, *36*, 140A-145A.
- (34) Taheran, M.; Naghdi, M.; Brar, S. K.; Verma, M.; Surampalli, R. Y. Emerging contaminants: Here today, there tomorrow! *Environmental Nanotechnology, Monitoring & Management* **2018**, *10*, 122-126.
- (35) Snow, D. D.; Bartelt-Hunt, S. L.; Saunders, S. E.; Cassada, D. A. Detection, Occurrence, and Fate of Emerging Contaminants in Agricultural Environments. *Water Environment Research* **2007**, *79*, 1061-1084.
- (36) Large, M. J. The Federal Environmental Pesticide Control Act of 1972: A Compromise Approach. *Ecology Law Quarterly* **1973**, *3*, 277-310.
- (37) Segura, P. A.; Yargeau, V.; Gagnon, C. On the Application of High-Resolution Mass Spectrometry to Environmental Analysis. *International Journal of Environmental Pollution and Remediation* **2012**, *1*, 82-89.
- (38) Barcelo, D.; Petrovic, M. Challenges and achievements of LC-MS in environmental analysis: 25 years on. *Trac-Trend Anal Chem* **2007**, *26*, 2-11.
- (39) Melin, J.; Guillon, A.; Enault, J.; Esperanza, M.; Dauchy, X.; Bouchonnet, S. How to select relevant metabolites based on available data for parent molecules: Case of neonicotinoids, carbamates, phenylpyrazoles and organophosphorus compounds in French water resources. *Environ Pollut* **2020**, *265*, 114992.
- (40) Boxall, A. B. A.; Sinclair, C. J.: Assessment of the environmental properties and effects of pesticide degradation products. In *Pesticide behaviour in soils and water. Symposium (Brighton 2001-11-13)*; BCPC, Farnham, 2001; pp 113-118.
- (41) Castillo, M.; Domingues, R.; Alpendurada, M. F.; Barcelo, D. Persistence of selected pesticides and their phenolic transformation products in natural waters using off-line liquid solid extraction followed by liquid chromatographic techniques. *Analytica Chimica Acta* **1997**, *353*, 133-142.
- (42) Bottaro, M.; Frascarolo, P.; Gosetti, F.; Mazzucco, E.; Gianotti, V.; Polati, S.; Pollici, E.; Piacentini, L.; Pavese, G.; Gennaro, M. C. Hydrolytic and photoinduced

degradation of tribenuron methyl studied by HPLC-DAD-MS/MS. *Journal of the American Society for Mass Spectrometry* **2008**, *19*, 1221-1229.

(43) Gosetti, F.; Bottaro, M.; Gianotti, V.; Mazzucco, E.; Frascarolo, P.; Zampieri, D.; Oliveri, C.; Viarengo, A.; Gennaro, M. C. Sun light degradation of 4-chloroaniline in waters and its effect on toxicity. A high performance liquid chromatography - Diode array - Tandem mass spectrometry study. *Environ Pollut* **2010**, *158*, 592-598.

(44) Gosetti, F.; Chiuminatto, U.; Zampieri, D.; Mazzucco, E.; Marengo, E.; Gennaro, M. C. A new on-line solid phase extraction high performance liquid chromatography tandem mass spectrometry method to study the sun light photodegradation of mono-chloroanilines in river water. *Journal of chromatography. A* **2010**, *1217*, 3427-3434.

(45) Gosetti, F.; Chiuminatto, U.; Mazzucco, E.; Mastroianni, R.; Bolfi, B.; Marengo, E. Ultra-high performance liquid chromatography tandem high-resolution mass spectrometry study of tricyclazole photodegradation products in water. *Environmental science and pollution research international* **2015**, *22*, 8288-8295.

(46) Aguera, A.; Martinez Bueno, M. J.; Fernandez-Alba, A. R. New trends in the analytical determination of emerging contaminants and their transformation products in environmental waters. *Environmental science and pollution research international* **2013**, *20*, 3496-3515.

(47) NORMAN Network. <https://www.norman-network.net/?q=Home> (accessed September 2020).

(48) Directive 98/83/CE du conseil du 3 novembre 1998. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1998L0083:20090807:FR:PDF> (accessed September 2020).

(49) Richard, C.; Sarakha, M.; Ter Halle, A.; Mazellier, P.; Chovelon, J.-M. La lumière solaire à l'assaut des polluants. *Actualite Chimique* **2007**, *308-309*, 71-75.

(50) Burrows, H. D.; Canle, L. M.; Santaballa, J. A.; Steenken, S. Reaction pathways and mechanisms of photodegradation of pesticides. *Journal of photochemistry and photobiology. B, Biology* **2002**, *67*, 71-108.

(51) Kiss, A.; Virag, D. Interpretation and modelling of environmental behaviour of diverse pesticides by revealing photodecomposition mechanisms. *Microchemical Journal* **2009**, *92*, 119-122.

(52) Vione, D.; Minella, M.; Maurino, V.; Minero, C. Indirect photochemistry in sunlit surface waters: photoinduced production of reactive transient species. *Chemistry* **2014**, *20*, 10590-10606.

(53) Guedira, F. Cours de Spectroscopie-Chapitre III : spectroscopie d'absorption dans l'UV-visible. *Faculté des Sciences de Rabat* **2013**.

(54) Katagi, T. Photodegradation of pesticides on plant and soil surfaces. *Rev Environ Contam Toxicol* **2004**, *182*, 1-189.

(55) *Surface water photochemistry*; Calza, P.; Vione, D., Eds.; Royal Society of Chemistry: Cambridge, U. K., 2016; Vol. 12, pp 297.

(56) Galindo, F. The photochemical rearrangement of aromatic ethers - A review of the Photo-Claisen reaction. *J Photoch Photobio C* **2005**, *6*, 123-138.

(57) Cazzaniga, N.; Varga, Z.; Nicol, E.; Bouchonnet, S. UV-visible photodegradation of naproxen in water - Structural elucidation of photoproducts and potential toxicity. *European journal of mass spectrometry (Chichester, England)* **2020**, *26*, 400-408.

- (58) Junot, C.; Madalinski, G.; Tabet, J. C.; Ezan, E. Fourier transform mass spectrometry for metabolome analysis. *The Analyst* **2010**, *135*, 2203-2219.
- (59) Witting, M.; Lucio, M.; Tziotis, D.; Wagele, B.; Suhre, K.; Voulhoux, R.; Garvis, S.; Schmitt-Kopplin, P. DI-ICR-FT-MS-based high-throughput deep metabotyping: a case study of the *Caenorhabditis elegans*-*Pseudomonas aeruginosa* infection model. *Analytical and bioanalytical chemistry* **2015**, *407*, 1059-1073.
- (60) Aharoni, A.; Ric de Vos, C. H.; Verhoeven, H. A.; Maliepaard, C. A.; Kruppa, G.; Bino, R.; Goodenowe, D. B. Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. *Omics : a journal of integrative biology* **2002**, *6*, 217-234.
- (61) Kendrick, E. A Mass Scale Based on CH₂ = 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. *Analytical chemistry* **1963**, *35*, 2146-2154.
- (62) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Analytical chemistry* **2001**, *73*, 4676-4681.
- (63) Marshall, A. G.; Rodgers, R. P. Petroleomics: the next grand challenge for chemical analysis. *Accounts of chemical research* **2004**, *37*, 53-59.
- (64) Stenson, A. C.; Landing, W. M.; Marshall, A. G.; Cooper, W. T. Ionization and fragmentation of humic substances in electrospray ionization Fourier transform-ion cyclotron resonance mass spectrometry. *Analytical chemistry* **2002**, *74*, 4397-4409.
- (65) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Analytical chemistry* **2003**, *75*, 1275-1284.
- (66) Kramer, R. W.; Kujawinski, E. B.; Hatcher, P. G. Identification of black carbon derived structures in a volcanic ash soil humic acid by Fourier transform ion cyclotron resonance mass spectrometry. *Environ Sci Technol* **2004**, *38*, 3387-3395.
- (67) Sleighter, R. L.; Hatcher, P. G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *Journal of mass spectrometry : JMS* **2007**, *42*, 559-574.
- (68) Kujawinski, E. B.; Hatcher, P. G.; Freitas, M. A. High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: improvements and comparisons. *Analytical chemistry* **2002**, *74*, 413-419.
- (69) D'Auria, M.; Emanuele, L.; Racioppi, R. FT-ICR-MS analysis of lignin. *Natural product research* **2012**, *26*, 1368-1374.
- (70) Haiying, Z.; Donglu, Z.; Ray, K.; Mingshe, Z. H. U. Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry. *Journal of mass spectrometry* **2009**, *44*, 999-1016.
- (71) Shumao, N.; Dawei, Q.; Duan, J.-A.; Jianming, G.; Shang, E.-X.; Yan, S.; Caifu, X. UPLC-QTOF/MS-based screening and identification of the constituents and their metabolites in rat plasma and urine after oral administration of *Glechoma longituba* extract. *Journal of chromatography. B* **2010**, *878*, 2741-2750.
- (72) Chu, F. L.; Pirastru, L.; Popovic, R.; Sleno, L. Carotenogenesis up-regulation in *Scenedesmus* sp. using a targeted metabolomics approach by liquid chromatography-high-resolution mass spectrometry. *Journal of agricultural and food chemistry* **2011**, *59*, 3004-3013.

(73) Helmer, P. O.; Korf, A.; Hayen, H. Analysis of artificially oxidized cardiolipins and monolyso-cardiolipins via liquid chromatography/high-resolution mass spectrometry and Kendrick mass defect plots after hydrophilic interaction liquid chromatography based sample preparation. *Rapid communications in mass spectrometry : RCM* **2020**, *34*, e8566.

(74) Bruce, C.; Shifman, M. A.; Miller, P.; Gulcicek, E. E. Probabilistic enrichment of phosphopeptides by their mass defect. *Analytical chemistry* **2006**, *78*, 4374-4382.

(75) Sleno, L. The use of mass defect in modern mass spectrometry. *Journal of Mass Spectrometry* **2012**, *47*, 226-236.

(76) Jobst, K. J.; Shen, L.; Reiner, E. J.; Taguchi, V. Y.; Helm, P. A.; McCrindle, R.; Backus, S. The use of mass defect plots for the identification of (novel) halogenated contaminants in the environment. *Analytical and bioanalytical chemistry* **2013**, *405*, 3289-3297.

(77) Taguchi, V. Y.; Nieckarz, R. J.; Clement, R. E.; Krolik, S.; Williams, R. Dioxin analysis by gas chromatography-Fourier transform ion cyclotron resonance mass spectrometry (GC-FTICRMS). *Journal of the American Society for Mass Spectrometry* **2010**, *21*, 1918-1921.

(78) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Analytical chemistry* **2003**, *75*, 5336-5344.

(79) Kind, T.; Fiehn, O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC bioinformatics* **2006**, *7*, 234.

(80) Ajaero, C.; McMartin, D. W.; Peru, K. M.; Bailey, J.; Haakensen, M.; Friesen, V.; Martz, R.; Hughes, S. A.; Brown, C.; Chen, H.; McKenna, A. M.; Corilo, Y. E.; Headley, J. V. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Characterization of Athabasca Oil Sand Process-Affected Waters Incubated in the Presence of Wetland Plants. *Energy and fuels* **2017**, *31*, 1731-1740.

(81) Ajaero, C.; Peru, K. M.; Hughes, S. A.; Chen, H.; McKenna, A. M.; Corilo, Y. E.; McMartin, D. W.; Headley, J. V. Atmospheric Pressure Photoionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Characterization of Oil Sand Process-Affected Water in Constructed Wetland Treatment. *Energy and fuels* **2019**, *33*, 4420-4431.

(82) Rodgers, R. P.; Schaub, T. M.; Marshall, A. G. Petroleomics: MS returns to its roots. *Analytical chemistry* **2005**, *77*, 20a-27a.

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

(84) Hatcher, P. G.; Lerch, H. E.; Bates, A. L.; Verheyen, T. V. Solid-state ¹³C nuclear magnetic resonance studies of coalified gymnosperm xylem tissue from Australian brown coals. *Organic Geochemistry* **1989**, *14*, 145-155.

(85) Bostick, N. H.; Daws, T. A. relationships between data from Rock-Eval pyrolysis and proximate, ultimate, petrographic, and physical analyses of 142 diverse US coal samples. *Organic geochemistry* **1994**, *21*, 35-49.

(86) Curiale, J. A.; Gibling, M. R. Productivity Control on Oil-Shale Formation - Mae-Sot-Basin, Thailand. *Organic Geochemistry* **1994**, *21*, 67-89.

(87) Visser, S. A. Application of Van Krevelen's graphical-statistical method for the study of aquatic humic material. *Environ Sci Technol* **1983**, *17*, 412-417.

- (88) Hertkorn, N.; Benner, R.; Frommberger, M.; Schmitt-Kopplin, P.; Witt, M.; Kaiser, K.; Kettrup, A.; Hedges, J. I. Characterization of a major refractory component of marine dissolved organic matter. *Geochimica Et Cosmochimica Acta* **2006**, *70*, 2990-3010.
- (89) Wozniak, A. S.; Bauer, J. E.; Sleighter, R. L.; Dickhut, R. M.; Hatcher, P. G. Technical Note: Molecular characterization of aerosol-derived water soluble organic carbon using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Atmospheric Chemistry and Physics* **2008**, *8*, 5099-5111.
- (90) D'Andrilli, J.; Foreman, C. M.; Marshall, A. G.; McKnight, D. M. Characterization of IHSS Pony Lake fulvic acid dissolved organic matter by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and fluorescence spectroscopy. *Organic Geochemistry* **2013**, *65*, 19-28.
- (91) Minor, E. C.; Swenson, M. M.; Mattson, B. M.; Oyler, A. R. Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environmental science. Processes & impacts* **2014**, *16*, 2064-2079.
- (92) Bianco, A.; Deguillaume, L.; Vaitilingom, M.; Nicol, E.; Baray, J. L.; Chaumerliac, N.; Bridoux, M. Molecular Characterization of Cloud Water Samples Collected at the Puy de Dome (France) by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environ Sci Technol* **2018**, *52*, 10275-10285.
- (93) Liger-Belair, G.; Cilindre, C.; Gougeon, R. D.; Lucio, M.; Gebefugi, I.; Jeandet, P.; Schmitt-Kopplin, P. Unraveling different chemical fingerprints between a champagne wine and its aerosols. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 16545-16549.
- (94) Kew, W.; Goodall, I.; Clarke, D.; Uhrin, D. Chemical Diversity and Complexity of Scotch Whisky as Revealed by High-Resolution Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 200-213.
- (95) Martins, N.; Jimenez-Morillo, N. T.; Freitas, F.; Garcia, R.; Gomes da Silva, M.; Cabrita, M. J. Revisiting 3D van Krevelen diagrams as a tool for the visualization of volatile profile of varietal olive oils from Alentejo region, Portugal. *Talanta* **2020**, *207*, 120276.
- (96) Wu, Z.; Rodgers, R. P.; Marshall, A. G. Two- and three-dimensional van krevelen diagrams: a graphical analysis complementary to the kendrick mass plot for sorting elemental compositions of complex organic mixtures based on ultrahigh-resolution broadband fourier transform ion cyclotron resonance mass measurements. *Analytical chemistry* **2004**, *76*, 2511-2516.
- (97) Bell, M.; Blais, J. M. "-Omics" workflow for paleolimnological and geological archives: A review. *The Science of the total environment* **2019**, *672*, 438-455.
- (98) Dei Cas, M.; Zulueta, A.; Mingione, A.; Caretti, A.; Ghidoni, R.; Signorelli, P.; Paroni, R. An Innovative Lipidomic Workflow to Investigate the Lipid Profile in a Cystic Fibrosis Cell Line. *Cells* **2020**, *9*, 1197-1204.
- (99) Winter, H.; Kaisaki, P. J.; Harvey, J.; Giacomuzzi, E.; Ferla, M. P.; Pentony, M. M.; Knight, S. J. L.; Sharma, R. A.; Taylor, J. C.; McCullagh, J. S. O. Identification of Circulating Genomic and Metabolic Biomarkers in Intrahepatic Cholangiocarcinoma. *Cancers (Basel)* **2019**, *11*, 1895-1913.
- (100) Palmieri, S.; Mascini, M.; Ricci, A.; Fanti, F.; Ottaviani, C.; Lo Sterzo, C.; Sergi, M. Identification of Cannabis sativa L. (hemp) Retailers by Means of Multivariate Analysis of Cannabinoids. *Molecules* **2019**, *24*, 3602-3611.

(101) Gaugain, M.; Mompelat, S.; Fourmond, M. P.; Manceau, J.; Rolland, J. G.; Laurentie, M.; Verdon, E.; Bellanger, L.; Hurtaud-Pessel, D. A non-targeted LC-HRMS approach for detecting exposure to illegal veterinary treatments: The case of cephalosporins in commercial laying Hens. *Journal of chromatography. A* **2019**, *1599*, 161-171.

(102) de Sain-van der Velden, M. G.; Rinaldo, P.; Elvers, B.; Henderson, M.; Walter, J. H.; Prinsen, B. H.; Verhoeven-Duif, N. M.; de Koning, T. J.; van Hasselt, P. The Proline/Citrulline Ratio as a Biomarker for OAT Deficiency in Early Infancy. *JIMD Rep* **2012**, *6*, 95-99.

(103) Kemsley, E. K. Discriminant analysis of high-dimensional data: A comparison of principal components analysis and partial least squares data reduction methods. *Chemometrics and Intelligent Laboratory Systems* **1996**, *33*, 47-61.

(104) Bylesjo, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J. K.; Holmes, E.; Trygg, J. OPLS discriminant analysis : combining the strengths of PLS-DA and SIMCA classification. *Journal of chemometrics* **2007**, *20*, 341-351.

(105) Trygg, J.; Holmes, E.; Lundstedt, T. Chemometrics in metabonomics. *J Proteome Res* **2007**, *6*, 469-479.

(106) Abdi, H.; Williams, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* **2010**, *2*, 433-459.

(107) Ringner, M. What is principal component analysis? *Nat Biotechnol* **2008**, *26*, 303-304.

(108) Callao, M. P.; Ruisanchez, I. An overview of multivariate qualitative methods for food fraud detection. *Food Control* **2018**, *86*, 283-293.

(109) Marti, M. P.; Busto, O.; Guasch, J. Application of a headspace mass spectrometry system to the differentiation and classification of wines according to their origin, variety and ageing. *Journal of Chromatography A* **2004**, *1057*, 211-217.

(110) Rubert, J.; Lacina, O.; Zachariasova, M.; Hajslova, J. Saffron authentication based on liquid chromatography high resolution tandem mass spectrometry and multivariate data analysis. *Food Chem* **2016**, *204*, 201-209.

(111) Roullier-Gall, C.; Boutegrabet, L.; Gougeon, R. D.; Schmitt-Kopplin, P. A grape and wine chemodiversity comparison of different appellations in Burgundy: vintage vs terroir effects. *Food Chem* **2014**, *152*, 100-107.

(112) Kenar, A.; Çiçek, B.; Arslan, F. N.; Akin, G.; Elmas, Ş. N. K.; Yilmaz, I. Electron Impact–Mass Spectrometry Fingerprinting and Chemometrics for Rapid Assessment of Authenticity of Edible Oils Based on Fatty Acid Profiling. *Food Analytical Methods* **2019**, *12*, 1369-1381.

(113) Corilo, Y. E.; Podgorski, D. C.; McKenna, A. M.; Lemkau, K. L.; Reddy, C. M.; Marshall, A. G.; Rodgers, R. P. Oil spill source identification by principal component analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Analytical chemistry* **2013**, *85*, 9064-9069.

(114) Karpuzcu, M. E.; Fairbairn, D.; Arnold, W. A.; Barber, B. L.; Kaufenberg, E.; Koskinen, W. C.; Novak, P. J.; Rice, P. J.; Swackhamer, D. L. Identifying sources of emerging organic contaminants in a mixed use watershed using principal components analysis. *Environmental science. Processes & impacts* **2014**, *16*, 2390-2399.

(115) Sleighter, R. L.; Liu, Z.; Xue, J.; Hatcher, P. G. Multivariate statistical approaches for the characterization of dissolved organic matter analyzed by ultrahigh resolution mass spectrometry. *Environ Sci Technol* **2010**, *44*, 7576-7582.

(116) Gaspari, M.; Verhoeckx, K. C.; Verheij, E. R.; van der Greef, J. Integration of two-dimensional LC-MS with multivariate statistics for comparative analysis of proteomic samples. *Analytical chemistry* **2006**, *78*, 2286-2296.

(117) Wang, X.; Chambers, M. C.; Vega-Montoto, L. J.; Bunk, D. M.; Stein, S. E.; Tabb, D. L. QC metrics from CPTAC raw LC-MS/MS data interpreted through multivariate statistics. *Analytical chemistry* **2014**, *86*, 2497-2509.

(118) Wang, C.; Kong, H.; Guan, Y.; Yang, J.; Gu, J.; Yang, S.; Xu, G. Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. *Analytical chemistry* **2005**, *77*, 4108-4116.

(119) Tsugawa, H.; Tsujimoto, Y.; Arita, M.; Bamba, T.; Fukusaki, E. GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC bioinformatics* **2011**, *12*, 131.

(120) Kiss, A.; Lucio, M.; Fildier, A.; Buisson, C.; Schmitt-Kopplin, P.; Cren-Olive, C. Doping control using high and ultra-high resolution mass spectrometry based non-targeted metabolomics-a case study of salbutamol and budesonide abuse. *PLoS One* **2013**, *8*, e74584.

(121) Dill, A. L.; Eberlin, L. S.; Zheng, C.; Costa, A. B.; Ifa, D. R.; Cheng, L.; Masterson, T. A.; Koch, M. O.; Vitek, O.; Cooks, R. G. Multivariate statistical differentiation of renal cell carcinomas based on lipidomic analysis by ambient ionization imaging mass spectrometry. *Analytical and bioanalytical chemistry* **2010**, *398*, 2969-2978.

(122) Alexandrov, T. MALDI imaging mass spectrometry: statistical data analysis and current computational challenges. *BMC bioinformatics* **2012**, *13*, S11.

(123) Peres, F. A. P.; Fogliatto, F. S. Variable selection methods in multivariate statistical process control: A systematic literature review. *Computers & Industrial Engineering* **2018**, *115*, 603-619.

(124) Zou, H.; Hastie, T.; Tibshirani, R. Sparse principal component analysis. *Journal of Computational and Graphical Statistics* **2006**, *15*, 265-286.

(125) Filzmoser, P.; Gschwandtner, M.; Todorov, V. Review of sparse methods in regression and classification with application to chemometrics. *Journal of Chemometrics* **2012**, *26*, 42-51.

(126) Sachse, D.; Sletner, L.; Morkrid, K.; Jenum, A. K.; Birkeland, K. I.; Rise, F.; Piehler, A. P.; Berg, J. P. Metabolic changes in urine during and after pregnancy in a large, multiethnic population-based cohort study of gestational diabetes. *PLoS One* **2012**, *7*, e52399.

(127) Saccenti, E.; Hoefsloot, H. C. J.; Smilde, A. K.; Westerhuis, J. A.; Hendriks, M. M. W. B. Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics* **2014**, *10*, 361-374.

Chapitre II - MATERIELS ET METHODES

Ce chapitre présente les outils et techniques utilisés lors de ce travail de recherche. La plupart d'entre eux sont décrits assez succinctement dans la mesure où ils sont largement connus au sein de la communauté scientifique. Certains sont également présentés dans les publications qui constituent le chapitre III. La présentation du logiciel de comparaison statistique de spectres (SPIX) ainsi que les résultats obtenus avec ce dernier sont décrits au chapitre IV.

Ce chapitre est divisé en quatre parties. La première présente les photoréacteurs utilisés dans le cadre de cette étude tandis que la seconde décrit les protocoles de préparation des échantillons. Les techniques utilisées pour l'évaluation de la toxicité des molécules-mères et des produits de dégradation sont détaillées dans la troisième partie. Enfin, la dernière partie porte sur les techniques physico-chimiques d'analyse utilisées pour la détection et la caractérisation des produits de dégradation.

II.1. PHOTOREACTEURS UTILISES DANS CETTE ETUDE

Deux types de photoréacteurs UV-Visible ont été utilisés pour ce travail : un photoréacteur à lampe à vapeur de mercure et une enceinte de vieillissement accélérée à lampe à arc de xénon.

II.1.1. Photoréacteur à lampe à vapeur de mercure

Ce photoréacteur est un prototype conçu au laboratoire pour permettre l'irradiation simultanée de six solutions (voir Figure II-1). Il est constitué d'un support pouvant accueillir jusqu'à sept tubes en quartz, placé dans bac à ultrasons pour permettre le dégazage et l'homogénéisation des échantillons. Le tube en quartz situé au centre du dispositif contient la lampe à vapeur de mercure. Les échantillons sont refroidis par une circulation d'eau de ville au sein du bac à ultrasons ; l'ensemble du montage est disposé dans un caisson de protection mécaniquement ventilé. Les échantillons sont placés dans les six tubes en quartz, à égale distance de la lampe à vapeur de mercure. Le quartz a été

retenu pour cette étude car, contrairement au pyrex, il ne filtre pas les rayonnements UV. Deux formats de tubes ont été utilisés au cours de ces études (2 ml et 100 ml) en fonction du volume de solution à irradier, de sorte que la surface d'échantillon exposée à la lampe soit maximale.



Figure II-1 : Prototype de photoréacteur à lampe à vapeur de mercure.

Ce photoréacteur est équipé d'une lampe à vapeur de mercure (HPL-N 125W/542 E27 SG, Philips, Ivry-sur-seine, France). Cette dernière est constituée d'un tube de quartz contenant deux électrodes et du mercure. Une décharge électrique entre les électrodes excite les atomes de mercure qui émettent des photons lors de leur retour à l'état fondamental. Le spectre d'émission obtenu couvre des longueurs d'onde comprises entre 200 et 650 nm, la raie la plus intense se trouvant à 360 nm. Le spectre de cette lampe est relativement éloigné du spectre solaire mais plusieurs études ont montré que les photoproduits obtenus avec cette lampe sont les mêmes que ceux obtenus avec une enceinte de vieillissement dont le spectre est très proche de celui du soleil.^{29,30,128,129} Le flux lumineux de cette lampe (6200 lm) permet de dégrader des polluants sur des temps d'irradiation très courts (quelques minutes à quelques heures) en comparaison des temps généralement utilisés avec les enceintes de vieillissement commerciales (quelques jours) comme celle décrite ci-après.

II.1.2. Enceinte de vieillissement accéléré à lampe à arc de Xénon

L'utilisation d'une enceinte de vieillissement accéléré (Q-SUN Xe-1, Q-LAB, Saarbrücken, Allemagne) permet une meilleure simulation du spectre solaire. Ce type d'enceinte (voir Figure II-2) a été initialement conçu pour simuler en accéléré l'altération climatique de matériaux sous l'effet de contraintes naturelles telles que le rayonnement solaire, l'humidité ou la chaleur.



Figure II-2 : Photographie du Q-SUN Xe-1

Cette enceinte est munie d'une lampe à arc de xénon (X-1800, Q-LAB, Saarbrücken, Allemagne) à laquelle on associe un filtre « lumière naturelle » (X-7640, Q-LAB, Saarbrücken, Allemagne) pour mimer au mieux la lumière solaire. Cette lampe est placée au-dessus des échantillons à irradier ; l'intérieur de l'enceinte est constitué de surfaces réfléchissantes pour que l'irradiation soit la plus homogène possible. Régulée par un capteur de température interne sur panneau noir, une ventilation permet de contrôler la température des échantillons. Ces derniers sont placés dans des tubes en quartz fermés afin d'éviter l'évaporation de la solution pendant la durée de l'irradiation (quelques heures à quelques jours).

Le principe de fonctionnement de la lampe à arc de xénon est similaire à celui de la lampe à vapeur de mercure mais son spectre d'émission est plus proche de la lumière du soleil comme illustré Figure II-3. La puissance d'irradiation a été fixée à $0,68 \text{ W/m}^2$ à la longueur d'onde de 340 nm afin de pouvoir contrôler la température de l'enceinte.¹³⁰

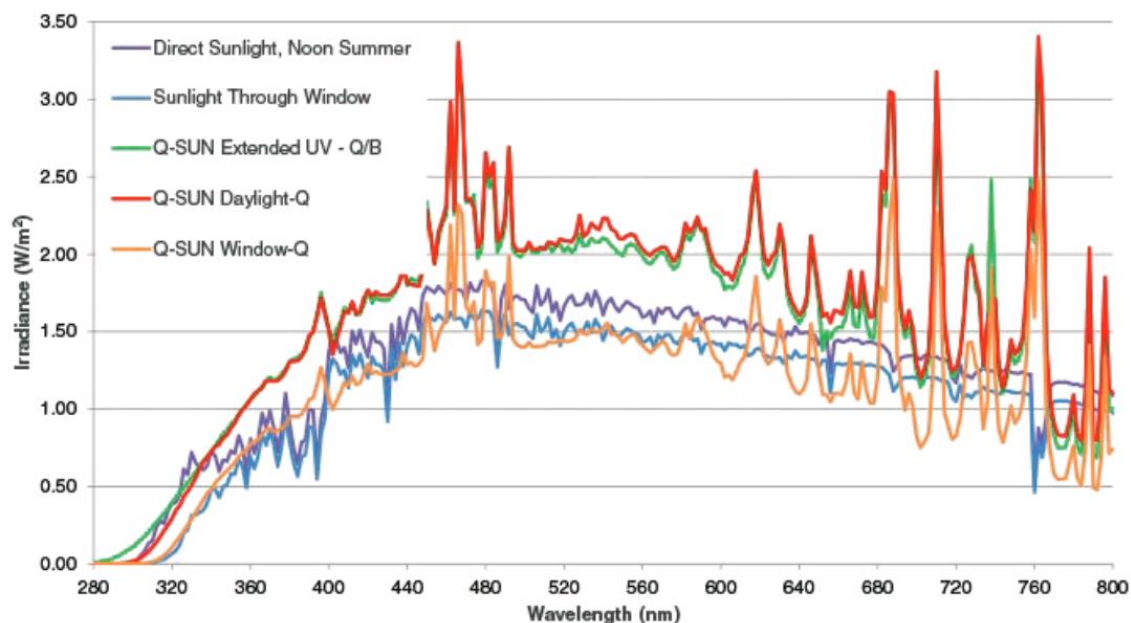


Figure II-3 : Spectre d'émission solaire à midi en été (violet) et comparaison avec la lampe à arc de Xénon équipée de filtre de simulation de la lumière solaire.¹³¹

II.2. PREPARATION DES ECHANTILLONS

Les produits de dégradation des polluants étudiés lors de ces travaux étant inconnus, la préparation d'échantillon a été limitée au strict minimum de manière à ne pas introduire trop de sélectivité et de subjectivité dans la démarche analytique. Nous avons dilué les échantillons dans l'eau ultra pure avec un minimum de solvant organique (nécessaire lorsque le produit est insoluble dans l'eau aux concentrations étudiées) en tentant de s'approcher des concentrations auxquelles sont généralement détectés les polluants en milieu naturel. Nous avons dû cependant travailler à des concentrations plus élevées que celles mesurées en milieu naturel car la photolyse induit généralement de nombreux photoproduits dont de nombreux isomères. Les concentrations individuelles des

photoproduits sont de fait susceptibles d'être très faibles, voire inférieures aux limites de détection des spectromètres de masse les plus sensibles. Pour les études de mélanges commerciaux, les principes actifs ont été irradiés isolément, préalablement à l'étude des mélanges eux-mêmes, afin de caractériser un éventuel effet de la matrice sur la dégradation photochimique des composés.

II.2.1. Préparation des solutions pour la caractérisation de photoproduits par LC-MS et GC-MS

Quatre polluants ont été étudiés lors de cette thèse : le cyprodinil, la benzisothiazolinone, l'acétamipride (seul et dans les mélanges commerciaux Roseclear Ultra® et Polysect Ultra SL®) et le naproxène. Ces études ont fait l'objet de publications qui sont présentées au chapitre III ; un protocole de préparation des échantillons est détaillé dans chacune de ces publications. Cette partie décrit donc une préparation d'échantillon globale.

Les contaminants - purs ou en mélange - et leur analogues deutérés ont été pesés (XP105DR, Mettler Toledo, Viroflay, France) pour réaliser des solutions à des concentrations finales de 2 à 500 *ppm* dans de l'eau ultrapure, additionnée d'acétonitrile (jusqu'à 0,1%) pour les composés les moins solubles dans l'eau. Certaines de ses solutions ont été irradiées après dégazage par ultrasons et sous bullage d'azote pour éliminer l'oxygène dissous et ainsi étudier le rôle de ce dernier par comparaison des résultats obtenus avec et sans dégazage. Des expériences d'irradiation ont été menées sur des solutions de polluants réalisées dans l'acétonitrile pour étudier l'effet du solvant (aqueux versus organique) sur la photodégradation. Durant les expériences de photodégradation, des prélèvements ont été effectués à différents temps afin de suivre l'évolution des concentrations en molécule-mère et en photoproduits au cours du temps. Les échantillons prélevés ont été stockés à 4 °C dans l'obscurité puis analysés en LC-MS après dilution entre 2 et 10 *ppm* dans le mélange de solvants constituant les conditions initiales de la phase mobile de la méthode chromatographique utilisée. Pour les analyses en GC-MS, les échantillons ont été séchés puis repris et injectés dans le dichlorométhane.

II.2.2. Préparation des échantillons pour l'identification de photoproduits par infusion directe

Les analyses relatives à la photodégradation de l'acétamipride ont été réalisées en infusion directe. Il s'agissait d'étudier la nature des produits de dégradation en fonction de la complexité du milieu d'irradiation. L'acétamipride a d'abord été dilué à 10 ppm dans de l'eau ultrapure ; une seconde solution a été réalisée dans un standard d'acide fulvique de la rivière Suwannee a une concentration de 40 µg/L.¹³² En vue des approches statistiques, chacune de ces solutions a été préparée six fois puis irradiée sur une durée de 2 h. Des prélèvements réguliers ont été effectués pour étudier l'évolution des produits de dégradation au cours du temps. Les échantillons ont été stockés à 4 °C dans l'obscurité puis analysés de façon aléatoire par un système d'injection automatique sans dilution ou concentration préalable.

II.3. TESTS DE TOXICITE

La toxicité des produits de dégradation obtenus a été estimée au moyen de tests *in silico* et *in vitro*. Les protocoles relatifs à ces tests sont décrits au chapitre III ; une brève description des logiciels et matériels utilisés est donnée ci-après.

II.3.1. Estimation *in silico* de la toxicité potentielle d'un composé

Les outils d'estimation de toxicité *in silico* sont plus rapides, plus éthiques et moins onéreux que les tests *in vivo* et *in vitro*. Les molécules testées n'ont pas besoin d'être obtenues pures, ni même synthétisées. Le logiciel T.E.S.T. (*Toxicity Estimation Software Tool*) a été utilisé dans cette étude. Ce logiciel, libre d'accès *via* l'agence de protection environnementale des Etats Unis (EPA), est régulièrement mis à jour.¹³³ Le logiciel T.E.S.T. utilise une méthodologie QSAR (*Quantitative Structure Activity Relationship*) pour prédire la toxicité d'une molécule à partir de descripteurs moléculaires, i.e. à partir de caractéristiques physico-physiques, et de modèles mathématiques qui comparent ces

descripteurs à ceux de molécules dont la toxicité est connue expérimentalement. 45 descripteurs sont ainsi considérés, tels que le poids moléculaire, le nombre d'atomes, le nombre d'insaturations, le nombre d'atomes halogènes, etc.¹³⁴ La méthodologie QSAR est basée sur une fonction linéaire décrite dans l'équation 9. Dans cette équation, x_1 et x_2 sont des variables descriptives indépendantes alors que a , b et c sont des paramètres ajustés.

$$\text{Toxicité} = ax_1 + bx_2 + c \quad (\text{eq. 9})$$

Le logiciel T.E.S.T. permet de prédire la toxicité de la molécule étudiée pour les modèles suivants :

- Estimation de la concentration létale après 96 h pour 50% (LC_{50} , *Lethal Concentration*) d'une population de poissons à tête de boule (*Fathead minnow*) ;
- Estimation de la LC_{50} après 48 h pour une population de *Daphnia magna* ;
- Estimation de la concentration inhibant la croissance de 50% d'une population de *Tetrahymena pyriformis* (IGC_{50} , *Growth Inhibition Concentration*) ;
- Estimation de la dose létale par voie orale pour 50% d'une population de rats (LD_{50} , *Lethal Dose*) ;
- Estimation du facteur de bioconcentration ;
- Estimation de la toxicité sur le développement embryonnaire ;
- Estimation de la mutagénicité cellulaire.

Le logiciel inclut également des algorithmes permettant de prédire les propriétés physiques d'une molécule à 25 °C : le point d'ébullition, le point éclair, la tension de surface, la viscosité, la densité, la solubilité dans l'eau, la conductivité thermique, la pression de vapeur et le point de fusion.¹³³

Les modèles QSAR sont fortement encouragés dans le cadre de la réglementation européenne REACH (*Registration, Evaluation and Authorisation of Chemicals*) pour la protection de la santé humaine et de l'environnement ; les spécialistes considèrent qu'ils sont de plus en plus fiables.¹³⁵ Si les tests *in silico* tels que présentés ci-dessus constituent une option plus rapide et plus accessible que les tests *in vivo* ou *in vitro*, il convient de

rappeler qu'il ne s'agit que de prédiction de toxicité et que les résultats obtenus n'ont pas autant de pertinence que des résultats expérimentaux.

II.3.2. Etude de toxicité *in vitro*

La méthode utilisée dans le cadre de ce travail pour les études de toxicité *in vitro* repose sur l'inhibition de la bioluminescence de la bactérie marine *Vibrio fischeri*. L'intensité de bioluminescence émise est directement corrélée à la quantité de bactérie vivantes au sein d'une population.¹³⁶ Le mécanisme de bioluminescence de *V. fischeri* a été décrit en détail par Parvez *et al.*¹³⁷ De nombreuses études ont montré que ce test présente une bonne sensibilité aux pesticides.^{138,139} Le protocole du test est normé (ISO 11348-3)¹⁴⁰ et prévoit des analyses en duplicata pour une meilleure reproductibilité de l'analyse ; il permet d'obtenir une évaluation de la toxicité globale d'un mélange ou d'un produit isolé ainsi que de l'évolution de cette dernière au cours du temps.

Les échantillons à tester sont amenés à pH 7 puis dilués à différentes concentrations dans une solution saline de NaCl à 2%. Les cellules de *V. fischeri* sont maintenues à 15 °C dans un incubateur (LUMIStherm, Hach Lange GmbH, Düsseldorf, Allemagne) où la bioluminescence est mesurée grâce à un luminomètre (LUMIStox 300, Hach Lange GmbH, Düsseldorf, Allemagne), d'abord sur les bactéries seules, puis après 5, 15 et 30 minutes de contact avec l'échantillon. Le pourcentage d'inhibition (I%) est calculé en comparant la bioluminescence de la solution de bactéries en contact avec l'échantillon à la référence constituée de bactéries seules. La concentration efficace pour laquelle 50% de l'effet maximum est observé (EC₅₀) est utilisée pour exprimer la toxicité de la solution : plus la valeur de l'EC₅₀ est faible, plus l'échantillon est toxique. Il est à noter que 30 minutes de contact suffisent généralement à déterminer une EC₅₀ avec ce test quand d'autres tests *in vitro* nécessitent un temps de contact de 96 h. De plus, la conservation des kits de bactéries est très simple et ne nécessite pas d'entretien ou de mise en culture.

II.4. METHODES PHYSICOCHIMIQUES D'ANALYSE

L'intérêt de la spectrométrie de masse à haute résolution pour l'analyse de mélanges complexes et la caractérisation de produits inconnus est aujourd'hui parfaitement établi. La chromatographie quant à elle permet une séparation - totale ou partielle - de mélanges complexes, apportant un avantage non négligeable dans le contexte de dégradation photochimique de molécules. Celle-ci conduit en effet souvent à la formation d'isomères ; ces derniers sont séparables par chromatographie alors que la spectrométrie de masse ne permet que rarement de les distinguer.

Les méthodes chromatographiques (HPLC et GC) utilisées lors de la présente étude sont détaillées au chapitre III. Il s'agit de protocoles très classiques qui ne nécessitent pas d'être introduits ici. Nous comparerons en revanche les avantages et limites du couplage LC-MS et de l'introduction directe par infusion. Les grands principes de la spectrométrie de masse FT-ICR seront également présentés dans la mesure où cet analyseur reste aujourd'hui assez peu connu du grand public.

II.4.1. Couplage LC-MS *versus* introduction directe, atouts et limites

Deux techniques de séparation ont été utilisées dans le cadre de ces études : la chromatographie en phase liquide et la chromatographie en phase gazeuse. Ces deux techniques sont complémentaires pour l'identification de produits inconnus polaires (pour la chromatographie liquide) et apolaires (pour la chromatographie gazeuse) dans un mélange complexe. Les séparations sont cependant parfois longues, l'analyse peut durer jusqu'à 1 h pour un unique échantillon, ce qui est peu compatible avec les approches statistiques qui nécessitent des répliques pour être menées à bien. Certains échantillons ont donc été analysés en introduction directe (DIMS). Cette technique consiste à injecter l'échantillon sans séparation chromatographique préalable, manuellement ou via un automate ; elle est aujourd'hui de plus en plus utilisée en métabolomique et en lipidomique^{59,60,141-143} car elle permet d'obtenir une empreinte globale de l'échantillon de façon rapide (quelques minutes d'analyse). L'introduction directe réduit fortement la subjectivité induite par le choix des paramètres chromatographiques en LC-MS lors de

l'analyse de mélanges complexes inconnus. En effet, aucune méthode chromatographique n'étant universelle, le choix d'une colonne et d'une phase mobile exclut de fait la détection de certaines molécules. Un grand nombre d'échantillons peuvent être analysés en peu de temps, incluant les répliques, blancs et contrôles nécessaires à une analyse statistique. Par ailleurs, l'utilisation d'une colonne séparative sur de grandes séries d'échantillons induit une variabilité non négligeable sur les temps de rétention.¹⁴⁴ Contrairement aux résultats obtenus en LC-MS, aucun réalignement en temps de rétention n'est nécessaire pour les analyses DIMS, ce qui limite les erreurs et les mauvaises attributions et améliore la répétabilité des analyses.

Cependant, lorsqu'il s'agit d'identifier des produits inconnus en mélanges complexes, le couplage à une méthode séparative demeure souvent nécessaire pour une analyse plus précise.⁸ En effet, le couplage avec la chromatographie fournit des informations complémentaires sur la polarité ou la volatilité des molécules et permet la séparation des composés isomères comme évoqué plus haut. Une séparation des analytes en amont de la source permet également de limiter les phénomènes de suppression spectrale pouvant être observés lors de l'analyse de mélanges complexes ou trop riches en sels. La suppression spectrale peut être éliminée ou réduite en fractionnant l'échantillon en amont de l'analyse par spectrométrie de masse,^{5,145} mais le choix de la technique de fractionnement induit ici encore de la subjectivité.

Il est à noter que les instruments offrant les meilleures résolutions (Orbitrap et FT-ICR), enregistrent les spectres de masse à des vitesses relativement lentes (jusqu'à quelques dizaines de secondes pour les plus hautes résolutions), ce qui limite la sensibilité de la méthode chromatographique en couplage. En LC-HRMS, la résolution est souvent inférieure à 70 000 pour permettre une bonne résolution chromatographique.¹⁴⁶⁻¹⁴⁸

II.4.2. La spectrométrie de masse

Cette technique d'analyse est basée sur la formation d'ions en phase gazeuse – ions moléculaires et ions fragments – et la mesure de leur rapport m/z (masse sur charge) pour caractériser voire quantifier des analytes. De manière générale, les spectromètres de

masse sont constitués d'une source où sont produits des ions gazeux à partir des analytes, de guides d'ions pour accélérer et guider ces ions, d'un ou plusieurs analyseurs permettant de les séparer, d'un détecteur pour les détecter et les compter et d'un système informatique permettant de traiter les signaux obtenus.

II.4.2.1) La source à électronébulisation ou electrospray

Il existe une grande variété de sources d'ions en spectrométrie de masse, choisies en fonction des caractéristiques physico-chimiques des composés à analyser : taille, polarité, volatilité, stabilité thermique, mais aussi selon le solvant d'injection et la quantité de produit disponible. Comme illustré Figure II-4, la source electrospray (ESI) est la mieux adaptée aux composés polaires de masses moléculaires élevées, tandis que l'ionisation électronique (EI) est dédiée à l'analyse de petites molécules volatiles.

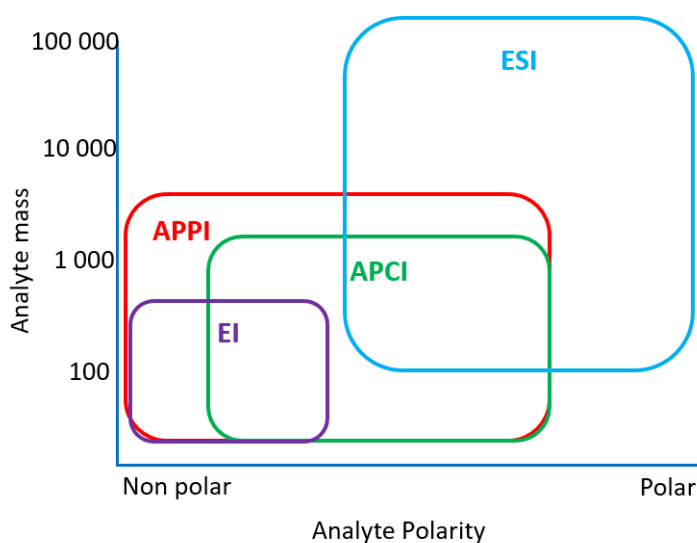


Figure II-4 : Domaines d'application des différentes sources d'ionisation en fonction des caractéristiques des molécules à étudier

De nos jours, la source electrospray est la plus couramment utilisée pour les composés polaires tels que les polluants organiques étudiés dans ce travail. Cette source à pression atmosphérique, mise au point par le groupe de John B Fenn^{149,150} permet d'ioniser des

molécules polaires de façon douce et sur une large gamme de masse. Elle fournit le plus souvent une grande sensibilité et l'énergie interne apportée aux ions peut être contrôlée, ce qui permet d'observer majoritairement les ions moléculaires, même pour des molécules thermolabiles.¹⁵¹

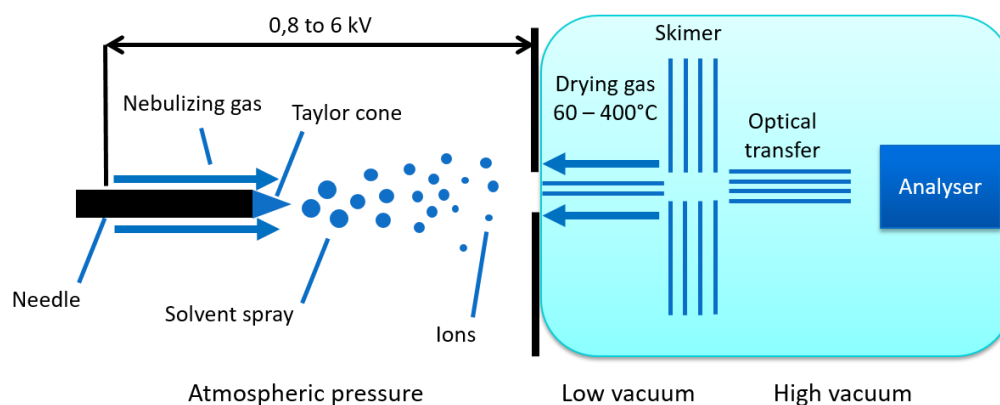


Figure II-5 : Schéma d'une source d'électronébullition

L'échantillon est injecté sous forme liquide dans une aiguille métallique (voir Figure II-5). Une différence de potentiel appliquée entre cette aiguille et une contre électrode permet, associée à la présence d'un gaz de nébulisation (N_2), la formation d'un cône de Taylor et le détachement de microgouttelettes électriquement chargées.¹⁵² Le solvant de ces microgouttelettes s'évapore et leur taille diminue, augmentant la densité de charge de chaque gouttelette. Lorsque la densité de charge d'une gouttelette est trop élevée (limite de Rayleigh), il se produit une explosion coulombienne provoquant la libération d'ions en phase gazeuse (voir Figure II-6).¹⁵³ Par la suite ces ions sont focalisés et guidés jusqu'à l'analyseur via une zone sous vide appelée « optique de transfert ». Très variable d'un appareil à l'autre, cette zone est généralement constituée de lentilles électrostatiques et/ou d'octopôles sur lesquels sont appliquées des tensions.

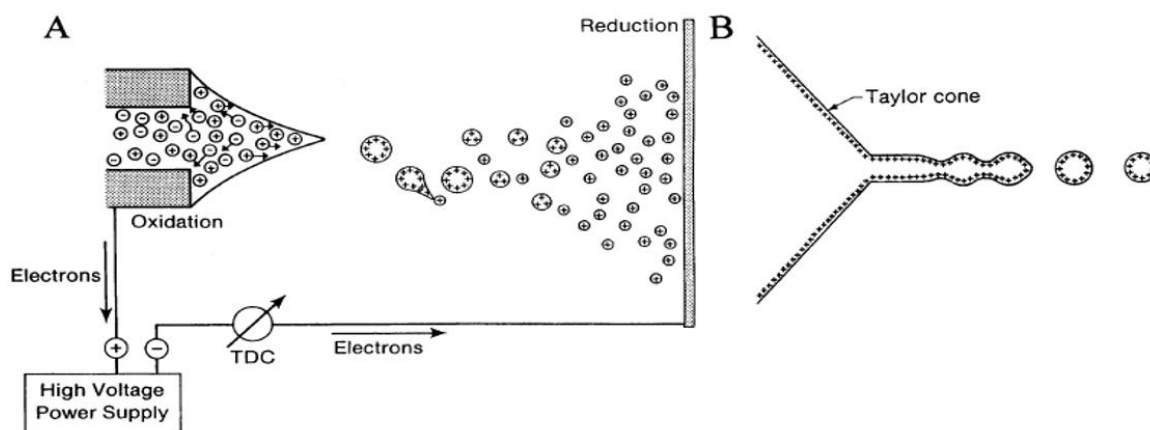


Figure II-6 : (A) Schéma d'émission des ions dans la source ESI. (B) Schéma du cône de Taylor avec émission de micro gouttelettes. (d'après Kebarle P., 2000)¹⁵⁴

Utilisée en mode positif, la source ESI produit des ions pseudo-moléculaires MH^+ , mais aussi des adduits MNa^+ , MK^+ ... En fonction de la taille et des groupements chimiques de l'analyte, il est également possible d'observer des ions moléculaires multichargés (MH_2^{2+} , MH_3^{3+} , MH_n^{n+}). En mode négatif, les polarités sont inversées et la source permet de détecter des ions négatifs tels que $[M-H]^-$, $[M-2H]^{2-}$ ou même $[M-2H+Na]^-$. Le mode positif est généralement privilégié car la chimie des ions positifs est beaucoup mieux connue que celles des ions négatifs ; la séparation chromatographique est par ailleurs souvent mieux maîtrisée avec une phase mobile acidifiée. Le mode négatif est réservé aux composés très acides, plus faciles à déprotoner qu'à protoner.

Cette source ne nécessite que peu d'entretien, elle est très sensible (détection de quelques femtomoles d'analyte), robuste et stable dans le temps. Le débit d'injection de l'échantillon peut varier de quelques $\mu L/min$ à quelques mL/min ce qui la rend utilisable aussi bien en injection directe qu'en couplage avec une chromatographie liquide. Cependant, l'efficacité de l'ESI dépend de la nature et de la quantité d'électrolytes dans la solution.¹⁵⁴ Ce point faible la rend extrêmement sensible à la présence de sels, agents ioniques... Elle est sujette à des phénomènes de suppression ionique, dus à une compétition lors de la protonation des espèces présentes et à la présence de composés moins volatiles qui peuvent changer l'efficacité de la formation des ions, d'autant plus lors de l'analyse de échantillons naturels complexes sans traitement préalable à l'infusion directe.¹⁵⁵

II.4.2.2) L'analyseur FT-ICR

En spectrométrie de masse, l'analyseur permet de séparer les ions en fonction de leur m/z . Les différents analyseurs disponibles sur le marché sont souvent comparés sur la base des paramètres suivants :

- la gamme de masse, qui correspond aux m/z minimum et maximum que l'analyseur peut détecter ;
- la sensibilité, qui traduit la quantité d'échantillon nécessaire pour obtenir un signal interprétable ;
- la vitesse d'acquisition, qui équivaut au temps nécessaire pour obtenir un spectre ;
- la résolution, qui traduit la capacité à distinguer deux pics de rapports m/z proches. La résolution peut s'exprimer comme indiqué dans l'équation 10, où m correspond à la masse de l'ion étudié et $\Delta m_{50\%}$ à la largeur du pic à mi-hauteur.¹⁵⁶

$$R = \frac{m}{\Delta m_{50\%}} \quad (\text{eq. 10})$$

- la précision en masse, qui traduit la différence entre le rapport m/z expérimentalement mesuré et le rapport m/z théorique de l'ion d'intérêt. Généralement cette valeur est indiquée en *part-per million (ppm)* ; elle est déterminée comme indiqué dans l'équation 11, où M_{exp} est le rapport m/z mesuré et M_{th} le rapport m/z théorique de l'ion d'intérêt.

$$\text{error}(ppm) = \frac{|M_{\text{exp}} - M_{\text{th}}| * 10^6}{M_{\text{th}}} \quad (\text{eq. 11})$$

De tous les instruments disponibles sur le marchés, le FT-ICR (analyseur à Résonance Cyclotronique Ionique à Transformée de Fourier) est l'instrument le plus précis sur la mesure de masse et le plus résolutif. Cet analyseur a été développé en 1974 par Comisarow et Marshall.¹⁵⁷ Le FT-ICR mesure la fréquence du mouvement cyclotronique des ions au sein d'un champ magnétique. Cette fréquence est convertie par transformée de Fourier en spectre de masse. Le FT-ICR SolarixXR 9.4T (Bruker Daltonics, Bremen, Allemagne) utilisé pour ces études (Figure II-7) est équipé d'une cellule de détection de type paracell dans laquelle les ions sont stockés, puis détectés. Cette cellule est placée au

sein d'un aimant supraconducteur de 9.4 Tesla. Cet instrument compte parmi les instruments les plus résolutifs du marché (jusqu'à 10 000 000 de résolution) et permet une mesure de masse très précise (quelques *ppm* d'erreur).



Figure II-7 : Photographie du spectromètre de masse FT-ICR SolarixXR 9.4T utilisé pour cette étude

Les ions formés dans la source ESI sont acheminés jusqu'au quadripôle permettant, si cela est souhaité, de sélectionner des ions en vue de les fragmenter dans la cellule de collision par dissociations induites par collisions (CID) pour obtenir des informations structurales. Les ions sont ensuite guidés par un octopôle jusqu'à la paracell. (Figure II-8). Cette cellule, placée dans un vide poussé (10^{-10} mbar) permet le stockage et la détection des ions par la mesure de leur fréquence cyclotronique induite par le champ magnétique.¹⁵⁸

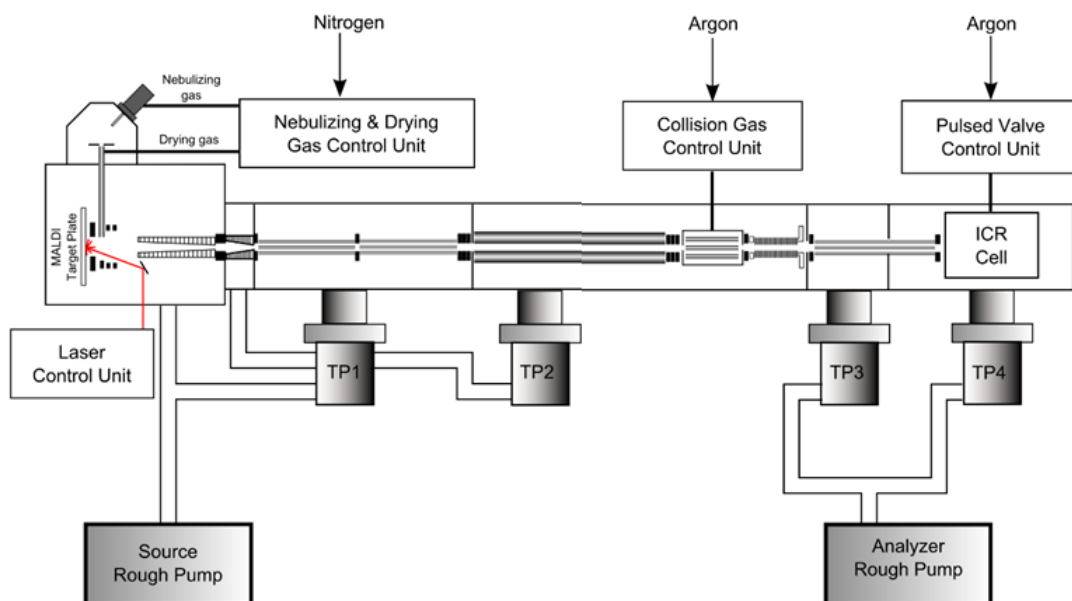


Figure II-8 : Schéma du FT-ICR SolarixXR 9.4T utilisé pour cette étude

Description de la cellule paracell

La cellule ICR (paracell) est composée de plaques d'excitation et de plaques de détection limitées par des plaques de piégeage des ions (Figure II-9). Elle est placée au centre d'un champ magnétique B créé par l'aimant supraconducteur.

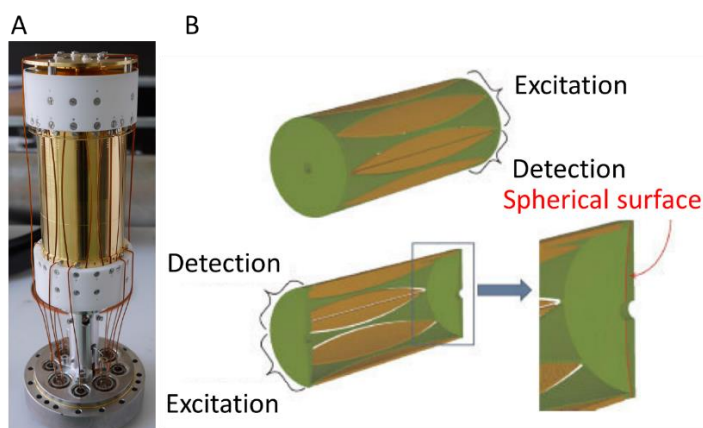


Figure II-9 : A) photographie de la cellule paracell installée au laboratoire. B) design de la paracell (d'après Boldin, I.A., et al. 2011)¹⁵⁹

Les ions présents dans la cellule ICR subissent la force de Lorentz (équation 12) sous l'effet du champ magnétique. Cette force induit un mouvement circulaire de l'ion, perpendiculaire au champ magnétique, appelé mouvement cyclotronique.

$$\vec{F} = q\vec{v} * \vec{B} \quad (\text{eq. 12})$$

La fréquence de ce mouvement dépend du champ magnétique et du rapport m/z des ions piégés comme indiqué par l'équation 13.¹⁵⁸ Dans cette équation, ν_c (fréquence cyclotronique) est exprimée en Hz, et B_0 (champ magnétique) en tesla.

$$\nu_c = \frac{\omega_c}{2\pi} = \frac{1.535611 * 10^7 B_0}{m/z} \quad (\text{eq. 13})$$

La fréquence de rotation cyclotron est proportionnelle au champ magnétique mais inversement proportionnelle au rapport m/z de l'ion.

Il faut aussi prendre en compte la fréquence de piégeage ν_r , qui consiste en une oscillation axiale (ν_t), formée par les tensions appliquées sur les plaques de piégeage de la cellule ICR afin de garder les ions dans la cellule sans qu'ils ne se déchargent sur les plaques. ν_r dépend du potentiel de piégeage V_{trap} appliqué sur les plaques de piégeage, du facteur géométrique de la cellule ICR α et de la distance entre les plaques de piégeage a . (équation 14)

$$\nu_r = \frac{1}{2\pi} \sqrt{\frac{2 q V_{\text{trap}} \alpha}{m a^2}} \quad (\text{eq. 14})$$

Le mouvement des ions

Les ions piégés dans la cellule possèdent un mouvement qui peut se décomposer en trois types d'oscillations (Figure II-10) :

- L'oscillation axiale ν_t dans l'axe du champ magnétique et du champ électrique induit par les plaques de piégeage;
- L'oscillation cyclotron ν_c , orbite circulaire de l'ion dans la cellule ;

- L'oscillation magnétron v_m qui correspond à la rotation de l'ion sur lui-même.

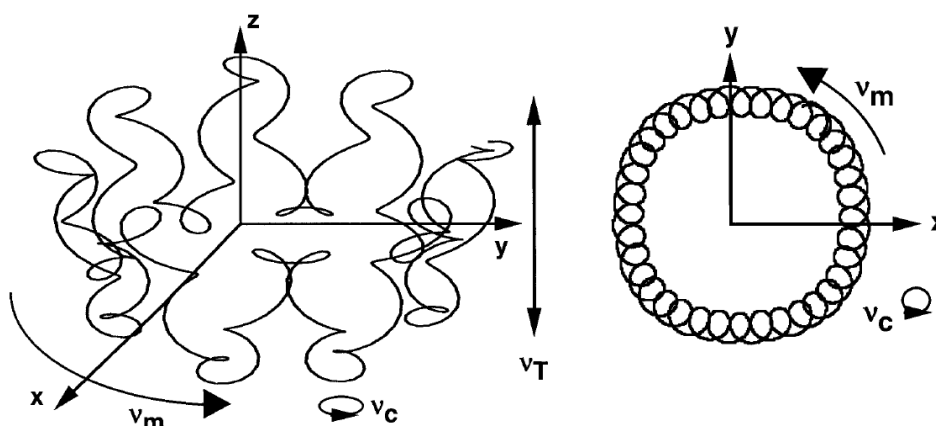


Figure II-10 : Mouvement d'un ion dans un champ électromagnétique. v_t : fréquence axiale, v_m : fréquence magnétron, v_c : fréquence cyclotronique (d'après Marshall, A.G., 1998)¹⁵⁸

Les instruments commerciaux mesurent la fréquence cyclotronique réduite décrite dans l'équation 15.

$$v_+ = \frac{v_c}{2} + \sqrt{\left(\frac{v_c}{2}\right)^2 - \frac{v_r^2}{2}} \quad (\text{eq. 15})$$

Le mouvement magnétron peut être négligé car il est de faible amplitude pour les ions piégés à proximité du centre de la cellule. La fréquence de piégeage est négligeable pour les ions de rapport m/z faibles, mais elle devient importante pour les ions dont le rapport m/z est élevé (> 8000). Pour les ions de rapport m/z faible, la fréquence cyclotronique réduite mesurée par l'instrument est ainsi très proche de la fréquence cyclotronique décrite en équation 13, qui ne dépend donc que du champ magnétique et du rapport m/z de l'ion.

Détection des ions

Tous les ions piégés dans la cellule ICR possèdent des positions et des rayons d'orbite différents. Afin de pouvoir mesurer leurs fréquences cyclotroniques, il faut donc exciter les ions grâce à une différence de potentiel appliquée sur les plaques d'excitation. Ceci

modifie l'orbite des ions et les approche des plaques de détection. Les ions se trouvent sur un même rayon, plus grand que le rayon initial, et oscillent à une fréquence propre à leur rapport m/z . La rotation des ions induit un courant détecté et mesuré par les plaques de détection. Ce signal, appelé transient, correspond à la superposition de tous les courants induits par les différents ions présents dans la cellule. Ce transient est ensuite converti en intensité de signal en fonction du temps puis en intensité de signal en fonction du rapport m/z par transformée de Fourier.

II.4.2.3) Intérêt de la haute résolution pour l'analyse de mélanges complexes

La résolution d'un spectromètre de masse est définie plus haut ; elle traduit la capacité de l'analyseur à séparer deux rapports m/z de valeurs proches. Plus elle est élevée, meilleure est la séparation des ions, en particulier des ions isobares. On différencie les analyseurs à basse résolution (quadripôles, trappes ioniques) et à haute résolution (temps de vol, Orbitrap, FT-ICR). Une résolution inférieure à 1000 ne permet pas de séparer deux ions isobares, mais seulement de déterminer l'état de charge d'un ion en se basant sur l'écart de m/z entre deux pics différant d'un isotope du carbone ^{12}C vs ^{13}C . Pour séparer deux masses nominales de faible poids moléculaire (inférieure à 2500 Da) une résolution de 10 000 ou plus est nécessaire. La résolution se situe aux alentours de 70 000 pour le Q-TOF près de 500 000 pour l'Orbitrap et jusqu'à 10 millions pour la dernière génération de FT-ICR. Le spectromètre de masse à transformée de Fourier constitue un outil de choix pour les analyses de mélanges complexes dans de nombreux domaines : métabolomique, lipidomique, pétroliomique, analyse d'échantillons alimentaires ou environnementaux.^{6,82,92,94,95,111,160-162} En effet, la précision associée à la haute résolution de cet instrument permet identifier des molécules au sein d'un mélange complexe sans interférence avec les autres espèces. La précision permet également de réduire fortement l'ambiguïté lors de l'assignation de formules brutes.¹⁶³ Néanmoins, au vu du grand nombre de formules brutes potentielles pour une masse donnée, il peut arriver qu'une seule et unique formule brute ne puisse être attribuée pour un ion, même avec une précision inférieure à 1 ppm.⁷⁹ La capacité des FT-ICR à fournir un massif isotopique correspondant à la structure isotopique fine permet d'obtenir une information

supplémentaire sur le nombre et la nature des atomes présents dans l'ion à identifier.^{136,137} En 2009, Kellmann et al. ont publié une étude sur le minimum de résolution nécessaire pour détecter sans ambiguïté une molécule connue, de faible poids moléculaire, dans des mélanges complexes.¹⁴⁷ Elle montre que plus le mélange est complexe, plus la résolution nécessaire pour séparer des molécules isobares est élevée, même en utilisant une méthode chromatographique en amont (Tableau II-1). Les auteurs concluent ainsi qu'un minimum de 50 000 de résolution est nécessaire pour l'analyse du mélange le moins complexe alors qu'une résolution de 100 000 ne suffit pas à assigner tous les contaminants connus avec moins de 2 ppm d'erreur sur les mélanges les plus complexes.

Tableau II-1 : Effet de la résolution sur l'identification de contaminants dans des matrices de complexité moyenne (miels) à élevées (nourriture animale) (d'après Kellmann et al., 2009)¹⁴⁷

Resolution	% of 151 analytes				
	Mass deviation (ppm)				
	<2	2-5	5-10	10-25	>25/ND
Honey					
10,000	83	12	5	0	0
25,000	100	0	0	0	0
50,000	100	0	0	0	0
100,000	100	0	0	0	0
Animal feed					
10,000	22	25	16	20	17
25,000	41	17	26	2	14
50,000	93	4	1	0	2
100,000	99	0	0	0	1

RÉFÉRENCES BIBLIOGRAPHIQUES

- (5) Shulaev, V.; Cortes, D.; Miller, G.; Mittler, R. Metabolomics for plant stress response. *Physiol Plant* **2008**, *132*, 199-208.
- (6) Ghaste, M.; Mistrik, R.; Shulaev, V. Applications of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap Based High Resolution Mass Spectrometry in Metabolomics and Lipidomics. *Int J Mol Sci* **2016**, *17*, 816.
- (8) Lin, L.; Yu, Q.; Yan, X.; Hang, W.; Zheng, J.; Xing, J.; Huang, B. Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabolomics? A serum metabolomic study of kidney cancer. *The Analyst* **2010**, *135*, 2970-2978.
- (29) Lassalle, Y.; Nicol, E.; Genty, C.; Bourcier, S.; Bouchonnet, S. Structural elucidation and estimation of the acute toxicity of the major UV-visible photoproduct of fludioxonil - detection in both skin and flesh samples of grape. *Journal of mass spectrometry : JMS* **2015**, *50*, 864-869.
- (30) Lassalle, Y.; Jellouli, H.; Ballerini, L.; Souissi, Y.; Nicol, E.; Bourcier, S.; Bouchonnet, S. Ultraviolet-vis degradation of iprodione and estimation of the acute toxicity of its photodegradation products. *Journal of chromatography. A* **2014**, *1371*, 146-153.
- (59) Witting, M.; Lucio, M.; Tziotis, D.; Wagele, B.; Suhre, K.; Voulhoux, R.; Garvis, S.; Schmitt-Kopplin, P. DI-ICR-FT-MS-based high-throughput deep metabolotyping: a case study of the *Caenorhabditis elegans*-*Pseudomonas aeruginosa* infection model. *Analytical and bioanalytical chemistry* **2015**, *407*, 1059-1073.
- (60) Aharoni, A.; Ric de Vos, C. H.; Verhoeven, H. A.; Maliepaard, C. A.; Kruppa, G.; Bino, R.; Goodenowe, D. B. Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. *Omics : a journal of integrative biology* **2002**, *6*, 217-234.
- (79) Kind, T.; Fiehn, O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC bioinformatics* **2006**, *7*, 234.
- (82) Rodgers, R. P.; Schaub, T. M.; Marshall, A. G. Petroleomics: MS returns to its roots. *Analytical chemistry* **2005**, *77*, 20a-27a.
- (92) Bianco, A.; Deguillaume, L.; Vaitilingom, M.; Nicol, E.; Baray, J. L.; Chaumerliac, N.; Bridoux, M. Molecular Characterization of Cloud Water Samples Collected at the Puy de Dome (France) by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environ Sci Technol* **2018**, *52*, 10275-10285.
- (94) Kew, W.; Goodall, I.; Clarke, D.; Uhrin, D. Chemical Diversity and Complexity of Scotch Whisky as Revealed by High-Resolution Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 200-213.
- (95) Martins, N.; Jimenez-Morillo, N. T.; Freitas, F.; Garcia, R.; Gomes da Silva, M.; Cabrita, M. J. Revisiting 3D van Krevelen diagrams as a tool for the visualization of volatile profile of varietal olive oils from Alentejo region, Portugal. *Talanta* **2020**, *207*, 120276.
- (111) Roullier-Gall, C.; Boutegrabet, L.; Gougeon, R. D.; Schmitt-Kopplin, P. A grape and wine chemodiversity comparison of different appellations in Burgundy: vintage vs terroir effects. *Food Chem* **2014**, *152*, 100-107.

(128) Lassalle, Y.; Kinani, A.; Rifai, A.; Souissi, Y.; Clavaguera, C.; Bourcier, S.; Jaber, F.; Bouchonnet, S. UV-visible degradation of boscalid--structural characterization of photoproducts and potential toxicity using in silico tests. *Rapid communications in mass spectrometry : RCM* **2014**, *28*, 1153-1163.

(129) Lassalle, Y.; Nicol, E.; Genty, C.; Bourcier, S.; Bouchonnet, S. Isomerization of fenbuconazole under UV-visible irradiation - chemical and toxicological approaches. *Rapid communications in mass spectrometry : RCM* **2015**, *29*, 1335-1342.

(130) Sunlight, Weathering & Light Stability Testing. <https://www.q-lab.com/documents/public/cd131122-c252-4142-86ce-5ba366a12759.pdf> (accessed 07 2020).

(131) Simulation de la lumière solaire. <https://www.q-lab.com/documents/public/fd2b69b6-b672-421b-81ce-10c11d63248a.pdf#page=8> (accessed 07 2020).

(132) Anderson, T. A.; Salice, C. J.; Erickson, R. A.; McMurry, S. T.; Cox, S. B.; Smith, L. M. Effects of landuse and precipitation on pesticides and water quality in playa lakes of the southern high plains. *Chemosphere* **2013**, *92*, 84-90.

(133) Toxicity Estimation Software Tool (TEST) <https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test> (accessed 07 2020).

(134) Molecular Descriptors Guide. <https://www.epa.gov/sites/production/files/2015-05/documents/moleculardescriptorsguide-v102.pdf> (accessed 07 2020).

(135) Gramatica, P. Principles of QSAR Modeling. *International Journal of Quantitative Structure-Property Relationships* **2020**, *5*, 61-97.

(136) Kaiser, K. L. Correlations of *Vibrio fischeri* bacteria test data with bioassay data for other organisms. *Environ Health Perspect* **1998**, *106*, 583-591.

(137) Parvez, S.; Venkataraman, C.; Mukherji, S.: A review on advantages of implementing luminescence inhibition test (*Vibrio fischeri*) for acute toxicity prediction of chemicals. 2006; Vol. 32; pp 265-268.

(138) Fernandez-Alba, A. R.; Guil, L. H.; Lopez, G. M.; Chisti, Y. Toxicity of pesticides in wastewater: a comparative assessment of rapid bioassays. *Analytica Chimica Acta* **2001**, *426*, 289-301.

(139) Fernandez-Alba, A. R.; Guil, M. D. H.; Lopez, G. D.; Chisti, Y. Comparative evaluation of the effects of pesticides in acute toxicity luminescence bioassays. *Analytica Chimica Acta* **2002**, *451*, 195-202.

(140) ISO 1998. Water Quality. Freshwater Algal Growth Inhibition Test with Unicellular Green Algae. *International Organization for Standardization, Switzerland*, 11348e11353.

(141) Allen, J.; Davey, H. M.; Broadhurst, D.; Heald, J. K.; Rowland, J. J.; Oliver, S. G.; Kell, D. B. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nature Biotechnology* **2003**, *21*, 692-696.

(142) Wang, M.; Wang, C.; Han, R. H.; Han, X. Novel advances in shotgun lipidomics for biology and medicine. *Prog Lipid Res* **2016**, *61*, 83-108.

(143) Southam, A. D.; Payne, T. G.; Cooper, H. J.; Arvanitis, T. N.; Viant, M. R. Dynamic range and mass accuracy of wide-scan direct infusion nano-electrospray fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Analytical chemistry* **2007**, *79*, 4595-4602.

(144) Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass spectrometry-based metabolomics. *Mass Spectrom Rev* **2007**, *26*, 51-78.

(145) Koch, B. P.; Ludwichowski, K. U.; Kattner, G.; Dittmar, T.; Witt, M. Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. *Marine Chemistry* **2008**, *111*, 233-241.

(146) Rajski, L.; Gomez-Ramos Mdel, M.; Fernandez-Alba, A. R. Large pesticide multiresidue screening method by liquid chromatography-Orbitrap mass spectrometry in full scan mode applied to fruit and vegetables. *Journal of chromatography. A* **2014**, *1360*, 119-127.

(147) Kellmann, M.; Muenster, H.; Zomer, P.; Mol, H. Full Scan MS in Comprehensive Qualitative and Quantitative Residue Analysis in Food and Feed Matrices: How Much Resolving Power is Required? *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 1464-1476.

(148) Makarov, A.; Scigelova, M. Coupling liquid chromatography to Orbitrap mass spectrometry. *Journal of chromatography. A* **2010**, *1217*, 3938-3945.

(149) Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Analytical chemistry* **1985**, *57*, 675-679.

(150) Fenn, J. B. Electrospray wings for molecular elephants (Nobel lecture). *Angew Chem Int Ed Engl* **2003**, *42*, 3871-3894.

(151) Choi, B. K.; Hercules, D. M.; Zhang, T. L.; Gusev, A. I. Comparison of quadrupole, time-of-flight, and Fourier transform mass analyzers for LC-MS applications. *Lc Gc N Am* **2001**, *19*, 514-+.

(152) Geoffrey, T. Disintegration of Water Drops in an Electric Field. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **1964**, *280*, 383-397.

(153) Fenn, J. B. Ion formation from charged droplets: Roles of geometry, energy, and time. *Journal of the American Society for Mass Spectrometry* **1993**, *4*, 524-535.

(154) Kebarle, P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *Journal of mass spectrometry : JMS* **2000**, *35*, 804-817.

(155) Annesley, T. M. Ion suppression in mass spectrometry. *Clin Chem* **2003**, *49*, 1041-1044.

(156) Marshall, A. G.; Hendrickson, C. L. High-resolution mass spectrometers. *Annu Rev Anal Chem (Palo Alto Calif)* **2008**, *1*, 579-599.

(157) Comisarow, M. B.; Marshall, A. G. Fourier-Transform Ion-Cyclotron Resonance Spectroscopy. *Chemical Physics Letters* **1974**, *25*, 282-283.

(158) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev* **1998**, *17*, 1-35.

(159) Boldin, I. A.; Nikolaev, E. N. Fourier transform ion cyclotron resonance cell with dynamic harmonization of the electric field in the whole volume by shaping of the excitation and detection electrode assembly. *Rapid communications in mass spectrometry : RCM* **2011**, *25*, 122-126.

(160) Rathahao-Paris, E.; Alves, S.; Junot, C.; Tabet, J. C. High resolution mass spectrometry for structural identification of metabolites in metabolomics. *Metabolomics* **2016**, *12*, 10.

(161) Brown, S. C.; Kruppa, G.; Dasseux, J. L. Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrom Rev* **2005**, *24*, 223-231.

(162) Bundy, J. G.; Davey, M. P.; Viant, M. R. Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **2009**, *5*, 3-21.

(163) Kim, S.; Rodgers, R. P.; Marshall, A. G. Truly "exact" mass: Elemental composition can be determined uniquely from molecular mass measurement at ~0.1mDa accuracy for molecules up to ~500Da. *International Journal of Mass Spectrometry* **2006**, *251*, 260-265.

Chapitre III - IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Ce chapitre présente, sous forme d'articles publiés dans des revues scientifiques, les résultats relatifs à la caractérisation de photoproduits de polluants organiques. Ces études ont été réalisées via une approche consistant à modéliser en laboratoire la photolyse des composés, seuls ou en mélange.

Ce chapitre se divise en 5 parties consacrées respectivement à la photodégradation des 4 contaminants étudiés au cours de ce travail (le cyprodinil, la benzisothiazolinone, l'acétamipride pure et au sein de mélanges commerciaux, et le naproxène) et à la synthèse des résultats obtenus pour l'ensemble des molécules.

III.1. PHOTODEGRADATION UV-VISIBLE DU CYPRODINIL – APPROCHES CHIMIQUE ET TOXICOLOGIQUE

III.1.1. Introduction des travaux portant sur la photodégradation du cyprodinil

Le cyprodinil (Figure III-1) est un fongicide à large spectre de la famille des anilino-pyrimidines, utilisé principalement sur des fruits pour lutter contre la pourriture grise (*Botrytis cinerea*) et la tavelure du poirier (*Venturia inaequalis* et *Venturia pirina*) en inhibant la biosynthèse de la méthionine et en perturbant la croissance mycélienne dans la plante.¹⁶⁴ Il s'agit d'un fongicide à action systémique ; il pénètre au sein du végétal et est redistribué à l'intérieur et en surface de la plante. Les fongicides systémiques sont généralement utilisés de façon préventive et curative. Le caractère lipophile du cyprodinil pourrait mener à sa bioaccumulation mais il se dégrade rapidement en milieu aqueux sous l'action du rayonnement ultraviolet (demi vie de 13,5 jours).¹⁶⁵

L'utilisation de ce fongicide est autorisée par l'Union Européenne depuis mai 2007 et son approbation a été prolongée jusqu'en avril 2021 par la directive (EU) 2020/421.¹⁶⁶ Cependant, ce produit est classé comme toxique pour les poissons et les algues, très toxique pour les crustacés.¹⁶⁵ En effet, de récentes études ont démontré son effet sur la toxicité développementale des poissons, notamment au plan cardiaque.¹⁶⁷ Le cyprodinil agit à la fois comme un perturbateur endocrinien et comme un perturbateur

physiologique cellulaire ; il peut augmenter la prolifération des cellules cancéreuses.^{168,169}
La dose journalière acceptable pour l'être humain a été fixée à 0,03 mg.kg⁻¹.j⁻¹.¹⁷⁰

L'utilisation de ce fongicide sur les cultures entraîne *de facto* une contamination des eaux et des sols. Le cyprodinil a été retrouvé à de fortes concentrations (jusqu'à 0,44 mg.L⁻¹) parmi d'autres pesticides au sein de produits alimentaires et eaux de surface, notamment dans les régions viticoles.¹⁷¹⁻¹⁷³

Une fois présent dans l'environnement le cyprodinil peut subir diverses voies de dégradation, naturelles ou suite au traitement des eaux. Nous avons étudié la dégradation par irradiation UV et cherché à identifier le photoproduits du cyprodinil obtenus par photolyse directe. Une fois ces photoproduits identifiés, nous avons évalué leur toxicité potentielle *in silico* ainsi que l'écotoxicité *in vitro* de la solution. Le premier objectif a été réalisé à l'aide d'analyses par LC-HRMS/MS et de l'utilisation d'un analogue deutéré du cyprodinil (le cyprodinil-D₅ voir Figure III-1) permettant d'éliminer un grand nombre d'isomères possibles et de valider les structures des photoproduits. Les toxicités potentielles *in silico* des photoproduits ont pu être évaluées individuellement grâce au logiciel T.E.S.T., tandis que l'écotoxicité *in vitro* de la solution a été déterminée, avant et après plusieurs temps d'irradiation par une lampe à vapeur de mercure, à l'aide de tests commerciaux de bioluminescence de la bactérie marine *Vibrio fischeri*. Les toxicités ainsi obtenues ont pu être comparées à celle du cyprodinil afin de corrélérer l'évolution de la toxicité à la durée d'irradiation. Les photoproduits toxiques ainsi caractérisés pourront être intégrés dans les analyses de routine de quantification de pesticides, dans les eaux ou des produits alimentaires par exemple.

Les résultats ont été publiés dans la revue *Rapid Communication in Mass Spectrometry*.¹⁷⁴
L'article est présenté ci-après.

PHOTODEGRADATION OF CYPRODINIL UNDER UV-VISIBLE IRRADIATION - CHEMICAL AND TOXICOLOGICAL APPROACHES

Edith Nicol, Houda Chayata, Christophe Genty, Stéphane Bouchonnet and Sophie Bourcier*

LCM, CNRS, Ecole polytechnique, Université Paris-Saclay, 91128 Palaiseau, France

Rapid Commun. Mass Spectrom. 2016, 30, 2201–2211

<https://doi.org/10.1002/rcm.7685>

ABSTRACT

RATIONALE: Cyprodinil is a fungicide active on grapes, strawberries, tomatoes, and many other fruits. Under UV–visible irradiation, it undergoes photodegradation through various processes to form transformation products (TPs) whose structures and potential toxicities are unknown. The structures of the TPs were elucidated by comparing the photodegradation of cyprodinil and cyprodinil-D₅. The potential toxicities of these compounds were compared with that of cyprodinil.

METHODS: Aqueous solutions of cyprodinil were irradiated in a reactor equipped with a mercury vapor lamp. Analyses were carried out using high-performance liquid chromatography coupled to a quadrupole time-of-flight (QTOF) mass spectrometer or to a SolarixXR 9.4 Tesla Fourier transform (FT) mass spectrometer. High-resolution mass measurements, MS/MS and isotopic labelling experiments allowed structural elucidation of the cyprodinil TPs. The toxicities were estimated by three tests *in silico* using the T.E.S.T. software and *in vitro* bioassays using *Vibrio fischeri* bacteria. These bioassays were carried out on irradiated solution for several exposure times and non-irradiated solutions.

RESULTS: The structures of 19 photoproducts were characterized by LC/HRMS/MS after 4 h of irradiation of a cyprodinil aqueous solution. The use of cyprodinil-D₅ allowed the TPs to be characterized with more confidence. Knowing the structure of the TPs allows the estimation of their potential toxicities by *in silico* tests. Most of the photoproducts are potentially more toxic than the parent compound, based on the oral rat LD₅₀ values, and most of them might induce more developmental and mutagenic toxicities. *In vitro* assays on *Vibrio fischeri* bacteria showed that the global ecotoxicity of the cyprodinil solution significantly increases with irradiation time.

CONCLUSIONS: Structures of photoproducts were characterized after irradiation of a cyprodinil aqueous solution combining LC/HRMS, LC/HRMS/MS and the use of a labelled compound. Their formations imply several photodegradation reactions, namely direct bond cleavages, cyclization, isomerization and hydroxylation. Most of the TPs exhibit a toxicity significantly higher than that of the parent molecule.

INTRODUCTION

Cyprodinil is an anilino-pyrimidine fungicide particularly active against gray mold (*Botrytis cinerea*) and pear scab (*Venturia inaequalis* and *Venturia pirina*) on grapes, strawberries, tomatoes, and many other fruits. The mechanism of action of anilinopyrimidines involves inhibiting methionine biosynthesis.¹⁶⁴ Cyprodinil belongs to a new generation of fungicides extensively employed worldwide in order to combat diseases affecting vineyards. This compound shows a high level of activity against *B. cinerea* strains resistant to commonly used fungicides. The use of fungicides in agriculture is considered essential to optimize yields, but leads to water and soil contamination. The non-degraded fungicides and their degradation products may be a source of soil contamination and pollution of rivers and lakes. Recently, several authors reported high levels of concentration of pesticides in wine regions (cyprodinil at 0.44 mg.L⁻¹).^{171,172}

All the common water treatment processes for these fungicides, such as hydrolysis, biodegradation, chemical oxidation and UV irradiation, are still under development.¹⁷⁵⁻¹⁷⁷ Although they are increasingly efficient, these treatments do not lead to full mineralization of pollutants and produce transformation products which can contaminate the environment.^{51,178} In a 2013 review on emerging pollutants, Aguera *et al.* stated that analytical trends are no longer limited to the quantitation of listed contaminants, but also involve identification of by-products generated by water treatment processes.⁴⁶ Several studies have established that some persistent transformation products (TPs) could prove to be more toxic than the parent molecules.⁴⁰⁻⁴⁵ These new contaminants are released into the environment, and their stability and persistency point to their potential as environmental contaminants. UV photolysis processes are commonly used as an alternative to chlorine disinfection. Some efforts have been devoted to the direct and catalyzed photodegradation of pesticides, including identification of photoproducts, kinetic studies and reaction pathway establishment, in order to integrate these TPs into pesticide control procedures.¹⁷⁹ These studies showed that photolysis, in combination with other techniques multiplying the number of degradation pathways, would target a large range of pollutants. Regardless of the treatment methods used, knowledge of the chemical structures of the TPs would permit analytical methods to be established that are

able to target those specific compounds and to predict their impact on the environment. Solar-radiation-induced degradation is an important route of pesticide dissipation in the environment, and it may even be the major process for some pesticides in surface water. Moreover, just after crop treatment and before having diffused inside the plants, pesticides deposited onto leaf surfaces are subject to solar light. Pesticide photodegradation has been extensively studied in water but only a few studies have focused on the degradation products.^{50,51,180-182}

Esteve-Turrillas *et al.* have recently published a survey of the occurrence of some fungicides in international wines, by using rapid antibody-based assays.¹⁷³ This study shows that contamination of commercial wines with pesticides is an issue of worldwide relevance, with potential implications for consumer health and international trade. Coleman *et al.* have explored the effects of a combination of three commonly used biocides (pyrimethanil, cyprodinil and fludioxonil) on cell viability, mitochondrial health and generation of oxidative stress, alone and in combination.¹⁸³ They showed that cyprodinil was the most toxic agent individually, although oxidative stress-related enzyme gene expression increases appeared to demonstrate some degree of synergy in the presence of a combination of agents. This report suggests that the impact of some pesticides, both individually and in combinations, merits further study in terms of their effect on human cellular health.

The present work had two goals; the first one was to characterize the photoproducts formed by direct photolysis of cyprodinil. Analyses were performed by liquid chromatography coupled with high-resolution mass spectrometry (HRMS). Characterization of the photoproducts was carried out performing HRMS in both full scan and collision-induced dissociation (CID) experiments. Isotopic labelling was used to ascertain the structures of the photoproducts. The second goal was to assess the potential toxicity of the photoproducts using T.E.S.T. software¹⁸⁴ and the *in vitro* ecotoxicity of the irradiated solution for several exposure times using *Vibrio fischeri* commercial test kits.

EXPERIMENTAL

Chemicals and reagents

Cyprodinil (4-cyclopropyl-6-methyl-N-phenylpyrimidin-2-amine; 99.99% purity) and cyprodinil-D₅ (≥95% purity) were purchased from Sigma Aldrich (St Quentin Fallavier, France). Their chemical structures are displayed in Figure III-1. Chromatographic grade solvents (99.99% purity), acetonitrile (ACN) and formic acid (FA) were also purchased from Sigma Aldrich. Solutions of both compounds were prepared at 12 mg.L⁻¹ in water. All solutions were degassed using nitrogen bubbling for 15 min and sonication for 10 min. A pH value of 5.9 was measured prior to irradiation; this pH remained the same whatever the irradiation time.

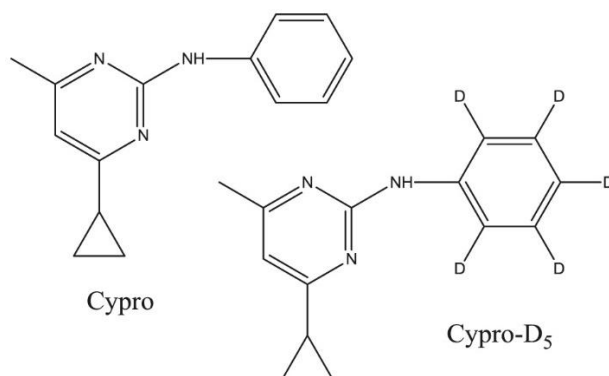


Figure III-1: Chemical structures of cyprodinil (Cypro) and cyprodinil-D₅ (Cypro-D₅).

Photolysis experiments

Photolysis experiments were carried out using a self-made reactor equipped with a high-pressure mercury lamp (HPL-N 125 W/542 E27 SC; Phillips, Ivry-sur-Seine, France) delivering radiation at wavelengths ranging from 200 to 650 nm. According to the manufacturer's data, the incident radiation flux was 6200 lm. The reactor consists of six 120-mL quartz tubes disposed in a circle around the lamp and immersed in a sonicator (AL04-12; Advantage-Lab, Vernier, Switzerland) filled with deionized water. During experiments, the reactor was regularly cooled by the addition of ice to avoid uncontrolled heating of the irradiated solutions and to maintain a constant temperature of 25 ± 3°C.

For the experiments devoted to the characterization of photoproducts, 30 mL of a solution of cyprodinil were used; 100 μL of solution were taken from the quartz tube, dried and dissolved in water with 0.1% FA for LC/MS analysis.

LC/HRMS operating conditions

All the chromatographic separations were carried out on a 2690 HPLC separation module (Waters Technologies, Saint Quentin en Yvelines, France). 10 μL of the sample were injected and separated on a C_{18} Pursuit XRs Ultra column (2.8 μm , 50 \times 2.0 mm; Agilent Technologies, Les Ulis, France). The effluent was introduced at a flow rate of 0.2 $\text{mL}\cdot\text{min}^{-1}$ into the Z spray interface with a gradient increasing from 20% of solvent B to 80% of solvent B in 8 min and until 10 min (A: water with 0.1% FA; B: ACN with 0.1% FA). The gradient was then set at 20% of solvent B for the last 10 min. For structural elucidation experiments, the HPLC system was coupled with a quadrupole time-of-flight (QTOF) Premier mass spectrometer (Waters Technologies) operated in positive ion electrospray ionization (ESI) mode. The ion source parameters were: cone voltage, 20 V; capillary voltage, 2.6 kV. Typical values for the other source parameters were: extraction cone voltage, 2 V; ion guide voltage, 2.4 V. The source and desolvation temperatures were set to 100°C and 450°C, respectively. Nitrogen was used as both the nebulizing and the desolvation gas, with gas flow rates ranging from 70 $\text{L}\cdot\text{h}^{-1}$ to 700 $\text{L}\cdot\text{h}^{-1}$. Argon was used as the collision gas at a flow rate of 0.28 $\text{mL}\cdot\text{min}^{-1}$, corresponding to a pressure of *ca* $4\cdot 10^{-3}$ mBar. The ESI mass spectra yielded $[\text{M} + \text{H}]^+$ ions for cyprodinil and its photoproducts, and these ions were used as the precursor ions for the MS/MS spectra. The MS/MS spectra were recorded at collision energies from 5 to 25 eV in order to obtain the main characteristic product ions and the decomposition pathways. When MS^n -like experiments were necessary, fragment ions were formed in the ion source by increasing the cone voltage, to subsequently record MS/MS spectra.

For HRMS experiments, the HPLC system was coupled to a Solarix XR 9.4 Tesla FT mass spectrometer (Bruker Daltonics, Bremen, Germany). In order to obtain a 200 $\mu\text{L}\cdot\text{min}^{-1}$ flow rate, a 1:60 split was used between the column and the ESI source. The ions were

accumulated for 0.5 s in the collision cell, and the time of flight in the optic transfer was set at 0.5 ms. Detection parameters were set using a 4 Mpts resolution in broadband mode to records ions between m/z 57 and 600, and one scan was acquired for each spectrum, corresponding to MS or MS/MS duty cycles of 1.2 s. In the MS/MS experiments, the precursor ion was selected with an isolation window of 1 m/z unit and the collision-induced dissociation (CID) was performed using collision energies ranging from 0 eV to 20 eV. Phosphoric acid was used for calibration. The elemental compositions were assigned with a tolerance of better than 3 ppm.

Computer-aided toxicity prediction

The Toxicity Estimation Software Tool (T.E.S.T.) is a US Environmental Protection Agency computerized predictive system available online with Quantitative Structure Activity Relationships (QSAR) mathematical models.¹⁸⁴ T.E.S.T. has a variety of toxicity endpoints used to predict chemical toxicity values from the physical properties of molecules. The QSAR model uses a simple linear function of molecular descriptors (such as octanol–water partition coefficient, molecular weight or the number of benzene rings):

The toxicity = $ax_1 + bx_2 + c$, where x_1 and x_2 are independent descriptor variables, and a , b , and c are fitted parameters. Models for assessing toxicity solely from molecular structure are based on information-rich structural descriptors that quantify the transport, bulk, and electronic attributes of a chemical structure. We were mainly interested in the oral rat LD₅₀, the developmental toxicity and the mutagenicity results.

Vibrio fischeri bioluminescence inhibition test

In vitro tests were carried out using *Vibrio fischeri* commercial test kits. The freeze-dried luminescent bacteria and the luminometer were purchased from Hach Lange GmbH (Düsseldorf, Germany). The experimental procedure for conducting the bacterial bioluminescence assay is based on the ISO11348–3 protocol (1998). The analysis was carried out with all dilutions and reagents maintained at 15 °C. A working solution of

luminescent bacteria was prepared by reconstituting a vial of frozen lyophilized *Vibrio fischeri* cells using 12.5 mL of the reagent diluent provided by the manufacturer. The reconstituted solution was equilibrated for 15 min at 4°C and the osmolality was adjusted in order to obtain 2% w/v NaCl in each solution or sample. Bacterial reagents were reconstituted just before analysis and the pre-incubation times followed standard protocols. The inhibition percentage (% I) was determined by comparing the relative responses of the control and the diluted sample. Each dilution was tested in duplicate. The sample pH values were approximately 6.3. A fixed amount of bacteria (100 µL of the reconstituted cell suspension) was added to the dilution vials. The luminescence was measured at time zero (before the addition of the test solution) and after 5, 15 and 30 min; it was compared with the value measured for the control solution containing only bacteria.

RESULTS AND DISCUSSION

Chemical analysis of cyprodinil photoproducts

The first step of our work consisted in comparing the chromatograms of irradiated and non-irradiated solutions of cyprodinil. Concomitant photolysis of the deuterated analogue was performed in order to obtain additional structural information on the photoproducts. We focused our interest on the most abundant by-products as they are likely to be present in environmental matrices. All the photoproducts were detected in the solution irradiated for 4 h; this solution was thus used to carry out CID and high-resolution experiments on the photoproducts. At this time, the degradation yield of cyprodinil - estimated on the basis of the total ion current - was about 22%. Figure III-2 shows the LC/MS chromatograms obtained for both reference and irradiated solutions.

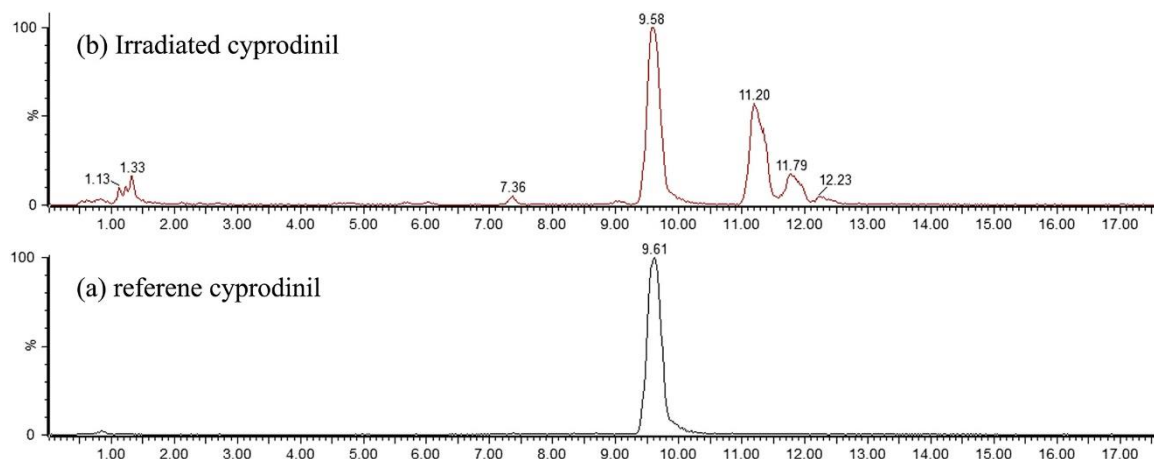


Figure III-2: LC/MS chromatograms recorded for the cyprodinil reference solution (a) and for the same solution irradiated for 4 h (b).

The fragmentation pathways of protonated cyprodinil were studied before investigating the chemical structures of the photoproducts. As cyprodinil has three nitrogen atoms, the $[M+H]^+$ decomposition pathways may follow three pathways (Figure III-3 to Figure III-5), corresponding to initial protonation on a different nitrogen atom. Figure III-3 shows the formation of two ions, m/z 108 and 210, obtained by cleavage of the pyrimidine ring after protonation on N1. The m/z 108 ion is not shifted in the cypro- D_5 MS/MS spectrum, in agreement with elimination of the phenyl ring. The formation of m/z 210 (m/z 215 for cypro- D_5) from $[M+H]^+$ could be explained by the elimination of methane, implying a four-center rearrangement.

Figure III-4 shows the proposed fragmentation of $[M + H]^+$ (protonated on N2), leading to four product ions at m/z 119, 184, 185 and 198. These ions all retain the aromatic ring, as confirmed by the ions at m/z 124, 189, 190 and 203 obtained for cypro- D_5 . Finally, Figure III-5 shows the fragmentation of $[M + H]^+$ protonated on the nitrogen atom of the amine function (N3). The only product ion obtained by this mechanism (m/z 133) is not shifted in the mass spectrum of cypro- D_5 and represents a substituted pyrimidine ring.

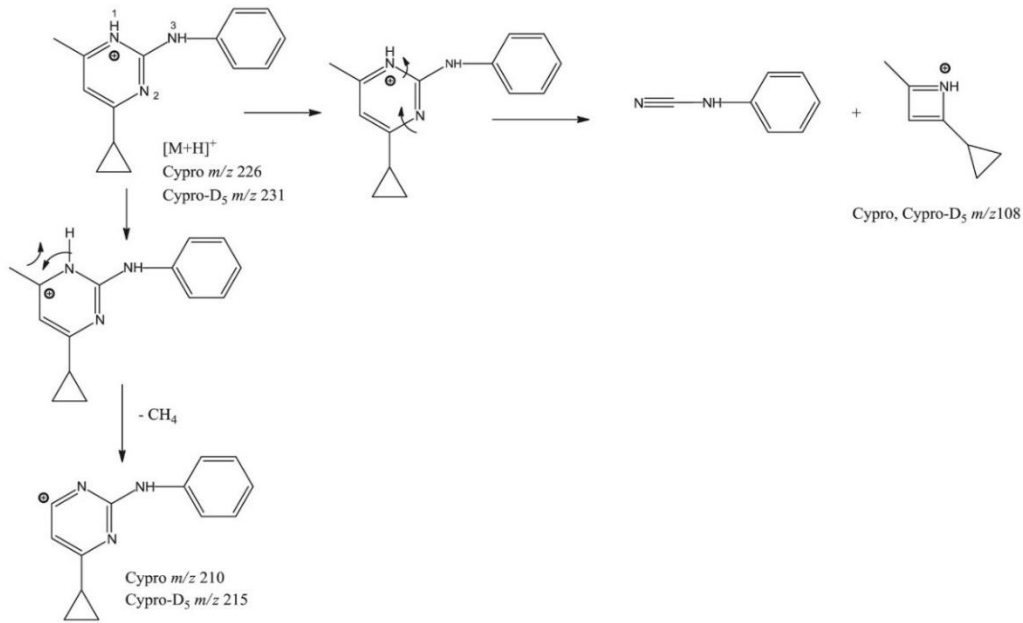


Figure III-3: Suggested dissociation pathways for cyprodinil protonated on N1

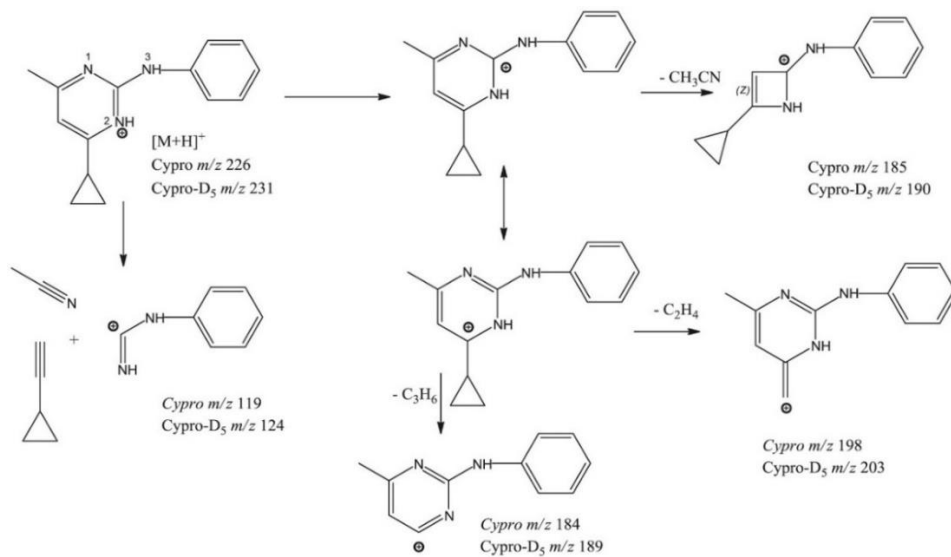


Figure III-4: Suggested dissociation pathways for cyprodinil protonated on N2

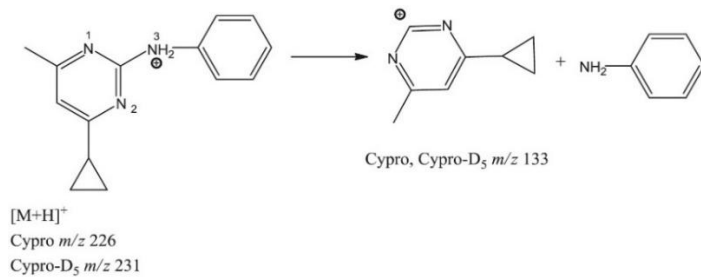


Figure III-5: Suggested dissociation pathways for cyprodinil protonated on N3

The irradiation of cyprodinil leads to 23 compounds with eight elemental compositions. For five of them, several isomers with different retention times are observed. Three compounds result from direct bond cleavage without (TP1) or with hydroxylation (TP2, TP14); several compounds result from cyclization of cyprodinil without (TP3–TP8) or with hydroxylation (TP9–TP13, TP15–TP19, TP 239 and TP 259). The hydroxylated compounds have one, two or three oxygen atoms. Table III-1 displays the elemental composition of TPH^+ , the double-bond equivalency (DBE), the names of the transformation products, their retention times, the m/z values of the protonated molecules, $[\text{TP} + \text{H}]^+$, the corresponding shift obtained with cyprodinil D_5 , and the CID mass spectra of the protonated molecules.

Table III-1: Elemental composition, double-bond equivalency (DBE), molecular weight, retention times (RT) and main CID mass spectra obtained in LC/ESI-MS for cyprodinil and its phototransformation products

Elemental composition of $[\text{TP} + \text{H}]^+$	DBE	Compound	RT (min)	$[\text{TP} + \text{H}]^+$ m/z (shift for cypro- D_5)	CID mass spectra of $[\text{TP} + \text{H}]^+$ (m/z)
$\text{C}_8\text{H}_{12}\text{N}_3$	5	TP-1	1.1	150.1024 (+0)	150–133-108
$\text{C}_7\text{H}_{11}\text{N}_2\text{O}_3$	4	TP-2	2.0	171.0762 (+0)	171–153–129-69
$\text{C}_{14}\text{H}_{16}\text{N}_3$	9	TP-3	1.1	226.1337 (+5)	See Table III-2
		Cypro	10.1	226.1337 (+5)	
		TP-4	11.6	226.1337 (+5)	
		TP-5	12.1	226.1337 (+5)	
		TP-6	12.4	226.1337 (+5)	
		TP-7	12.6	226.1337 (+5)	
		TP-8	12.8	226.1337 (+5)	
		$\text{C}_{14}\text{H}_{14}\text{N}_3\text{O}$	10	TP239–1	
TP239–2	11.5			240.1127 (+5)	
$\text{C}_{14}\text{H}_{16}\text{N}_3\text{O}$	9	TP-9	1.1	242.1283 (+5)	See Table III-3
		TP-10	2.5	242.1283 (+4)	
		TP-11	6.2	242.1283 (+5)	
		TP-12	7.1	242.1283 (+4)	
		TP-13	9.1	242.1283 (+5)	
$\text{C}_{13}\text{H}_{16}\text{N}_3\text{O}_2$	8	TP-14	9.8	246 (+5)	246–204–178-161-136-119-94
$\text{C}_{14}\text{H}_{16}\text{N}_3\text{O}_2$	9	TP-15	1.0	258.1233 (+5)	See Table III-4
		TP-16	2.1	258.1233 (+5)	
		TP-17	6.5	258.1233 (+5)	
		TP-18	7.4	258.1233 (+5)	
		TP19	9.0	258.1233 (+5)	
$\text{C}_{14}\text{H}_{18}\text{N}_3\text{O}_2$	8	TP259–1	3.0	260.1389 (+5)	
		TP259–2	5.0	260.1389 (+5)	

Table III-2 to III-4 give the CID mass spectra of $[\text{TP} + \text{H}]^+$ for each isomer family, namely $\text{C}_{14}\text{H}_{16}\text{N}_3$, $\text{C}_{14}\text{H}_{14}\text{N}_3\text{O}$ and $\text{C}_{14}\text{H}_{16}\text{N}_3\text{O}_2$. They also provide the elemental compositions obtained in the high-resolution mode for the neutral losses observed and the diagnostic product ions.

CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Table III-2: Retention times and main CID mass spectra of $[TP + H]^+$ obtained in LC/ESI-MS for $C_{14}H_{16}N_3$ and its isomers

	RT (min)	Ions m/z										
		226	210	198	185	184	168	159	144	133	119	108
Neutral loss from $[TP + H]^+$		-CH ₄ -C ₂ H ₄ -C ₂ H ₃ N -C ₃ H ₆ -C ₂ H ₆ N ₂ -C ₄ H ₅ N										
Elemental composition of product ions		C ₁₀ H ₁₀ N ⁺ C ₈ H ₉ N ₂ ⁺ C ₇ H ₇ N ₂ ⁺ C ₇ H ₁₀ N ⁺										
TP3	1.1	MH ⁺	+		+				+	++		++
Cyprodinil	10.1	MH ⁺	++	+	+	+				++	+	+
TP4	11.6	MH ⁺	++	+	+	+	+	+	++	++	++	++
TP5	12.1	MH ⁺	++	+	+	+	+	+	++	++	++	++
TP6	12.4	MH ⁺	++	+	+	+	+	+	++	++	++	++
TP7	12.6	MH ⁺	++	+	+	+	+	+	++	++	++	++
TP8	12.8	MH ⁺	++	+	+	+	+	+	++	++	+	++

++: relative intensity $\geq 50\%$

Table III-3: Retention times and main CID mass spectra of $[TP + H]^+$ obtained in LC/ESI-MS for $C_{14}H_{15}N_3O$ isomers

	RT min	Ions m/z									
		242	224	214	214	212	200	198	186	124	119
Neutral loss from $[TP + H]^+$		-H ₂ O -C ₂ H ₄ -CO -CH ₂ O -C ₂ H ₂ O -C ₂ H ₄ O -C ₃ H ₄ O									
Elemental composition of product ions		C ₇ H ₁₀ NO ⁺ C ₇ H ₇ N ₂ ⁺									
TP9	1.1	+	++	+				+	+		
TP10	2.5	+	++					+			+
TP11	6.2	+	++				+	+			
TP12	7.1	+	+		++	++	++	+			++
TP13	9.1	+	++		++				+		

++: relative intensity $\geq 50\%$

Table III-4: Retention times and main CID mass spectra of $[TP + H]^+$ obtained in LC/ESI-MS for $C_{14}H_{15}N_3O_2$ isomers

	RT min	Ions m/z										
		258	240	228	214	212	210	200	190	186	184	119
Neutral loss from $[TP + H]^+$		-H ₂ O -CH ₂ O -C ₂ H ₄ O -CO -H ₂ O -CH ₂ O -CH ₂ O -CO -C ₄ H ₄ O -C ₃ H ₄ O ₂ C ₃ H ₆ O ₂										
Elemental composition of product ions		C ₇ H ₇ N ₂ ⁺ C ₆ H ₈ N ⁺										
TP15	1.0	+	+		++	+				+		
TP16	2.1	+	+		++	++			++			+
TP17	6.5	+	+			++						
TP18	7.4	+	+	+	+	++		+		+	+	+
TP19	9.0	+	+	+		++	+			+		+

++: relative intensity $\geq 50\%$

Table III-5–8 summarizes the proposed structures for TPs and their associated *in silico* predicted toxicity values for oral rat LD₅₀. Decomposition pathways of [TP + H]⁺ for TP9 to TP19 are proposed in the Supporting Information (files Supplementary Information III-1 to Supplementary Information III-12).

The fragmentation pathways of [cyprodinil + H]⁺ have been used for the identification of the TPs, assuming that the photoproducts generated have kept the structural skeleton of the parent compound or may share common molecular moieties with the original compound. The presence of characteristic neutral losses (16 Da for CH₄, 18 for H₂O, 26 for C₂H₂, 28 for CO or C₂H₄, 41 for CH₃CN, 42 for C₃H₆ or CH₂CO, 93 for C₆H₇N...) and that of diagnostic product ions (*m/z* 133, 119, 108, 94...) observed for the degradation of cyprodinil and cypro-D₅ facilitated the structural assignments.

The identification of photodegradation products included several experimental steps. The first one consisted in recording full scan mass spectra at low cone voltage and low collision energy to preserve the protonated molecules [M + H]⁺. The comparison of the two chromatograms allowed direct determination of the molecular weights of the photoproducts. The molecular weights of the compounds were often confirmed by the presence of both [M + H]⁺ and [M + Na]⁺ ions. A confident structural elucidation of TPs has been achieved using (1) isotopic labelling, (2) high-resolution mass measurements, and (3) CID mass spectra for each [TP + H]⁺ ion. Among the identified TPs, two (TP1 and TP2) have the same molecular weight after irradiating both cyprodinil and cypro-D₅. This means that these compounds have lost the aromatic ring during photolysis. The elemental composition of [TP1 + H]⁺ is C₈H₁₂N₃, and TP1 was characterized as 4-cyclopropyl-6-methylpyrimidin-2-amine based on its CID mass spectrum. The corresponding TP has also been observed for pyrimethanil¹⁸⁵ and mepanipyrim.¹⁷⁹ TP1 loses either NH₃ from the form protonated on the amino function to give *m/z* 133, or NCNH₂ from the structure protonated on another nitrogen atom to give *m/z* 108. These two ions are also observed in the fragmentation of the [cyprodinil + H]⁺. The elemental composition of [TP2 + H]⁺ is C₇H₁₁N₂O₃: one carbon atom and one nitrogen atom fewer than TP1. We assume that the methyl group and probably the nitrogen atom of the amine function have been replaced by two hydroxyl groups, based on the fragmentation of [TP2 + H]⁺ which leads to the

formation of m/z 153 (loss of H_2O) and m/z 129 (loss of $OHCCCH$). This latter ion gives m/z 69 by elimination of $NH_2C(OH)NH$ (see Supplementary Information III-13).

Six compounds (TP3 to TP8) yield a $[M + H]^+$ ion at m/z 226 and have the same elemental composition as cyprodinil. Isomerization reactions have been obtained previously by irradiation of fungicides.^{129,186} The corresponding compounds in the irradiated cypro-D₅ solution are all shifted by +5 mass units. Moreover, the six CID mass spectra look very similar and provide the same product ions as those observed for cyprodinil, although with different relative intensities. One isomer (TP3) is eluted very quickly (retention time (RT) = 1.3 min), while the other five (TP4 to TP8) are eluted after cyprodinil. We aimed at investigating if those isomers might result from a cyclization process. Several mechanisms have been considered involving the cyclisation between a N=C bond of the pyrimidine ring and one of the C=C bonds of the aromatic ring. As the pyrimidine ring is not symmetric, two cyclizations have been considered which lead to the formation of two biradicals A and B. These biradicals allow the formation of cyprodinil isomers (see Supplementary Information III-14). We assumed that irradiation of the C=C bond of the pyrimidinic ring leads to two other isomers of cyprodinil (see Supplementary Information III-15). The intermediates C and D lead to the formation of two isomers including a five-membered ring, implying that the carbon atoms of the isopropyl ring are bonded to either a carbon atom or a nitrogen atom of the pyridine ring. Only the structures that have five deuterium atoms, when photolysis occurs on cypro-D₅, have to be considered. The more intense ions of five CID mass spectra are given in Table III-2. The CID mass spectra of the isomers are very similar, making it difficult to assign a compound to a given chromatographic peak. Only the structures 225-2a, 225-3a, 225-4a and 225-4b allow the formation of the m/z 119 ion from the protonated molecule.

The relative intensity of m/z 210, corresponding to the elimination of CH_4 from $[M + H]^+$, is lower for TP3. The CID mass spectrum of $TP3H^+$ does not display m/z 119, and this ion is of low intensity for TP4. Structures that include the aniline function of cyprodinil give an ion at m/z 119 with low relative abundance; this ion is shifted to m/z 124 for the irradiated cypro-D₅ solution, indicating that the aromatic ring has been retained. Except for $[TP3 + H]^+$, the ion at m/z 184 corresponding to neutral loss of C_3H_6 from $[M + H]^+$ is observed in all the CID mass spectra of the $[TP + H]^+$ ions. The C_3H_6 neutral corresponds to

the isopropyl ring. It can be reformed by protonation of the nitrogen atom of the pyrimidine ring for the compounds having a five-membered ring including a nitrogen atom. TP3 is the only isomer which can be attributed to the structure 225–1. Consequently, all the others structures (shifted +5 mass units for cypro-D₅) have been evaluated by *in silico* tests.

The biradicals may, in the presence of water, be oxidized one or two times. The proposed structures arise from the addition of a water molecule to the structures A, B, C and D and their mesomeric forms (see Supplementary Information III-16 and Supplementary Information III-17). The compounds with one oxygen atom have C₁₄H₁₅N₃O or C₁₄H₁₃N₃O as their elemental composition. After 4 h of irradiation, the solution of cyprodinil provides a dozen compounds with C₁₄H₁₅N₃O as their elemental composition. The relative abundances of five of these compounds allowed us to record the CID mass spectra given in Table III-1. These CID spectra provide differences that were used to assign a structure to each compound. Three compounds (TP9, TP11 and TP13) have a structure shifted by 5 mass units in the corresponding irradiated cypro-D₅ solution, and two (TP10 and TP12) that are shifted by +4. For the five compounds, the protonated molecule MH⁺ loses water to give an ion at *m/z* 224. Three compounds lose 28 Da to give the ion at *m/z* 214. The high-resolution measurement allowed this loss of 28 Da to be assigned as the elimination of C₂H₄ from [TP9 + H]⁺, and CO from [TP12 + H]⁺ and [TP13 + H]⁺. As for TP3, the retention time of TP9 is shorter than those of its analogue, suggesting that its structure corresponds to hydroxylated TP3. [TP + H]⁺ can expel CO if the hydroxyl group is on a tertiary carbon atom bonded to two proton-attracting atoms, and not bonded to a free alkyl chain. The location of the OH group in TP3 is determined by the fragmentation of [TP9 + H]⁺. The *m/z* 198 and 186 product ions are obtained by consecutive eliminations from [TP3 + H]⁺ at *m/z* 242 of H₂O and C₂H₂ for *m/z* 198, and CO and C₂H₄ for *m/z* 186.

The structure of TP13 has been assigned based upon the absence of the diagnostic ion at *m/z* 119 and the presence of *m/z* 124 whose elemental composition is C₇H₁₀NO⁺. This ion is not shifted for irradiated cypro-D₅. The hydroxyl group is on a tertiary carbon, near a nitrogen atom and a double bond; it promotes the decarboxylation of [TP13 + H]⁺.

As for $[TP13 + H]^+$, the decomposition of $[TP11 + H]^+$ does not lead to the formation of the m/z 119 ion. This absence may be explained by oxidation of the aromatic ring or by involvement of the latter in a fully delocalized structure. As for $[TP9 + H]^+$ and $[TP11 + H]^+$, the m/z 198 ion is observed: the elimination of H_2O from $[TP + H]^+$ allows subsequent elimination of C_2H_2 . In the case of TP13, this ion is shifted by 5 mass units, meaning that C_2H_2 is not eliminated from the aromatic ring. Moreover, $[TP11 + H]^+$ gives an ion at m/z 200 corresponding to the elimination of CH_2CO . These results led us to propose the structure displayed in Table III-7.

CID mass spectra of $[TP10 + H]^+$ and $[TP12 + H]^+$ display the m/z 119 ion. For $[TP10 + H]^+$, $[TP10H-H_2O]^+$ loses C_2H_2 to give m/z 198 according the proposed structure. Finally, the CID mass spectra of $[TP12 + H]^+$ shows an ion at m/z 212 obtained by elimination of CH_2O from the protonated molecule. This neutral loss is specific of the presence of a primary alcohol or aldehyde function. Considering that photolysis may not lead to direct hydroxylation of a methyl group, we propose the formation of an aldehyde on the terminal carbon of the alkyl chain of the pyrimidine substituent (see Supplementary Information III-4).

Two compounds provide a MH^+ ion at m/z 240. They have an elemental composition of $C_{14}H_{13}N_3O$ and are shifted to either m/z 243 or m/z 245 with cypro- D_5 . The addition of an oxygen atom occurs on the aromatic ring for the first one and on the pyrimidine part for the second. In both cases, the protonated molecule is of too low abundance for a CID mass spectrum to be recorded. The structures of these compounds have not been fully elucidated.

The elemental composition of TP14 is $C_{13}H_{15}N_3O_2$; the compound has formally lost a carbon atom and gained two oxygen atoms from the parent molecule. During irradiation, the $CH_3\cdot$ radical has been probably replaced by a hydroxyl function, as indicated by the formation of the m/z 204 ion, obtained from $[M + H]^+$ by elimination of $OHCCH$. Both the presence of the m/z 119 ion and the formation of m/z 136 by elimination of C_4H_4O from m/z 204 allow us to locate the position of the second hydroxyl function in TP14. The ion at m/z 178 is obtained from $[M + H]^+$ by elimination of C_4H_4O . The decomposition pathway

of [TP14 + H]⁺ is proposed in Supplementary Information III-6 and Supplementary Information III-7.

The elemental composition of TP15 to TP19 is C₁₄H₁₅N₃O₂. These TPs have a hydroxyl and a carbonyl function. All compounds are shifted by 5 mass units when irradiating the cyprodinil-D₅ solution. All the protonated molecules lose H₂O to generate the *m/z* 240 ion, and give an ion at *m/z* 212 obtained by a subsequent loss of CO. Since the ion at *m/z* 119 is observed in the CID mass spectra of these protonated molecules, the possibility of oxidation on the aromatic ring for these TPs is precluded. [TP15 + H]⁺, [TP16 + H]⁺ and [TP18 + H]⁺ lose C₂H₄O, corresponding to the successive eliminations of H₂O and CHCH, to give *m/z* 214. In the irradiated cypro-D₅ solution, this ion is shifted by 3 mass units for TP15 and TP16, and by 5 mass units for TP18. The elimination of HCCH in TP15 and TP16 implies these are carbon atoms of the cyprodinil aromatic ring. The presence of the *m/z* 190 ion for TP16 and TP18, obtained by elimination of C₄H₄O from MH⁺, allows us to locate the hydroxylation onto the isopropyl alkyl chain. As for TP12 and TP13, an elimination of CH₂O from TPH⁺ is observed for TP 18 and TP19, indicating oxidation at the end of the alkyl chain. The elimination of CH₂CO implies that one aldehyde has been fixed during the irradiation process. Alternatively, a primary hydroxyl function may be present on the molecule.

The CID mass spectrum of [TP17 + H]⁺ does not include the ions obtained by characteristic neutral eliminations from MH⁺, such as 26 Da for C₂H₂, 68 Da for C₄H₄O and 72 Da for C₃H₄O₂, meaning that oxidation has not taken place on the pyrimidine ring. Moreover, the diagnostic ion at *m/z* 119 is not observed. Two structures are proposed for this compound on the basis of these observations.

Finally, two other photoproducts with C₁₄H₁₇N₃O₂ as elemental composition are observed. The relative intensity of these compounds is very low and does not allow their structural characterization. The two [M + H]⁺ ions lose one then two H₂O molecules, suggesting the presence of two hydroxyl groups for these photoproducts.

In silico toxicity prediction

In order to evaluate the toxicity of cyprodinil and its 19 photoproducts, *in silico* tests were performed using the toxicity simulation programme described above. In this study we chose to report only the most significant results, i.e. the potential oral rat lethal dose (LD₅₀), mutagenicity and developmental toxicity including endocrine perturbation and reprotoxicity signs. The potential oral rat lethal dose (LD₅₀) was assessed with QSAR tests for cyprodinil and each photoproduct (see Table III-5–8). The LD₅₀ indicates the amount of chemical in mg.kg⁻¹ body weight that would cause the death of 50% of a test population of rats after oral ingestion. The oral rat LD₅₀ estimated for cyprodinil (909.2 mg.kg⁻¹) is coherent with the reporting values by the food and agriculture organization of the UN.¹⁶⁵ This value corresponds to a slight toxicity according to the Hodge and Sterner toxicity classification scale (LD₅₀ ranging from 500 to 5000 mg.kg⁻¹).¹⁸⁷ Among the 19 structures whose toxicity has been tested, only three exhibit a potential toxicity lower than that of cyprodinil and five photoproducts (TP225-4a, TP225-4b, TP17a, TP18 and TP19) exhibit a potential toxicity lower than or similar to that of cyprodinil. Oral values of 15.24 and 31.51 mg.kg⁻¹ were estimated for TP13 and TP15, respectively. These values correspond to a high toxicity. All the other compounds seem to be slightly or moderately toxic.

Recent studies showed that cyprodinil enhanced dexamethasone-induced gene expression and activated ERK activity, and, therefore, acts as both a potential endocrine and a cellular physiological disrupter.¹⁶⁸ However, developmental toxicity test results, displayed in Table III-9, show that no developmental toxicity was observed for cyprodinil, TP1 and TP2 although all the other photoproducts exhibited a potential risk for development with estimated values around 0.5 and 0.95 A.U.

CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Table III-5: Proposed structures and associated *in silico* predicted toxicity values for oral rat LD₅₀ for cyprodinil and its photoproducts TP1, TP2 and TP14

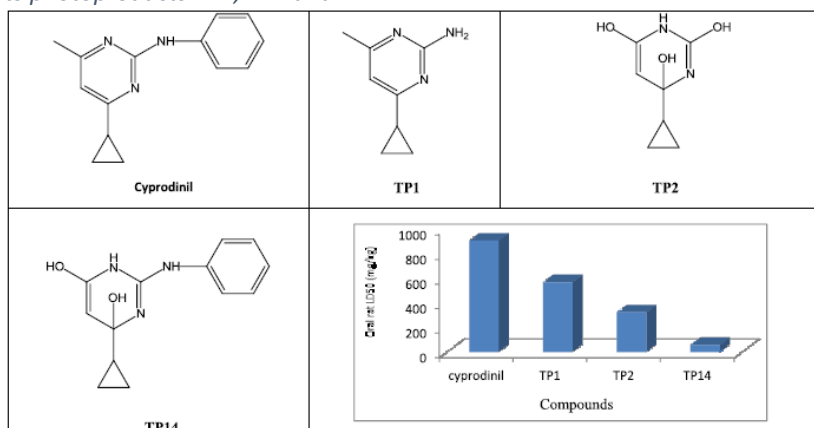


Table III-6: Proposed structures and associated *in silico* predicted toxicity values for oral rat LD₅₀ for cyprodinil and isomers of TP3

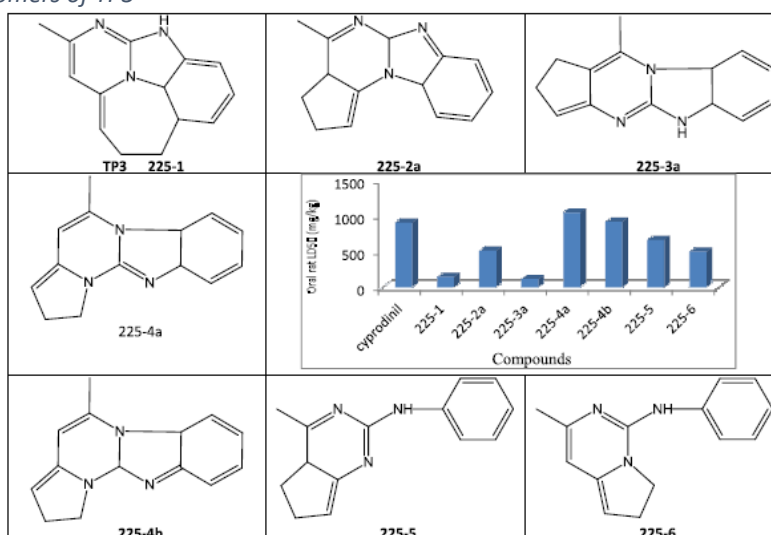


Table III-7: Proposed structures and associated *in silico* predicted toxicity values for oral rat LD₅₀ for cyprodinil and its photoproducts TP9 to TP13

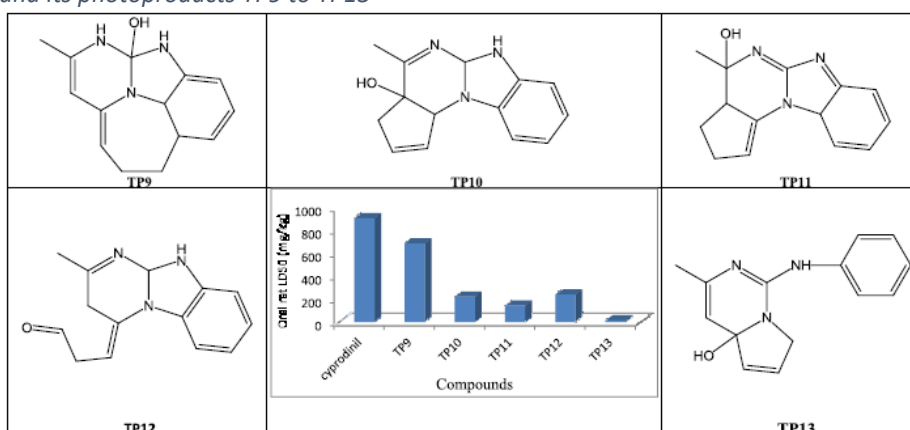
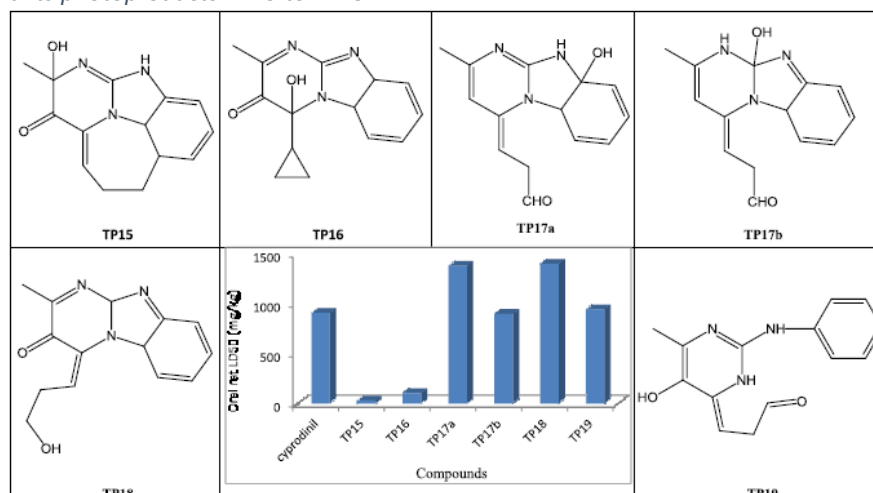


Table III-8: Proposed structures and associated *in silico* predicted toxicity values for oral rat LD₅₀ for cyprodinil and its photoproducts TP15 to TP19



The results of the mutagenicity estimation test, displayed in Table III-9, indicated that cyprodinil and six of its degradation products (TP2, TP225-2a, TP225-4a, TP225-4b, TP9 and TP18) were expected to be mutagenic. This is in agreement with the results of a previous study reporting that cyprodinil promotes carcinogenesis by increasing the transactivation activity of the cytochrome P450. Cyprodinil may increase cancer cell proliferation by stimulating the protein expressions of cyclin D1 and cyclin E. These compounds can also promote metastasis by increasing the protein expression of cathepsin D via an ER-dependent pathway.¹⁶⁹ Furthermore, other studies report that cyprodinil induces the phosphorylation of the extracellular signalling regulated kinase (ERK) that phosphorylates growth and transcription factors, and regulates various processes including proliferation, differentiation, survival, and migration of mammalian cells.^{188,189}

Table III-9: *In silico* predicted toxicity values and results for developmental toxicity and mutagenicity tests for cyprodinil and its photoproducts

Compound	Developmental toxicity value	Developmental toxicity result	Mutagenicity value	Mutagenicity result
Cyprodinil	0.45	No	0.82	Yes
TP1	0.5	No	0.3	No
TP2	0.41	No	0.52	Yes
TP3	0.56	Yes	0.04	No
225-2a	0.88	Yes	0.66	Yes
225-3a	0.81	Yes	0.26	No
225-4a	0.86	Yes	0.69	Yes
225-4b	0.91	Yes	0.51	Yes
225-5	0.82	Yes	0.29	No
225-6	0.54	Yes	0.08	No
TP9	0.95	Yes	0.57	Yes
TP10	1.00	Yes	0.39	No
TP11	0.63	Yes	0.19	No
TP12	0.77	Yes	0.36	No
TP13	0.64	Yes	0.35	No
TP14	0.69	Yes	-0.34	No
TP15	0.75	Yes	0.12	No
TP16	0.58	Yes	0.46	No
TP17a	0.81	Yes	0.43	No
TP17b	0.87	Yes	0.36	No
TP18	0.68	Yes	0.52	Yes
TP19	0.57	Yes	0.06	No

Whatever the chemical modification (hydroxylation after direct bond rupture, isomerization by cyclization without and with oxidation), most of the photoproducts are expected to be much more toxic than cyprodinil, which is, according to the toxicity scale classification of Hodge and Sterner, moderately toxic.¹⁸⁷ These results remain QSAR estimations but such a difference in predicted toxicity between the parent compound and most of its degradation products suggest that 'real' bioassays should be carried out.

***In vitro* bioassay results**

According to the procedure detailed in the Experimental section, a reference solution of cyprodinil and a 4 h photolyzed solution at 130 mg.L⁻¹ in water/acetonitrile (95:5) were tested. The inhibition percentage after 30 min of incubation, presented in Figure III-6, shows that the irradiated solution toxicity for the *Vibrio fischeri* bacteria is higher than that of the reference solution.

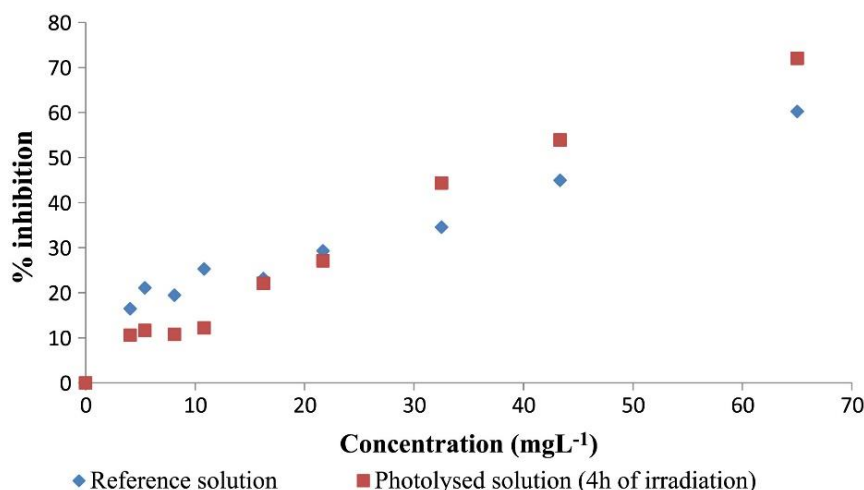


Figure III-6: Response-dose curves of reference and 4 h photolyzed solutions of cyprodinil with the *Vibrio fischeri* bioluminescence inhibition test.

The EC₅₀ values of those solutions were determined for the reference and the photolyzed solution at 63.18 mg.L⁻¹ and 40.43 mg.L⁻¹, respectively. UV-vis irradiation leads to an increase in the toxicity of the solution. Due to the complexity of the chromatographic separation, the photoproducts could not be tested separately. According to the *in silico* approach, this global increase in toxicity could be mainly attributed to the photoproducts which result from the cyclization of cyprodinil.

CONCLUSIONS

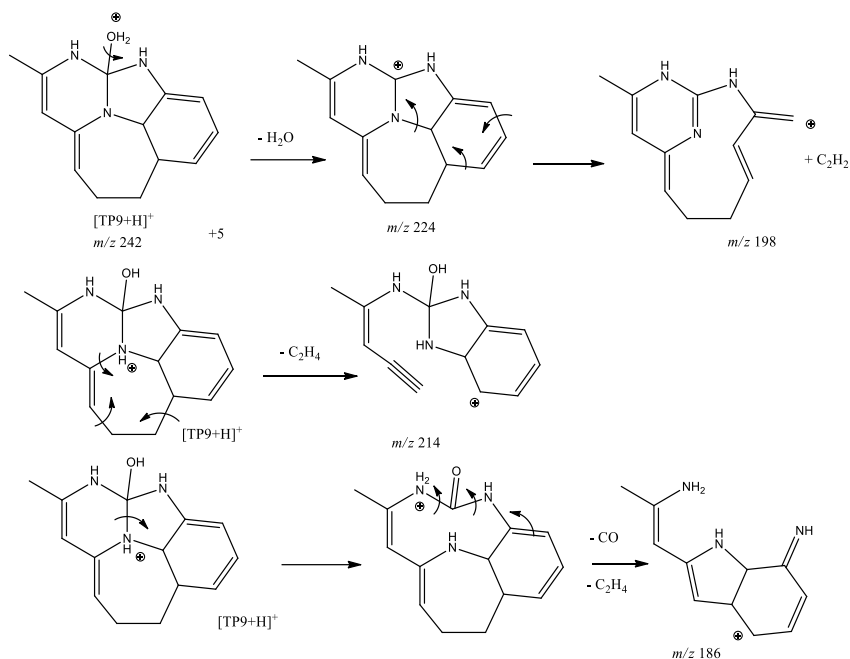
Nineteen structures of photoproducts were characterized after irradiation of a cyprodinil aqueous solution, by combining HRMS/MS and the use of a labelled compound. Only TP1 has been reported elsewhere. The photoproducts arise from several degradation reactions, namely direct bond cleavages, and bond and cyclization mechanisms with and without hydroxylation. Cyclization mechanisms involve one of the three double bonds of the pyrimidine ring with the aromatic ring or the cyclopropyl ring. The combined use of CID experiments and labelled compounds allowed the identification of new cyclic structures. The potential toxicities for these compounds were estimated by *in silico* tests with TEST software. Most of them exhibit a toxicity significantly higher than that of the parent molecule. *In vitro* assays on *Vibrio fischeri* bacteria showed that the global

ecotoxicity of the cyprodinil solution significantly increases with irradiation time, in agreement with the results provided by *in silico* toxicity predictions. Among many others, this study contributes to clearly showing that routine water analysis should also take into account some degradation products and not only the listed environmental pollutants.

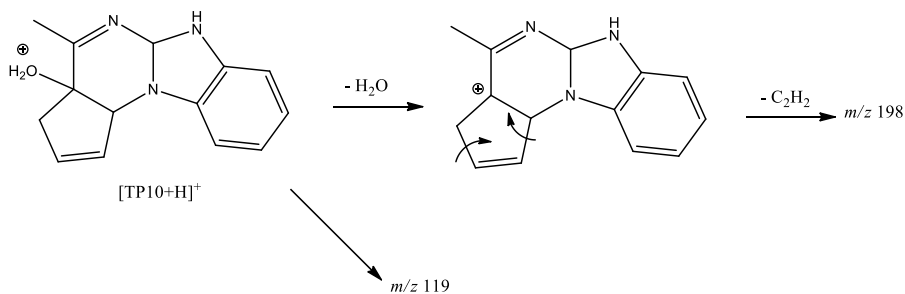
ACKNOWLEDGMENT

Financial support from the National FT-ICR network (FR 3624 CNRS) for conducting the research is gratefully acknowledged.

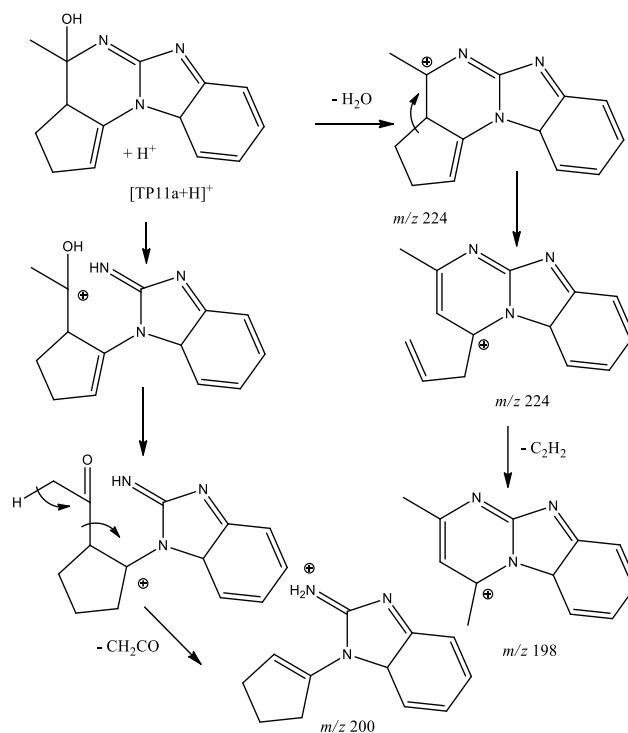
SUPPLEMENTARY INFORMATION



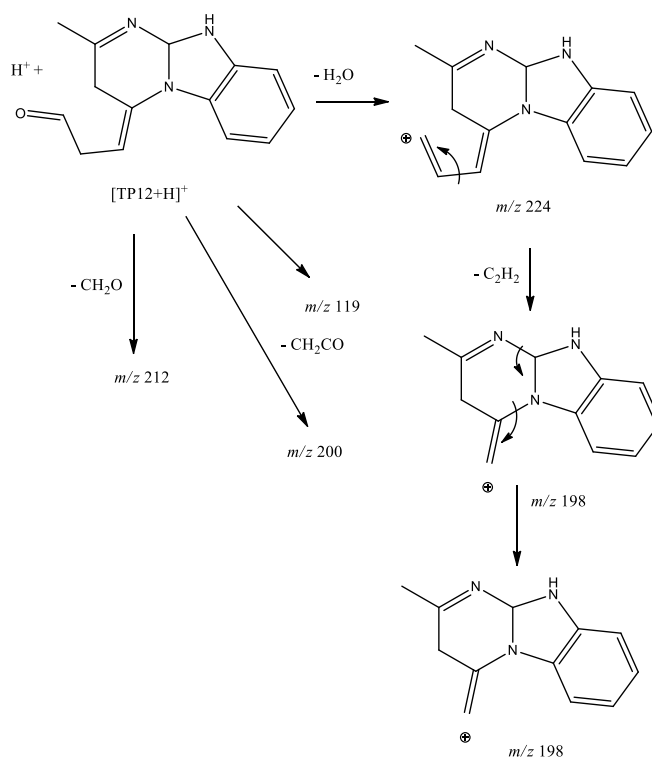
Supplementary Information III-1: Suggested dissociation pathways for $[TP9 + H]^+$



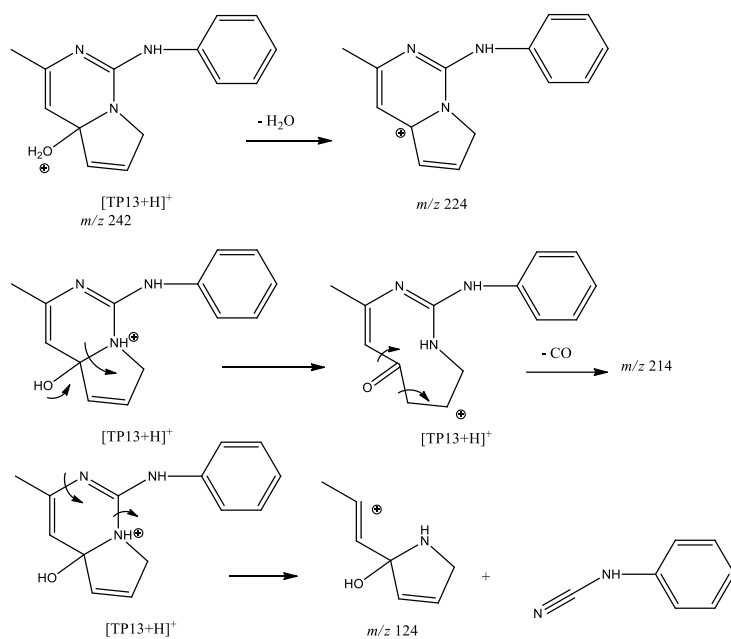
Supplementary Information III-2: Suggested dissociation pathways for $[TP10 + H]^+$



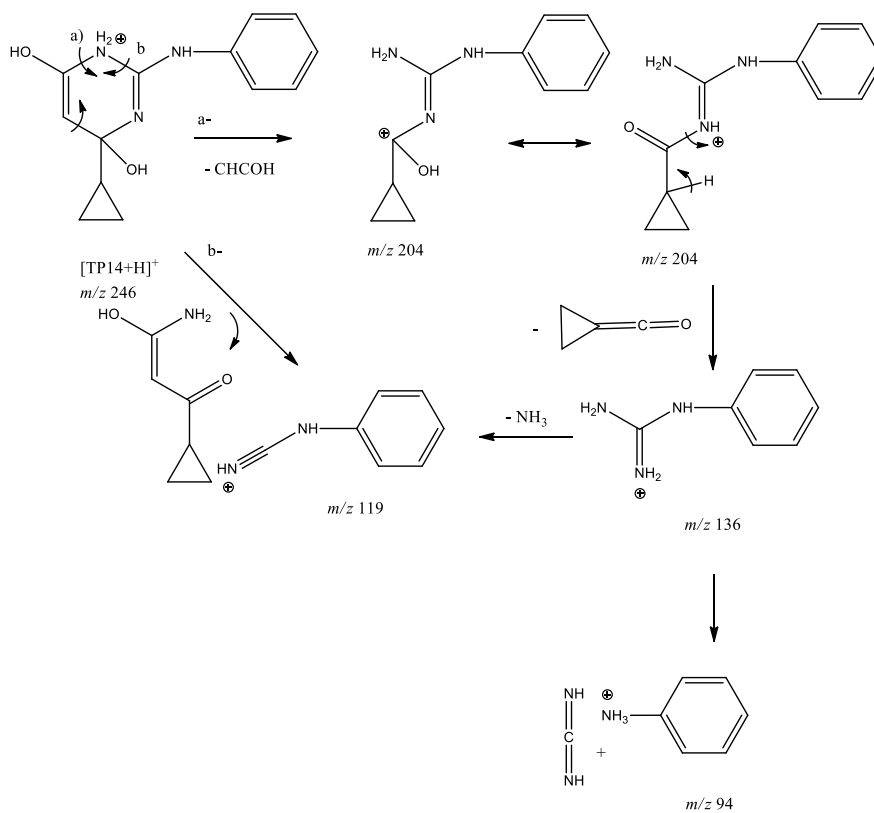
Supplementary Information III-3: Suggested dissociation pathways for $[TP11 + H]^+$



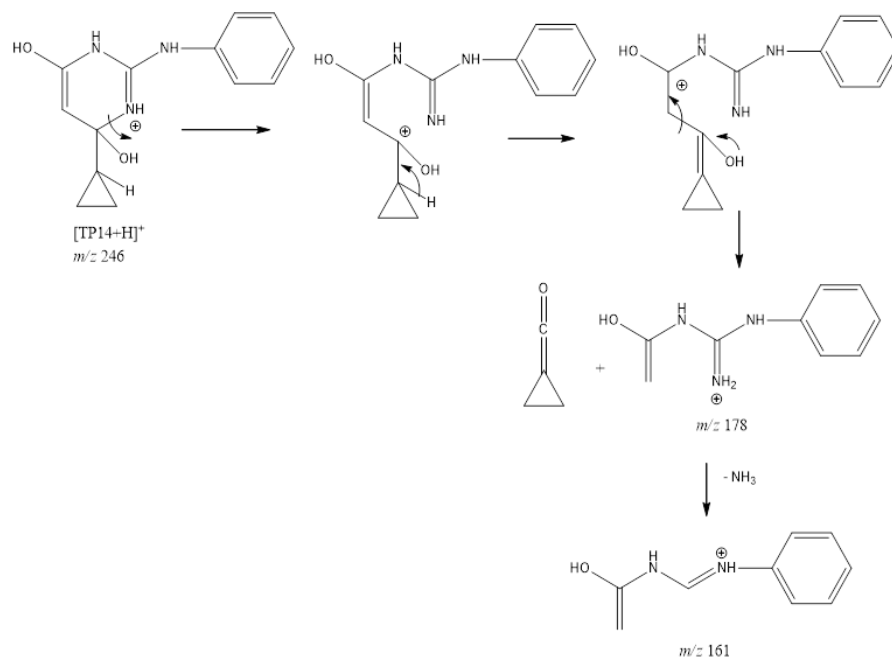
Supplementary Information III-4: Suggested dissociation pathways for $[TP12 + H]^+$



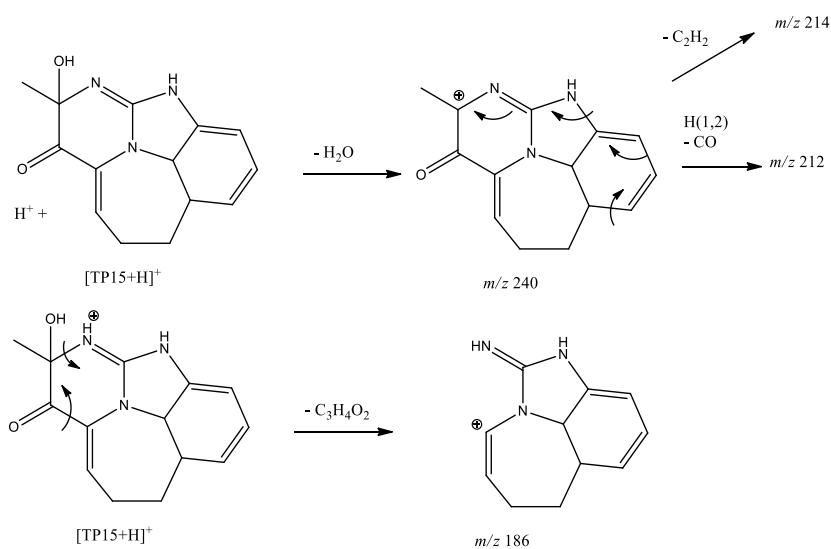
Supplementary Information III-5: Suggested dissociation pathways for $[TP13 + H]^+$



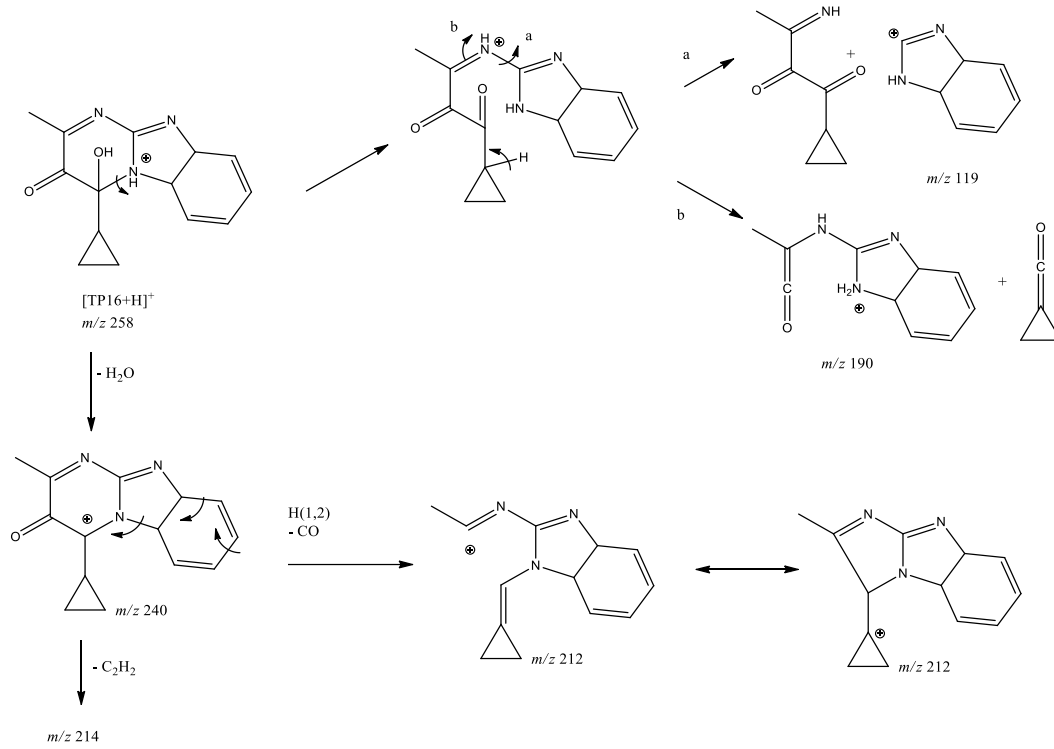
Supplementary Information III-6: Suggested dissociation pathways for $[TP14 + H]^+$



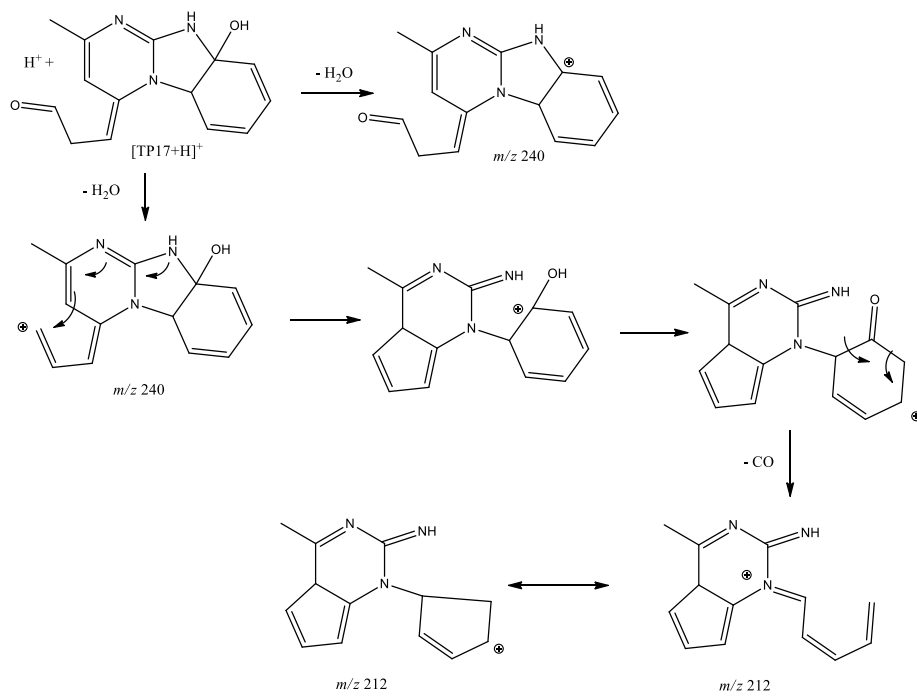
Supplementary Information III-7: Suggested dissociation pathways for $[TP14 + H]^+$



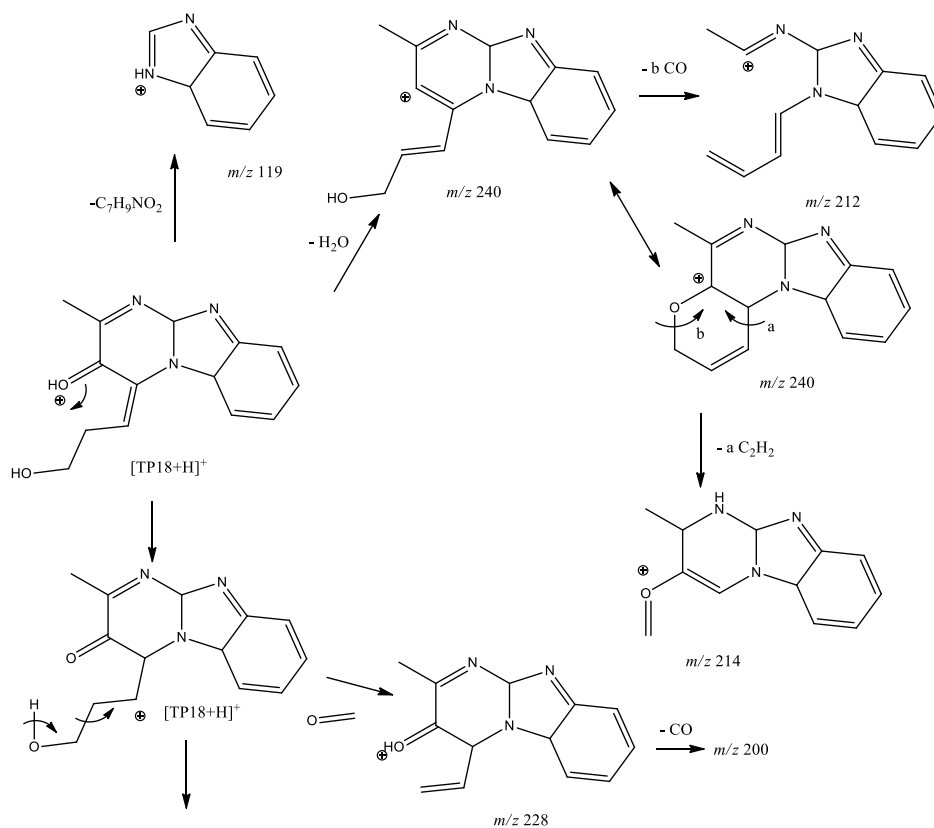
Supplementary Information III-8: Suggested dissociation pathways for $[TP15 + H]^+$



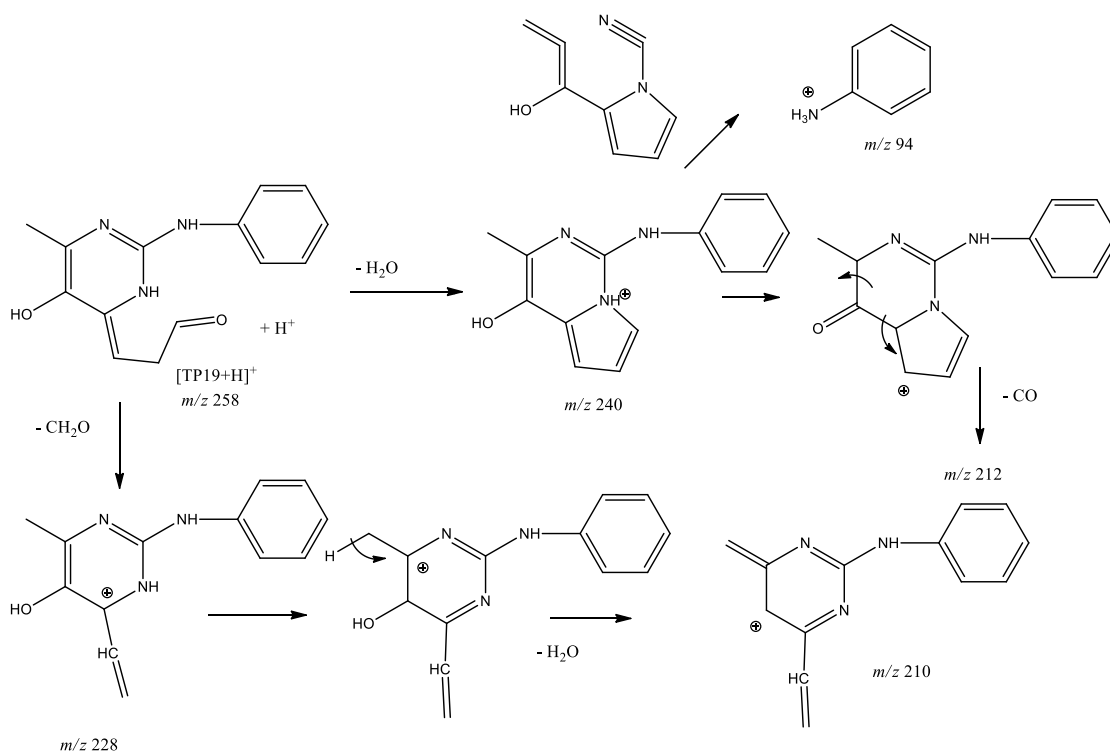
Supplementary Information III-9: Suggested dissociation pathways for $[TP16 + H]^+$



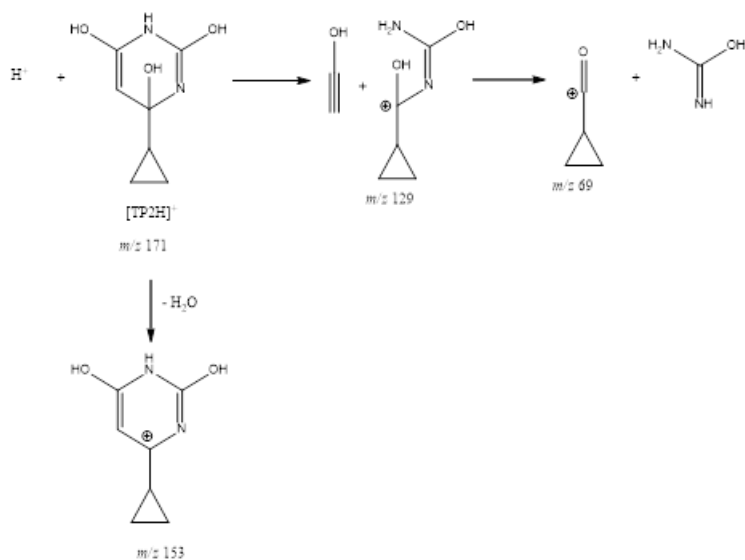
Supplementary Information III-10: Suggested dissociation pathways for $[TP17 + H]^+$



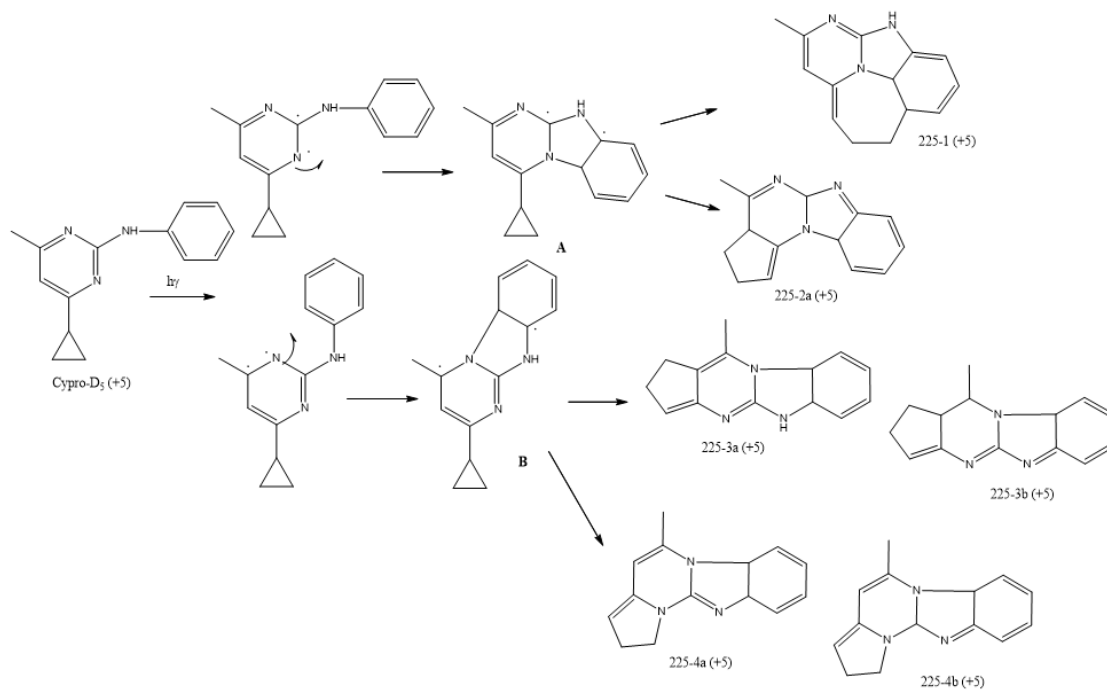
Supplementary Information III-11: Suggested dissociation pathways for $[TP18 + H]^+$



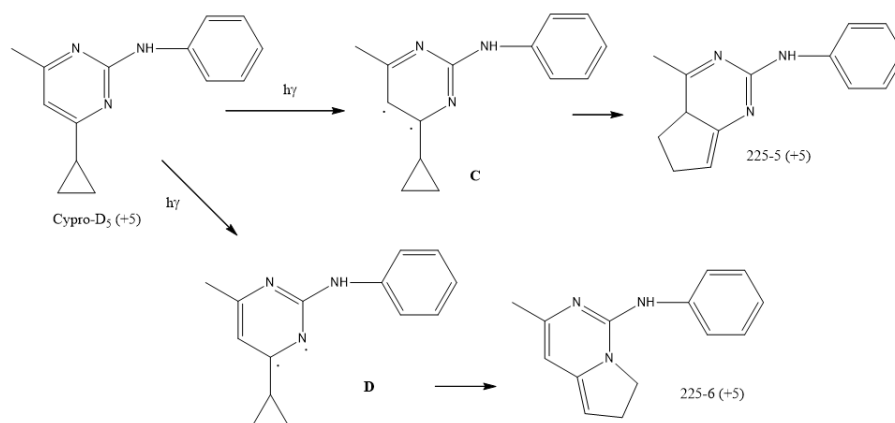
Supplementary Information III-12: Suggested dissociation pathways for $[TP19 + H]^+$



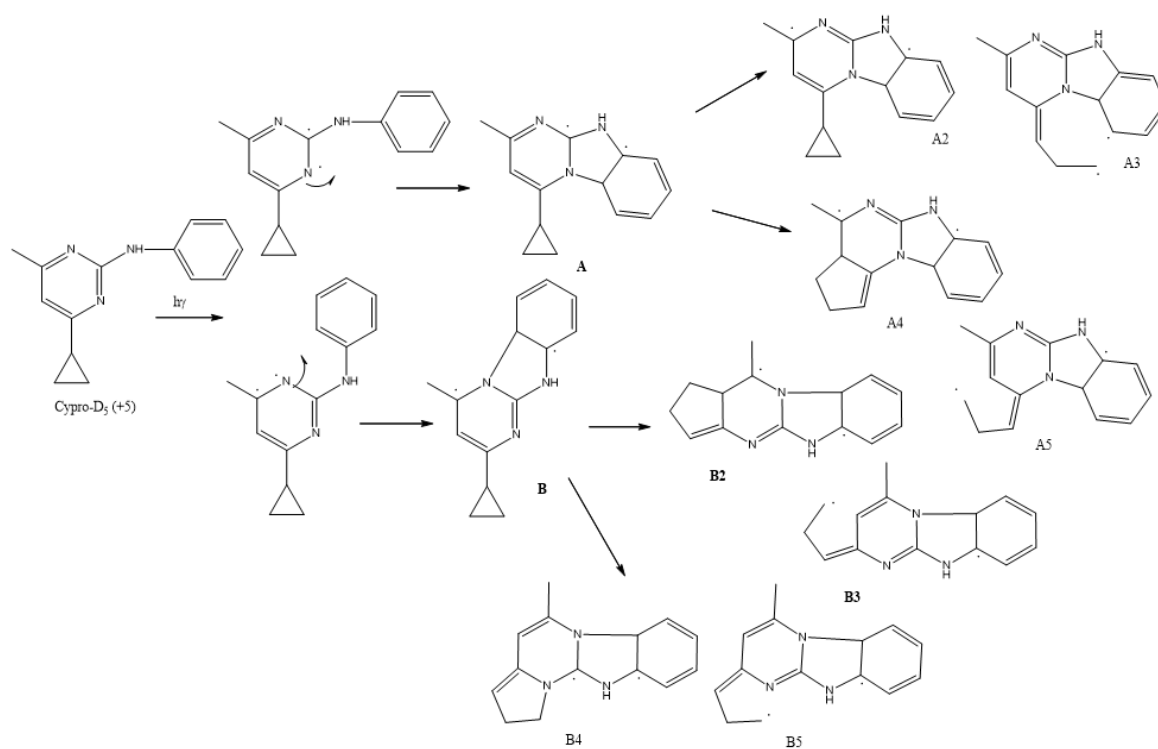
Supplementary Information III-13: Suggested dissociation pathways for $[\text{TP2} + \text{H}]^+$



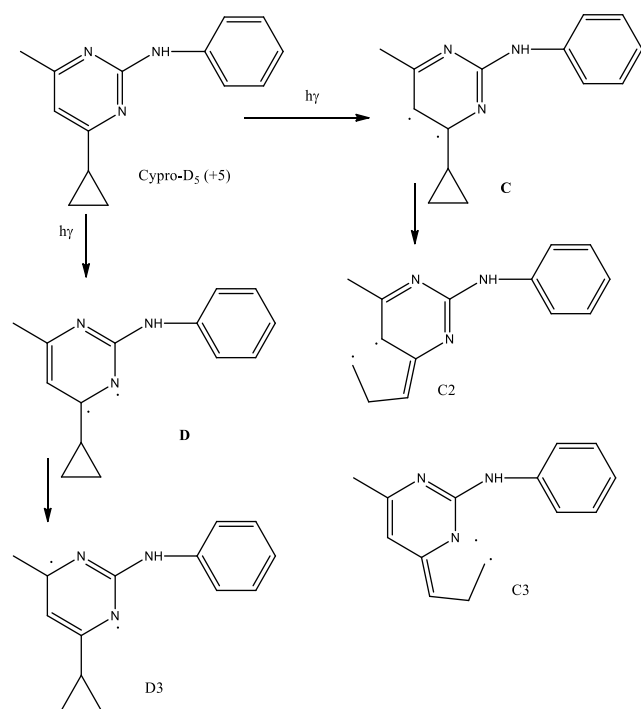
Supplementary Information III-14: Suggested phototransformation pathways of cyprodinil



Supplementary Information III-15: Suggested phototransformation pathways of cyprodinil



Supplementary Information III-16: Suggested phototransformation pathways of cyprodinil



Supplementary Information III-17: Suggested phototransformation pathways of cyprodinil

III.1.2. Synthèse des principaux résultats de cette étude

Les structures de 19 photoproduits du cyprodinil ont été caractérisées par LC-HRMS/MS et confirmées par les résultats obtenus pour le cyprodinil-D₅. Un seul de ces composés avait déjà été rapporté dans la littérature. La toxicité potentielle de chacun de ces photoproduits a été estimée par des tests *in silico* démontrant que la plupart d'entre eux sont significativement plus toxiques que le cyprodinil lui-même, que ce soit pour la LD₅₀ orale chez le rat, la toxicité développementale ou la mutagénicité. L'écotoxicité de la solution, mesurée par des tests *in vitro* sur *Vibrio fischeri*, augmente significativement avec la durée de la photolyse, conformément aux résultats obtenus par les tests *in silico*.

L'ajout de ces produits de dégradation à ceux déjà ciblés lors des analyses réglementaires de routine de produits alimentaires s'avérerait donc très pertinent.

Pour ce qui est du deuxième objectif du laboratoire (c'est-à-dire de mieux comprendre les mécanismes réactionnels liés à la photolyse de polluants), nous avons pu caractériser des transformations impliquant de nombreuses réactions telles que la rupture homolytique de liaison, la cyclisation, l'isomérisation et l'hydroxylation.

III.2. PHOTODEGRADATION DE LA BENZISOTHIAZOLINONE : IDENTIFICATION ET ACTIVITE BIOLOGIQUE DES PRODUITS DE DEGRADATION

III.2.1. Présentation des travaux publiés

La benzisothiazolinone est un fongicide et un agent antimicrobien de la famille des isothiazolinones, utilisé en tant que conservateur. Elle agit contre divers champignons et bactéries grâce à la présence d'un atome de soufre capable de former des ponts disulfure en oxydant les groupes thiols des bactéries (groupe thiol de la cystéine par exemple).¹⁹⁰ Son manque de spécificité d'action lui confère un caractère toxique : elle peut en effet interagir avec toutes les protéines intracellulaires contenant du soufre, quel que soit l'organisme considéré. Elle est ainsi considérée comme fortement toxique vis-à-vis des organismes aquatiques, mais aussi pour l'être humain.¹⁹¹⁻¹⁹³ Des sensibilisations de la peau, des dermatites de contact allergique, des irritations pour la peau et les yeux ont été reportés par l'Agence Européenne des produits chimiques ; des cas de dermatites de contact par voie aérienne ont aussi été signalés.^{191,192,194} Malgré sa proscription dans les produits cosmétiques, elle reste largement présente dans les produits de consommation domestique, notamment en tant qu'agent de conservation dans les peintures, les vernis, les produits d'entretien ménager, les lessives, les assouplissants, les solutions de traitement de cuirs, mais aussi en mélange dans les formulations de produits pesticides.¹⁹⁵⁻¹⁹⁸ De manière générale, la concentration en benzisothiazolinone dans ces produits varie entre 200 et 400 ppm en fonction de l'utilisation et de la nature des biocides avec lesquels elle est associée. On retrouve la benzisothiazolinone dans les eaux de station d'épuration et dans l'environnement en forte concentration.¹⁹⁹ Afin de réduire sa concentration dans l'environnement, plusieurs techniques de dégradation sont utilisées, notamment l'ozonolyse et la photodégradation. Malheureusement l'efficacité de ces techniques est estimée sur la base de la décroissance des concentrations en isothiazolinones et ne prend pas en compte la toxicité des produits de dégradation obtenus.

Nous avons étudié la dégradation photochimique afin d'identifier les photoproduits de la benzisothiazolinone dans l'eau et d'évaluer leur toxicité. La caractérisation des principaux photoproduits a été effectuée par LC-HRMS/MS et GC-MS. Une étude *in silico* de la toxicité potentielle de chacun des photoproduits a permis d'estimer leur mutagénicité, leur LC₅₀ vis-à-vis de *Fathead minnow* et leur LD₅₀ orale vis-à-vis du rat. Ces différentes toxicités ont été comparées à celle de la benzisothiazolinone afin d'évaluer la toxicité globalement induite par photolyse de la molécule mère. L'étude du comportement photochimique de cette molécule a également contribué à mieux comprendre les mécanismes de photolyse de polluants organiques.

Les résultats ont été publiés dans la revue *Chemosphere*.²⁰⁰ L'article est présenté ci-après.

PHOTODEGRADATION OF BENZISOTHIAZOLINONE: IDENTIFICATION AND BIOLOGICAL ACTIVITY OF DEGRADATION PRODUCTS

Zsuzsanna Varga, Edith Nicol and Stéphane Bouchonnet*

Laboratoire de Chimie Moléculaire – CNRS / Ecole polytechnique, Institut Polytechnique
de Paris – 91128 Palaiseau, France

Keywords:

Photolysis, Photodegradation products, Benisothiazolinone, Structural elucidation, *In silico* tests

Chemosphere, 240, (2020), 124862

<https://doi.org/10.1016/j.chemosphere.2019.124862>

HIGHLIGHTS

- The UV-VIS photodegradation of benzisothiazolinone led to fourteen photoproducts.
- Degradation mechanism has been proposed based on photoproducts' structures.
- Some photoproducts are potentially more toxic than benzisothiazolinone.

ABSTRACT

The photodegradation of benzisothiazolinone was studied in water under UV-Vis irradiation and led to fourteen photoproducts. Chemical structures of these compounds were elucidated using GC-MS, LCMS/MS, and FT-ICR-MS experiments. Based on the chemical structures determined and their appearance order, a photo induced-degradation mechanism of benzisothiazolinone has been proposed, which combines isomerization, oxidation, hydroxylation, hydrolysis, and elimination processes. *In silico* tests on mutagenicity, *Fathead minnow* LC₅₀ and oral rat LD₅₀ were carried out to estimate the toxicity of the photoproducts. Compared with experimental data, the calculated oral rat LD₅₀ values were found to be the most relevant and thus used for toxicity estimation. The photoproducts including a phenolic or a sulfinio group or both functions were found potentially more toxic than benzisothiazolinone

1. INTRODUCTION

1,2-benzisothiazol-3(2H)-one (referred to as benzisothiazolinone or BIT), is an antimicrobial agent used as a preservative in mainly aqueous-based chemical chemicals. Banned from cosmetic products, BIT can be found in paints, lacquers, polishes, printing inks, emulsions, cleaning agents, disposable gloves; it is also widely used as a preservative for pesticides.¹⁹⁵⁻¹⁹⁸ Due to environmental considerations, water-based paints recently replaced solvent-based paints, which require the addition of biocides; isothiazolinones are widely used for this purpose.²⁰¹ BIT is also used in paper-based jointing but it degrades during sunlight exposure losing its biocidal efficiency.²⁰²

Isothiazolinone-type biocides act against various fungi and bacteria by the active sulfur moiety able to oxidize thiol functional groups to form disulfides (*e.g.* with cysteine).¹⁹⁰ Besides the biocidal effects on unwanted compounds by reacting with intracellular sulfur-containing proteins,^{190,203} isothiazolinones have also lethal effects on living organisms, the median lethal concentration (LC₅₀) of BIT for zebrafish embryos being 1.03 mg/L.²⁰⁴ Growth inhibition effect on *Escherichia coli* and *Schizosaccharomyces pombe* was measured; this inhibitory activity was quenched by thiol-containing materials (glutathione, cysteine), confirming the mechanism of interaction.¹⁹⁰ According to the European Chemicals Agency, BIT is very toxic to aquatic life, causes eye damage and skin irritation.¹⁹¹ Toxicity tests on zebrafish embryos were performed during ozonation processes and a decrease was found in the overall toxicity of the mixture, due to oxidation of the reductive sulfur in the parent compound.²⁰⁴ BIT is a developmental toxicant in rats but shows negative mutagenicity.²⁰⁵ Isothiazolinones were shown to cause severe contact dermatitis and allergy in humans mainly with a high level of occupational exposure.^{192,193} It has been shown that isothiazolinones degrade rapidly in soil.²⁰⁶ Indeed, BIT is easily degraded in the environment but there is no extensive study on the structure of its transformation products, their persistence, and their toxicity. Conventional wastewater treatment plant efficiently removes isothiazolinones at low concentrations.¹⁹⁶ In reverse-osmosis (RO) wastewater treatment plants, BIT is used thanks to its compatibility and non-oxidizing nature and it prevents biofilm growth on membranes.²⁰⁷ Nevertheless, studies demonstrated that some isothiazolinones increase the genotoxicity of the RO

concentrates.²⁰⁸ BIT is added at high concentrations (around 160 mg/l) for which the removal efficiencies of existing approaches are low; ozonation was proposed as an efficient way to degrade BIT.²⁰⁴ Considering the photochemical degradation, the kinetics was studied under irradiation with a low-pressure UV-C lamp. The process was described to follow pseudo-first order kinetics. Radical species quenching studies were carried out, where the photodegradation of BIT was concluded to be both direct, due to UV absorbance of BIT and indirect, induced by radical oxygen species formation, a widespread mechanism of micropollutant photodegradation.²⁰⁹

Different analytical methods were developed for the quantitation of BIT in various matrices. Ultrasonic assisted extraction combined with LC-MS was used to determine BIT at 9.9 mg/kg in hygienic consumer products, specifically in liquid detergents.²¹⁰ BIT was analyzed by matrix solid-phase dispersion followed by LC-MS/MS in household products where it is used as a substitute to chlorinated isothiazolinone; it was found at quite high concentrations in laundry detergent (0.0255% w/w), surface cleaner cream (0.0116% w/w) and liquid surface cleaner (0.00377% w/w).²¹¹ BIT was found in 94% of 47 paints analyzed by LC-MS at concentrations ranging from 28.6 to 1110.74 ppm. These analyses also showed that these products were sometimes labeled incorrectly, e.g. one paint labeled “preservative-free” contained BIT at 71.5 ppm.²¹² Isothiazolinones were monitored and not detected in environmental waters using derivatization of BIT followed by pre-concentration and GC-MS analysis, with limits of detection ranging from 0.01 to 0.1 µg/L.¹⁹⁶

This work aimed at performing accurate structural elucidation of the photodegradation products of BIT using chromatography and tandem mass spectrometry approaches. Benzisothiazolinone was degraded using two different types of photoreactors in water. Structural elucidation of BIT photoproducts allowed to suggest a coherent degradation pathway on one hand, and to carry out *in silico* toxicity tests to estimate the toxicity of the photoproducts on another hand.

2. MATERIALS AND METHODS

2.1. Chemicals, reagents and sample preparation

1,2-benzisothiazol-3(2H)-one (CAS: 2634-33-5), benzamide, salicylamide, 4-hydroxybenzamide and chromatography grade solvents: acetonitrile (ACN), methanol, dichloromethane and formic acid (FA) were purchased from Sigma Aldrich (Saint-Quentin-Fallavier, France). A Purelab Chorus 1 water purification system (Veolia, Wissous, France) was used to produce ultra-pure water with a conductance of 18.2 S/cm. BIT solutions were prepared at different concentrations: 5, 10, 25, 50 and 100 ppm in water (solubility 1.1 g/l at 20 °C and in acetonitrile).²¹³ Some experiments were carried out with degassing of the initial solution using nitrogen bubbling for 30 minutes. Different irradiation times were tested to determine the appearance order of photoproducts. Solutions were irradiated for 6 hours. Sampling was carried out every 10 minutes during the first hour and then every hour until the end. For GC-MS analysis, all samples (ACN solutions and aqueous samples) were dried using an Xcelvap automated evaporation/concentration system (Horizon Technologies, Bullion, France) and 1 ml of dichloromethane was added. For LC-MS measurements, the samples of different BIT starting concentrations were diluted, in a mixture made of 50% ACN, 50% H₂O and 0.1% formic acid to reach a final concentration of 5 ppm. For HRMS analysis in the infusion mode, samples were diluted in a 90% water / 10% ACN / 0.1% FA mixture, for a starting concentration of BIT of 5 ppm. All the samples were stored away from light at 4 °C; HPLC-MS data showed that degradation was negligible under these conditions, in agreement with literature data where the hydrolytic stability of the compound is mentioned.²⁰⁵

2.2. Irradiation experiments

The UV-Vis absorption spectrum of BIT was recorded in the range of 190 to 500 nm, using a 6800 UV-Vis Jenway spectrophotometer (Cole-Parmer, Villepinte, France), with a 1 cm length quartz cell. Based on this spectrum, two irradiation setups were used for photodegradation experiments. The photodegradation of BIT was mainly studied by a Philips HPL-N 125W/542 E27 SC high-pressure mercury lamp (France-Lampes, Saint-Cirq,

France) with a wavelength range of 200-650 nm, a maximum irradiation wavelength at 254 nm, and a radiation flux of 6200 lm. The setup consisted of one lamp, circularly surrounded by 6 quartz tubes of 120 ml, all of the elements being in a water bath with water circulation so that the reaction temperature was kept below 28 °C. The samples were also irradiated using a Q-sun test chamber Xe-1-B/S (Q-Lab Saarbücken, Germany) equipped with a xenon arc lamp; a natural light filter X-7640 from the same manufacturer has been used for the reproduction of full sunlight spectrum. The lamp power was 1800 W and the irradiation 0.68 W.m⁻² at a black-standard temperature of 55 °C.

2.3. GC-MS analysis

Gas chromatographic separations were carried out on an Agilent 450-GC instrument equipped with a 60 m “FactorFour VF-Xms” (10% phenyl, 90% methylpolysiloxane) capillary column, with an internal diameter of 0.25 mm and film thickness of 0.25 µm (Agilent Technologies, Les Ulis, France), coupled with an Agilent 240-MS ion trap mass spectrometer. Samples were injected in splitless mode using an Agilent CP-8400 autosampler, the injection volume was 1 µl and injector temperature was 280 °C. A gradient temperature program was used with an initial temperature of 50 °C, held for 0.5 min then increased to 280 °C with a ramp of 10 °C/min. The carrier gas was high purity helium, with a flow of 1.4 ml/min. Measurements were carried out in electron ionization and positive chemical ionization using methanol as the reagent gas. The filament emission current was set to 10 µA and automatic gain control was used for all experiments. The electron multiplier voltage was automatically optimized at 2200 V for a gain value of 10⁵. In the full scan mode, ions were scanned on the range of m/z 50 to m/z 500. Tandem experiment measurements were performed with an ionization storage level of m/z 35 and an isolation window of 3 m/z , using a resonant waveform type with an excitation storage level of m/z 50 and an excitation amplitude of 0.8 V.

2.4. LC-MS analysis

LC separations were carried out on an Acquity HPLC system (Waters Technologies, Guyancourt, France) coupled with a SolarixXR FT-ICR mass spectrometer equipped with a 9.4 T superconducting magnet (Bruker Daltonics, Bremen, Germany). A Pursuit XRs^{Ultra} C₁₈ column (2.8 μm x 50 mm x 2 mm) (Agilent Technologies, Les Ulis France) was used in gradient mode. Water + 0.1% FA (solvent A) and acetonitrile + 0.1% FA (solvent B) were used as solvents with a flow rate of 0.2 ml/min. The gradient started at 95% of solvent A for 3 min, then changed linearly to 50% of solvent A in 9 min and then to 5% of solvent A after 12 min of LC run. It was kept at 5% of solvent A for 5 minutes and then set back to the initial percentage of solvent A (95%) for 5 min of equilibrium. An electrospray ionization source was used in both positive and negative modes with a sample flow of 200 μL/min. For both positive and negative mode the nebulizer and drying gas was nitrogen with a flow of 8 L/min at 250 °C for drying gas and 1 bar for nebulizer gas. The capillary voltage was set at - 4500 V and 4000 V in both positive and negative modes, respectively; the spray shield was set at -500 V. The detection mode was broadband with a 4 Mpt resolution in the range of 57.5 to 700 *m/z* with a data reduction of noise of 95%. For tandem MS experiments, precursor ions were isolated with an isolation window of 1 *m/z* and dissociated with collision voltages of 10, 15, 20 and 25 V to study the fragmentation mechanisms. In direct infusion mode, the flow rate was 120 μL/h. The spectra were acquired with 4 M data points on a range of 57.7 to 1000 *m/z*. Accumulation time was set to 0.02 s. 100 scans were recorded for each spectrum. The Bruker DataAnalysis software was used to process the chromatograms and spectra.

2.5. *In silico* bioassays

Preliminary toxicity assessments were performed using the T.E.S.T. software developed by the US Environmental Protection Agency (EPA). This software incorporates different endpoints and approaches, with various individual predictions. The different individual approaches include hierarchical clustering, the Food and Drug Administration (FDA) method, single model, group contribution and nearest neighbor.²¹⁴ The reliability of the

results in the T.E.S.T. software should be assessed based on the capability of each individual prediction to calculate the toxicity values and then study concordance of the results. A complete description of T.E.S.T. and the included methods can be found on the EPA website.²¹⁴ Different QSAR methodologies are available to model mutagenicity. The biological testing background of most of them, including T.E.S.T., is the Ames test, which is based on a bacterial mutation of *Salmonella typhimurium*, reverting the bacteria from histidine dependence to histidine independence.²¹⁵ The correlation between carcinogenicity in animals and positive mutagenicity in the Ames test was found to be high.²¹⁶ The QSAR models for mutagenicity of the compounds of interest were found to be generally of high reliability.

3. RESULTS AND DISCUSSION

3.1. Physicochemical analysis and photoproducts structural elucidation

The UV/Vis absorption spectrum of BIT shows a high absorbance in the UV region and an additional absorption band with a maximum at 319 nm (see Supplementary Information III-19). The photodegradation was evidenced by the decrease of the molecule's absorption bands according to time (Figure III-7). A strong decrease in concentration can be observed during the first 15 minutes of irradiation, whereas the degradation slows down later on. This is in good agreement with the exponential decrease evidenced by LC-MS experiments (see below).

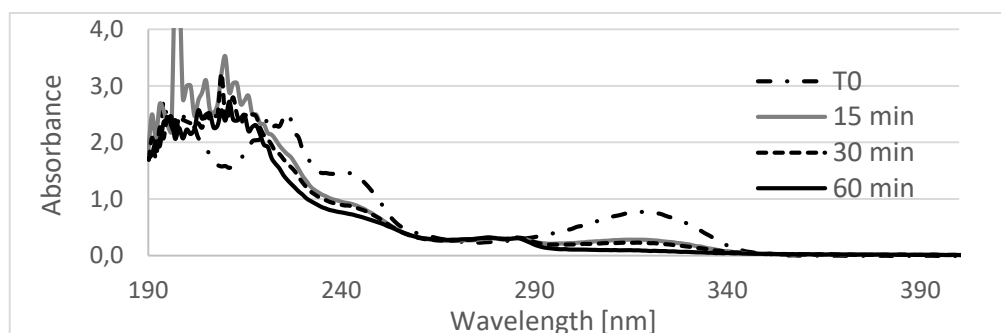


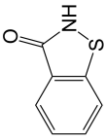
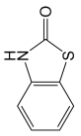
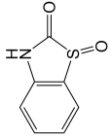
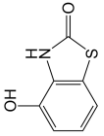
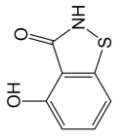
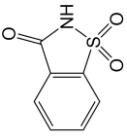
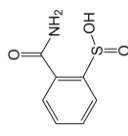
Figure III-7: UV-Vis absorption spectrum of BIT (1.2 mM) in ultrapure water at different irradiation times ranging from 0 to 60 minutes

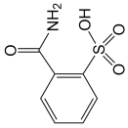
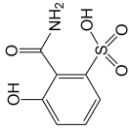
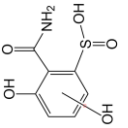
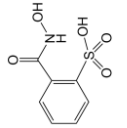
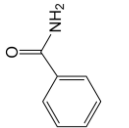
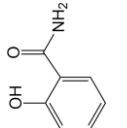
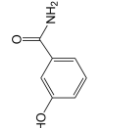
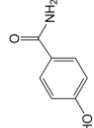
Whatever the chromatograph, using chromatography introduces some subjectivity in the detection process. Actually, the choice of stationary and mobile phases induces some selectivity regarding elution and thus detection of molecules. Therefore, direct infusion MS was used prior to LC-MS coupling, to carry out non-targeted investigation of transformation products in ESI-HRMS. The irradiated solutions were analyzed in both positive and negative modes for the same reasons. The ions detected in direct ESI-HRMS were the same as those detected in LC-MS. Retention times in LC-MS and GC-MS couplings, appearance times of photoproducts during the irradiation process, exact m/z values of $[M+H]^+$ and $[M-H]^-$ ions and corresponding formulae are listed in Table III-10.

Exact m/z values and corresponding formulae of collision-induced product ions from $[M+H]^+$ or $[M-H]^-$ ions are given in Supplementary Information III-20.

PP1 is among the first appearing photoproducts. It is eluted at 19.3 min in GC-MS; it has the same mass as the parent compound (eluted at 17.8 min) and its EI mass spectrum shows a very similar fragmentation pattern. Its exact formula (C_7H_5NOS) has been determined by HR-MS and confirmed that PP1 corresponds to an isomer of BIT. 1,3-benzothiazol-2(3H)-one was postulated on the basis of PP1 mass spectrum. It was synthesized (the synthesis protocol is reported in Supplementary material - Supplementary Information III-18) and analyzed using GC-MS; both its retention time and EI mass spectrum confirmed that PP1 corresponds to 1,3-benzothiazol-2(3H)-one (Table III-10). Light-induced isomerization of BIT has been reported in an old study in which three possible isomerization mechanisms were postulated but none was demonstrated.²¹⁷ Considering the rapid isomerization of BIT into PP1, consequent oxidation processes will have to be considered from both isomers.

Table III-10: Photoproducts of BIT detected in LC-MS and GC-MS couplings

Compound	Time of first detection (min)	Time when not present anymore (min)	Retention time in LC-MS (min)	Retention time in GC-MS (min)	m/z measured for $[M+H]^+$ in ESI+	m/z measured for $[M-H]^-$ in ESI-	Theoretical mass of the neutral	Empirical formula of the neutral	Chemical structure
BIT	0	120	7.2	17.8	152.01642	150.00203	151.00918	C ₇ H ₅ NOS	
PP1	10	480	9.3	19.3	152.01636	150.00206	151.00918	C ₇ H ₅ NOS	
PP2a	10	120	5.4	20.3	168.01123	165.99698	167.00410	C ₇ H ₅ NO ₂ S	
PP2b	50	420	7.0	-	168.01120	165.99700	167.00410	C ₇ H ₅ NO ₂ S	
PP2c	50	420	8.3	-	168.01122	165.99699	167.00410	C ₇ H ₅ NO ₂ S	
PP3	10	120	3.8	-	-	181.99191	182.99901	C ₇ H ₅ NO ₃ S	
PP4	10	120	1.2	-	168.01123	184.00756	185.01466	C ₇ H ₇ NO ₃ S	

Compound	Time of first detection (min)	Time when not present anymore (min)	Retention time in LC-MS (min)	Retention time in GC-MS (min)	m/z measured for $[M+H]^+$ in ESI+	m/z measured for $[M-H]^-$ in ESI-	Theoretical mass of the neutral	Empirical formula of the neutral	Chemical structure
PP5	10	480	1.8	-	202.01665	200.00250	201.00957	$C_7H_7NO_4S$	
PP6a	30	480	0.9	-	-	215.99742	217.00449	$C_7H_7NO_5S$	
PP6b	30	480	1.5	-	-	215.99741	217.00449	$C_7H_7NO_5S$	
PP6c	30	480	2.3	-	-	215.99742	217.00449	$C_7H_7NO_5S$	
PP7	50	480	5.6	14.9	122.05992	-	121.05276	C_7H_7NO	
PP8a	50	480	7.3	16.5	138.05482	-	137.04768	$C_7H_7NO_2$	
PP8b	50	480	2.5	-	138.05480	-	137.04768	$C_7H_7NO_2$	
PP8c	50	360	7.1	-	138.05483	-	137.04768	$C_7H_7NO_2$	

Three photoproducts (referred as PP2a, PP2b, and PP2c) were detected with retention times of 5.4, 6.9 and 8.3 min in LC-MS coupling. They all have a molecular weight of 167 (confirmed by both EI and CI mass spectra in GC-MS) and have been assumed to result from oxygen addition onto the parent compound. For the first one, the concerted losses of CO and SO (m/z 168.01196 \rightarrow m/z 96.04471 in CID measurements, Supplementary Information III-20) from the protonated molecule allowed to conclude that the oxygen atom was added onto the sulfur one. The fact that the nitrogen atom was retained in the m/z 96.04471 product ion permitted to conclude in favor of the structure depicted in Table III-10, which results from oxidation of PP1 (and not from BIT). Both other PP2 isomers appear after 3 h of irradiation; they display exactly the same CID dissociation pathways. Eliminations of HNCO and HS \cdot from protonated molecules show that the oxygen atom has been added on the aromatic ring. Hydroxylation of aromatic rings under UV-Vis irradiation has been widely reported in the literature.^{218,219} Because it can play an important role regarding toxicity, the exact position of the hydroxyl group has been determined based on CID experiments as depicted in Supplementary Information III-21. All the possible structures resulting from hydroxyl addition on the aromatic ring either from BIT or PP1 have been considered in a systematic way: only two of them can undergo two consecutive eliminations of carbon monoxide from the protonated species in positive ESI; their formulae are presented in Table III-10.

The photoproduct referred as PP3 appears after only 10 min of irradiation. Its exact formula corresponds to the addition of two oxygen atoms on BIT or PP1. In CID experiments, the elimination of SO₂ from the deprotonated species (m/z 181.99191 \rightarrow m/z 118.02998, Supplementary Information III-20) leaves no doubt that both oxygen atoms have been added to the sulfur atom. Three mechanisms could be responsible for the formation of PP3: straight addition of a dioxygen molecule onto the sulfur atom from either BIT or PP1, or oxidation of PP2a. As this photoproduct is no more observed when irradiating a degassed solution, it can be concluded in favor of direct dioxygen addition on the sulfur atom. CID results do not permit to discriminate between oxidation of either BIT or PP1 but the structures elucidated for PP4 to PP8 photoproducts are strongly in favor of BIT oxidation, as discussed in the next section. It is to be noted that the structure

suggested for PP3 corresponds to saccharin, which has been proved to be an oxidation product of BIT in a study dedicated to BIT ozonation.²⁰⁴

PP4 has been detected in negative ESI, at a very short retention time (1.2 min) in reverse-phase chromatography suggesting a highly polar structure. Its mass spectrum exhibits a major ion resulting from deprotonation of the molecule. For these reasons, the structure suggested for PP4 in Table III-10 includes a SOOH group, which allows easy removal of a proton in negative ESI. The SO₂ elimination in CID experiments is in good agreement with the proposed structure.

PP5 has been detected in both positive and negative ESI. Its elemental composition corresponds to that of PP4 plus a hydroxyl group. In the negative mode, observation of the deprotonated molecule indicates an acidic function (in agreement with the short eluting time in LC-MS). In the positive mode, the consecutive losses of ammonia and SO₂ imply that the amido function has been kept and that the sulfur atom carries two oxygen atoms. This implies that the additional hydroxyl group is bound to the aromatic ring. It has been located in ortho position with regard to the amido function since the formation of both PP2b and PP2c photoproducts shows that hydroxylation in this position is favored over additions onto other positions of the ring. The consecutive losses of SO₂ and H₂O from [PP5 - H]⁻ in CID experiments (m/z 200.00250 → m/z 118.02997) confirm this hypothesis (see the dissociation pathways of [PP5 - H]⁻ in Supplementary Information III-22).

Three photoproducts referred as PP6a to PP6c were detected at short retention times (between 0.9 and 2.3 min) in negative ESI. Their exact formulae correspond to the addition of an oxygen atom onto PP5. In CID experiments, the isomer eluting at 0.9 min undergoes losses of NH₃ and O=C=NH indicating that it still has the amido function; it demonstrates that the aromatic ring has been hydroxylated. The ammonia elimination, barely observed in negative ESI, has to be charge-induced. By analogy with other photoproducts and to rationalize the NH₃ loss (see Supplementary Information III-23), the hydroxyl group was assumed to be in ortho position relative to the amido group. In the negative mode, the losses of HCNO and SO₂ from [PP6b - H]⁻ indicate that two hydroxyl groups are carried by the aromatic ring in the PP6b structure. Unfortunately, CID experiments provided two few

ions to allow the location of these hydroxyl groups. Both losses of HCNO_2 and SO_3 from $[\text{PP6c} - \text{H}]^-$ allowed the easy determination of the PP6c structure, as depicted in Supplementary Information III-23.

PP7 has been easily identified as benzamide based on GC-MS and LC-MS analysis of the corresponding standard. As for PP7, PP8 isomers (PP8a to PP8c, see Table III-10) were easily identified since salicylamide and 4-hydroxybenzamide were commercially available and were used to check mass spectra and retention times in GC-MS.

3.2. Mechanistic approach of BIT photodegradation

The photodegradation of BIT and the appearance of photoproducts were monitored using LC-MS in both positive and negative modes, at various concentrations ranging from 5 to 100 *ppm*. The detected photoproducts were the same regardless the initial BIT concentration, although the reaction kinetics slightly differed, the photodegradation being slower at higher concentrations, as displayed in Supplementary Information III-24 and Supplementary Information III-25. At 5 *ppm*, BIT is undetectable after 30 minutes of irradiation while the initial amount of BIT has been reduced by 60% at 100 *ppm* for the same irradiation time. For each photoproduct, Table III-10 gives the time at which it is first detected and the time at which it is no longer detected. Based on the kinetics of the appearance and disappearance of photoproducts, stepwise photodegradation mechanisms were proposed (Figure III-8). Structures between brackets were not observed, likely because of too short lifetimes. It is to be noted that the formation of sulfite ions, reported by a previous study, was not evidenced in the present work but is in good agreement with the observed decrease of the mixture pH.²⁰⁴

Isomerization of BIT into PP1 is very fast; it is not possible to state if this isomerization is partially reversible. Isomerization of other compounds including five-membered ring structures (PP2 isomers) could also not be established but could explain the formation of PP4 from PP2a (see Figure III-8). From BIT or PP1, the formation of all the photoproducts can be rationalized by successive or competitive mechanisms including direct oxygen addition onto the sulfur atom (so that it reaches its highest oxidation state), hydroxyl

addition, hydrolysis, and SO₃ elimination. The comparison of results between degassed and non-degassed solutions showed that the same photoproducts are produced in both cases. They are less abundant in degassed solutions, for which BIT dimers are also observed. This is likely because a reduced concentration in oxygen disfavors mechanisms involving dissolved oxygen and thus favors intermolecular reactions between BIT excited species.

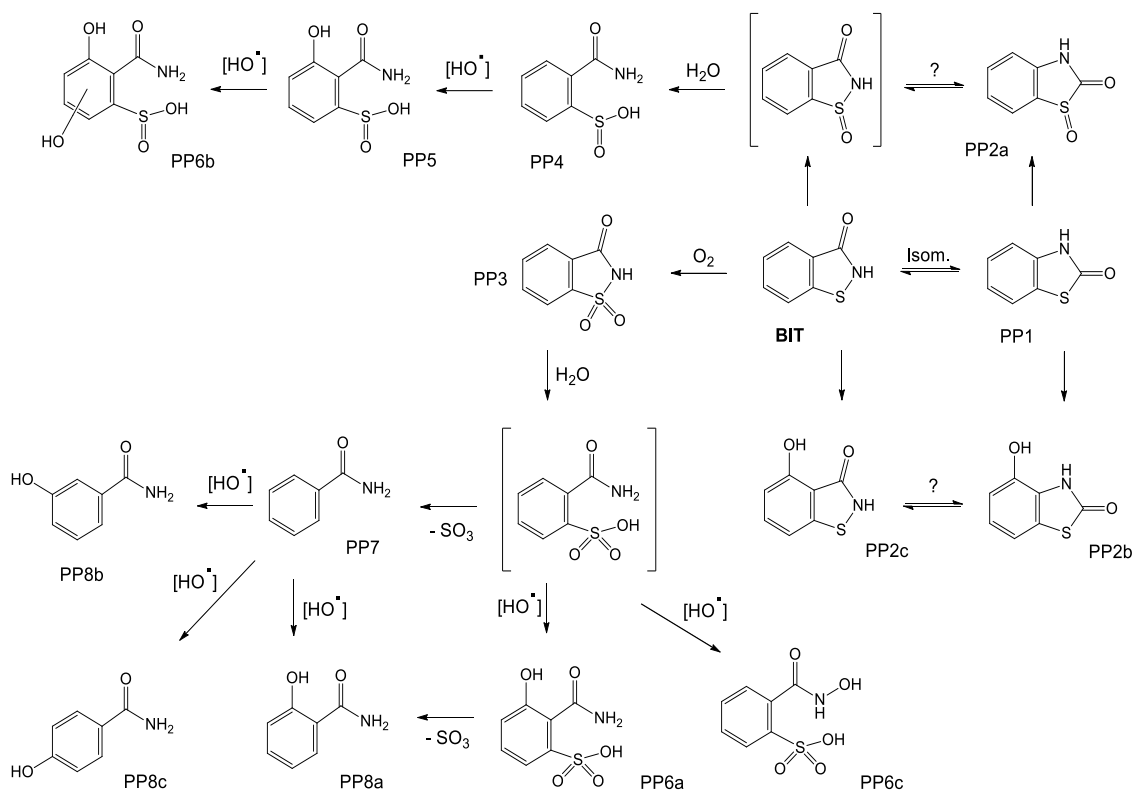


Figure III-8: Suggested pathways for the photodegradation of BIT.

3.3. In silico toxicity estimations

For each photoproduct, QSAR results on mutagenicity, oral rat and Fathead minnow toxicity were considered to be compared to that of the parent compound. The results of *in silico* toxicity estimations are summarized in Table III-11. Since the structure of PP6b was not accurately determined (the position of the second hydroxyl group on the aromatic ring could not be stated, see Table III-10), toxicity estimation could not be performed on this compound. Regarding mutagenicity, T.E.S.T. works with a dataset of 5743 chemicals and provides a binary result (positive/negative). Experimental values were provided by

the software for BIT, PP3, PP7, and PP8a; they were all in accordance with the calculated results. *In silico* estimations showed non-mutagenicity for all the photoproducts except PP4. For this compound, only 2 out of 3 individual predictions provided any result and both results were contradictory, therefore the mutagenicity positive result for PP4 is of low reliability; it might be a false positive result and should be treated accordingly. Fathead minnow LC₅₀ (96 h), represents the concentration that kills half of the fish population (*Pimephales promelas*) in 4 days. The training set was obtained from the ECOTOX database, with a dataset containing 823 chemicals. The parent compound showed to be the most toxic but LC₅₀ values could not be estimated for PP1, PP2b, PP3, and PP6c: none of the individual models provided a result since the requirements for reliable toxicity assessment were not fulfilled for these compounds. For PP8a (salicylamide) the proposed experimental toxicity value was in good agreement with the calculated one but for PP7, for which an experimental result was also available, the experimental value was 3.5 times higher than the calculated one. The oral rat LD₅₀ endpoint gives the amount of chemical, in terms of the mass of the chemical over weight of the rat, which kills half of the rats when orally ingested. In this case, the ChemIDplus database was used to calculate toxicity data of interest with a final dataset of 7413 chemicals. For this endpoint, there was no prediction for photoproduct PP6a but the others proved to be of high reliability. Experimental values showed good accordance with the calculated ones for the parent compound. Photoproducts PP2a, PP2c, PP4, PP5, PP8b, and PP8c exhibited slightly higher or similar toxicity as the parent compound. Compound PP5 fell into the moderately toxic category (50-500 mg/kg) on the Hodge and Sterner scale.¹⁸⁷ All the other compounds were categorized as slightly toxic. For PP8a different oral rat toxicity values were reported in the literature but generally, the experimental values were approximately 2 times lower, closer to the toxicity values determined for PP8b and PP8c (Table III-11).^{220,221} It could explain the discrepancy among the PP8 compound family, in which photoproducts display isomeric structures. Given that oral rat toxicity calculations proved to be of adequate reliability, toxicophores of the critical compounds are further discussed. The potential toxicity of PP2a cannot be compared to those of other compounds since its structure includes a very specific and uncommon –SO-CO-NH- structure (Table III-10). It is interesting to note that, at the exception of PP2a, all the photoproducts exhibiting potential toxicity include either a phenolic (PP2c, PP8b and

PP8c) or sulfinio (-SO₂H) group (PP4). Furthermore, the most toxic compound (PP5) is the one including both functions. The toxicity of phenolic compounds has been widely reported but that of the sulfinio group is not discussed in the literature, therefore no conclusion could be drawn but it might be interesting to investigate further this class of compounds.

Table III-11: Toxicity values estimated by the T.E.S.T. software for BIT and its photoproducts except PP6b

Compounds	Mutagenicity (Ames test)	Fathead minnow LC50 (96) [mg/L]	Oral rat LD50 [mg/kg]
BIT	Negative	25	894
PP1	Negative	NA ^a	1169
PP2a	Negative	148	527
PP2b	Negative	NA	3235
PP2c	Negative	36	941
PP3	Negative	NA	1685
PP4	Positive	140	649
PP5	Negative	79	412
PP6a	Negative	55	NA
PP6c	Negative	NA	3385
pp7	Negative	186	1123
PP8a	Negative	166	1619
PP8b	Negative	144	897
PP8c	Negative	126	608

NA^a: Non applicable

4. CONCLUSION

The UV-VIS irradiation of an aqueous solution of benzisothiazolinone led to the formation of fourteen photoproducts, their chemical structures were elucidated based on GC-MS, LC-MS/MS, and FT-ICR-MS experiments. Based on the chemical structures determined and their appearance order, a photo induced-degradation mechanism of benzisothiazolinone was proposed, which combines isomerization, oxidation, hydroxylation, hydrolysis, and elimination processes. Regarding to toxicity estimations, the absence of mutagenicity of photoproducts and the moderate oral rat LD₅₀ values determined for some of them (while most of them are less toxic than the parent compound) leads to consider direct photodegradation as a potential powerful tool to remove benzisothiazolinone, even when it is used in large amounts in some wastewater

treatment plants. This is even more interesting considering that these photoproducts were not persistent after a few hours of UV-light irradiation.

ACKNOWLEDGMENTS

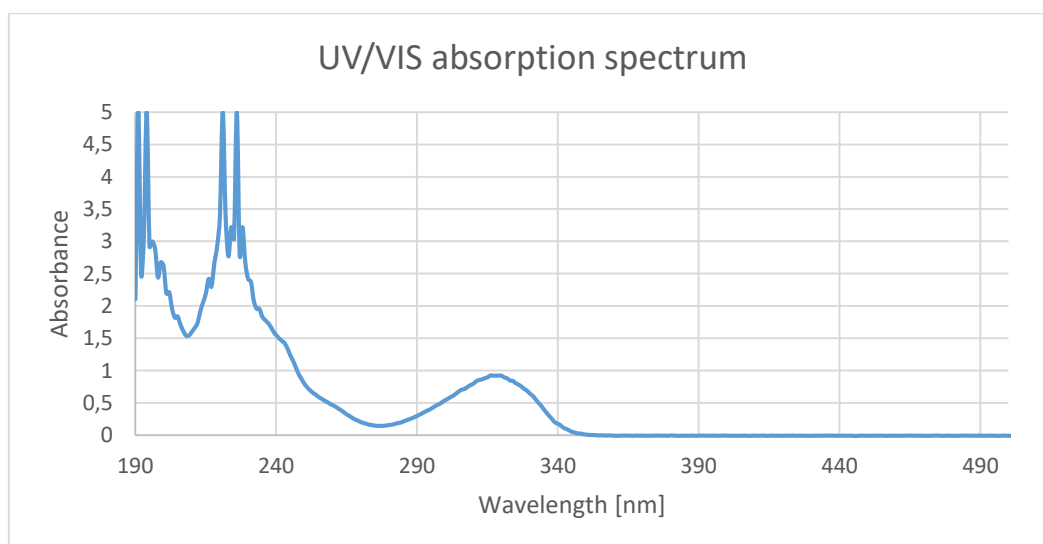
We are very thankful to Dr Gregory Danoun, Researcher at the LCM, who performed the synthesis of BIT.

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No 765860 (AQUALity). Financial support from the National FT-ICR network (FR3624 CNRS) and from the Ile de France Region are also gratefully acknowledged.

SUPPLEMENTARY MATERIAL

Supplementary Information III-18: Protocol for the synthesis of PP1 (Benzothiazolinone)

To a solution of 2-aminothiophenol (1 g, 8 mMol) in THF (20 mL) was added carbonyldiimidazole (1.4 g, 8.6 mmol) at room temperature and the resulting solution was stirred for 12 hours. Upon disappearance of starting material, the reaction solution was poured into 1 N HCl (15 mL) and extracted with ethyl acetate (15 mL). The organic layer was dried over sodium sulfate and concentrated to give 1,3-benzothiazol-2(3H)-one as a white solid.

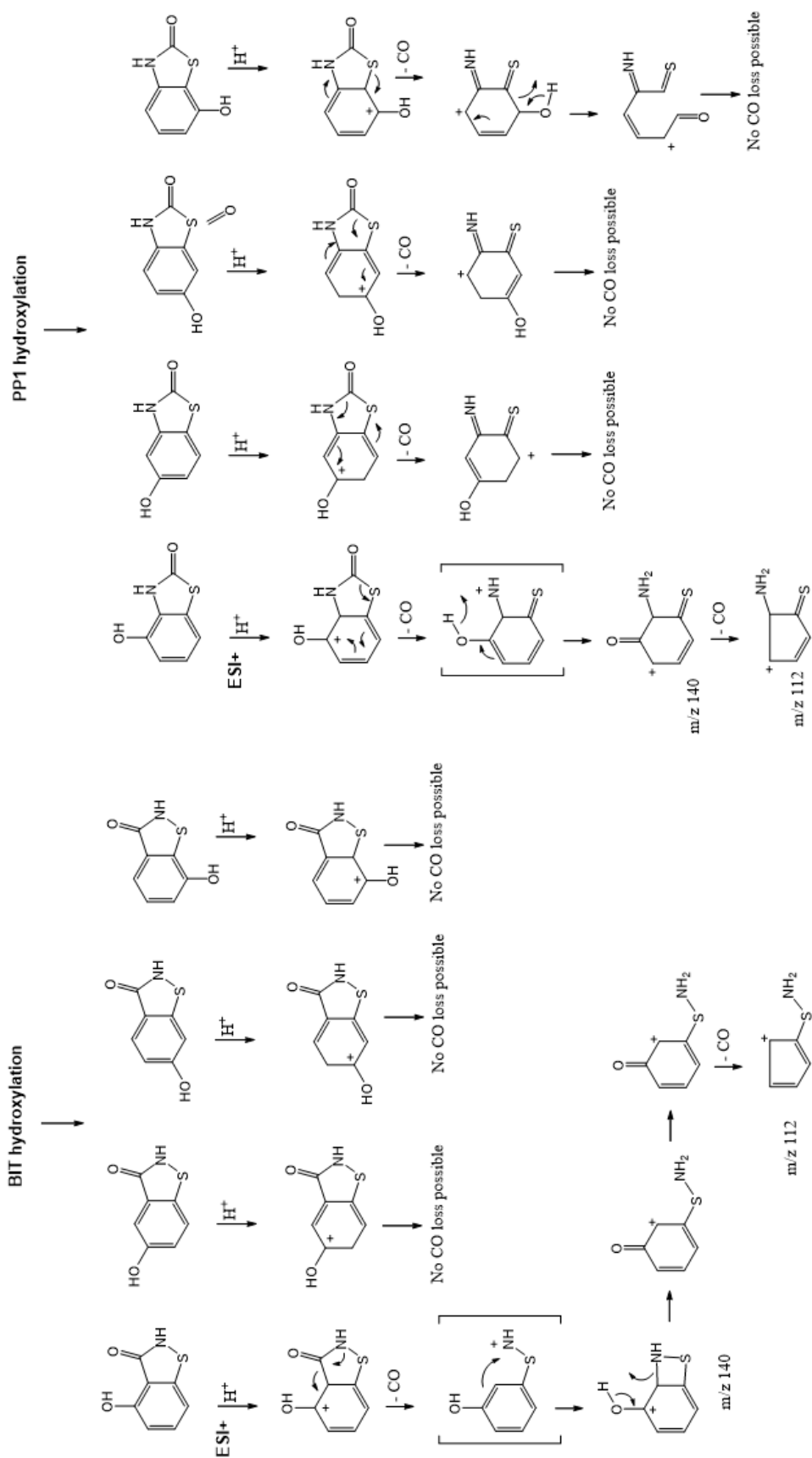


Supplementary Information III-19: UV-Vis absorption spectrum of BIT (1.2 mM) in ultrapure water

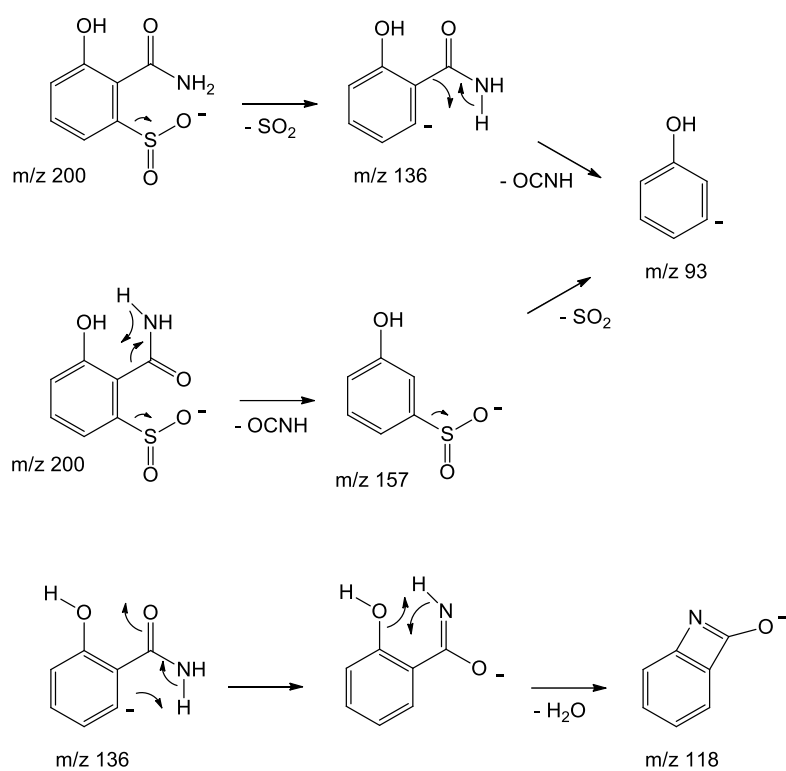
CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Supplementary Information III-20: Exact m/z values and corresponding formulae of collision-induced product ions from [M+H]⁺ and [M-H]⁻ ions in LC-MS/MS

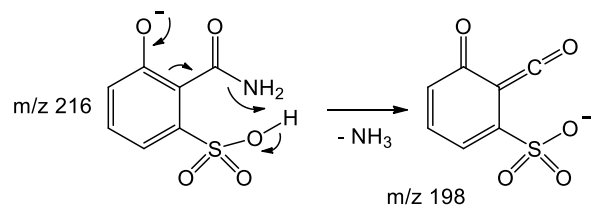
Compound	Retention time in LC-MS (min)	[M+H] ⁺ in ESI+ (m/z)	Daughter ions of [M+H] ⁺ (m/z)	Chemical formula of the daughter ion	[M-H] ⁻ in ESI- (m/z)	Daughter ions of [M-H] ⁻ (m/z)	Chemical formula of the daughter ion
BIT	7.2	152.01642	-	-	-	-	-
PP1	9.3	152.01636	-	-	-	-	-
PP2a	5.4	168.01196	151.00915 150.00133 122.00631 96.04471	C ₇ H ₅ NOS (- HO·) C ₇ H ₄ NOS (- H ₂ O) C ₆ H ₄ NS (- H ₂ O, - CO) C ₅ H ₆ NO (- CO, - SO)	-	-	-
PP2b	7.0	168.01196	140.01695 135.03148 125.00600 112.02193 108.04477	C ₆ H ₆ NOS (- CO) C ₇ H ₅ NO ₂ (- HS·) C ₆ H ₅ OS (- CHNO) C ₅ H ₆ NS (- CO, - CO) C ₆ H ₆ NO (- COS)	-	-	-
PP2c	8.3	168.01196	140.01695 135.03148 125.00600 122.00632 112.02193	C ₆ H ₆ NOS (- CO) C ₇ H ₅ NO ₂ (- HS·) C ₆ H ₅ OS (- CHNO) C ₆ H ₄ NS (- H ₂ O, - CO) C ₅ H ₆ NS (- CO, - CO)	-	-	-
PP3	3.8	-	-	-	181.99191	118.02998	C ₇ H ₄ NO (- SO ₂)
PP4	1.2	-	-	-	184.00756	120.04494	C ₇ H ₆ NO (- SO ₂)
PP5	1.8	202.01665	184.99094 121.02882	C ₇ H ₅ O ₄ S (- NH ₃) C ₇ H ₅ O ₂ (- NH ₃ , - SO ₂)	200.00250	156.99666 136.04055 118.02997 93.03468	C ₆ H ₅ O ₃ S (- CHNO) C ₇ H ₆ NO ₂ (- SO ₂) C ₇ H ₄ NO (- SO ₂ , - H ₂ O) C ₆ H ₅ O (- CHNO, - SO ₂)
PP6a	0.9	-	-	-	215.99742	198.07082 172.99153	C ₇ H ₃ O ₅ S (- NH ₃) C ₆ H ₅ O ₄ S (- CHNO)
PP6b	1.5	-	-	-	215.99741	172.99153 152.03542	C ₆ H ₅ O ₄ S (- CHNO) C ₇ H ₆ NO ₃ (- SO ₂)
PP6c	2.3	-	-	-	215.99742	156.99660 136.04048	C ₆ H ₅ O ₃ S (- CHNO ₂) C ₇ H ₆ NO ₂ (- SO ₃)
PP7	5.6	122.05992	105.03384 79.05447	C ₇ H ₅ O C ₆ H ₇	-	-	-
PP8a	7.3	138.05482	121.02882 111.04443 95.04946	C ₇ H ₅ O ₂ C ₆ H ₇ O ₂ C ₆ H ₇ O	-	-	-
PP8b	2.5	138.05480	121.02882 111.04443 95.04946	C ₇ H ₅ O ₂ C ₆ H ₇ O ₂ C ₆ H ₇ O	-	-	-
PP8c	7.1	138.05483	120.04480 92.04979	C ₇ H ₆ NO C ₆ H ₆ N	-	-	-



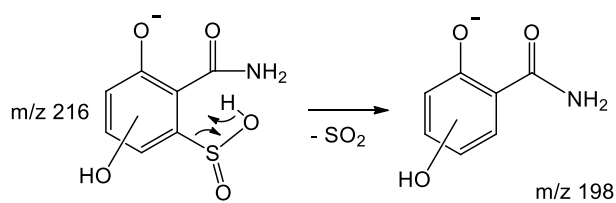
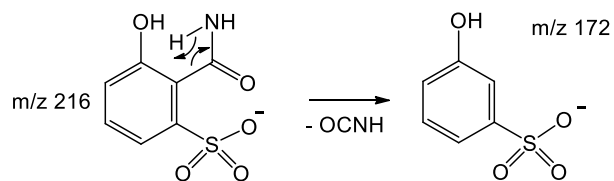
Supplementary Information III-21: Consecutive CO eliminations from hydroxylated isomers of BIT and PP1



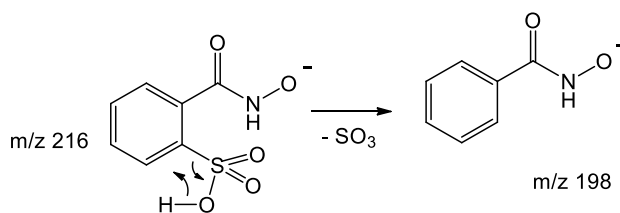
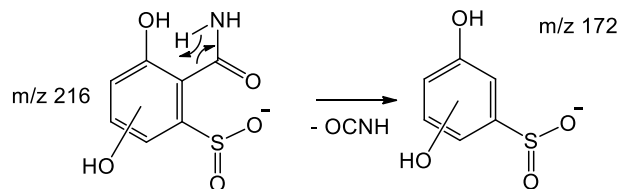
Supplementary Information III-22: Dissociation pathways $[PP5 - H]^-$ under collisional activation



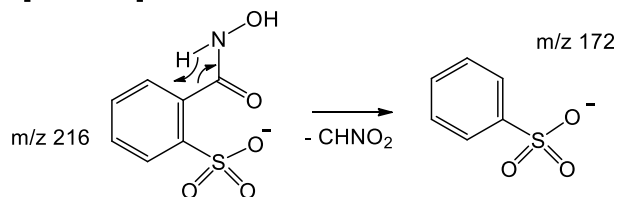
[PP6a - H]⁻



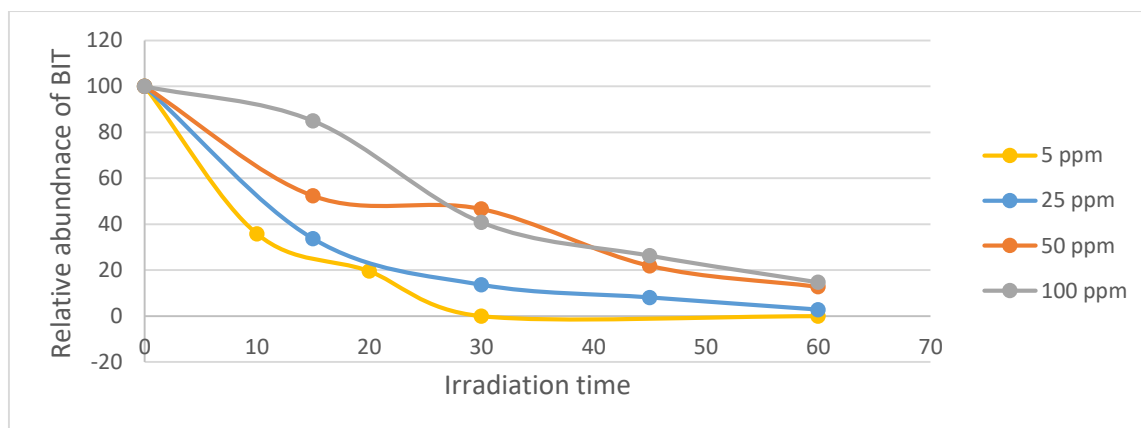
[PP6b - H]⁻



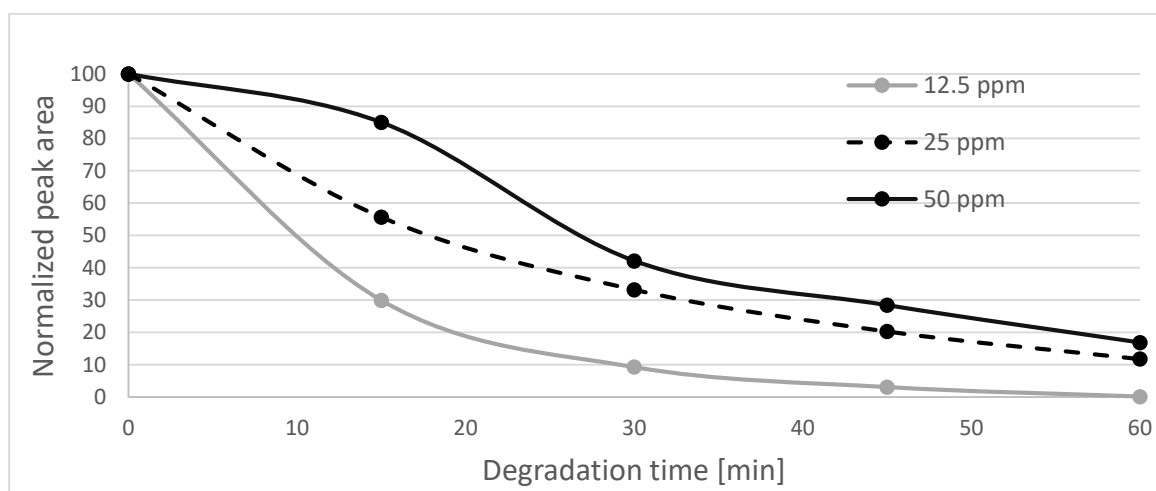
[PP6c - H]⁻



Supplementary Information III-23: Dissociation pathways of PP6 isomers



Supplementary Information III-24: Relative amount of BIT as a function of irradiation time for concentrations of BIT ranging from 5 to 100 ppm in ACN



Supplementary Information III-25: Relative amount of BIT as a function of irradiation time for concentrations of BIT ranging from 12.5 to 50 ppm in water

III.2.2. Synthèse des principaux résultats de cette étude

Au cours de cette étude, 14 photoproduits de la benzisothiazolinone ont été mis en évidence par LC-HRMS/MS et GC-MS. L'utilisation de standards tels que la 1,3-benzothiazol-2(3H)-one (PP1), la saccharine (PP3), la benzamide (PP7), la salicylamide et la 4-hydroxybenzamide (isomères de PP8), a permis la confirmation des structures pour certains de ces photoproduits. La saccharine a été également rapportée comme produit de dégradation par ozonolyse.²⁰⁴ Sur la base des valeurs estimées pour la LD₅₀ orale du rat, certains photoproduits notamment ceux incluant un groupe phénolique et/ou un groupe sulfino, sont potentiellement plus toxiques que la benzisothiazolinone elle-même. Néanmoins, l'absence de mutagénicité et les valeurs modérées de LD₅₀ orale du rat pour la plupart des photoproduits ainsi que le fait que la plupart des photoproduits ne persistent pas après quelques heures d'irradiation tendent à démontrer que la dégradation de la benzisothiazolinone par traitement UV-Vis est un procédé efficace et pertinent pour les stations de traitement de l'eau.

Cette étude a par ailleurs permis de mettre en évidence des mécanismes de photodégradation impliquant des isomérisations, des oxydations, des hydroxylations, des hydrolyses et des éliminations.

III.3. DEGRADATION UV-VISIBLE DE L'ACETAMIPRIDE DANS DES MATRICES AQUEUSES COMMERCIALES – ELUCIDATION STRUCTURALE DES PHOTOPRODUITS ET CONSEQUENCES TOXICOLOGIQUES

III.3.1. Présentation des travaux relatifs à la dégradation photochimique de l'acétamipride

L'acétamipride est un insecticide organochloré de la famille des néonicotinoïdes incluant la clothianidine, l'imidaclopride, le thiaclopride et le thiméthoxame. Il s'agit d'un insecticide systémique imprégnant toute la plante pour la protéger des parasites. Les néonicotinoïdes sont destinés à éliminer les insectes suceurs dans de nombreuses cultures telles que les betteraves, les agrumes, la vigne, le coton, la cerise mais aussi les plantes ornementales et les feuillus. Ils agissent en tant que neurotoxiques pour les insectes en faisant dysfonctionner les canaux calciques des neurones. L'acétamipride est classé comme potentiellement carcinogène pour l'être humain avec une dose journalière admissible (DJA) de 0,025 mg/kg/j, une dose de référence aiguë (ARfD) de 0,025 mg/kg et un niveau acceptable d'exposition (AOEL) de 0,025 mg/kg/j. Plusieurs pesticides organochlorés présentent de faibles doses létales (LD₅₀), notamment une LD₅₀ de 22,20 µg/abeille pour l'acétamipride, ce qui a conduit à les classer comme hautement toxiques pour les écosystèmes.²²²⁻²²⁵ Au vu de leur nocivité pour l'être humain et l'environnement, les néonicotinoïdes ont été bannis de l'agriculture européenne.^{226,227} Cependant, malgré l'interdiction d'utilisation de l'acétamipride en France dans les produits phytopharmaceutiques depuis le 1^{er} septembre 2018 - conformément à l'article L. 253-8 du Code rural et de la pêche maritime - l'ANSES a délivré une dérogation pour les cultures pour lesquelles les alternatives à cet insecticide demeurent insuffisantes. Au niveau européen, l'acétamipride reste autorisé pour l'agriculture contrairement aux autres néonicotinoïdes. En raison de la grande solubilité de l'acétamipride dans l'eau (4,25 g/L à 25°C), la plupart des résidus associés persistent dans le sol des cultures traitées, et pénètrent au final dans les eaux de surface et/ou les eaux souterraines. Lors d'une étude menée sur des eaux de surface de plusieurs pays, l'acétamipride a été retrouvé dans plus

de 80% des échantillons à des concentrations allant jusqu'à 18 µg/L et même 44 µg/L au niveau des cultures.^{51,132,228} Il a aussi été récemment retrouvé parmi les polluants présents dans des colonies d'abeilles.²²⁷

En raison de son utilisation sur les cultures et de sa présence dans les eaux de surface, l'acétamipride peut être dégradé sous l'effet de l'irradiation solaire. Lors de son épandage sur les cultures, il se retrouve à la surface de la plante en présence des excipients inclus dans le produit commercial. La présence de ces composés peut faire varier les voies de dégradation et mener à de nouveaux photoproduits.^{186,229} Le but de ce travail est donc d'identifier les photoproduits de l'acétamipride, seul et au sein de deux mélanges commerciaux : Polysect Ultra SL® et Roseclear Ultra®. L'élucidation structurale a été effectuée par des analyses LC-FT-ICR-MS/MS associées à des approches cinétiques et à l'utilisation de l'analogue deutéré de l'acétamipride, l'acétamipride-d₃ (structure décrite ci-après). L'écotoxicité de l'échantillon avant, pendant et après photolyse a été déterminée à l'aide de tests *in vitro* conduits sur la bactérie *Vibrio fischeri* pour l'échantillon d'acétamipride seul ainsi que pour les deux mélanges commerciaux. Les toxicités individuelles des différents photoproduits identifiés ont été estimées par des tests *in silico* réalisés *via* le logiciel T.E.S.T. Ces toxicités *in silico* et *in vitro* ont été comparées à celles de l'acétamipride, seul ou en mélange, avant photolyse.

Les résultats ont été publiés dans la revue *Chemosphere*.²³⁰ L'article est présenté ci-après.

LABORATORY SCALE UV-VISIBLE DEGRADATION OF ACETAMIPRID IN AQUEOUS MARKETED MIXTURES - STRUCTURAL ELUCIDATION OF PHOTOPRODUCTS AND TOXICOLOGICAL CONSEQUENCES

Edith Nicol, Zsuzsanna Varga, Svetlana Vujovic, Stéphane Bouchonnet *

Laboratoire de Chimie Moléculaire – CNRS / Ecole polytechnique, Institut Polytechnique
de Paris – 91128 Palaiseau, France

Keywords:

acetamiprid, photodegradation, structural elucidation, toxicity, commercial mixtures, LC-
HR-MS/MS

Chemosphere, 248, (2020), 126040

<https://doi.org/10.1016/j.chemosphere.2020.126040>

HIGHLIGHTS

- UV photodegradation of acetamiprid was investigated, alone and in marketed mixtures.
- Ten photoproducts were characterized using LC-HR-MS/MS analysis.
- Degradation pathways have been proposed based on photoproducts' structures.
- *In vitro* tests showed that marketed mixtures are more toxic after UV-light exposure.

ABSTRACT

Acetamiprid is a neonicotinoid pesticide, which is extensively used on agricultural crops, but has a high toxic effect on beneficial insects and the human body. It is exposed to sunlight irradiation on crops but also in surface waters where it is found at a high level due to its resistance to common water treatments. The aim of the present work was to study the UV-visible photodegradation of acetamiprid, alone and in two marketed mixtures (Polysect Ultra SL® and Roseclear Ultra®). Ten photoproducts were characterized using LC-HR-MS/MS analysis. Photodegradation pathways were proposed based on the chemical structures of photoproducts and kinetic measurements; a matrix effect has been evidenced for commercial mixtures. Most photoproducts exhibit potential developmental toxicity twice higher than that of the parent compound. Regarding potential mutagenicity, all photoproducts are less toxic than acetamiprid. Estimated oral rat LD₅₀ values show that the potential toxicities of photoproducts are similar or lower than that of acetamiprid. *In vitro* tests on *Vibrio fischeri* bacteria showed that the ecotoxicities of marketed mixtures are significantly higher than that of acetamiprid in aqueous solution; they slightly increase after UV-light exposure.

INTRODUCTION

(E)-N-(6-chloro-3-pyridylmethyl)-N0-cyano-N-methylacetamide, also known as acetamiprid (ACT), is a systemic, contact and digestive insecticide used for agricultural crops. It belongs to the neonicotinoid family, which also includes clothianidin, imidacloprid, nitenpyram, nithiazine, thiacloprid, and thiamethoxam and is one of the most employed class of pesticides.²³¹ Neonicotinoid pesticides have been introduced in crops since 1990. They act as neuro-active insecticides on nicotinic acetylcholine receptors in the same way as nicotine.²³² Acetamiprid has a high water solubility of 4250 mg.L⁻¹ at 25 °C and a half-life of 5-25 days in the environment.^{233,234} It has been detected worldwide in surface and wastewaters at concentrations ranging from 2.10⁻⁶ mg.L⁻¹ to 0.041 mg.L⁻¹^{132,235-237} and recently in the urine of young children in Japan.²³⁸ In 2013, the European Food Safety Authority (EFSA) delivered a scientific opinion explaining that chronic exposure to acetamiprid could affect human neural functions and development.²³⁹ It has been proved that this insecticide is also toxic for the environment^{223,240}, mice²⁴¹ and beneficial insects such as bees.^{224,225} Despite the ability of neonicotinoids to be metabolized, they have been detected in songbird after 6 h of ingestion.²⁴² Acetamiprid is resistant to most of the water and wastewater treatment processes.²⁴³⁻²⁴⁶ Several studies were conducted for accelerating its removal. Sun *et al.* (2019)²⁴⁷ worked on removing ACT using biological processes while low-temperature heat plasma processes were employed by Li *et al.* (2014).²⁴⁴ John *et al.* (2016) studied the effects of coagulation and flocculation processes on the removal of acetamiprid.²⁴⁸ The photocatalyzed degradation of the compound was studied by Khan *et al.* (2010)²⁴⁹ and optimized by Toolabi *et al.* (2017)²⁵⁰ with Fe₃O₄ and SiO₂ addition to increase degradation efficiency. More recently, Cruz-Alcalde *et al.* (2017) worked on the removal of ACT by ozonation.²⁵¹ Only a few of these studies achieved structural elucidation of the degradation products.^{46,252} Nevertheless, it has been shown that the nature and relative amounts of photoproducts are strongly dependent on the irradiation medium.^{186,229} Since pesticides are marketed in mixtures with other chemical compounds (whose nature is not provided by the manufacturer), it is meaningful to study their photodegradation in such matrices and not only in aqueous or organic solutions as predominantly reported in the literature. The present work is devoted to studying the UV-Vis photodegradation of acetamiprid, in water and in two marketed

mixtures (Polysect Ultra SL[®] and Roseclear Ultra[®]). The main objectives were to elucidate the chemical structures of photoproducts in order to suggest degradation pathways and to compare matrix effects on the photodegradation process. Structural elucidation was achieved using LC-FT-ICR-MS/MS based on exact mass measurements, tandem mass spectrometry experiments and kinetic approaches. The global environmental toxicity was estimated using *in vitro* bioassays on *Vibrio fischeri* bacteria for an aqueous solution of acetamiprid and both commercial mixtures, before and after photolysis. *In silico* tests were also carried out to estimate the potential toxicity (mutagenicity, developmental toxicity, *Daphnia magna* LC₅₀ 48 hours and oral rat LD₅₀) of the parent molecule and its photoproducts.

2 EXPERIMENTAL

2.1 Chemicals and reagents

Acetamiprid, acetamiprid-D₃ (purity > 98% for both), acetonitrile (ACN) and formic acid (FA) (chromatographic grade purity > 99.99% for both) were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). The chemical structures of acetamiprid and acetamiprid-D₃ are displayed in Figure III-9. Two commercial solutions, containing acetamiprid, were purchased: Polysect Ultra SL[®] and Roseclear Ultra[®] from Evergreen Garden Care (Ecully, France). Polysect Ultra SL[®] contains acetamiprid at 5000 mg.L⁻¹ while Roseclear Ultra[®] contains acetamiprid at 2500 mg.L⁻¹ in mixture with triticonazole at 7500 mg.L⁻¹. No other information was provided by manufacturers. Solutions of acetamiprid and acetamiprid-D₃ were prepared at 10 mg.L⁻¹ in ultrapure water. Polysect Ultra SL[®] and Roseclear Ultra[®] were diluted in ultrapure water to reach a final concentration in acetamiprid of 10 mg.L⁻¹ for the photolysis experiments. It is to be noted that the concentrations used in the present study are greater than those usually reported in real samples. Using concentration values higher than those frequently detected in real samples is very common in experiments aiming at MS structural elucidation. This is because photodegradation usually leads to tens of products (including many isomers), leading to a kind of "signal dilution". In such conditions, even the most performant mass

spectrometers do not provide suitable detection thresholds since they have to be first used in scanning mode. Furthermore, structural elucidation requires MS/MS or MSⁿ measurements than cannot be properly achieved if the precursor ion abundance is too weak.

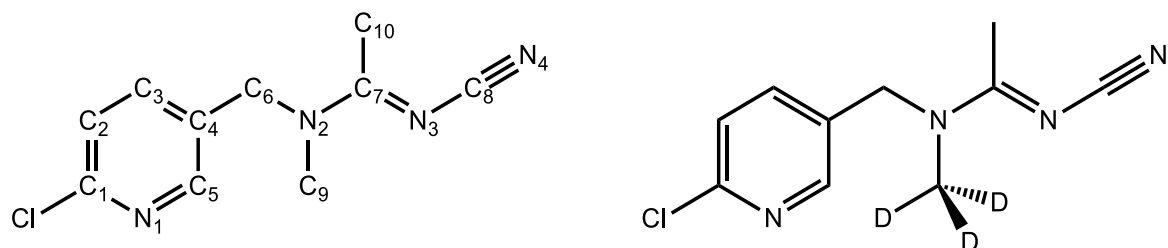


Figure III-9: Chemical structures of acetamiprid (left-hand) - with numbering of carbon and nitrogen atoms - and acetamiprid-D₃ (right-hand)

2.2 UV-Vis irradiation experiments

Solutions were irradiated in a laboratory-made reactor equipped with a UV-Vis high-pressure mercury lamp HPL-N125W/542 E27 SC (Philips, Ivry-sur-Seine, France) emitting light on wavelengths ranging from 200 nm to 650 nm, with a maximum irradiation wavelength at 254 nm and a radiation flux of 6200 lm. The spectrum of the Hg lamp is provided in the Supplementary Information III-26. Pesticide solutions were irradiated in quartz tubes immersed in a sonicator AL04-12 (Advantage-Lab GmbH, Darmstadt, Germany) cooled with water recirculation in order to maintain the solutions at a temperature of 30 ± 5 °C. Irradiations were conducted on 50 mL solutions on 4 h. 1 mL samples were collected at 0, 1, 2, 5, 15, 30, 60, 120, 180, 240 and 360 min and analyzed by LC-HRMS.

2.3 LC-MS operating conditions

Chromatographic separations were carried out using an Acquity HPLC system (Waters Technologies, Saint Quentin en Yvelines, France) equipped with a C₁₈ Pursuit XRs Ultra 2.8

μm , 50 x 2.0 mm column (Agilent Technologies, Les Ulis, France) and coupled with a FT-ICR SolarixXR 9.4 T (Bruker Daltonics, Bremen, Germany). The injection volume was 10 μL . Elution was performed using a 0.1 mL/min flow with a mixture of solvents A (water, formic acid 0.1%) and B (acetonitrile, formic acid 0.1%). After 10 minutes, percentage of solvent B was increased from an initial value of 5% to 40% in 15 min, then to 90% in 5 min. The gradient was then set back to initial conditions (5% of B) during the last 20 min, for a total method duration of 50 min. An electrospray source was used in positive mode with a capillary voltage and an end plate offset value at - 4500 V and - 4000 V, respectively. Nitrogen was used as nebulizer and drying gas, the former being set at 1 bar and the latter with a flow of 4 L/min. The drying gas temperature was set at 250 °C. For MS/MS experiments, argon was used as the collision gas with collision voltages ranging from 5 to 20 V. Ions were accumulated in the collision cell for 0.2 s in MS experiments and 2 s in MS/MS experiments. Ion detection was set at 4 Mpt with a 0.8389 s transient duration in broadband mode from m/z 57.7 to m/z 1000 to obtain a resolution higher than 350 000 at m/z 200. In order to reduce the data file size, 97% of the data (contributing to noise) were not recorded to get a data file size below 2 Go per analysis. Phosphoric acid was used for mass calibration and exact chemical formulae were attributed with a tolerance below 3 ppm.

2.4 Toxicity estimations

In vitro toxicity tests were carried out using a *Vibrio fischeri* bioluminescence inhibition commercial test kit (Hach Lange GmbH, Dusseldorf, Germany) which is based on the ISO 11348-3 protocol.¹⁴⁰ 500 μL of frozen lyophilized bacteria were reconstituted just before analysis by adding 1.3 mL of a saline solution provided by the manufacturer. The reconstituted solution was equilibrated for 15 minutes at 4 °C before testing. A working solution was prepared by adding 12.5 mL of dilution reagent to the hydrated bacteria. The percentage of inhibition (I%) was determined by comparing the bioluminescence of the sample solution to that of the reference solution at different concentrations. In what follows, the concentration that causes 80% of bioluminescence inhibition is designated as EC₈₀ value; it was determined for exposure times of 5, 15, and 30 min. Stock solutions of

acetamiprid, Polysect Ultra SL[®] and Roseclear Ultra[®] were prepared at an acetamiprid concentration of 300 mg.L⁻¹. Each diluted test solution was prepared in 2% NaCl solution and tested in duplicates at 15 °C. *In vitro* toxicity measurements were carried out on the reference solution (before irradiation) and on solutions after 30, 60, 120 and 240 min of irradiation.

In silico toxicity predictions were carried out with the Toxicity Estimation Software Tool (T.E.S.T.). This free-downloadable software is a US Environmental Protection Agency predictive system that uses quantitative structure-activity relationship (QSAR) mathematical models. It is used to predict acute toxicity values from chemical and physical properties of a given molecular structure. T.E.S.T. uses a simple linear function of molecular descriptors such as the octanol-water partition coefficient and parameters related to the presence or absence of some chemical groups.²⁵³ In the present study, we focused on mutagenicity, developmental toxicity, oral rat LD₅₀ and *Daphnia magna* LC₅₀ 48 hours.

3 RESULTS AND DISCUSSION

3.1 *Kinetics data and photoproducts characterization*

The relative acetamiprid concentration in ultrapure water, with a starting concentration of 10 mg.L⁻¹, has been followed as a function of irradiation time, by integrating the area of the corresponding chromatographic peak on the [M+H]⁺ ionic current in ESI+ (See Supplementary Information III-27 in the supplementary information data file). With the mercury lamp, the initial acetamiprid amount is halved after about 10 min and the compound is no longer detected after 40 min of irradiation. It is to be noted that these results are appropriate at laboratory scale but cannot be extrapolated to real conditions, as this irradiation does not reach the earth's surface. Possibly, degradation under sunlight irradiation would be different. Figure III-10 shows the LC-MS chromatograms of the aqueous solution of acetamiprid at t₀ (10 mg.L⁻¹) and after 60 min of irradiation. The chromatogram of the reference solution (t₀) confirms the purity of the marketed product while that of the irradiated one shows the formation of seven photoproducts, which have

been divided into four groups – referred as PP1 to PP4 – on the basis on their exact mass. With the aim of correlating the appearance of some photoproducts with the disappearance of some others, their relative abundances (normalized data) have been plotted as a function of irradiation time (See the Supplementary Information III-28). Ascertaining such correlations may allow an easier structural elucidation. The exact m/z values of the $[M+H]^+$ ions and those of the product ions from collision-induced dissociation experiments are provided for each photoproduct; the corresponding tables are referred as Supplementary Information III-29 to Supplementary Information III-33 in the supplementary information data file.

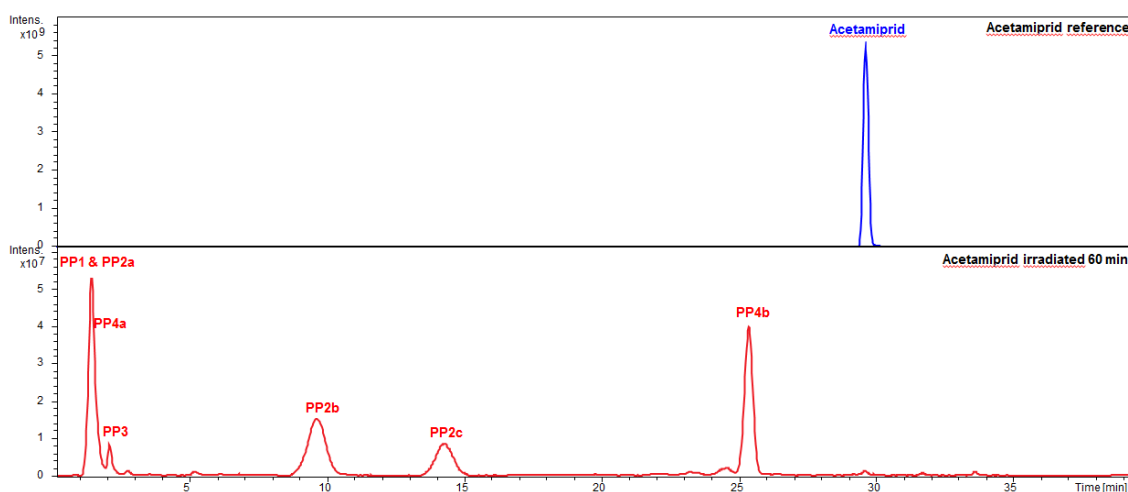


Figure III-10: LC-MS chromatograms of the aqueous solution of acetamiprid ($10 \text{ mg}\cdot\text{L}^{-1}$) at t_0 and after 60 min of irradiation

Exact formulae of photoproducts show the absence of the chlorine atom in all cases, suggesting that cleavage of the C-Cl bond constitutes the first step of the decomposition process, as already reported for the typical photodegradation mechanism of compounds containing a halogenated phenyl group.¹²⁸ Given that the radical species resulting from Cl. elimination leads to a non-conjugated species that does not allow delocalization of the unpaired electron, the following step of the process may likely consist in hydroxyl or hydrogen abstraction from a water molecule to lead to A or B, as illustrated in Figure III-11. The formation of A is expected to be favored over that of B because, in aqueous solution,

the Cl^\cdot radical will abstract H \cdot from a water molecule to release a HO^\cdot radical able to react with the [acetamiprid - Cl] \cdot species. That is probably the reason why A is detected from the beginning of irradiation while B is not detected in the aqueous solution. A has been considered as the direct precursor of most of the photoproducts since its relative amount starts decreasing after only five minutes of reaction. According to the Hamilton rule, the short retention times of photoproducts – compared to that of acetamiprid – suggest that cyclization mechanisms are involved in the phototransformation of A.²⁵⁴

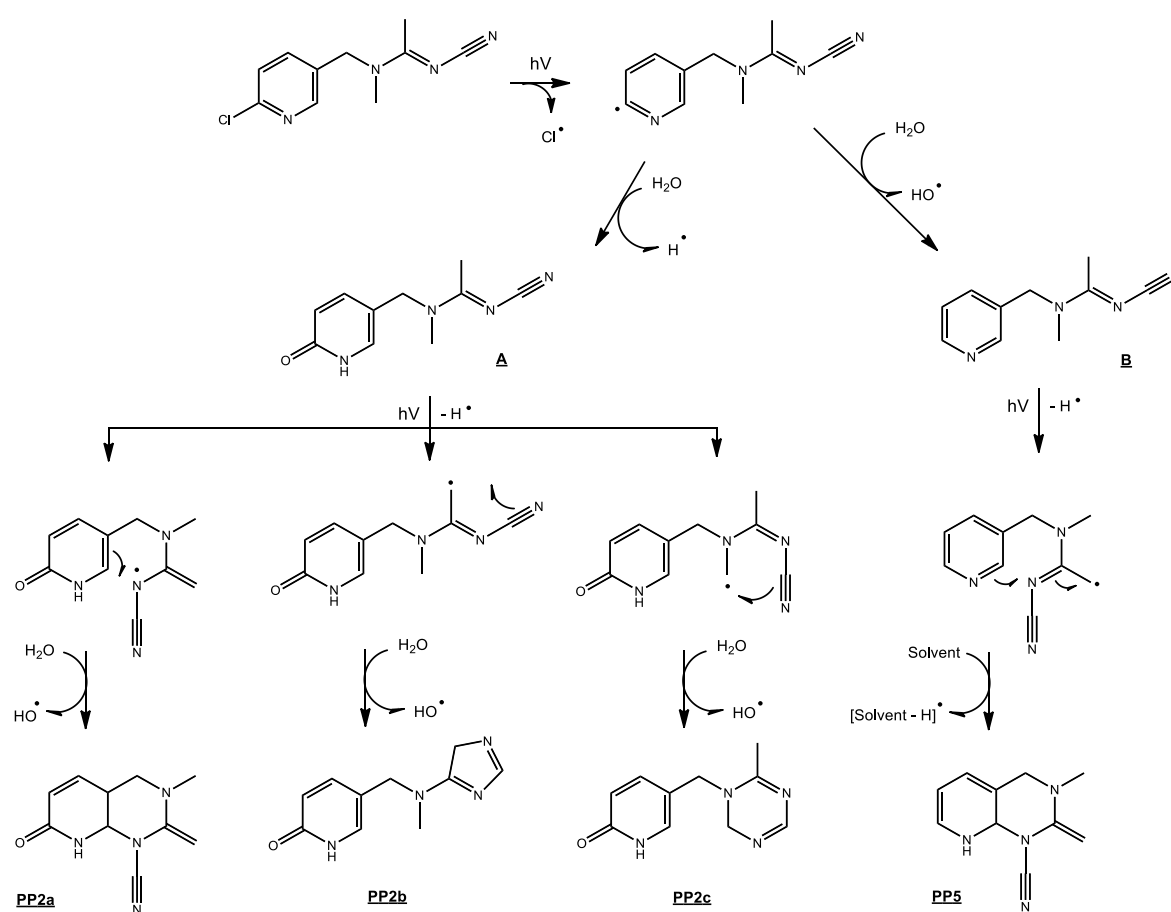


Figure III-11: Mechanisms proposed for the photo-induced formation of A, B, PP2 isomers and PP5

Photoproducts PP2a to PP2c

With an exact formula of $\text{C}_{10}\text{H}_{12}\text{N}_4\text{O}$, PP2 isomers were assumed to result from cyclisation of A. Cyclization implies photo-induced stripping of a hydrogen atom, pairing of the resulting lone electron with a π electron and hydrogen recapture from the solvent.

Irradiation of acetamiprid-D₃ in water leads to a 1 mass unit shift for PP2c, showing evidence that a deuterium atom stripping from C9 was at the origin of the process, in agreement with the mechanism suggested for cyclization. Photo-induced cleavages of C-H and C-D bonds have been recently demonstrated in a study by De Vaugelade *et al.* reporting the UV-Vis photodegradation of alpha-tocopherol.²²⁹ Cleavage of such bonds is only observed if the resulting radical is strongly stabilized by mesomeric and/or inductive effects. In the case of structure A, the hydrogen atoms that may likely be abstracted under irradiation are those carried by the carbon atoms C6, C9 and C10 (see Figure III-9 for the numbering of atoms). All the possibilities involving the formation of a 5- or 6-membered ring from A, after C-H bond cleavage, to provide a photoproduct carrying 3 deuterium atoms were systematically considered; the 11 resulting structures, referred as A1 to A11, are depicted in the supplementary data file Supplementary Information III-34. As an example, the following shows how structural elucidation has been achieved in the case of PP2a. The loss of CH₃NO from [PP2a+H]⁺, providing the ion at *m/z* 160, confirms the presence of the amide function postulated for products A1 to A11. The corresponding dissociation mechanism is displayed in Figure III-12 (mechanism 1). The loss of HCN from [PP2a+H]⁺ suggests that the nitrile group has been kept intact and that there is a hydrogen atom in alpha position regarding C8, in accordance with the dissociation mechanism 2 in Figure III-12. Consequently, structures A1, A2, A3, A6, A8, and A9 cannot correspond to PP2a. The loss of C₄H₄N₂O has to be considered as subsequent eliminations of HCN (see above) and C₃H₃NO. The latter implies the presence of a double bond between C2 and C3 carbon atoms (see mechanism 3 in Figure III-12); this allows to discard structures A5 and A7. The elimination of C₂H₅N (*m/z* 162 for acetamiprid and acetamiprid-D₃) permits to conclude that the N2-C9H₃ moiety has been kept intact and that there is a single bond between C6 and N2 (mechanism 4), while that of CH₃N (not shifted for the deuterated analogue) implies that there is a single bond between C4 and C6 (mechanism 5 in Figure III-12); this allows to eliminate A10 and A11 structures. Structure A4 is thus the only one that permits to rationalize all the collision-induced dissociations reported for [PP2a+H]⁺. The losses of C₄H₄N₂O and NH₃ (leading to *m/z* 109 and *m/z* 188, respectively, and both non-shifted for the deuterated analogue) are in agreement with the structure established for PP2 (see Figure III-12). Structural elucidation of PP2a illustrates how exact mass measurements combined with CID experiments and isotopic labeling may allow the

characterization of an unknown structure without the help of additional analytical technics. In the present study, the other photoproducts of acetamiprid have been characterized using the same approach.

Under collision activation, [PP2b+H]⁺ loses C₄H₇N₃ (C₄H₄D₃N₃ from the deuterated analogue) and C₆H₅N₂O (idem for the deuterated analogue). Both dissociations correspond to the cleavage of the C6-N2 bond after protonation on N2 (see Supplementary Information III-29). The only structure allowing such cleavages in an easy way corresponds to that depicted in Figure III-11.

PP2c carries two deuterium atoms, indicating that cyclization occurred after cleavage of a C9-D bond. The presence of C₆H₆NO⁺ and C₄H₈N₃⁺ ions in the CID spectrum of [PP2c+H]⁺ means that the oxygen atom is carried by the pyridine ring (as for A) and that cyclization did not occur between the C5 carbon atom and the resulting ring. PP2c was assumed to come from cyclization of A because the neutral loss of C₂H₅N (consecutive losses of HCN and CH₄) can only be rationalized from a cyclic pattern (see Supplementary Information III-29). The only structure allowing both eliminations of HCN and C₂H₅N from [PP2c+H]⁺ is that displayed in Figure III-11.

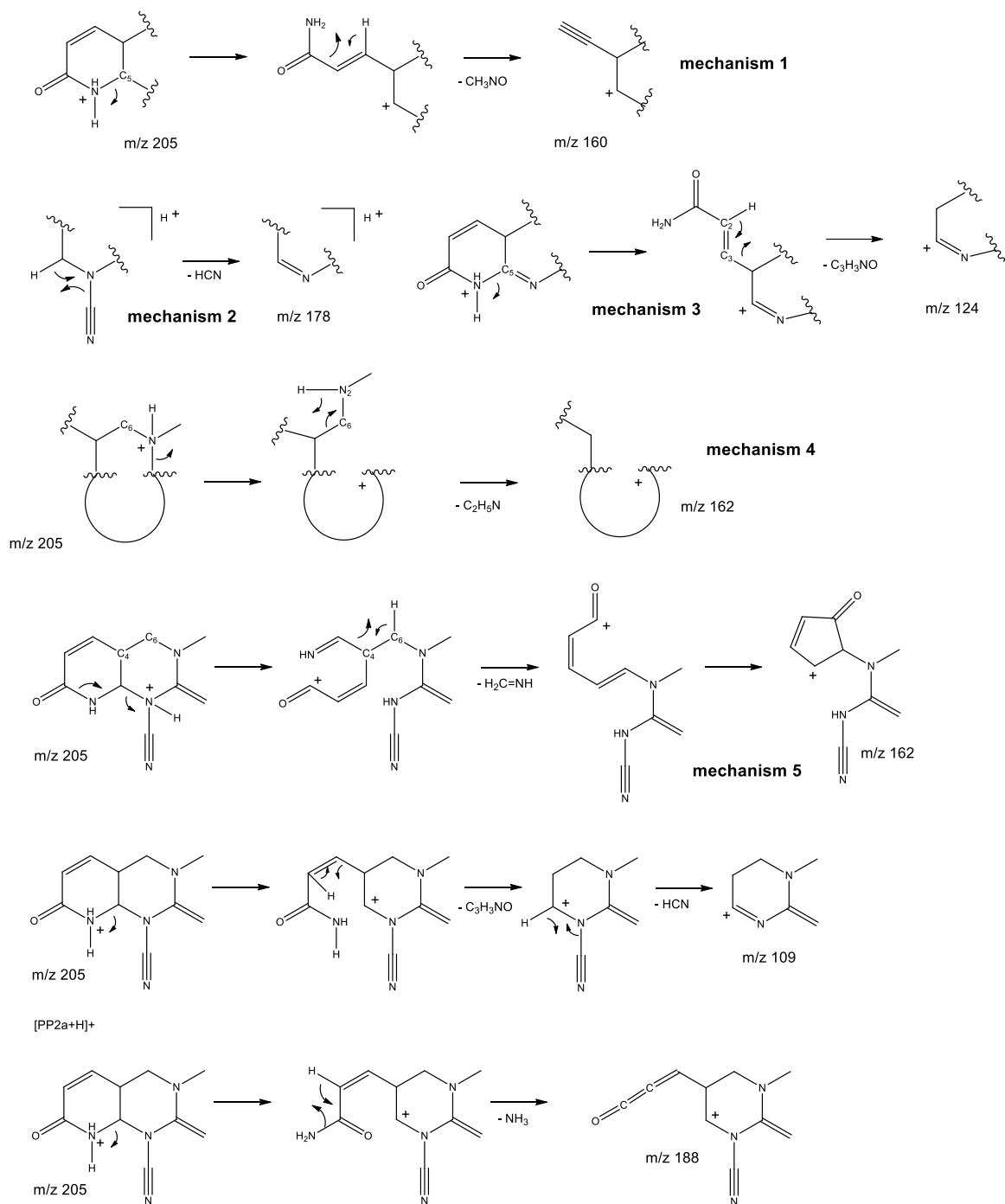


Figure III-12: Mechanisms involved in the collision-induced dissociation of $[PP2a + H]^+$ ions

Photoproduct PP1

Kinetics data suggest that PP1 comes from phototransformation of PP2a since its appearance is correlated to the disappearing of the latter. Compared to that of PP2a, the chemical formula of PP1 ($C_{10}H_{15}N_3O_4$) correspond to the loss of one nitrogen atom and the addition of three hydrogen and three oxygen atoms; the three deuterium atoms were

kept in the case of the transformation of acetamiprid-D₃. Considering that the number of carbon atoms remains unchanged, the loss of a nitrogen atom has been assumed to result from NH₃ elimination. PP1 is thus expected to result from NH₃ elimination and addition of three water molecules. The loss of ammonia implies the preliminary formation of a primary amine. As demonstrated by Abuin *et al.*, Norrish type I cleavages in alpha position of carbonyl groups are frequently reported under UV irradiation.²⁵⁵ Such a mechanism has been considered from PP2a to break the C1-N1 bond and form a carboxylic acid and a primary amine after water addition (see Figure III-13). Considering the dissociations of [PP1+H]⁺, the losses of NH₃ and CHNO suggest the presence of an amide function while the easy elimination of CH₄N₂O (observed in the source spectrum) suggests that this amino function is carried by a nitrogen atom. Eliminations of CH₃NO₂ and CH₆N₂O₂ show that C1 carries an oxygen atom. Given that pH is less than 3, acidic hydrolysis is expected to replace the primary amine by an alcohol one and the cyano group by an amino function, as depicted in Figure III-13. The transition m/z 181 \rightarrow m/z 137 traduces a loss of CO₂ and thus the presence of a carboxylic function, in agreement with the Norrish-type reaction suggested above. All the daughter ions of [PP1+H]⁺ are in good agreement with the suggested structure (see Supplementary Information III-30).

Photoproduct PP3

Kinetics data suggest that PP3 also comes from phototransformation of PP2a (see Supplementary Information III-28). A mechanism analogous to that proposed for the formation of PP1 but starting by a Norrish cleavage between C1 and C2 has been suggested in Figure III-13. There is no proof of this mechanism but the numerous fragmentations of the protonated molecule [PP3+H]⁺ are in good agreement with the proposed structure, particularly the consecutive eliminations of NH₃ and CO₂ (see Supplementary Information III-31).

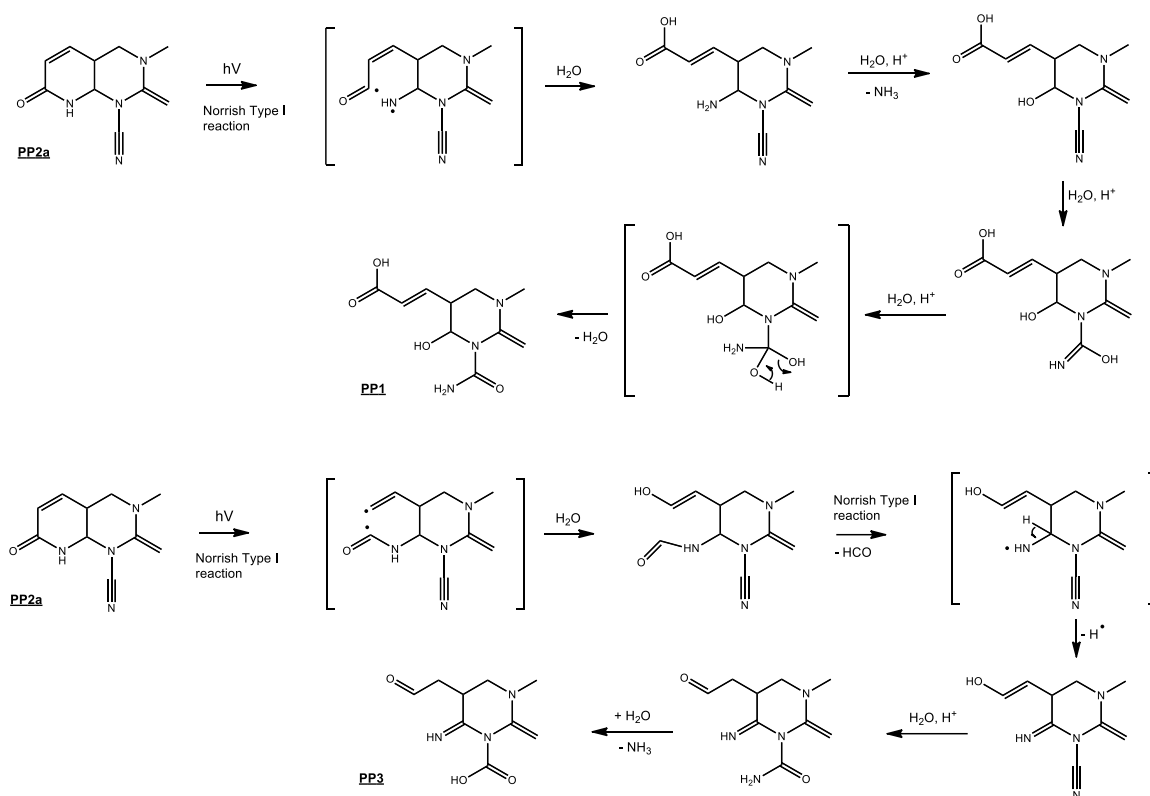


Figure III-13: Photo-induced mechanisms postulated for the formation of PP1 and PP3

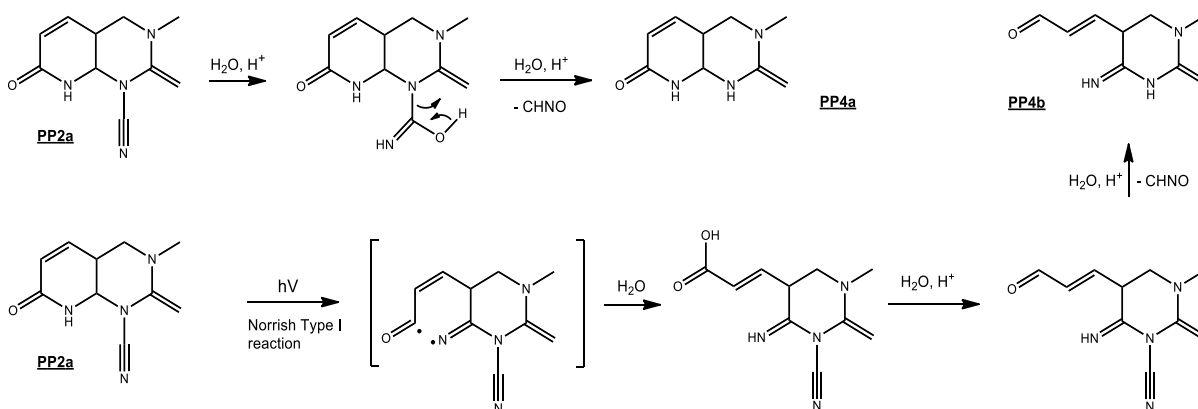


Figure III-14: Mechanisms postulated for the formation of PP4a and PP4b isomers

Photoproducts PP4a and PP4b

Considering their retention times (1.4 min for PP4a and 25.0 min for PP4b), both isomers are expected to be structurally very different, i.e. with a very low cross section for PP4a and a much larger one for PP4b. Kinetics data suggest that PP4a and PP4b also result from

the transformation of PP2a. For PP4a, the loss of CHNO (and in a lesser extent that of CO) from [PP4a+H]⁺ shows the presence of an amide function while that of CH₂N₂ implies the presence of a NH=CHR=NH pattern. Conservation of the 3 deuterium atoms for the deuterated analogue of acetamiprid led to propose the structure depicted in Figure III-14 which can be easily formed from PP2a after simple hydrolysis of the latter. Its very short retention time and all the collision-induced fragment ions of [PP4a+H]⁺ are in good agreement with the suggested structure (see Supplementary Information III-32). Concerning PP4b, its retention time suggests a monocyclic structure rather than a bicyclic one. Like for PP4a, the loss of CH₂N₂ testifies the presence of the NH=CHR=NH pattern while those of H₂O and CO reveal the presence of a carbonyl group. The fragmentation pathways of PP4b are reported in Supplementary Information III-32.

Analysis of marketed mixtures

As indicated in the experimental section, two marketed mixtures of acetamiprid: Polysect Ultra SL[®] and Roseclear Ultra[®] were diluted so that the final concentration of acetamiprid was 10 mg.L⁻¹, same as for the solution of acetamiprid in ultrapure water. Both solutions were irradiated up to 360 minutes. For each one, the relative amount of acetamiprid – estimated by integrating the area of the corresponding chromatographic peak on the [M+H]⁺ ionic current in positive ESI – has been plotted as a function of the irradiation time (see Supplementary Information III-27 in the supplementary information data file). The comparison between kinetics data shows that photodegradation of acetamiprid is significantly slower in commercial mixtures in comparison with the aqueous solution: total degradation of the insecticide is observed after about 120 min and 360 min for Polysect Ultra SL[®] and Roseclear Ultra[®], respectively, versus 40 min for the aqueous solution at the same concentration in acetamiprid. Given that the major part of the mixture remains unknown for both marketed products, it is difficult to draw conclusions about this strong matrix effect. However, two plausible explanations based on filter-effect and quenching of radical species can be suggested. The first one considers that an important part of incident photons is absorbed by other molecules in the mixture with acetamiprid so that the number of photons available for the excitation of the insecticide is reduced. Another

possible explanation lies in the photodegradation mechanisms pointed out in section 3.2. It can be considered that mechanisms induced by photocleavage of a C-H bond and leading to a radical species could be quenched by some compounds in the mixture. As a matter of fact, compounds including a good hydrogen-donor group could transfer a hydrogen atom onto the intermediate radical species displayed in Figure III-11 to instantly reform acetamiprid. Both phenomena could be responsible for the strong decrease in photodegradation rate of acetamiprid in complex mixtures in comparison with the aqueous solution.

Supplementary Information III-35 (see the supplementary information file) displays the LC-MS chromatograms of the Polysect Ultra SL[®] and Roseclear Ultra[®] mixtures diluted in water at t₀ and after irradiation for 240 min. In the case of Roseclear, the chromatographic peaks marked with an asterisk correspond to isomers of triticonazole. Isomerization of the analogous fungicide fenbuconazole under UV-Vis irradiation has been reported in a previous study.¹²⁹ For the Polysect Ultra SL[®] solution, all the photoproducts characterized after irradiation of the aqueous solution are detected, in relative amounts comparable to those in Figure III-10. For the Roseclear Ultra[®] solution, only the major ones, i.e. PP2b, PP2c, and PP4b, are detected at 240 min. A photoproduct, which has not been detected in the aqueous solution of acetamiprid, is detected at 2.3 min for both commercial mixtures; it is referred to as PP5. Its formation is easily rationalized by cyclization of B after photo-induced stripping of a hydrogen atom (see Figure III-11). It is to be remarked that the formation of A and that of B are in competition: the former requires hydroxyl addition onto the [acetamiprid - Cl][·] species and will be favored in an aqueous medium while the second requires hydrogen addition onto the [acetamiprid - Cl][·] and will be favored by the presence of a hydrogen donor. This is in good agreement with the fact that PP5, issued from B, is only observed in marketed mixtures that usually contain hydrocarbons to prevent the mixture from being rinsed by the rain. MS/MS dissociation pathways of PP5H⁺ are given in Supplementary Information III-33; they are all in good agreement with the structure postulated for PP5.

3.2 Toxicity predications

3.2.1 *In silico* toxicity estimation

QSAR *in silico* tests were carried out with the T.E.S.T. software (see section 1.4) for mutagenicity, developmental toxicity (including endocrine perturbation and reprotoxicity) and oral rat LD₅₀. Results for acetamiprid and its photoproducts are presented in Figure III-15; detailed values are given in the Supplementary Information III-36. Regarding developmental toxicity, all compounds but acetamiprid and B were found to be toxicant but the values estimated for the latter (0.49 and 0.48, respectively) correspond to the limit value below which molecules are considered as toxic (0.50). The correlation between the developmental toxicity of a molecule and the presence of chlorine atoms in its chemical structure has been reported in the literature.²⁵⁶ In the present study, the elimination of the chlorine atom constitutes undoubtedly the first step of the photochemical process so that none of the photoproducts is chlorinated. Considering the diversity of functional groups (cyano, hydroxyl, carboxyl, amido, etc.) on photoproducts, it would be unwise attempting to correlate the potential developmental toxicity to the presence of one or several chemical functions. It has to be kept in mind that most photoproducts exhibit potential developmental toxicity twice higher than that of the parent compound. Regarding potential mutagenicity, the results displayed in Figure III-15 show that all photoproducts are less toxic than acetamiprid. For oral rat LD₅₀, which is the amount of a chemical product in mg/kg bodyweight required to cause the death of 50% of a rat population, a value of 700 mg/kg was estimated in the present study versus 217 mg/kg in a recent work by Mosbah et al. questioning the accuracy of *in silico* approaches.²⁵⁷ In a relative way, the results displayed in Figure III-15 show that all the potential toxicities of photoproducts are similar or lower than that of acetamiprid. Here again, there is no evident correlation between the determined values and the chemical functions of the products. According to the Hodge and Sterner toxicity scale, all the photoproducts and acetamiprid can be considered as slightly toxic at the exception of PP2a, which can be considered as moderately toxic.¹⁸⁷ There are very few literature data on human and animals to allow comparison with the present results; acetamiprid has been reported to be potentially genotoxic, neurotoxic and hematotoxic at sublethal doses.^{258,259}

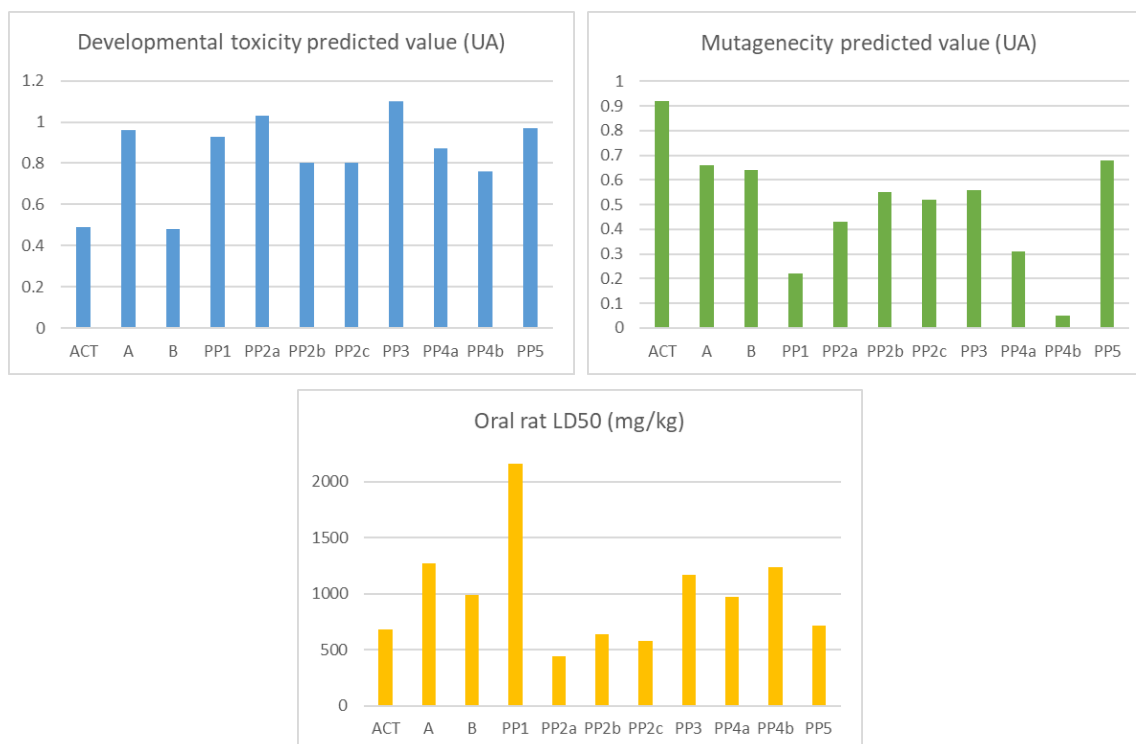


Figure III-15: Results of the *in silico* toxicity estimations for acetamiprid and its photoproducts

3.2.2 *In vitro* bioassays

Reference solutions of acetamiprid in ultrapure water, Polysect Ultra SL[®] and Roseclear Ultra[®] were irradiated following the protocol described in section 1.2. EC₈₀ values measured after 30 min of incubation are reported in Figure III-16 for irradiation times between 0 and 240 minutes. For the aqueous solution of acetamiprid, the evolution of the EC₈₀ value as a function of the irradiation time indicates that the toxicity increases by about 30% during the first 2 hours of irradiation before remaining stable until the end of the experiment. Considering the total disappearance of acetamiprid after 40 min of irradiation (see Supplementary Information III-27), the potential toxicity of the mixture of all the photoproducts except PP4a, can be assumed similar to that of acetamiprid. PP4a would be slightly more toxic since it appears after 40 min of irradiation and is the only photoproduct remaining in significant amount after 200 minutes of irradiation. Thus, these photoproducts should be monitored when analyzing food-products treated with acetamiprid and exposed to UV-Vis light. With a contact time of 30 minutes, the EC₈₀ values determined for the reference solutions of acetamiprid, Polysect Ultra SL[®], and

Roseclear Ultra® are 27 mg.L⁻¹, 14 mg.L⁻¹, and 5 mg.L⁻¹, respectively. It means that Polysect Ultra SL® and Roseclear Ultra® are potentially twice and five times more toxic for *Vibrio fischeri* than acetamiprid in water solutions. This higher toxicity can be imputed to the other constituents of the commercial mixtures whose composition is not provided by manufacturers. Regarding the evolution of EC₈₀ values with irradiation time, the same tendency is observed for Polysect Ultra SL® and Roseclear Ultra® as for acetamiprid: the potential toxicities of the irradiated solutions of marketed mixtures slightly increase. Even if the acetamiprid concentration decreases with the irradiation time in all cases (see Supplementary Information III-27), the solution gets more and more toxic when exposed to UV-Vis light. Regarding *in vitro* tests on bacteria, there is no literature available to permit comparison with those of the present work. The reliability of *in vitro* test results can be questioned with regards to the concentrations used, which are greater than those usually reported for real samples, as previously discussed in section 1.1. Actually, *in vitro* tests were conducted in a relative way, meaning that we were mainly interested in comparing toxicities of the solutions with each other, before and after irradiation. The protocol followed for these tests cannot be properly achieved if the toxicity of the pesticide solution - related to its concentration - is too high, because EC₈₀ values cannot be determined if too many bacteria are killed when adding the pesticide to the reference solution. In the present study, using high concentrations for acetamiprid and related commercial solutions was possible because the toxicity of acetamiprid is not so high.

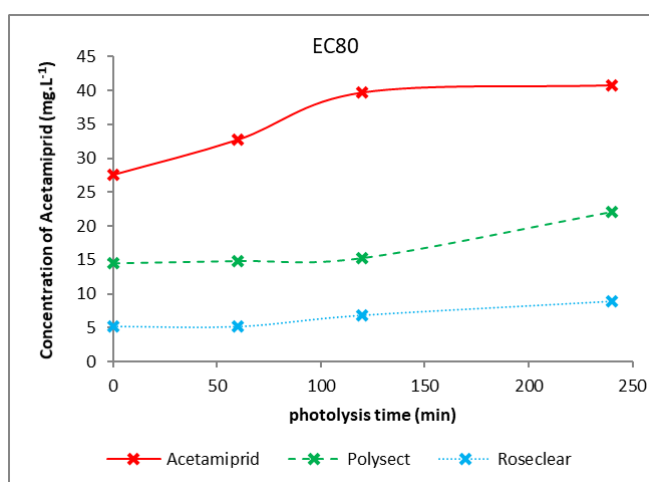


Figure III-16: EC₈₀ values determined for aqueous solutions of acetamiprid, Polysect Ultra SL® and Roseclear Ultra® at different irradiation times

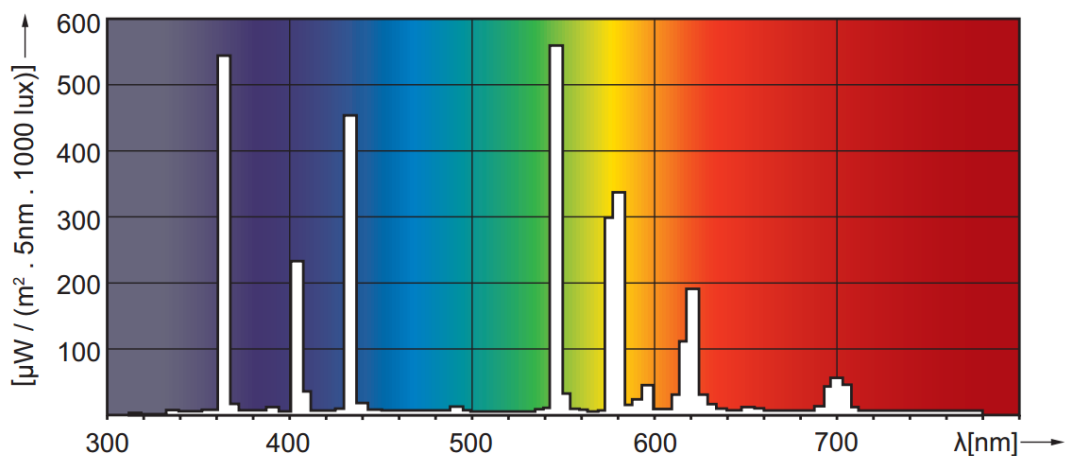
3 CONCLUSION

The UV-visible photodegradation of acetamiprid, in aqueous solution and in two marketed mixtures (Polysect Ultra SL® and Roseclear Ultra®) was investigated. Ten photoproducts were characterized using LC-HR-MS/MS analysis and degradation pathways were proposed based on their chemical structures and order of appearance/disappearance. The decomposition of acetamiprid is slower in commercial products where it is in mixture with unknown compounds that probably quench some intermediate radical species; one photoproduct is formed only in the presence of those excipients. Toxicities of the characterized compounds were estimated by QSAR *in silico* predictions. Most photoproducts exhibit potential developmental toxicity twice higher than that of the parent compound. Regarding potential mutagenicity, all photoproducts are less toxic than acetamiprid. Regarding oral rat LD₅₀, the potential toxicities of photoproducts are lower than or equal to that of acetamiprid. *In vitro* tests on *Vibrio fischeri* bacteria showed that the ecotoxicities of marketed mixtures are significantly higher than that of acetamiprid in aqueous solution; they slightly increase after UV-light exposure.

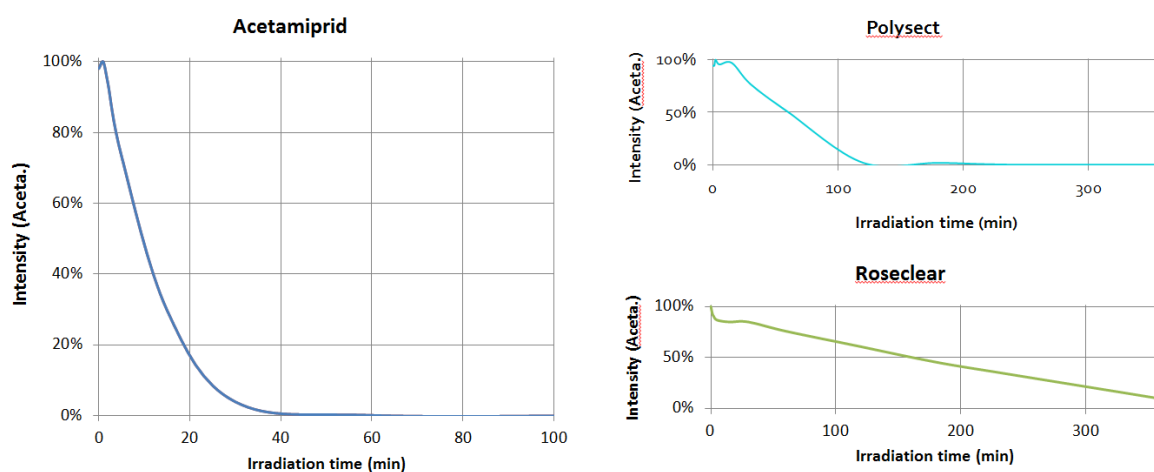
ACKNOWLEDGEMENTS

Financial support from the National FT-ICR network (FR 3624 CNRS) for conducting the research is gratefully acknowledged. This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No 765860 (AQUALity). Cindy Karmakian, Ségolène de Vaugelade and Sophie Bourcier are thanked for their scientific support.

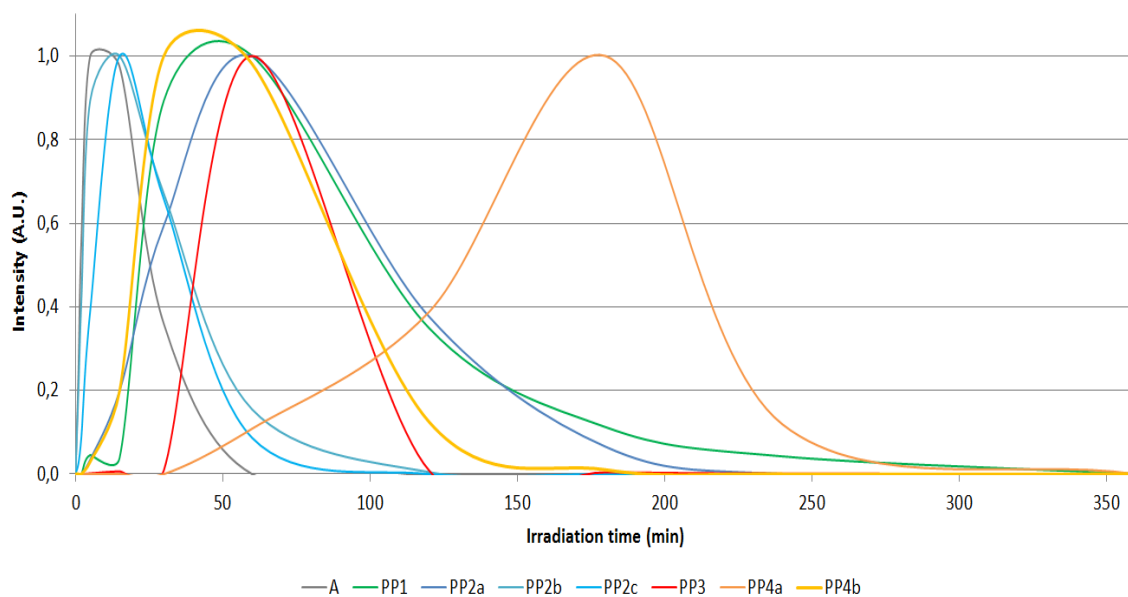
SUPPLEMENTARY INFORMATION FILE



Supplementary Information III-26: Emission spectrum of the mercury lamp HPL-N125W/542 E27 SC (Philips, Ivry-sur-Seine, France)



Supplementary Information III-27: Relative amount of acetamiprid (estimated by integrating the area of the corresponding chromatographic peak on the $[M+H]^+$ ionic current in positive ESI) as a function of the irradiation time in pure water, Polysect® and Roseclear® solutions

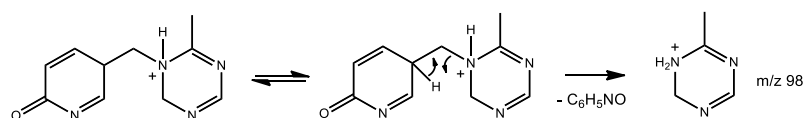
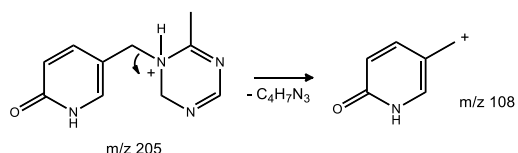
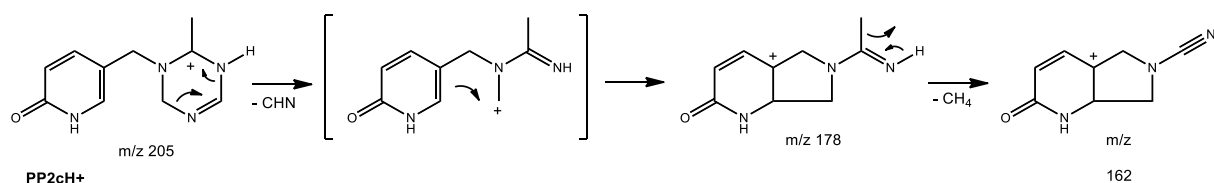
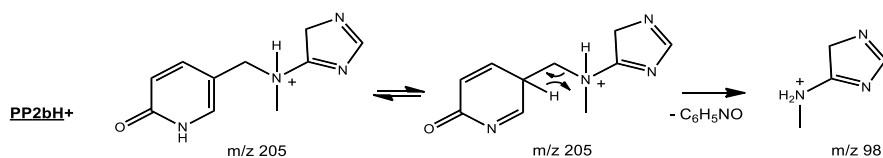
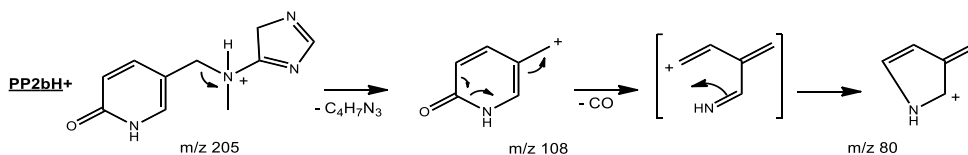
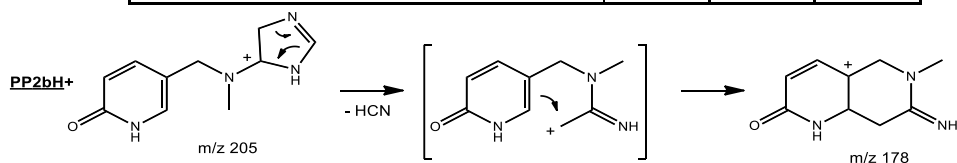


Supplementary Information III-28: Relative abundances of photoproducts of acetamiprid as a function of irradiation time (normalized data)

CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Supplementary Information III-29: CID data for [PP2+H]⁺ isomers and suggested dissociation pathways for PP2b and PP2c (for PP2a see Figure III-12 in the main document)

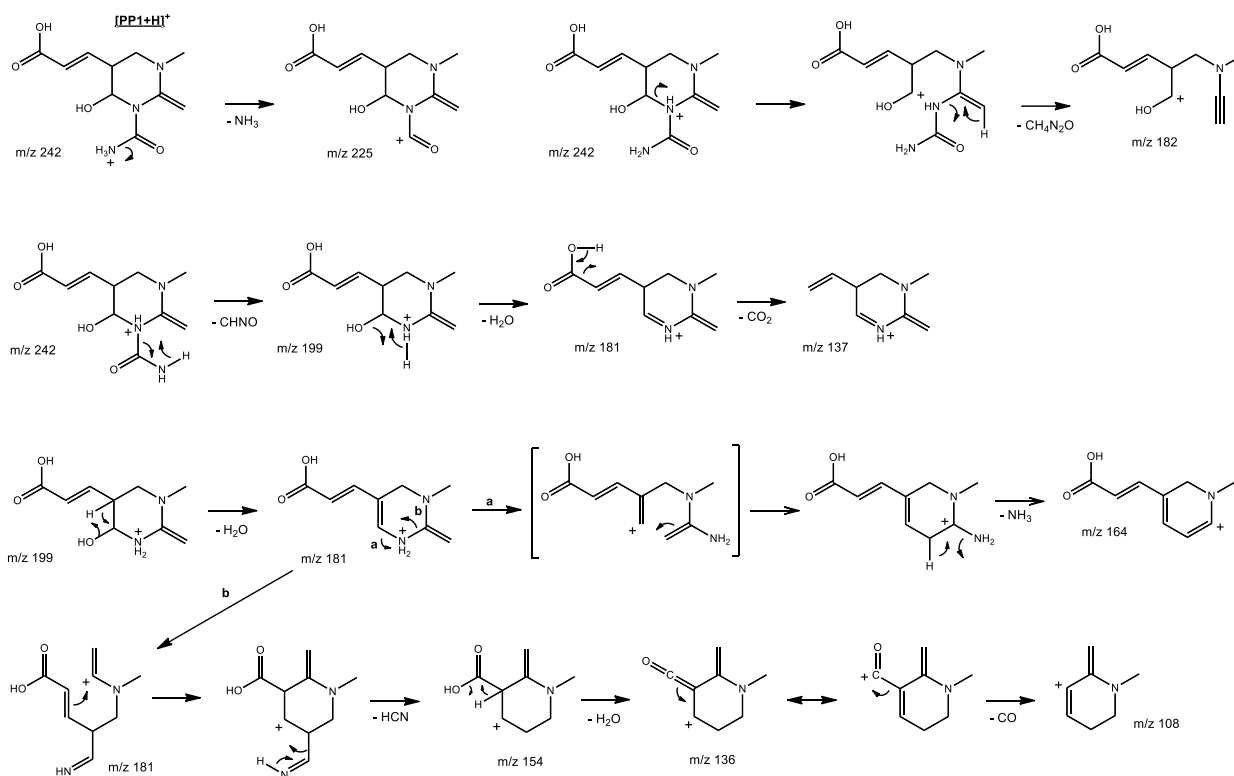
Ion Formula	<i>m/z</i>	PP2a	PP2b	PP2c
C ₁₀ H ₁₃ N ₄ O [PP2+H] ⁺	205.10839	D ₃	D ₃	D ₂
C ₁₀ H ₁₀ N ₃ O - NH ₃	188.08184	D ₃		
C ₉ H ₁₂ N ₃ O - CHN	178.09749		D ₃	D ₂
C ₉ H ₁₀ N ₃ O - CH ₃ N	176.08184	D ₃		
C ₈ H ₈ N ₃ O - C ₂ H ₅ N	162.06619	D ₀		D ₂
C ₉ H ₁₀ N ₃ - CH ₃ NO	160.08692	D ₃		
C ₈ H ₁₀ N ₃ - C ₂ H ₃ NO	148.08692	D ₃		
C ₆ H ₁₀ N ₃ - C ₄ H ₃ NO	124.08692	D ₃		
C ₆ H ₉ N ₂ - C ₄ H ₄ N ₂ O	109.07602	D ₃		
C ₆ H ₆ NO - C ₄ H ₇ N ₃	108.04439		D ₀	D ₀
C ₄ H ₈ N ₃ - C ₆ H ₅ NO	98.07127		D ₃	D ₂
C ₅ H ₆ N - C ₅ H ₇ N ₃ O	80.04948		D ₀	



CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

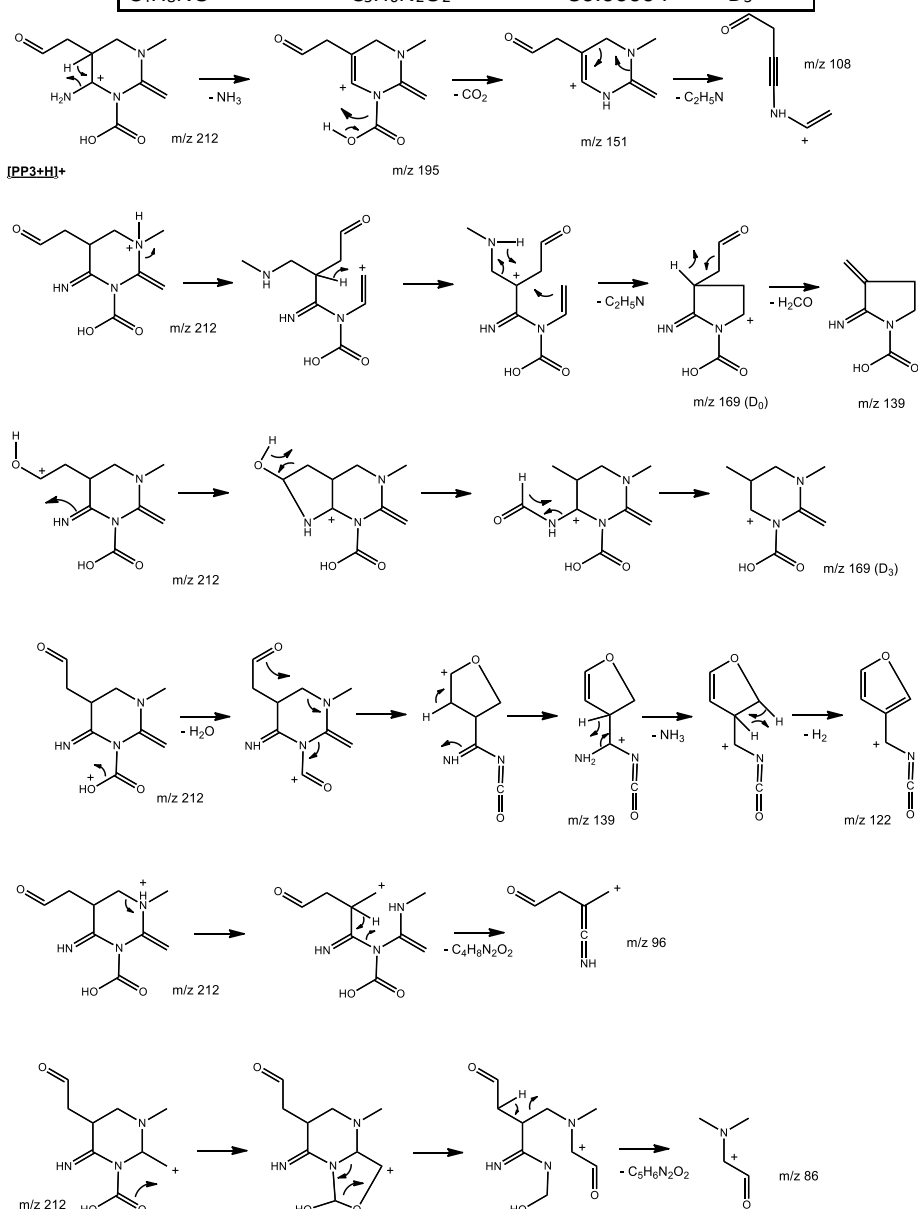
Supplementary Information III-30: CID data for $[PP1+H]^+$ and suggested dissociation pathways

Ion Formula		m/z	m/z shift for the D_3 analogues
$C_{10}H_{16}N_3O_4$	$[PP1+H]^+$	242.11353	D_3
$C_9H_{12}NO_3$	- CH_4N_2O	182.08117	D_3
$C_{10}H_{13}N_2O_4$	- NH_3	225.08698	D_3
$C_9H_{13}N_2O_2$	- CH_3NO_2	181.09715	D_3
$C_9H_{15}N_2O_3$	- $CHNO$	199.10772	D_3
$C_8H_{12}NO_2$	- $C_2H_4N_2O_2$	154.08625	D_3
$C_9H_{10}NO_2$	- $CH_6N_2O_2$	164.07060	D_3
$C_8H_{10}NO$	- $C_2H_6N_2O_3$	136.07569	D_3
$C_8H_{13}N_2$	- $C_2H_3NO_4$	137.10732	D_3
$C_7H_{10}N$	- $C_3H_6N_2O_4$	108.08078	D_3



Supplementary Information III-31: CID data for [PP3+H]⁺ and suggested dissociation pathways

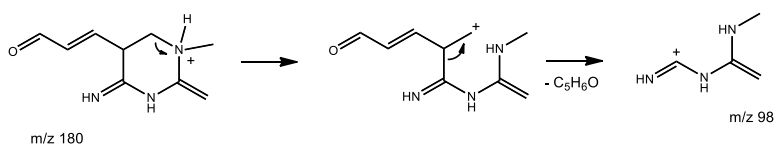
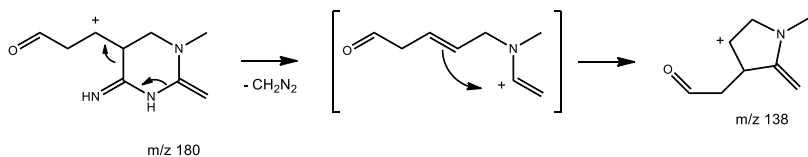
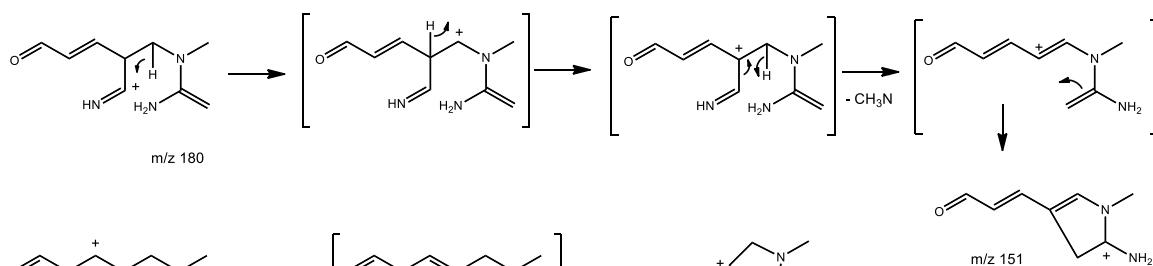
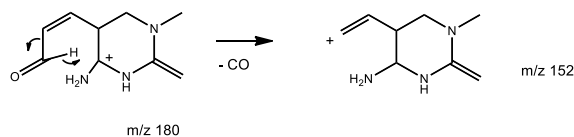
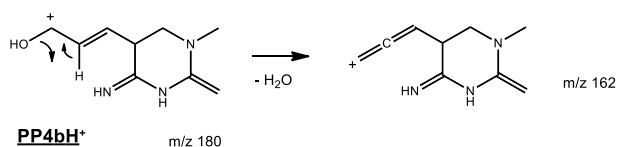
Ion Formula		<i>m/z</i>	<i>m/z</i> shift for the D ₃ analogous
C ₉ H ₁₄ N ₃ O ₃	[PP3+H] ⁺	212.10297	D ₃
C ₉ H ₁₁ N ₂ O ₃	- NH ₃	195.07642	D ₃
C ₈ H ₁₃ N ₂ O ₂	- CHNO	169.09715	D ₃
C ₇ H ₉ N ₂ O ₃	- C ₂ H ₅ N	169.06077	-
C ₈ H ₁₁ N ₂ O	- CH ₃ NO ₂	151.08659	D ₃
C ₆ H ₇ N ₂ O ₂	- C ₃ H ₇ NO	139.05020	-
C ₆ H ₄ NO ₂	- C ₃ H ₁₀ N ₂ O	122.02365	-
C ₆ H ₆ NO	- C ₃ H ₈ N ₂ O ₂	108.04439	-
C ₅ H ₆ NO	- C ₄ H ₈ N ₂ O ₂	96.04439	-
C ₄ H ₈ NO	- C ₅ H ₆ N ₂ O ₂	86.06004	D ₃



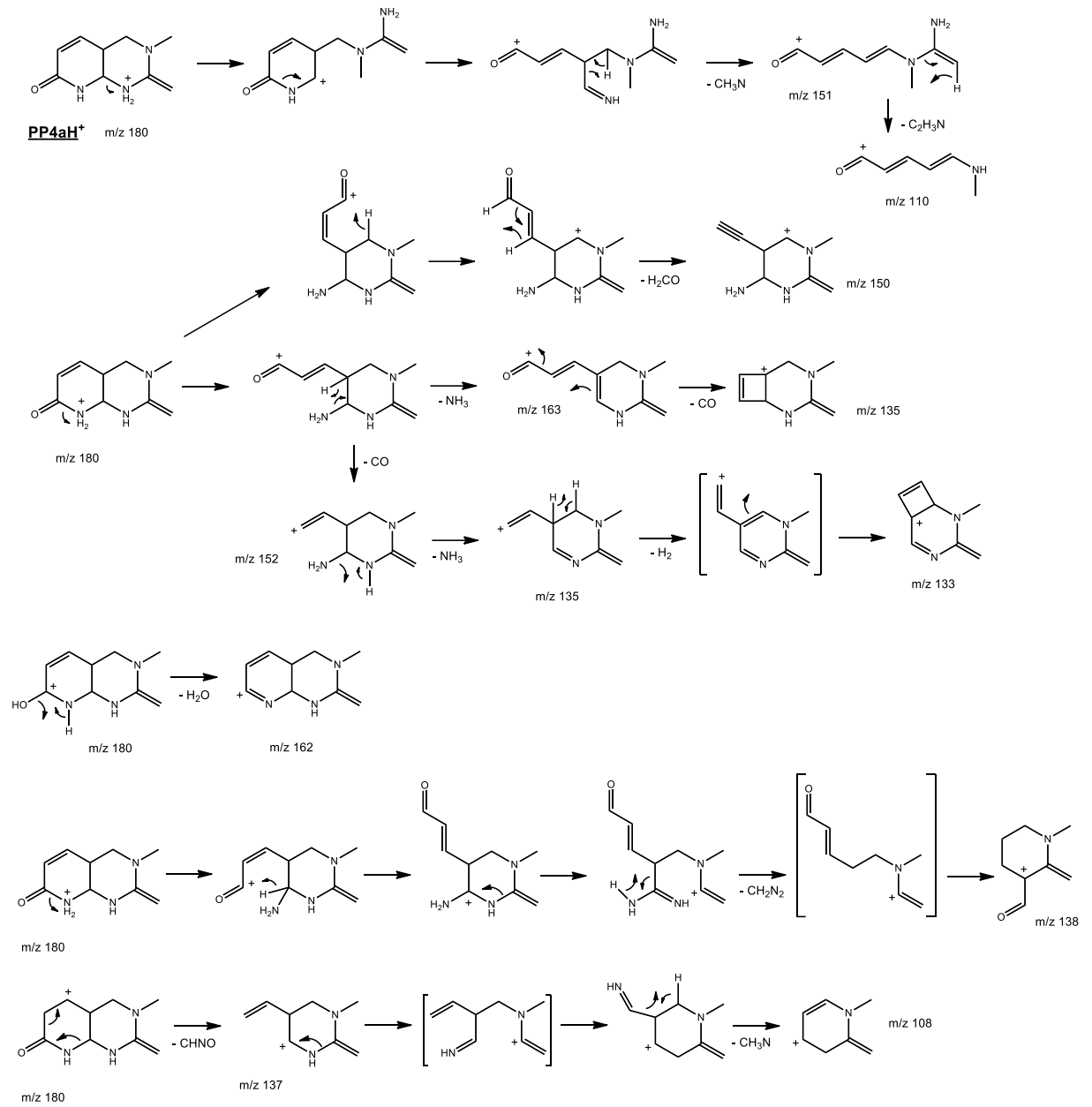
CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Supplementary Information III-32: CID data for [PP4+H]⁺ isomers and suggested dissociation pathways

Ion Formula	<i>m/z</i>	PP4a	PP4b
C ₉ H ₁₄ N ₃ O [PP4+H]⁺	180.11314	D ₃	D ₃
C ₉ H ₁₁ N ₂ O - H ₃ N	163.08659	D ₃	
C ₉ H ₁₂ N ₃ - H ₂ O	162.10257	D ₃	D ₃
C ₈ H ₁₄ N ₃ - CO	152.11822	D ₃	D ₃
C ₈ H ₁₁ N ₂ O - CH ₃ N	151.08659	D ₃	D ₃
C ₈ H ₁₂ N ₃ - CH ₂ O	150.10257	D ₃	
C ₈ H ₁₂ NO - CH ₂ N ₂	138.09134	D ₃	D ₃
C ₈ H ₁₃ N ₂ - CHNO	137.10732	D ₃	
C ₈ H ₁₁ N ₂ - CH ₃ NO	135.09167	D ₃	
C ₈ H ₉ N ₂ - CH ₅ NO	133.07602	D ₃	
C ₇ H ₁₂ N - C ₂ H ₂ N ₂ O	110.09643	D ₃	
C ₇ H ₁₀ N - C ₂ H ₄ N ₂ O	108.08078	D ₃	
C ₄ H ₈ N ₃ - C ₅ H ₆ O	98.07127		D ₃

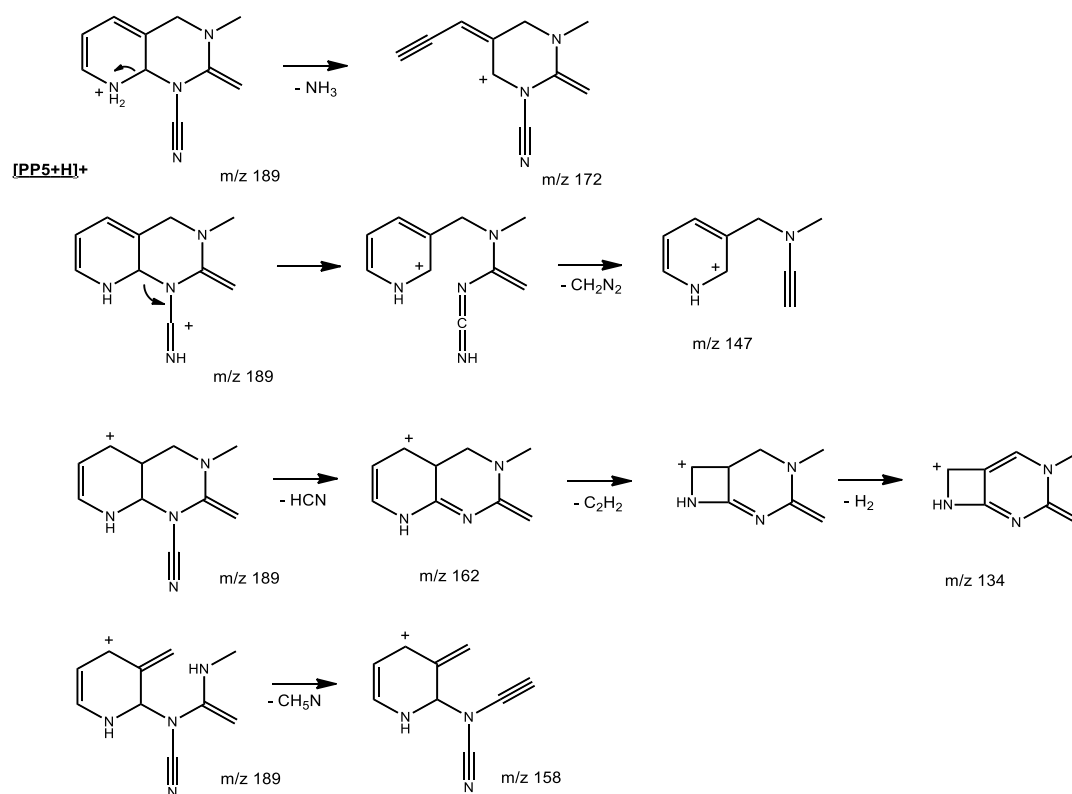


CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

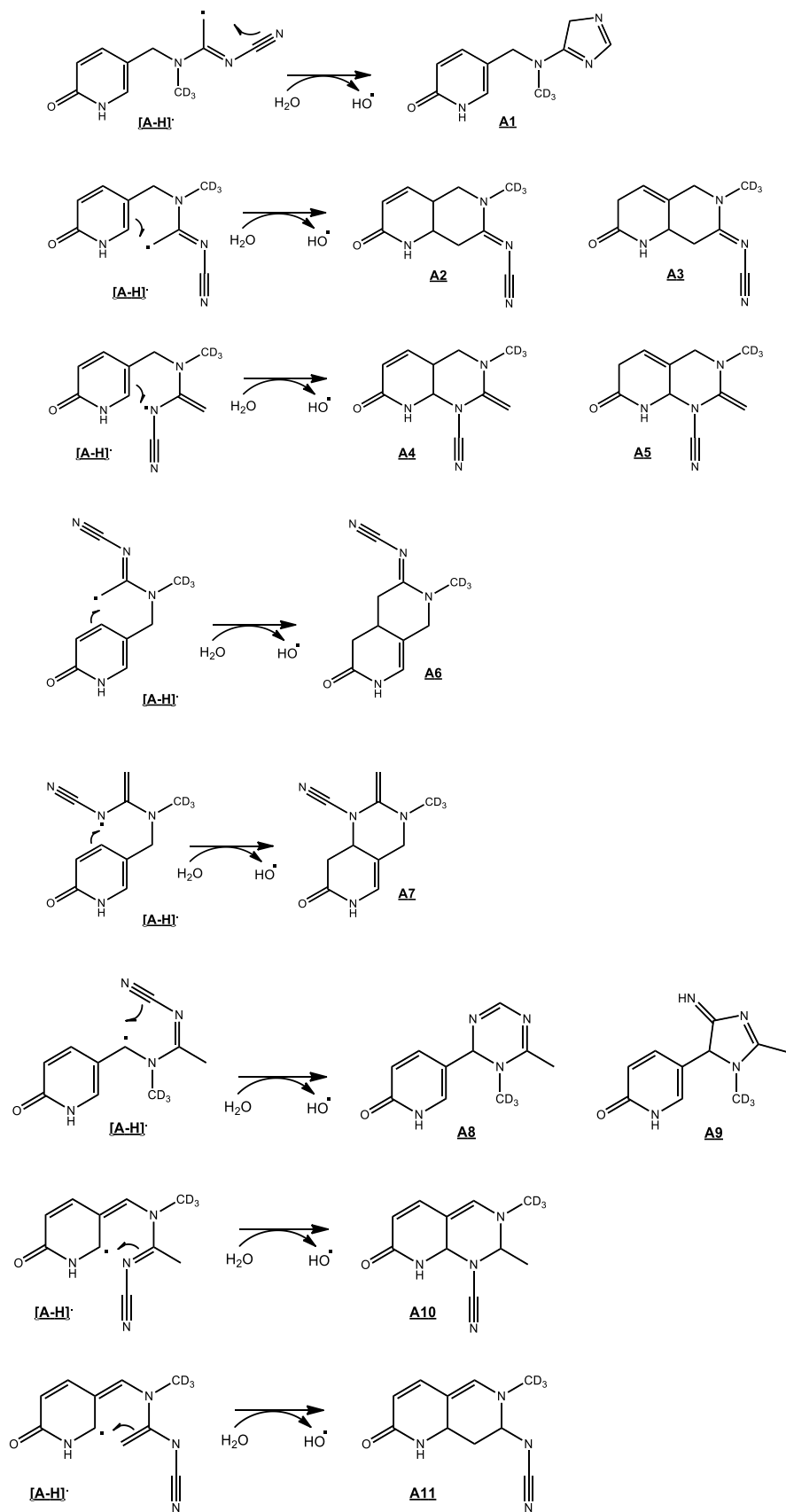


Supplementary Information III-33: CID data for [PP5+H]⁺ and suggested dissociation pathways

Ion Formula		<i>m/z</i>	<i>m/z</i> shift for the D ₃ analogous
C ₁₀ H ₁₃ N ₄	[PP5+H] ⁺	189.113473	D ₃
C ₁₀ H ₁₀ N ₃	- NH ₃	172.086924	D ₃
C ₉ H ₁₁ N ₂	- CH ₂ N ₂	147.091675	D ₃
C ₇ H ₈ N ₃	- C ₃ H ₅ N	134.071274	
C ₉ H ₁₂ N ₃	- CHN	162.102574	D ₃
C ₉ H ₈ N ₃	- CH ₅ N	158.071274	

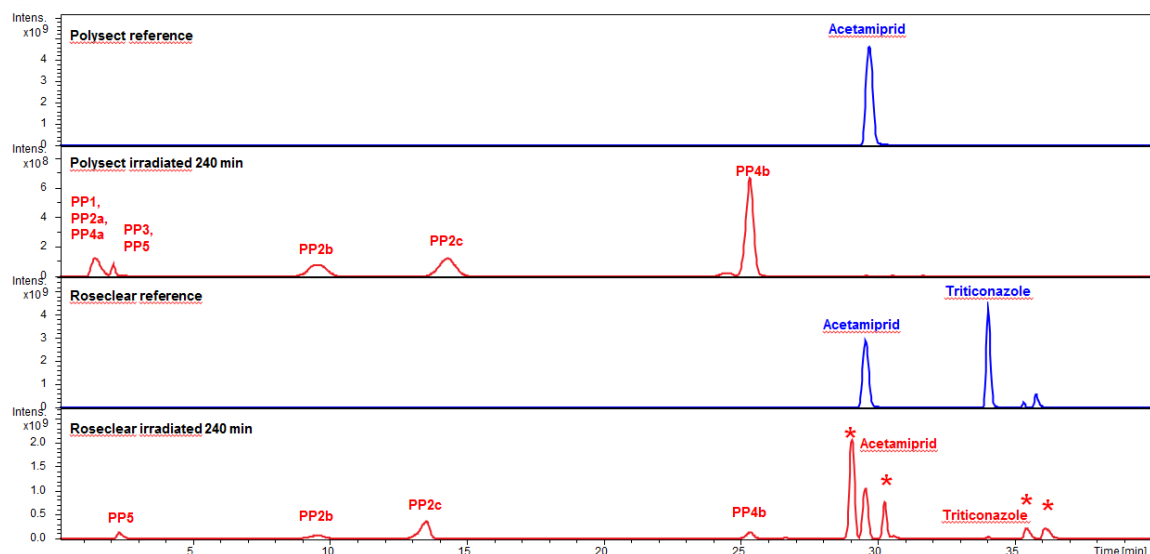


Supplementary Information III-34: Possible mechanisms involving the formation of a 5- or 6-membered ring from A to provide a photoproduct carrying 3 Deuterium atoms



CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Supplementary Information III-35: LC-MS chromatograms of the Polysect Ultra SL® and Roseclear Ultra® mixtures diluted in water at t0 and after irradiation for 240 min. Peaks marked with a star correspond to isomers of triticonazole.



Supplementary Information III-36: Detailed results of *in silico* tests

Compound	Mutagenicity (predicted value)	Developmental toxicity	Developmental toxicity predicted value	Oral rat LD50 (mg/kg)	Daphnia magna LC50 48 hours (mg/L)
acetamidiprid	Positive (0.92)	Non-toxicant	0.49	678.93	6.80
A	Positive (0.66)	Toxicant	0.96	1269.70	93.04
B	Positive (0.64)	Non-toxicant	0.48	983.94	43.06
PP1	Negative (0.22)	Toxicant	0.93	2156.19	43.41
PP2a	Negative (0.43)	Toxicant	1.03	444.59	39.21
PP2b	Positive (0.55)	Toxicant	0.80	633.11	NC
PP2c	Positive (0.52)	Toxicant	0.80	573.52	NC
PP3	Positive (0.56)	Toxicant	1.10	1168.02	NC
PP4a	Negative (0.31)	Toxicant	0.87	972.57	23.97
PP4b	Negative (0.05)	Toxicant	0.76	1237.94	NC
PP5	Positive (0.68)	Toxicant	0.97	717.31	7.34

III.3.2. Synthèse des principaux résultats de cette étude

Lors de cette étude de photodégradation de l'acétamipride, seule et au sein de mélanges commerciaux, 10 photoproduits ont été identifiés à l'aide d'analyses conduites par LC-FT-ICR-MS/MS et de l'utilisation d'un analogue deutéré de l'acétamipride. Un effet de matrice a pu être mis en évidence lors de la photolyse des mélanges commerciaux avec l'apparition d'un photoproduit (PP5) qui n'a pas été identifié lors de la photolyse de l'acétamipride seule. La formation de ce photoproduit est favorisée par la présence de groupements donneurs d'hydrogène au sein des mélanges commerciaux (par exemple les hydrocarbures permettant d'éviter au produit d'être rincé dès la première pluie). Des voies de dégradation ont pu être proposées sur la base des structures des photoproduits caractérisés.

Les tests *in silico* ont mis en évidence que, pour la plupart des photoproduits, la toxicité estimée pour le développement est deux fois plus élevée que celle de l'acétamipride. En revanche, les photoproduits ne s'avèrent pas plus toxiques que l'acétamipride en termes de LD₅₀ orale pour le rat et de mutagénicité. Les tests *in vitro* portant sur l'écotoxicité des échantillons d'acétamipride seul et en mélange commercial montrent que les solutions commerciales sont toutes deux plus toxiques pour l'environnement que l'acétamipride seul et que les toxicités de ces trois solutions augmentent légèrement après exposition à la lumière UV.

Des mécanismes de photodégradation tels que des ruptures homolytiques, des hydrogénations, des hydroxylations, des cyclisations, des éliminations ainsi que des réactions de Norrish ont pu être caractérisés lors de la dégradation par irradiation UV-Vis de l'acétamipride.

III.4. PHOTODEGRADATION UV-VISIBLE DU NAPROXENE DANS L'EAU – ELUCIDATION STRUCTURALE DES PHOTOPRODUITS ET DETERMINATION DE LEUR TOXICITE POTENTIELLE

III.4.1. Présentation des travaux publiés

Le naproxène est un anti-inflammatoire non stéroïdien (AINS) utilisé pour le traitement de douleurs faibles à modérées, fièvre ou inflammations.²⁶⁰ Il agit par inhibition d'une enzyme (COX ou cyclooxygénase) importante dans le mécanisme de l'inflammation et de la douleur. Le naproxène n'est pas spécifique et inhibe aussi la COX1 qui est une enzyme essentielle de la protection gastrique, induisant ainsi un risque d'ulcère à l'estomac. Le naproxène fait partie des AINS les plus utilisés et détectés dans le monde. Les AINS se retrouvent excrétés par l'homme en forte quantité dans les eaux de station d'épuration par le biais des excréments urinaires et fécales humaines, de rejets industriels mais aussi d'effluents hospitaliers.²⁶¹⁻²⁶³ Comme d'autres résidus de médicaments et leurs métabolites, le naproxène est régulièrement identifié dans les eaux de surfaces, les eaux usées, les sols, etc.^{264,265} Des études portant sur l'élimination de ces anti-inflammatoires non stéroïdiens dans les stations d'épurations montrent que l'efficacité de ces dernières n'est pas complète et que ces composés peuvent se retrouver dans les boues de station d'épuration ainsi que dans les effluents.^{260,262} Le naproxène, comme une grande partie des anti-inflammatoires est considéré comme fortement toxique pour les écosystèmes aquatiques.²⁶⁶ Une étude a démontré que les produits de dégradation du naproxène peuvent avoir des effets néfastes sur l'environnement et les écosystèmes aquatique ; certains seraient même plus toxiques que la molécule mère.²⁶⁷ Des études sur ce composé ont fourni des valeurs de LD₅₀ pour l'homme allant de 10 à 1000 mg/L.^{265,268} Il fait donc partie de la liste prioritaire des contaminants émergents définie par le réseau NORMAN.²⁶⁹ N'étant pas correctement éliminé par les traitements communs, le naproxène se retrouve en quantité élevée dans les eaux de surface et où il est soumis à l'irradiation solaire.¹⁷⁵

L'objectif de ce travail était donc de caractériser par LC-HR-MS/MS et GC-MS/MS les produits de photodégradation du naproxène en milieu aqueux. Les mécanismes de dégradation par irradiation UV des photoproduits ainsi identifiés ont été postulés. Les toxicités *in silico* de chacun de ces photoproduits ont été estimées grâce au logiciel T.E.S.T. L'écotoxicité de la solution avant et après photolyse a été déterminée par l'utilisation de tests *in vitro* sur la bactérie marine *Vibrio fischeri*, prenant en compte de potentiels effets de mélange. Les valeurs de toxicité déterminées pour les photoproduits, comparées à celles du naproxène, ont permis d'étudier la pertinence de la photodégradation par irradiation UV pour l'élimination du naproxène dans les effluents, notamment dans les effluents hospitaliers.

Les résultats ont été publiés dans la revue *European Journal of Mass Spectrometry* ; l'article correspondant est présenté ci-après.⁵⁷

UV-VISIBLE PHOTODEGRADATION OF NAPROXEN IN WATER – STRUCTURAL ELUCIDATION OF PHOTOPRODUCTS AND POTENTIAL TOXICITY

Noemi Cazzaniga, Zsuzsanna Varga, Edith Nicol, Stéphane Bouchonnet*

Laboratoire de Chimie Moléculaire – CNRS/Ecole polytechnique, Institut Polytechnique de
Paris, Palaiseau, France

Keywords:

Naproxen, photodegradation, structural elucidation, toxicity, mass spectrometry

European Journal of Mass Spectrometry 2020, Vol. 26(6) 400–408

<https://doi.org/10.1177/1469066720973412>

ABSTRACT

The UV-visible photodegradation of Naproxen (6-methoxy- α -methyl-2-naphthaleneacetic acid, CAS: 22204-53-1), one of the most used and detected non-steroidal anti-inflammatory drugs (NSAIDs) in the world, and its ecotoxicological consequences were investigated in an aqueous medium. The photo-transformation products were analyzed and the structures of photoproducts were elucidated using gas chromatography coupled with tandem mass spectrometry (GC-MS/MS) and high-performance liquid chromatography coupled with ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry (LC-FTICR-MS). Seven photoproducts were detected and characterized, photo-transformation mechanisms have been postulated to rationalize their formation under irradiation. *In silico* Q.S.A.R. (Quantitative Structure-Activity Relationship) toxicity predictions were performed with the Toxicity Estimation Software Tool (T.E.S.T.) and *in vitro* assays were carried out on *Vibrio fischeri* bacteria. Some of the obtained photoproducts exhibit higher potential toxicity than Naproxen itself but the whole toxicity of the irradiated solution is not of major concern.

INTRODUCTION

As the human body does not break them down entirely, pharmaceuticals and their metabolites represent a significant source of environmental contamination and their ever-increasing consumption is a major issue for the environment and human safety. The volume of medicines used globally will reach 4.5 trillion doses by the end of 2020 on a 1.4 trillion dollar market.²⁷⁰ Frequent use and active consumption of pharmaceuticals led to increased levels of drug-related residues in the environment. More than 600 active pharmaceutical ingredients (APIs) or their metabolites and photo-degradation products have been identified, mainly in surface water and wastewater, but also in groundwater, soil, and other environmental matrices.²⁶⁴ High concentrations of pharmaceutical products reach wastewater treatment plants (WWTP), mainly through human urinary, fecal excretion, and industrial discharges, and a considerable variety of drugs is not

mineralized by conventional treatment processes. It is now turned out that hospital effluents are also significant sources of aquatic environmental pollution.²⁶³ Consequently, more and more research is devoted to design specific pilot plants suitable for the degradation of biocides and drugs in hospital effluents, aiming at reducing the pharmaceutical burden of wastewater treatment plants and the associated risk.

In the past years, different prioritization approaches, based on consumption data, experimental toxicological and physicochemical data, and *in silico* predictions, were carried out to screen and identify the most environmentally hazardous pharmaceuticals. A recent study using a QSAR (Quantitative Structure-Activity Relationship) approach has shown that pharmaceutical products constitute a new generation of contaminants of emerging concern that can affect wildlife and the ecosystem even at low concentrations.²⁷¹ The majority (83%) of the drugs screened were predicted to be non-PBT (Persistent, Bioaccumulative, Toxic) and therefore, may not be considered dangerous for the ecosystem. 35 drugs have been identified as potential PBTs, for which an experimental evaluation will have to be made to confirm the expected PBT behavior.^{272,273}

Contrary to chronic toxicity, the acute toxicity of nonsteroidal anti-inflammatory drugs (NSAIDs) has not been widely studied, except in the case of propranolol, which shows high acute toxicity. 17% of the medicines exhibited human LC₅₀ values lower than 100 mg/L, which according to EU Directive 93/67/EEC²⁷⁴ correspond to acute toxicity, 38% exhibited LC₅₀ values above 100 mg/L and were thus classified as not harmful to aquatic organisms. The remaining 45% showed a considerable variability of acute toxicity values depending on the way they were evaluated. In the case of Naproxen, the provided EC₅₀ values fall in a range as wide as 10-1000 mg/L.^{265,268}

According to statistics conducted in 2017 on clinical drugs, Naproxen was the 71st most prescribed medication in the United States, with more than eleven million prescriptions. It belongs to the priority list of substances defined by the Global Water Research Coalition (GWRC, 2008);²⁷⁵ it is also on the list of contaminants of emerging concern compiled by the NORMAN network, specialized in monitoring and biomonitoring of emerging environmental substances.²⁶⁹ Naproxen is one of the most commonly detected pharmaceutical compounds in the environment, with concentrations in surface waters up

to 32 µg/L.²⁶⁵ It is mainly excreted in urine (95% of the quantity ingested) and is not effectively removed by current wastewater treatments. A Naproxen concentration of 6 mg/L has been reported in raw urban wastewater and effluent from a system of activated sludge, a common treatment adopted for urban wastewaters prior to final discharge into surface water bodies.¹⁷⁵ In another study conducted on raw urban wastewater and effluent from an activated sludge system, Ibuprofen was detected at the highest concentration (373 µg/L), followed by Naproxen (53 µg/L). No significant reduction was found for Ibuprofen and Naproxen in the pre-treatment and sedimentation step; because of their acidic structures and their very low solid-liquid partition K_d coefficients, both molecules undergo very poor sorption onto sludge and mainly remain in the aqueous phase.²⁷⁶

It has been demonstrated that Naproxen, by itself or through its metabolites, can affect organisms that inhabit ecosystems and one of the most important aspects of studies on the degradation of pollutants is the toxicity of degradation products, which in some cases can be higher than that of the parent molecule.²⁶⁷ The main objective of the present work was to characterize the photoproducts generated from direct photolysis of Naproxen, based on GC-MS and LCMS investigations. Ultra-high-resolution measurements permitted direct assignment of exact formulae while tandem experiments allowed structural elucidation for the major part of photoproducts. The second aim of this work was to estimate the potential toxicity of the photoproducts using both *in silico* (QSAR calculations) and *in vitro* (luminescence inhibition tests on *Vibrio fischeri* marine bacteria) approaches, to determine whether UV-visible photodegradation could be an efficient solution for the safe removal of Naproxen, for instance from hospital effluents.

EXPERIMENTAL

Chemicals and reagents

Naproxen (6-methoxy-a-methyl-2-naphthaleneaceticacid, CAS: 22204-53-1) was provided by Sigma Aldrich (Madrid, Spain). Acetonitrile (ACN), methanol and dichloromethane (Sigma Aldrich, Saint Quentin Fallavier, France) were of chromatographic grade, i.e. with

purity > 99.99%. BSFTA (N,O-Bis(trimethylsilyl)trifluoroacetamide) and formic acid were also from Sigma Aldrich (Saint Quentin Fallavier, France). Ultra-pure water (conductance of 18.2 S/cm) was obtained purifying tap water with a Purelab Chorus 1 system (Veolia, Wissous, France).

Sample preparation

Naproxen solutions were prepared at four concentrations: 2, 10, 100, and 500 ppm in water to which 0.1% acetonitrile was added regarding the poor solubility of Naproxen in pure water close to neutral pH (around 4mg/L at 25 °C according to the Henderson-Hasselbach equation). Different irradiation times were tested, to determine the appearance order of photoproducts. The solutions were irradiated up to a maximum of 6 hours with sampling at t 0-2-5-10-15-30-45-60-80-100-120 minutes, then every thirty minutes up to the fourth hour and every hour until the end of the analysis. For GC-MS analysis, all samples were dried using an Xcelvap automated evaporation/concentration system (Biotage, Uppsala, Sweden) for an hour, then put in an oven (WTC Binder 7200, Tuttlingen, Germany) for 10 minutes at 80 °C to remove any trace of water. Silylation was achieved adding 100 mL of BSTFA (N,O-Bis (trimethylsilyl)trifluoroacetamide) and placing the sample in the oven for 30 minutes. 100 µL of dichloromethane (DCM) was added prior to GC-MS analysis to reconstitute the sample. For LC-MS coupling measurements, samples were dissolved in ultrapure water to reach a final concentration of 2 ppm for nondegraded Naproxen at t₀. For HRMS analysis in direct infusion mode, samples were diluted in a mixture of H₂O/ACN/FA 90/10/0.1%, to reach a 2 ppm concentration for Naproxen at t₀. All samples were stored sheltered from light at 4 °C.

Photolysis experiments

Photodegradation experiments were conducted using a home-made photoreactor equipped with a high-pressure mercury lamp HPL-N 125 W/542 E27 SC (Phillips, Ivry-sur-Seine, France) that emits light on wavelengths between 200nm and 650 nm, with maximum irradiation at 254nm and a radiation flow of 6200 lm. The lamp is located in a quartz tube at the center of an AL04-12 sonicator (Advantage-Lab GmbH, Darmstadt, Germany) and samples are placed in quartz tubes around, at equal distance to the lamp.

The sonicator is used to achieve the solubilisation of compounds before irradiation. To avoid thermal degradation, the temperature was manually adjusted to 30 ± 5 °C flowing cold-water through the sonicator manifold.

GC-MS operating conditions

Gas chromatographic separations were carried out on an Agilent 7890B GC instrument equipped with a 30m HP-5MS column (5% Phenylmethylpolysiloxane) with an internal diameter of 0.25mm and film thickness of 0.25 μ m, coupled with an Agilent 7000D GC/TQ triple quadrupole mass spectrometer (Agilent Technologies, Les Ulis, France). Samples were injected in splitless mode using an Agilent 7693 A autosampler, the injection volume was 1 μ L and injector temperature was 280 °C. A gradient temperature program was used with an initial temperature of 50 °C, held for 0.5 min then increased to 320 °C with a ramp of 20 °C/min. The carrier gas was high purity helium, with an autoregulated flow of 1.2 ml/min. Measurements were carried out in electron ionization (70 eV) with a filament emission current of 35 μ A. The electron multiplier voltage was automatically optimized at 1090 V, for a gain value of 10^5 . In the full scan mode, ions were scanned from m/z 50 to m/z 600. Tandem experiment measurements were performed, using argon as the collision gas, with collision energies of 15, 30 and 45 eV. The source temperature was set to 230 °C.

LC-MS operating conditions

Chromatographic separations were carried out using an Acquity HPLC system (Waters Technologies, Guyancourt, France) equipped with a C₁₈ Pursuit XRs Ultra 2.8 μ m, 50 x 2.0mm column (Agilent Technologies, Les Ulis, France) and coupled with an FT-ICR SolarixXR 9.4 T (Bruker Daltonics, Bremen, Germany). The injection volume was 10 μ L, elution was performed using a 0.1 mL/min flow with a mixture of solvents A (water, formic acid 0.1%) and B (acetonitrile, formic acid 0.1%). After 10 minutes, the percentage of solvent B was increased from an initial value of 5% to 40% in 15 min, then to 90% in 5min. The gradient was then set back to initial conditions (5% of B) during the last 20 min, for a total method duration of 50 min. An electrospray ionization source was used both in positive and negative mode, the nebulizer and drying gas was nitrogen with a flow of 8

L/min at 250 °C for drying gas and 1 bar for nebulizer gas. The capillary voltage and end plate offset values were set to - 4500V and - 4000 V, respectively for positive mode, and 3000V and 2500 V, respectively for negative mode. Argon was used as the collision gas for tandem experiments. Ions were accumulated in the collision cell for 0.2 s in MS experiments and 2 s in MS/MS experiments. The ion detection mode was set to 4 Mpt with a 0.8389 s transient duration in broadband mode from m/z 57.7 to m/z 1000 to obtain a resolution higher than 350 000 at m/z 200. To reduce the size of the data file, 97% of the data (contributing to noise) were not recorded to get a data file size below 2 GB per analysis. Phosphoric acid was used for external mass calibration and exact chemical formulae were attributed with tolerance below 3 ppm. For tandem MS experiments, precursor ions were isolated with an isolation window of 1 m/z and dissociated, using argon as the collision gas, with collision energies of 10, 15, 20 and 25V to study the fragmentation mechanisms.

In silico toxicity estimations

The toxicities of Naproxen and its photoproducts were *in silico* evaluated regarding developmental toxicity, oral rat LD₅₀ (mg/kg), and Fathead minnow LC₅₀ (96 h) (mg/L). The Toxicity Estimation Software Tool (T.E.S.T.) - based on Quantitative Structure-Activity Relationship (QSAR) mathematical models - was used for this purpose.¹³³ T.E.S.T. has been developed by the U.S. Environmental Protection Agency, its main features are briefly described in Supplementary Information III-37.

In vitro bioassays

Vibrio fischeri commercial *in vitro* test kit was used to evaluate the global ecotoxicity of solutions before (initial concentration of 50 µg/L in Naproxen) and after two times of irradiation: 30 and 80 min. The freeze-dried luminescent bacteria and the luminometer were purchased from Hach Lange (GmbH, Düsseldorf, Germany). The experimental method used in this study is based on the ISO 11348-3 protocol.²⁷⁷ Luminescence was measured in duplicate before the addition of the test solution and after contact times of 5, 15 and 30 min; results were corrected using the values measured with control samples.

The effective nominal concentration leading to 20% inhibition of bioluminescence is designated as the EC₂₀.

RESULTS AND DISCUSSION

Kinetics of photoproducts and degradation efficiency

The relative concentration of Naproxen in ultrapure water was followed by LC-MS over the irradiation time (6 hours), integrating the area of the corresponding chromatographic peak on the [M+H]⁺ ion. The initial amount of Naproxen (initial concentration of 2 ppm) was halved after about 30 minutes and the compound was no longer detected after 45 minutes. In order to correlate the appearance of some photoproducts with the disappearance of some others, their relative abundances (normalized data from GC-MS and LC-MS signals) were plotted as a function of irradiation time (Figure III-17). Photoproducts are referred as PP1 to PP7; their structures are given in Table III-12 and Table III-13.

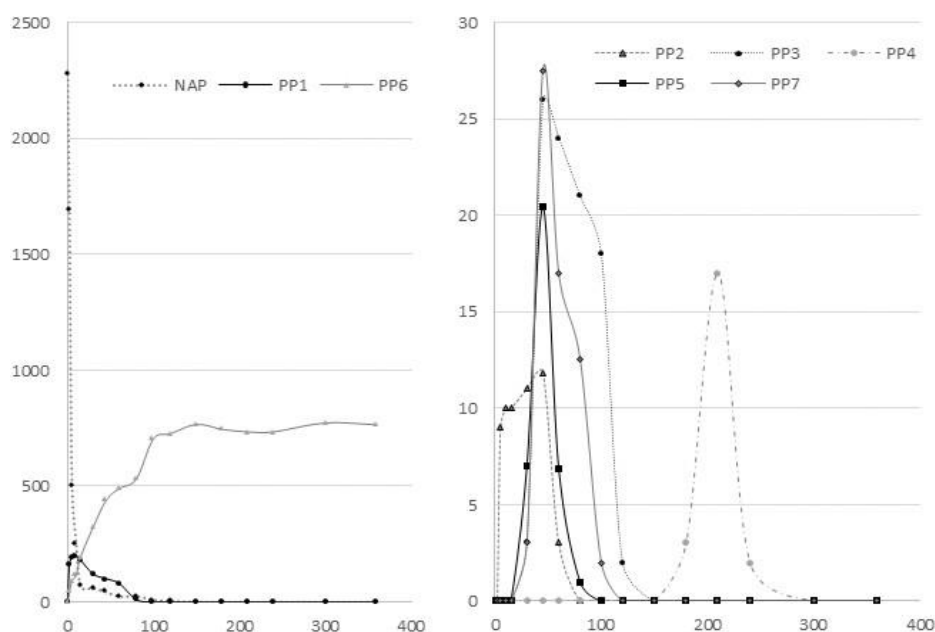
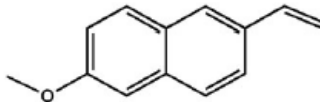
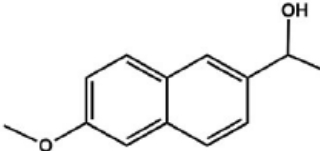
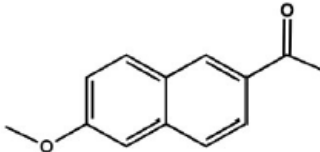
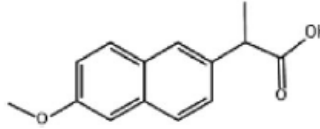


Figure III-17: Relative abundances of Naproxen and its photoproducts over time: compounds of major abundances (NAP, PP1 and PP6) on left side and compounds of minor abundances (PP2, PP3, PP4, PP5 and PP7) on right side. On the y-axis, abundances are normalized to that of Naproxen at t₀. The x-axis corresponds to the irradiation time in minutes.

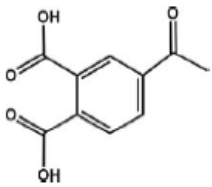
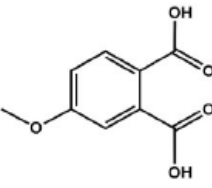
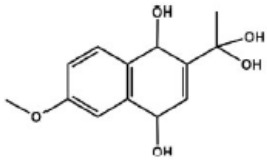
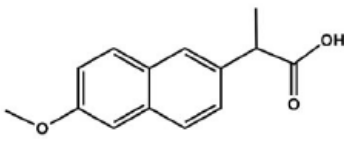
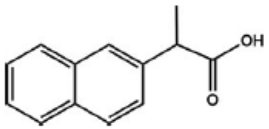
Table III-12: Retention times, main ions, and elucidated structures of the photoproducts **PP1** to **PP3** detected in GC-MS/MS (electron ionization).

Ret. time (min)	Compound	Silylated	Main ions in MS/MS (m/z) – precursor in bold ^a	Photoproduct
9.08	PP1	No	184 , 169, 141, 115	
10.12	PP2	Yes	274 , 259, 185, 141, 73	
10.18	PP3	No	200 , 185, 158, 157, 128	
11.00	Naproxen	Yes	302 , 287, 243, 185, 73	

^aNote that the provided m/z values correspond to ions from the silylated structure in the case of derivatized compounds, i.e. PP2 and Naproxen. For both compounds, the molecular mass of the photoproduct and thus that of the molecular ion are shifted by 72 amu since the OH function has been replaced by OSi(CH₃)₃.

To compare the relative amounts estimated by GC-MS and LC-MS couplings, LC-MS peak areas were corrected by a factor determined on the basis of the peak areas of Naproxen detected by both techniques at t₀. This empirical approach aimed at making possible a rough comparison between ion abundances from both techniques; it did not serve a quantification purpose, but allowed to discriminate reaction intermediates from photoproducts likely to be found under real sunlight degradation. Seven photoproducts were determined. After a few minutes of irradiation, PP6 is by far the most abundant before PP1. Furthermore, PP6 appears as the only persistent compound at the end of the experiment while all the other photoproducts appear as reaction intermediates. Correlating their relative appearance and disappearance order helped to assign their chemical structures (see section 3.2).

Table III-13: Retention times, main ions, and elucidated structures of the photoproducts **PP4** to **PP7** detected by LC-ESI-MS.

Ret. time (min)	Compound	Detection in ESI+		Detection in ESI-		Structure of the compound
		MH ⁺ ion (m/z)	MH ⁺ formula	MH ⁻ ion (m/z)	MH ⁻ formula	
7.51	PP4	209.04502	C ₁₀ H ₉ O ₅ ⁺	207.02990	C ₁₀ H ₇ O ₅ ⁻	
8.26	PP5	197.04475	C ₉ H ₉ O ₅ ⁺	195.03025	C ₉ H ₇ O ₅ ⁻	
8.72	PP6	253.10742	C ₁₃ H ₁₇ O ₅ ⁺	251.09250	C ₁₃ H ₁₅ O ₅ ⁻	
13.00	Naproxen	231.10198	C ₁₄ H ₁₅ O ₃ ⁺	-	-	
13.14	PP7	201.09152	C ₁₃ H ₁₃ O ₂ ⁺	-	-	

Structural elucidation of photoproducts

GC-MS experiments allowed the structural elucidation of photoproducts PP1 to PP3; retention times, ions, and chemical structures are given in Table III-12. Dissociations of molecular ions were interpreted; the silylated compounds were easily identified due to the presence of a fairly intense ion at m/z 73 ($\text{Si}(\text{CH}_3)_3^+$) in their mass spectrum. Dissociation mechanisms of molecular ions are provided in Supplementary Information III-38.

The retention times and m/z measured for protonated and deprotonated species in LC-ESI-MS are given in Table III-13. Collision induced experiments permitted the structural

elucidation of the photoproducts. Transitions and associated mechanisms are provided in Supplementary Information III-39 and Supplementary Information III-40.

Based on the chemical structures of photoproducts, a photodegradation mechanism of naproxen has been proposed, which involves classical photochemistry pathways such as Norrish-type cleavage, hydrogen abstraction, oxygen addition (from water dissolved oxygen), hydrolysis, and concerted eliminations (see Figure III-18). As is often the case with ketones, the photodegradation process begins by the Norrish-type cleavage of the C-CO bond of naproxen, leading to a radical which, in contact with a water molecule, can abstract a hydrogen atom or a hydroxyl group to lead to PP1 and PP2, respectively. Photo-induced cleavages of C-H bonds have been recently demonstrated in a study by De Vaugelade *et al.* reporting the UV-Vis photodegradation of alpha-tocopherol.²²⁹ Cleavage of such bonds is only observed if the resulting radical is strongly stabilized by mesomeric and/or inductive effects, which is the case when removing the hydrogen atom carried by the same carbon atom than the hydroxyl group in PP2. The resulting radical can lose a second hydrogen atom to lead to the ketone PP3. Successive oxygen additions on PP3, combined with hydrolysis reactions lead to PP6, then to PP5. By analogous mechanisms, the consecutive oxidation and hydrolysis of PP3 leads to PP4. Quite surprisingly, PP7 results from methanol elimination from naproxen, followed by hydrogen addition, the hydrogen atom is abstracted from a water molecule. This type of mechanism is not “classical” in photochemistry but both the structure of naproxen and the MS dissociation pathways of PP7H⁺ allow unambiguous identification of the latter.

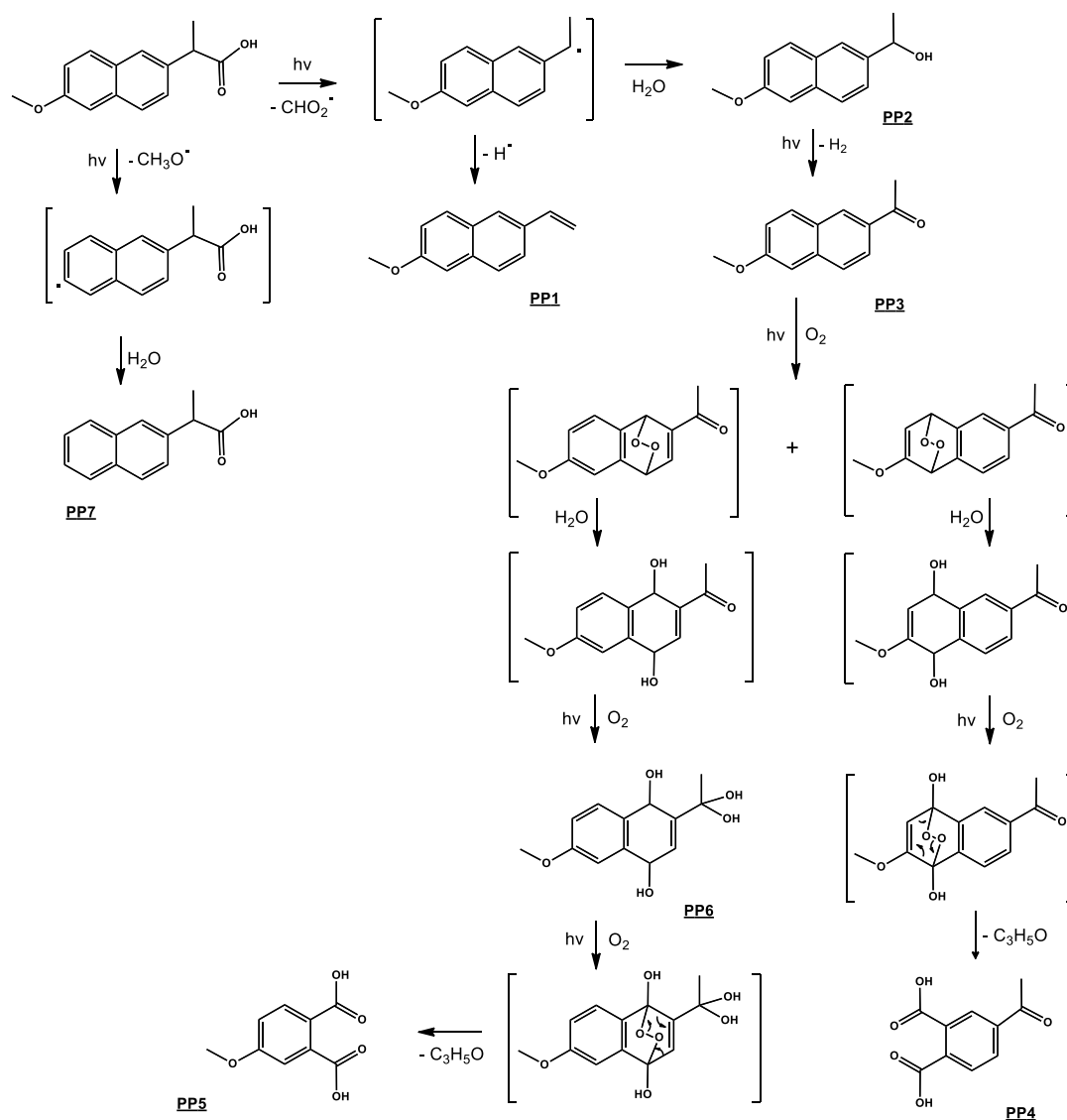


Figure III-18: Mechanisms proposed for the photodegradation of Naproxen in water

In silico and in vitro toxicity estimations

In silico toxicity estimation.

QSAR *in silico* tests were carried out with the T.E.S.T. software (see Supplementary Information III-37) regarding development toxicity, oral rat LD_{50} (mg/kg), and Fathead minnow LC_{50} (96 hours) (mg/L); results are presented in Table III-14.

Table III-14: Toxicity values estimated by the T.E.S.T. software for Naproxen and its photoproducts.

Compound	Fathead		Developmental toxicity value ^a	Developmental toxicity
	Minnow LC50 (96 h) (mg/L)	oral rat LD50 (mg/kg)		
NAP	5.3	515.5	0.89	Toxicant
PP1	0.2	667.8	0.53	Toxicant
PP2	5.5	1817.9	0.85	Toxicant
PP3	6.6	1594.0	0.51	Toxicant
PP4	22.8	4230.5	0.37	Non-toxicant
PP5	69.9	3913.9	0.69	Toxicant
PP6	19.4	1661.8	0.52	Toxicant
PP7	6.8	311.7	0.96	Toxicant

^aArbitrary unit.

Considering developmental toxicity, all compounds except PP4 are potentially toxic, the latter having a toxicity value coefficient of 0.37, which is below the threshold of toxicity value (0.50); they are expected to be at least as toxic as Naproxen. Regarding oral rat toxicity, the values reported in Table III-14 show that all photoproducts are in category III (slightly toxic) at the exception of except PP7 which belongs to category II corresponding to moderate toxicity. The higher toxicity of PP7 compared to other photoproducts is quite close to that of Naproxen; this is consistent with the fact that the chemical structure of PP7 is the closest to that of the parent molecule. The lowest LC₅₀ values were determined for PP5 and PP6, which exhibit the most hydrophilic structures and thus the lowest potentials for bioaccumulation. An oral rat LD₅₀ experimental value of 248 mg/kg was provided for Naproxen, which is comparable but below the estimated one (516 mg/kg).²⁷⁸ Nevertheless, it suggests that the *in silico* approach used could underestimate the toxicity of Naproxen and its photoproducts due to the lack of experimental data on similar structures. Regarding the toxic effect on Fathead minnow, PP1 exhibits by far the highest potential toxicity, in agreement with previous results.²⁷⁹ PP2, PP3, and PP7 provide LC₅₀ values close to that of the parent compound. The specificity of PP1 is that it does not include any CO nor OH group, which makes it much less polar than other compounds. Here again, the most hydrophilic compounds exhibit the lowest potential toxicity. In a previous study on UV photodegradation of Naproxen by Cory *et al.*, only PP2 and PP3 were detected though the experimental conditions were quite similar to those used in the present work.²⁸⁰ Like in the present study, both phototransformation products of

Naproxen were found to be more stable than Naproxen itself when exposed to simulated sunlight; their toxicity was investigated towards toad tadpoles and they have been shown to be much more toxic than the parent molecule.

In vitro toxicity estimation.

In vitro bioassays based on the bioluminescence extinction of *Vibrio fischeri* were carried out on the solution of Naproxen at different irradiation times according to the protocol described in section 1.7. Unlike *in silico* testing, this approach takes into account potential mixture effects. In the above-mentioned protocol, estimation of an EC₅₀ value requires a minimum of three measurements for which the inhibition is between 10% and 90%, with a value above 50% for at least one of the three measurements. A Naproxen solution at 50 ppm was submitted to photodegradation; sampling was performed at t₀, t₃₀, and t₈₀ (minutes). For each sample, the conditions were not met for EC₅₀ determination, indicating a low ecotoxicity; an EC₂₀ was thus evaluated, based on linear regression analysis of the dose-effect relationship. For a duration exposure of 15 minutes, the EC₂₀ value decreases when increasing the irradiation time, from 11.7 at t₀ to 7.2 at 80 minutes, which traduces a slight increase in toxicity of the mixture. As shown in Figure III-17, PP6 is by far the most abundant photoproduct in the mixture at 80 minutes. Regarding *in silico* estimations, PP6 exhibits Fathead Minnow LC₅₀ and oral rat LD₅₀ values higher than those estimated for Naproxen so this photoproduct is not a priori expected to contribute to the increase in toxicity of the mixture. Different conclusions can be drawn from these estimations: (i) insufficient parametrization of the T.E.S.T. software regarding the molecules of interest makes the *in silico* estimations inaccurate in the present case, especially in the case of PP6, which displays a quite complex and unusual structure; (ii) with a Fathead Minnow LC₅₀ estimated to be 25 times lower than that of Naproxen (see Table III-14), PP1, which is the second photoproduct in terms of abundance, could be responsible for the slight increase in toxicity of the mixture; (iii) a limited “cocktail effect” due to the concomitant presence of several photoproducts could also be suspected. Whatever the right explanation, both *in silico* and *in vitro* evaluations lead to the outcome that the ecotoxicity induced by Naproxen photoproducts is not a major issue compared to that of Naproxen itself. However, the long-term environmental effects of Naproxen exposure at concentration levels close to those found in the environment remain poorly

studied although detection of the drug in wastewater and river samples in the ng/L - µg/L range demonstrated the need for such knowledge.²⁸¹⁻²⁸³

CONCLUSION

The in-lab photodegradation of Naproxen under UV-Visible irradiation led to the formation of seven photoproducts, which were characterized using LC-HR-MS/MS and GC-MS techniques. A photodegradation pathway of Naproxen has been suggested based on the elucidated structures and kinetics considerations. *In silico* and *in vitro* tests were conducted, on isolated molecules in the first case, on an irradiated aqueous solution of Naproxen in the second one. The general conclusion is that photoproducts do not increase the global toxicity in an alarming way. It is interesting to note that Naproxen was 90% degraded after only a few minutes and that all the photoproducts are eliminated after 5 hours of irradiation with the exception of PP6, which remains in the mixture in significant amounts and whose toxicity would deserve to be further investigated after synthesis. Aside from the consequences of sunlight-induced degradation of Naproxen in the environment, these results suggest that direct UV light irradiation – a technique gradually replacing the more classical photocatalytic processes – could be an efficient tool for Naproxen removal from contaminated waters.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

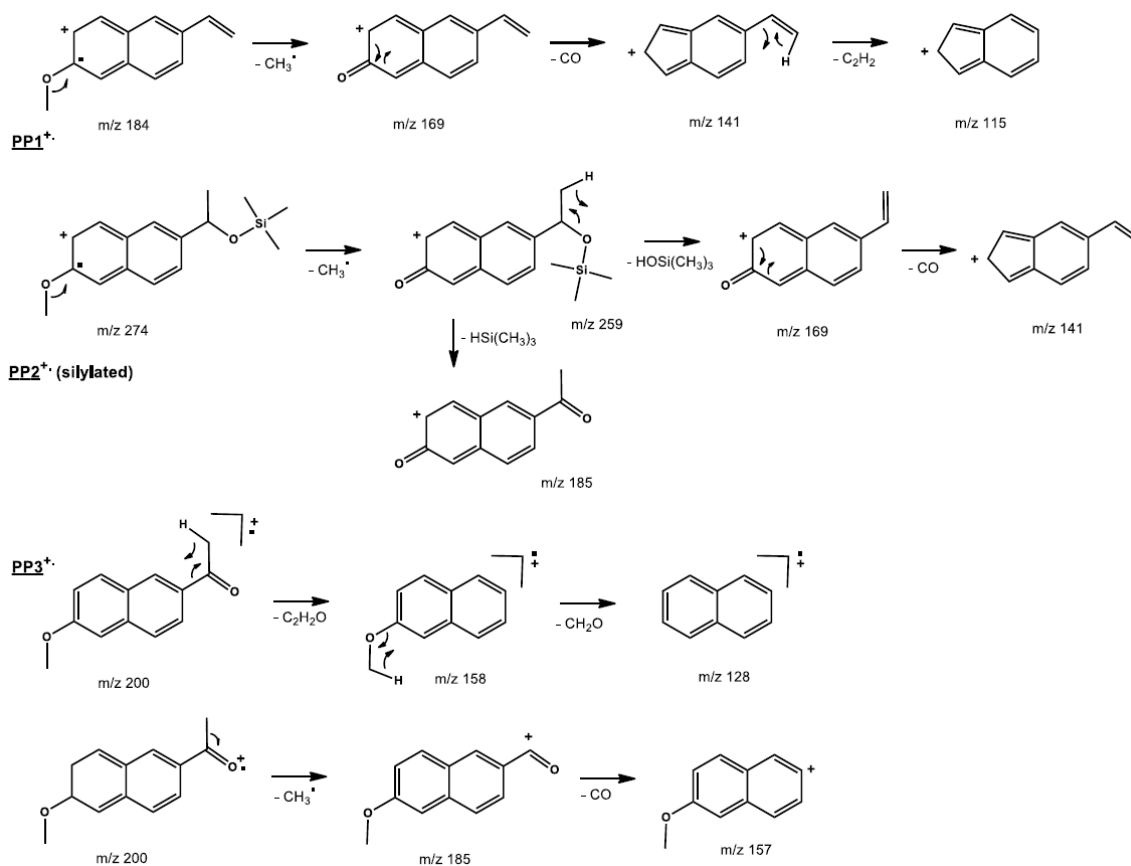
The author(s) disclosed the receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support from the National FT-ICR network (FR 3624 CNRS) for conducting the research is gratefully acknowledged. This

work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No 765860 (AQUALity).

SUPPORTING INFORMATION

Supplementary Information III-37: In silico toxicity estimations using the T.E.S.T. software

The Toxicity Estimation Software Tool (T.E.S.T.) is an Environmental Protection Agency online available computerized predictive system with Quantitative Structure Activity Relationships (QSAR) mathematical model.²⁵³ T.E.S.T. has a variety of toxicity endpoints used to predict chemical toxicity values from the physical properties of the molecular structure. T.E.S.T. model uses a simple linear function of molecular descriptors such as the octanol-water partition coefficient, the molecular weight or the number of benzene rings. Models for assessing toxicity solely from molecular structure are based on information-rich structural descriptors that quantify transport, bulk, and electronic attributes of a chemical structure. Besides molecular weight, the QSAR model employs size-corrected E-values for quantification of molecular bulk. The size-corrected E-values are computed from a rescaled count of valence electrons. Electrotopological state values (E-values) as numerical quantifiers of molecular structure encode information about the electron content (valence, sigma, pi and lone-pair), topology and environment of an atom or a group of atoms in a molecule.²⁸⁴ The predicted toxicity is estimated by taking an average of the predicted toxicities from the above QSAR methods, provided the predictions are within the respective applicability domains.

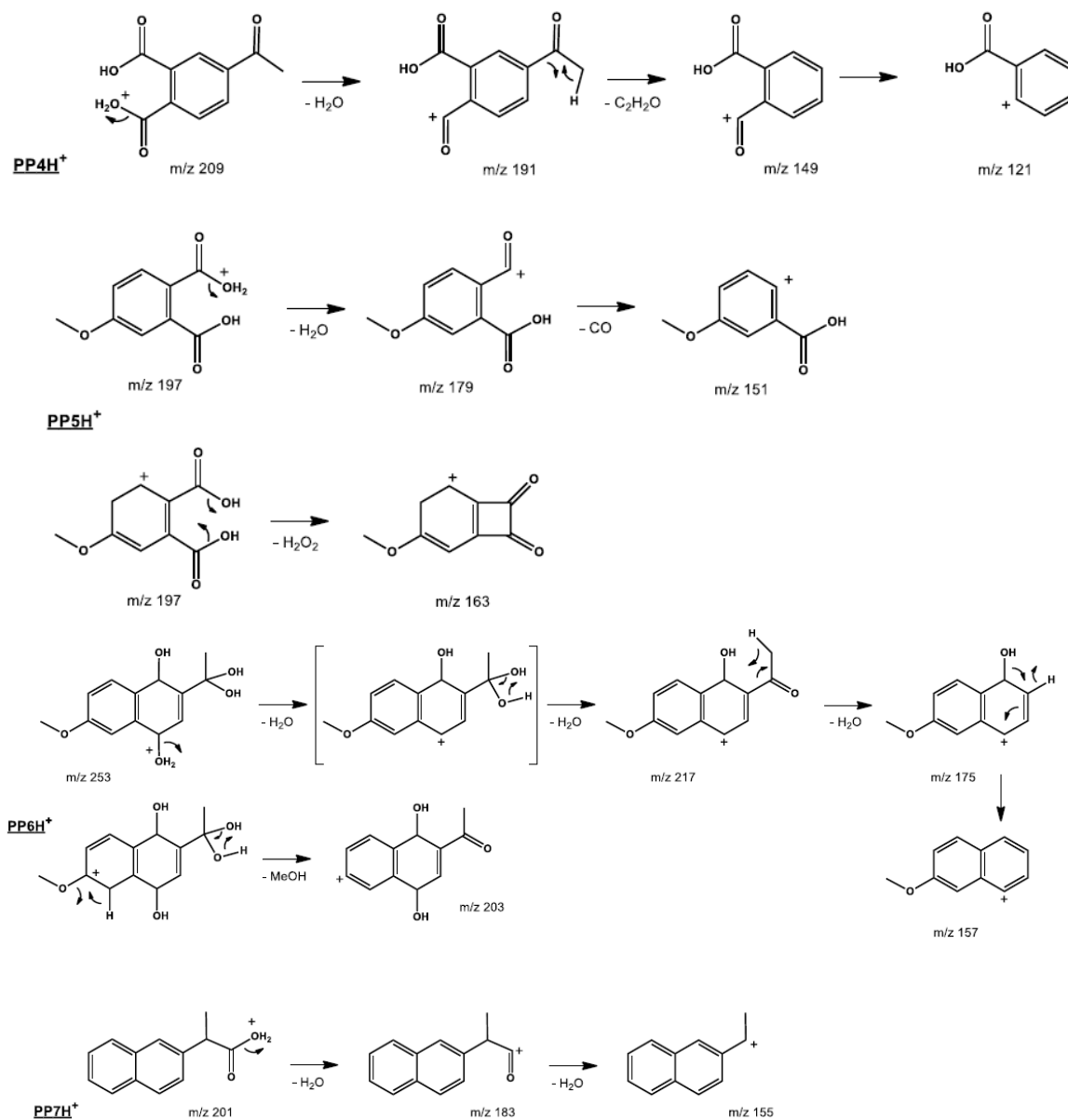


Supplementary Information III-38: Dissociations of molecular ions of the photoproducts detected by GC-EI-MS

Supplementary Information III-39: Collision-induced ESI⁺ mass spectra of naproxen and its photoproducts

Compound	MH ⁺ precursor ion		Fragment ions	
	m/z	Formula	m/z	Formula
PP4	209	C ₁₀ H ₉ O ₅ ⁺	191	C ₁₀ H ₇ O ₄ ⁺
			149	C ₈ H ₅ O ₃ ⁺
			121	C ₇ H ₅ O ₂ ⁺
PP5	197	C ₉ H ₉ O ₅ ⁺	179	C ₉ H ₇ O ₄ ⁺
			163	C ₉ H ₇ O ₃ ⁺
			151	C ₈ H ₇ O ₃ ⁺
PP6	253	C ₁₃ H ₁₇ O ₅	217	C ₁₃ H ₁₃ O ₃ ⁺
			203	C ₁₂ H ₁₁ O ₃ ⁺
			175	C ₁₁ H ₁₁ O ₂ ⁺
			157	C ₁₁ H ₉ O ⁺
PP7	201	C ₁₃ H ₁₃ O ₂ ⁺	183	C ₁₃ H ₁₁ O ⁺
			155	C ₁₂ H ₁₁ ⁺

CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS



Supplementary Information III-40: Dissociations of the pseudomolecular ions (MH^+) of the photoproducts detected by LC-ESI-MS

III.4.2. Synthèse des principaux résultats de cette étude

Cette étude de photodégradation du naproxène a conduit à l'identification et la caractérisation de 7 photoproduits par LC-HRMS/MS et la GC-MS/MS. Le naproxène est dégradé à 90% en moins d'une heure et tous les photoproduits sont éliminés après 5 h d'irradiation, à l'exception du produit PP6 qui est le plus abondant et semble être persistant.

Les tests *in silico* réalisés *via* le logiciel T.E.S.T. ont permis d'évaluer la toxicité des composés individuellement. Certains des photoproduits présentent une toxicité estimée plus élevée que le naproxène lui-même. Néanmoins, la LD₅₀ orale chez le rat est estimée par le logiciel T.E.S.T. à 516 mg/kg tandis que la valeur expérimentale disponible est de 248 mg/kg.²⁷⁸ Ces résultats suggèrent que l'approche *in silico* utilisée pourrait, en raison du manque de données expérimentales sur des structures proches, sous-estimer la toxicité du naproxène et de ces photoproduits. Des tests de toxicité *in vitro* sur l'extinction de bioluminescence de la bactérie marine *Vibrio fischeri* ont permis quant à eux d'observer que l'écotoxicité de la solution augmente légèrement avec le temps de photolyse. Le seul photoproduit encore présent (PP6) après photolyse possède une toxicité *in silico* estimée proche du naproxène, ne contribuant pas, *a priori*, à une augmentation de la toxicité du mélange. Différentes théories peuvent expliquer cette augmentation : (1) le logiciel T.E.S.T. pourrait ne pas être suffisamment paramétré pour les molécules d'intérêt, rendant les estimations *in silico* inexactes pour cette étude, en particulier pour la structure du produit PP6 dont la structure est complexe et inhabituelle ; (2) la LD₅₀ estimée pour le poisson *fathead minnow* étant 25 fois plus faible que celle du naproxène, cela pourrait expliquer la légère augmentation de toxicité si des traces de ce composé, non détectables avec les outils utilisés pour cette étude, persistent après photolyse ; (3) un effet cocktail de la présence - même à l'état de traces - de certains photoproduits pourrait être suspecté. Des études de toxicité individuelle *in vitro* du composé PP6 sont envisagées.

Les photoproduits ne semblent pas augmenter de façon alarmante la toxicité globale, suggérant par conséquent que la dégradation du naproxène par irradiation directe à la lumière UV-Visible pourrait être un outil efficace pour l'élimination de ce composé dans

des eaux contaminées, remplaçant ainsi les procédés photocatalytiques plus classiques et plus coûteux.

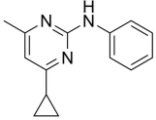
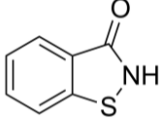
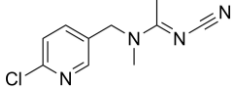
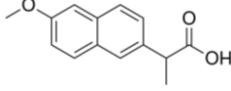
Sur la base des structures chimiques et de la cinétique de dégradation des différents photoproduits, des mécanismes de photodégradation ont pu être proposés répondant ainsi au second objectif du laboratoire. Les mécanismes classiques ont pu être identifiés, tels que des réactions de Norrish, des oxydations, des déshydrogénations mais aussi des hydrolyses et des éliminations concertées. Des clivages de liaisons C-H ont également été observés.

III.5. CONCLUSION SUR LES TRAVAUX PUBLIES

Lors de ces travaux, quatre polluants émergents ont été dégradés par photolyse UV-Visible. Les photoproduits ainsi obtenus ont été identifiés et caractérisés par spectrométrie de masse en tandem avec l'apport de la haute résolution et de différentes techniques chromatographiques. Les toxicités des photoproduits ainsi obtenus ont été estimées par une méthode *in silico* utilisant le logiciel T.E.S.T. et l'écotoxicité des mélanges avant et en cours de photolyse a été déterminée par des tests *in vitro* d'extinction de bioluminescence de la bactérie marine *Vibrio fischeri*. Les résultats, publiés dans des revues internationales, sont synthétisés dans le Tableau III-15.

CHAPITRE III : IDENTIFICATION DE PRODUITS DE PHOTODEGRADATION DE CECs PAR LC-MS

Tableau III-15 : Synthèse des résultats publiés

Molécules	Cyprodinil	Benzisothiazolinone	Acétamipride (et mélanges commerciaux)	Naproxène
Structures				
Outils d'analyse	LC-HRMS/MS	LC-HRMS/MS GC-MS/MS	LC-HRMS/MS	LC-HRMS/MS GC-MS/MS
Analogue deutéré	Oui cyprodinil-D ₅	Non Standards pour certains PP	Oui Acétamipride-D ₃	Non
Nombre de PP identifiés	19	14	10 (dont un uniquement présent dans les mélanges commerciaux)	7
Toxicités estimées <i>in silico</i> des PP	La plupart plus toxiques que la molécule mère	Certains plus toxiques que la molécule mère. (toxicité modérée)	La plupart de toxicité proches ou supérieure à celle de la molécule mère	La plupart proches de la toxicité de la molécule mère
Ecotoxicité du mélange au cours de la photolyse	Augmente significativement avec le temps de photolyse	Non effectué	Mélanges commerciaux plus toxiques que l'acétamipride. La toxicité des solutions augmente légèrement au cours de l'irradiation.	Augmente légèrement au cours de l'irradiation
Mécanismes de photolyse identifiés	- Ruptures directe de liaison - Cyclisations - Isomérisations - Hydroxylations	- Isomérisations - Oxydations - Hydroxylations - Hydrolyses - Eliminations	- Ruptures homolytiques - Hydrogénations - Hydroxylations - Cyclisations - Eliminations - Réactions de Norrish	- Réactions de Norrish - Oxydations - Déshydrogénations - Hydrolyses - Eliminations concertées - Clivages de liaison C-H

PP : Photoproduits

Peu de ces produits de dégradations étaient déjà connus et caractérisés avant ces études et certains d'entre eux s'avèrent être plus toxiques pour les êtres vivants et les écosystèmes que les molécules mères. Ceci démontre la nécessité de poursuivre ces

études et de les étendre au maximum de polluants d'intérêt émergent, le mieux étant d'étudier les différentes voies de dégradation avant la mise sur le marché d'un composé. Cela permettra d'ajouter ces produits de dégradation à ceux déjà répertoriés lors des analyses de routines pour déterminer le niveau de contamination d'un milieu ou d'un produit alimentaire. L'étude portant sur le naproxène a démontré les limitations de l'utilisation de logiciels *in silico* pour évaluer les toxicités des photoproduits avec une toxicité estimée deux fois moins importante que la toxicité déterminée expérimentalement. Cela peut être expliqué par un paramétrage insuffisant du logiciel T.E.S.T. pour certaines structures et cela doit conduire à une certaine vigilance quant à l'interprétation des résultats qui demeurent des estimations. Toutefois, ces valeurs de toxicités estimées restent rapides à obtenir, peu coûteuses et limitent les problèmes déontologiques. Les analyses des mélanges commerciaux ont montré l'intérêt de la prise en compte de l'effet matrice lors de la recherche de produits de dégradation. Cela concerne aussi bien à la matrice « commerciale » - comme dans le cas de l'étude portant sur la dégradation de l'acétamipride – que les matrices environnementales ou alimentaires.

Ces analyses en mélanges complexes restent difficiles à mettre en œuvre et l'interprétation des résultats est souvent longue et complexe. La quantité de matrices commerciales et environnementales disponibles ainsi que la grande variété de voies de dégradation connues rendent la tâche impossible pour une seule équipe. La subjectivité des utilisateurs, les problèmes de reproductibilité et de répétabilité sont également à considérer lors de l'identification de produits de dégradation de polluants émergents, comme nous allons le voir au chapitre suivant.

REFERENCES BIBLIOGRAPHIQUES

- (40) Boxall, A. B. A.; Sinclair, C. J.: Assessment of the environmental properties and effects of pesticide degradation products. In *Pesticide behaviour in soils and water. Symposium (Brighton 2001-11-13)*; BCPC, Farnham, 2001; pp 113-118.
- (41) Castillo, M.; Domingues, R.; Alpendurada, M. F.; Barcelo, D. Persistence of selected pesticides and their phenolic transformation products in natural waters using off-line liquid solid extraction followed by liquid chromatographic techniques. *Analytica Chimica Acta* **1997**, *353*, 133-142.
- (42) Bottaro, M.; Frascarolo, P.; Gosetti, F.; Mazzucco, E.; Gianotti, V.; Polati, S.; Pollici, E.; Piacentini, L.; Pavese, G.; Gennaro, M. C. Hydrolytic and photoinduced degradation of tribenuron methyl studied by HPLC-DAD-MS/MS. *Journal of the American Society for Mass Spectrometry* **2008**, *19*, 1221-1229.
- (43) Gosetti, F.; Bottaro, M.; Gianotti, V.; Mazzucco, E.; Frascarolo, P.; Zampieri, D.; Oliveri, C.; Viarengo, A.; Gennaro, M. C. Sun light degradation of 4-chloroaniline in waters and its effect on toxicity. A high performance liquid chromatography - Diode array - Tandem mass spectrometry study. *Environ Pollut* **2010**, *158*, 592-598.
- (44) Gosetti, F.; Chiuminatto, U.; Zampieri, D.; Mazzucco, E.; Marengo, E.; Gennaro, M. C. A new on-line solid phase extraction high performance liquid chromatography tandem mass spectrometry method to study the sun light photodegradation of mono-chloroanilines in river water. *Journal of chromatography. A* **2010**, *1217*, 3427-3434.
- (45) Gosetti, F.; Chiuminatto, U.; Mazzucco, E.; Mastroianni, R.; Bolfi, B.; Marengo, E. Ultra-high performance liquid chromatography tandem high-resolution mass spectrometry study of tricyclazole photodegradation products in water. *Environmental science and pollution research international* **2015**, *22*, 8288-8295.
- (46) Aguera, A.; Martinez Bueno, M. J.; Fernandez-Alba, A. R. New trends in the analytical determination of emerging contaminants and their transformation products in environmental waters. *Environmental science and pollution research international* **2013**, *20*, 3496-3515.
- (50) Burrows, H. D.; Canle, L. M.; Santaballa, J. A.; Steenken, S. Reaction pathways and mechanisms of photodegradation of pesticides. *Journal of photochemistry and photobiology. B, Biology* **2002**, *67*, 71-108.
- (51) Kiss, A.; Virag, D. Interpretation and modelling of environmental behaviour of diverse pesticides by revealing photodecomposition mechanisms. *Microchemical Journal* **2009**, *92*, 119-122.
- (57) Cazzaniga, N.; Varga, Z.; Nicol, E.; Bouchonnet, S. UV-visible photodegradation of naproxen in water - Structural elucidation of photoproducts and potential toxicity. *European journal of mass spectrometry (Chichester, England)* **2020**, *26*, 400-408.
- (128) Lassalle, Y.; Kinani, A.; Rifai, A.; Souissi, Y.; Clavaguera, C.; Bourcier, S.; Jaber, F.; Bouchonnet, S. UV-visible degradation of boscalid--structural characterization of photoproducts and potential toxicity using in silico tests. *Rapid communications in mass spectrometry : RCM* **2014**, *28*, 1153-1163.

(129) Lassalle, Y.; Nicol, E.; Genty, C.; Bourcier, S.; Bouchonnet, S. Isomerization of fenbuconazole under UV-visible irradiation - chemical and toxicological approaches. *Rapid communications in mass spectrometry : RCM* **2015**, *29*, 1335-1342.

(132) Anderson, T. A.; Salice, C. J.; Erickson, R. A.; McMurry, S. T.; Cox, S. B.; Smith, L. M. Effects of landuse and precipitation on pesticides and water quality in playa lakes of the southern high plains. *Chemosphere* **2013**, *92*, 84-90.

(133) Toxicity Estimation Software Tool (TEST) <https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test> (accessed 07 2020).

(140) ISO 1998. Water Quality. Freshwater Algal Growth Inhibition Test with Unicellular Green Algae. *International Organization for Standardization, Switzerland*, 11348e11353.

(164) Masner, P.; Muster, P.; Schmid, J. Possible methionine biosynthesis inhibition by pyrimidinamine fungicides. *Pesticide Science* **1994**, *42*, 163-166.

(165)

http://www.fao.org/fileadmin/templates/agphome/documents/Pests_Pesticides/Specs/Cyprodinil09.pdf.

(166) Règlement d'exécution (UE) 2020/421 de la commission. <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1585141442408&uri=CELEX:32020R0421> (accessed september 2020).

(167) Tang, C.; Shen, C.; Zhu, K.; Zhou, Y.; Chuang, Y. J.; He, C.; Zuo, Z. Exposure to the AhR agonist cyprodinil impacts the cardiac development and function of zebrafish larvae. *Ecotoxicol Environ Saf* **2020**, *201*, 110808.

(168) Fang, C. C.; Chen, F. Y.; Chen, C. R.; Liu, C. C.; Wong, L. C.; Liu, Y. W.; Su, J. G. Cyprodinil as an activator of aryl hydrocarbon receptor. *Toxicology* **2013**, *304*, 32-40.

(169) Go, R. E.; Kim, C. W.; Choi, K. C. Effect of fenhexamid and cyprodinil on the expression of cell cycle- and metastasis-related genes via an estrogen receptor-dependent pathway in cellular and xenografted ovarian cancer models. *Toxicol Appl Pharmacol* **2015**, *289*, 48-57.

(170) AGRITOX - Database on active plant protection substances. <http://www.agritox.anses.fr/php/data-criteria.php> (accessed september 2020).

(171) Cus, F.; Cesnik, H. B.; Bolta, S. V.; Gregoric, A. Pesticide residues and microbiological quality of bottled wines. *Food Control* **2010**, *21*, 150-154.

(172) Herrero-Hernandez, E.; Andrades, M. S.; Alvarez-Martin, A.; Pose-Juan, E.; Rodriguez-Cruz, M. S.; Sanchez-Martin, M. J. Occurrence of pesticides and some of their degradation products in waters in a Spanish wine region. *Journal of Hydrology* **2013**, *486*, 234-245.

(173) Esteve-Turrillas, F. A.; Agullo, C.; Abad-Somovilla, A.; Mercader, J. V.; Abad-Fuentes, A. Fungicide multiresidue monitoring in international wines by immunoassays. *Food Chem* **2016**, *196*, 1279-1286.

(174) Nicol, E.; Chayata, H.; Genty, C.; Bouchonnet, S.; Bourcier, S. Photodegradation of cyprodinil under UV-visible irradiation - chemical and toxicological approaches. *Rapid communications in mass spectrometry : RCM* **2016**, *30*, 2201-2211.

(175) Petrovic, M.; Gonzalez, S.; Barcelo, D. Analysis and removal of emerging contaminants in wastewater and drinking water. *Trac-Trend Anal Chem* **2003**, *22*, 685-696.

(176) Fenoll, J.; Ruiz, E.; Hellin, P.; Flores, P.; Navarro, S. Heterogeneous photocatalytic oxidation of cyprodinil and fludioxonil in leaching water under solar irradiation. *Chemosphere* **2011**, *85*, 1262-1268.

(177) Brosillon, S.; Ludovic, L.; Dominique, W. Modelling of a falling thin film deposited photocatalytic step reactor for water purification: Pesticide treatment. *Chemical Engineering Journal* **2011**, *169*, 216-225.

(178) Sanches, S.; Barreto Crespo, M. T.; Pereira, V. J. Drinking water treatment of priority pesticides using low pressure UV photolysis and advanced oxidation processes. *Water Res* **2010**, *44*, 1809-1818.

(179) Anfossi, L.; Sales, P.; Vanni, A. Degradation of anilinopyrimidine fungicides photoinduced by iron(III)-polycarboxylate complexes. *Pest Manag Sci* **2006**, *62*, 872-879.

(180) Fadaei, A. M.; Dehghani, M. H. Photocatalytic Oxidation of Organophosphorus Pesticides using Zinc Oxide. *Research Journal of Chemistry and Environment* **2012**, *16*, 104-109.

(181) Navarro, S.; Fenoll, J.; Vela, N.; Ruiz, E.; Navarro, G. Photocatalytic degradation of eight pesticides in leaching water by use of ZnO under natural sunlight. *Journal of Hazardous Materials* **2009**, *172*, 1303-1310.

(182) Texier, I.; Giannotti, C.; Malato, S.; Richter, C.; Delaire, J. Solar photodegradation of pesticides in water by sodium decatungstate. *Catalysis Today* **1999**, *54*, 297-307.

(183) Coleman, M. D.; O'Neil, J. D.; Woehrling, E. K.; Ndunge, O. B.; Hill, E. J.; Menache, A.; Reiss, C. J. A preliminary investigation into the impact of a pesticide combination on human neuronal and glial cell lines in vitro. *PLoS One* **2012**, *7*, e42768.

(184) <http://www.epa.gov/nrmrl/std/qsar/qsar.html>.

(185) Kinani, A.; Rifai, A.; Bourcier, S.; Jaber, F.; Bouchonnet, S. Structural characterization of photoproducts of pyrimethanil. *Journal of mass spectrometry : JMS* **2013**, *48*, 983-987.

(186) Chastain, J.; ter Halle, A.; de Sainte Claire, P.; Voyard, G.; Traikia, M.; Richard, C. Phototransformation of azoxystrobin fungicide in organic solvents. Photoisomerization vs. photodegradation. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **2013**, *12*, 2076-2083.

(187) Hodge, H. C.; Sterner, J. H. Tabulation of toxicity classes. *Am Ind Hyg Assoc Q* **1949**, *10*, 93-96.

(188) Dunn, K. L.; Espino, P. S.; Drobic, B.; Shihua, H.; Davie, J. R. The Ras-MAPK signal transduction pathway, cancer and chromatin remodeling. *Biochemistry & Cell Biology* **2005**, *83*, 1-14.

(189) Yoon, S.; Seger, R. The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. *Growth Factors* **2006**, *24*, 21-44.

(190) Collier, P. J.; Ramsey, A. J.; Austin, P.; Gilbert, P. Growth Inhibitory and Biocidal Activity of Some Isothiazolone Biocides. *Journal of Applied Bacteriology* **1990**, *69*, 569-577.

(191) European Chemicals Agency. (2019). 1,2-benzisothiazol-3(2H)-one <https://echa.europa.eu/substance-information/-/substanceinfo/100.018.292>.

(192) Bregnbak, D.; Lundov, M. D.; Zachariae, C.; Menne, T.; Johansen, J. D. Five cases of severe chronic dermatitis caused by isothiazolinones. *Contact Dermatitis* **2013**, *69*, 57-59.

(193) Pedersen, N. B. Occupational allergy from 1,2-benzisothiazolin-3-one and other preservatives in plastic emulsions. *Contact Dermatitis* **1976**, *2*, 340-342.

(194) Andersson, A. M.; Opstrup, M. S.; Zachariae, C.; Friis, U. F.; Thyssen, J. P.; Johansen, J. D. The importance of a complete declaration of isothiazolinones in products beyond cosmetics. *Contact Dermatitis* **2017**, *77*, 171-172.

(195) Schwensen, J. F.; Lundov, M. D.; Bossi, R.; Banerjee, P.; Gimenez-Arnau, E.; Lepoittevin, J. P.; Liden, C.; Uter, W.; Yazar, K.; White, I. R.; Johansen, J. D. Methylisothiazolinone and benzisothiazolinone are widely used in paint: a multicentre study of paints from five European countries. *Contact Dermatitis* **2015**, *72*, 127-138.

(196) Rafoth, A.; Gabriel, S.; Sacher, F.; Brauch, H. J. Analysis of isothiazolinones in environmental waters by gas chromatography-mass spectrometry. *Journal of chromatography. A* **2007**, *1164*, 74-81.

(197) Flyvholm, M. A. Preservatives in registered chemical products. *Contact Dermatitis* **2005**, *53*, 27-32.

(198) Aalto-Korte, K.; Ackermann, L.; Henriks-Eckerman, M. L.; Valimaa, J.; Reinikka-Railo, H.; Leppanen, E.; Jolanki, R. 1,2-benzisothiazolin-3-one in disposable polyvinyl chloride gloves for medical use. *Contact Dermatitis* **2007**, *57*, 365-370.

(199) Paijens, C.; Bressy, A.; Frere, B.; Tedoldi, D.; Mailler, R.; Rocher, V.; Neveu, P.; Moilleron, R. Urban pathways of biocides towards surface waters during dry and wet weathers: Assessment at the Paris conurbation scale. *Journal of Hazardous Materials* **2021**, *402*.

(200) Varga, Z.; Nicol, E.; Bouchonnet, S. Photodegradation of benzisothiazolinone: Identification and biological activity of degradation products. *Chemosphere* **2020**, *240*, 124862.

(201) Amsler, E.; Aerts, O.; Raison-Peyron, N.; Debons, M.; Milpied, B.; Giordano-Labadie, F.; Waton, J.; Ferrier-Le Bouedec, M. C.; Lartigau, I.; Pecquet, C.; Assier, H.; Avenel-Audran, M.; Bernier, C.; Castelain, F.; Collet, E.; Crepy, M. N.; Genillier, N.; Girardin, P.; Pralong, P.; Tetart, F.; Vital-Durand, D.; Soria, A.; Barbaud, A.; Dermatology Allergy Group of the French Society of, D. Airborne allergic contact dermatitis caused by isothiazolinones in water-based paints: a retrospective study of 44 cases. *Contact Dermatitis* **2017**, *77*, 163-170.

(202) Lugg, M. J. Photodegradation of the biocide 1,2-benzisothiazolin-3-one used in a paper-based jointing material (Reprinted). *International Biodeterioration & Biodegradation* **2001**, *48*, 252-254.

(203) Morris, S. L.; Walsh, R. C.; Hansen, J. N. Identification and Characterization of Some Bacterial-Membrane Sulfhydryl-Groups Which Are Targets of Bacteriostatic and Antibiotic Action. *J Biol Chem* **1984**, *259*, 3590-3594.

(204) Li, A.; Chen, Z.; Wu, Q. Y.; Huang, M. H.; Liu, Z. Y.; Chen, P.; Mei, L. C.; Hu, H. Y. Study on the removal of benzisothiazolinone biocide and its toxicity: The effectiveness of ozonation. *Chemical Engineering Journal* **2016**, *300*, 376-383.

(205) U.S. Environmental Protection Agency. (2005) Reregistration Eligibility Decision (RED) for Benzisothiazolin-3-one. https://archive.epa.gov/pesticides/reregistration/web/pdf/benzisothiazolin_red.pdf.

(206) Bollmann, U. E.; Fernandez-Calvino, D.; Brandt, K. K.; Storgaard, M. S.; Sanderson, H.; Bester, K. Biocide Runoff from Building Facades: Degradation Kinetics in Soil. *Environ Sci Technol* **2017**, *51*, 3694-3702.

(207) Tang, F.; Sun, Y.-x.; Shi, Y.; Li, X.; Hu, H.-y. Chemicals consumption and cost analysis of a microfiltration-reverse osmosis process for wastewater reclamation. *China Environmental Science* **2012**, *32*, 1613-1619.

(208) Tang, F.; Hu, H. Y.; Wu, Q. Y.; Tang, X.; Sun, Y. X.; Shi, X. L.; Huang, J. J. Effects of chemical agent injections on genotoxicity of wastewater in a microfiltration-reverse osmosis membrane process for wastewater reuse. *J Hazard Mater* **2013**, *260*, 231-237.

(209) Wang, T.; Wu, Q. Y.; Wang, W. L.; Chen, Z.; Li, B. T.; Li, A.; Liu, Z. Y.; Hu, H. Y. Self-sensitized photodegradation of benzisothiazolinone by low-pressure UV-C irradiation: Kinetics, mechanisms, and the effect of media. *Separation and Purification Technology* **2017**, *189*, 419-424.

(210) Heo, J. J.; Kim, U. J.; Oh, J. E. Simultaneous quantitative analysis of four isothiazolinones and 3-iodo-2-propynyl butyl carbamate in hygienic consumer products. *Environmental Engineering Research* **2019**, *24*, 137-143.

(211) Alvarez-Rivera, G.; Dagnac, T.; Lores, M.; Garcia-Jares, C.; Sanchez-Prado, L.; Lamas, J. P.; Llompert, M. Determination of isothiazolinone preservatives in cosmetics and household products by matrix solid-phase dispersion followed by high-performance liquid chromatography-tandem mass spectrometry. *Journal of chromatography. A* **2012**, *1270*, 41-50.

(212) Goodier, M. C.; Siegel, P. D.; Zang, L. Y.; Warshaw, E. M. Isothiazolinone in Residential Interior Wall Paint: A High-Performance Liquid Chromatographic-Mass Spectrometry Analysis. *Dermatitis* **2018**, *29*, 332-338.

(213) SCCS (Scientific Committee on Consumer Safety), (2012), Opinion on benzisothiazolinone. https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_099.pdf.

(214) User's Guide for T.E.S.T. (version 4.2) (Toxicity Estimation Software Tool). <https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test%0AUser's>.

(215) Mortelmans, K.; Zeiger, E. The Ames Salmonella/microsome mutagenicity assay. *Mutat Res* **2000**, *455*, 29-60.

(216) Hengstler, J. G.; Oesch, F. Ames Test. *Encyclopedia of Genetics (pp. 51–54)*. <https://doi.org/10.1006/rwgn.2001.1543> **2001**.

(217) Darlage, L. J. Thermal, photochemical, and electron impact induced transformations of 1,2-benzisoxazolin-3-ones and related heterocyclic compounds. *Iowa State University, Iowa, USA* **1971**.

(218) Matsuura, T.; Omura, K. Photochemical Hydroxylation of Aromatic-Compounds. *Synthesis-Stuttgart* **1974**, *3*, 173-184.

(219) Kinani, S.; Souissi, Y.; Kinani, A.; Vujovic, S.; Ait-Aissa, S.; Bouchonnet, S. Photodegradation of fluorene in aqueous solution: Identification and biological activity testing of degradation products. *Journal of chromatography. A* **2016**, *1442*, 118-128.

(220) Goldenthal, E. I. A compilation of LD50 values in newborn and adult animals. *Toxicol Appl Pharmacol* **1971**, *18*, 185-207.

(221) Boxill, G. C.; Nash, C. B.; Wheeler, A. G. Comparative pharmacological and toxicological evaluation of N-acetyl-p-aminophenol, salicylamide, and acetylsalicylic acid. *Journal of the American Pharmaceutical Association. American Pharmaceutical Association* **1958**, *47*, 479-487.

(222) Sanchez-Bayo, F.; Goka, K. Pesticide residues and bees--a risk assessment. *PLoS One* **2014**, *9*, e94482.

(223) Yao, X.-h.; Min, H.; Lü, Z.-h.; Yuan, H.-p. Influence of acetamiprid on soil enzymatic activities and respiration. *European Journal of Soil Biology* **2006**, *42*, 120-126.

(224) El Hassani, A. K.; Dacher, M.; Gary, V.; Lambin, M.; Gauthier, M.; Armengaud, C. Effects of sublethal doses of acetamiprid and thiamethoxam on the behavior of the honeybee (*Apis mellifera*). *Arch Environ Contam Toxicol* **2008**, *54*, 653-661.

(225) Zaworra, M.; Koehler, H.; Schneider, J.; Lagojda, A.; Nauen, R. Pharmacokinetics of Three Neonicotinoid Insecticides upon Contact Exposure in the Western Honey Bee, *Apis mellifera*. *Chem Res Toxicol* **2019**, *32*, 35-37.

(226) Blacquiere, T.; Smagghe, G.; van Gestel, C. A.; Mommaerts, V. Neonicotinoids in bees: a review on concentrations, side-effects and risk assessment. *Ecotoxicology* **2012**, *21*, 973-992.

(227) Kasiotis, K. M.; Zafeiraki, E.; Kapaxidi, E.; Manea-Karga, E.; Antonatos, S.; Anastasiadou, P.; Milonas, P.; Machera, K. Pesticides residues and metabolites in honeybees: A Greek overview exploring *Varroa* and *Nosema* potential synergies. *The Science of the total environment* **2021**, *769*, 145213.

(228) Ma, X.; Li, H.; Xiong, J.; Mehler, W. T.; You, J. Developmental Toxicity of a Neonicotinoid Insecticide, Acetamiprid to Zebrafish Embryos. *Journal of agricultural and food chemistry* **2019**, *67*, 2429-2436.

(229) De Vaugelade, S.; Nicol, E.; Vujovic, S.; Bourcier, S.; Pirnay, S.; Bouchonnet, S. UV-vis degradation of alpha-tocopherol in a model system and in a cosmetic emulsion- Structural elucidation of photoproducts and toxicological consequences. *Journal of chromatography. A* **2017**, *1517*, 126-133.

(230) Nicol, E.; Varga, Z.; Vujovic, S.; Bouchonnet, S. Laboratory scale UV-visible degradation of acetamiprid in aqueous marketed mixtures - Structural elucidation of photoproducts and toxicological consequences. *Chemosphere* **2020**, *248*, 126040.

(231) Cimino, A. M.; Boyles, A. L.; Thayer, K. A.; Perry, M. J. Effects of Neonicotinoid Pesticide Exposure on Human Health: A Systematic Review. *Environ Health Perspect* **2017**, *125*, 155-162.

(232) Crossthwaite, A. J.; Bigot, A.; Camblin, P.; Goodchild, J.; Lind, R. J.; Slater, R.; Maienfisch, P. The invertebrate pharmacology of insecticides acting at nicotinic acetylcholine receptors. *J Pestic Sci* **2017**, *42*, 67-83.

(233) Fasnabi, P.; Remya, C.; Madhu, G. Degradation of acetamiprid in wastewater by ozonation. *Int J Eng Res Appl* **2012**, *2*, 1672-1675.

(234) Carra, I.; Sirtori, C.; Ponce-Robles, L.; Sanchez Perez, J. A.; Malato, S.; Aguera, A. Degradation and monitoring of acetamiprid, thiabendazole and their transformation products in an agro-food industry effluent during solar photo-Fenton treatment in a raceway pond reactor. *Chemosphere* **2015**, *130*, 73-81.

(235) Kreuger, J.; Graaf, S.; Patring, J.; Adielsson, S. Pesticides in surface water in areas with open ground and greenhouse horticultural crops in Sweden. *Tech. Rep. Swedish University of Agricultural Sciences, Division of Water Quality Management* **2010**, 49.

(236) Bernabeu, A.; Vercher, R. F.; Santos-Juanes, L.; Simon, P. J.; Lardin, C.; Martinez, M. A.; Vicente, J. A.; Gonzalez, R.; Llosa, C.; Arques, A.; Amat, A. M. Solar photocatalysis as a tertiary treatment to remove emerging pollutants from wastewater treatment plant effluents. *Catalysis Today* **2011**, *161*, 235-240.

(237) Struger, J.; Grabuski, J.; Cagampan, S.; Sverko, E.; McGoldrick, D.; Marvin, C. H. Factors influencing the occurrence and distribution of neonicotinoid insecticides in surface waters of southern Ontario, Canada. *Chemosphere* **2017**, *169*, 516-523.

(238) Ikenaka, Y.; Miyabara, Y.; Ichise, T.; Nakayama, S.; Nimako, C.; Ishizuka, M.; Tohyama, C. Exposures of children to neonicotinoids in pine wilt disease control areas. *Environ Toxicol Chem* **2019**, *38*, 71-79.

(239) EFSA, 2013. Scientific opinion on the developmental neurotoxicity potential of acetamiprid and imidacloprid. *EFSA J* **11**, 3471. <https://doi.org/10.2903/j.efsa.2013.3471> (accessed 11 October 2019).

(240) Li, L.; Chen, X.; Zhang, D. Y.; Pan, X. L. Effects of insecticide acetamiprid on photosystem II (PSII) activity of *Synechocystis* sp (FACHB-898). *Pesticide Biochemistry and Physiology* **2010**, *98*, 300-304.

(241) Kagawa, N.; Nagao, T. Neurodevelopmental toxicity in the mouse neocortex following prenatal exposure to acetamiprid. *Journal of applied toxicology : JAT* **2018**, *38*, 1521-1528.

(242) Hao, C.; Eng, M. L.; Sun, F.; Morrissey, C. A. Part-per-trillion LC-MS/MS determination of neonicotinoids in small volumes of songbird plasma. *The Science of the total environment* **2018**, *644*, 1080-1087.

(243) Miguel, N.; Ormad, M. P.; Mosteo, R.; Ovelleiro, J. L. Photocatalytic Degradation of Pesticides in Natural Water: Effect of Hydrogen Peroxide. *International Journal of Photoenergy* **2012**, *2012*, 1-11.

(244) Li, S.; Ma, X.; Jiang, Y.; Cao, X. Acetamiprid removal in wastewater by the low-temperature plasma using dielectric barrier discharge. *Ecotoxicol Environ Saf* **2014**, *106*, 146-153.

(245) Sahithya, K.; Das, N. Remediation of Pesticides using Nanomaterials : An overview. *International Journal of ChemTech Research* **2015**, *8*, 86-91.

(246) Jafari, S. J.; Moussavi, G.; Hossaini, H. Degradation and mineralization of diazinon pesticide in UVC and UVC/TiO₂ process. *Desalination and Water Treatment* **2016**, *57*, 3782-3790.

(247) Sun, S.; Fan, Z.; Zhao, Y.; Guo, L.; Dai, Y. A Novel Nutrient Deprivation-Induced Neonicotinoid Insecticide Acetamiprid Degradation by *Ensifer adhaerens* CGMCC 6315. *Journal of agricultural and food chemistry* **2019**, *67*, 63-71.

(248) John, S.; Soloman, P. A.; Fasnabi, P. A. Study on removal of Acetamiprid from wastewater by electrocoagulation. *Proc Tech* **2016**, *24*, 619-630.

(249) Khan, A.; Haque, M. M.; Mir, N. A.; Muneer, M.; Boxall, C. Heterogeneous photocatalysed degradation of an insecticide derivative acetamiprid in aqueous suspensions of semiconductor. *Desalination* **2010**, *261*, 169-174.

(250) Toolabi, A.; Malakootian, M.; Ghaneian, M. T.; Esrafil, A.; Ehrampoush, M. H.; Tabatabaei, M.; AskarShahi, M. Optimization of photochemical decomposition acetamiprid pesticide from aqueous solutions and effluent toxicity assessment by *Pseudomonas aeruginosa* BCRC using response surface methodology. *AMB Express* **2017**, *7*, 159.

(251) Cruz-Alcalde, A.; Sans, C.; Esplugas, S. Priority pesticides abatement by advanced water technologies: The case of acetamiprid removal by ozonation. *The Science of the total environment* **2017**, *599-600*, 1454-1461.

(252) Agüera, A.; Ramos, M. d. M. G.; Fernández-Alba, A. R. Chemical Evaluation of Water Treatment Processes by LC-(Q)TOF-MS. *Comprehensive Analytical Chemistry* **2012**, *58*, 61-109.

(253) U. S. E. P. Agency, 2012. In: Risk Management Sustainable Technology. <http://www.epa.gov/ordntrnt/ORD/NRMRL/std/qsar/qsar.html>. (Accessed 11 October 2019).

(254) Hamilton, R. J.: *Lipid analysis in oils and fats*; Blackie Academic & Professional, London; New York., 1998.

(255) Abuin, E. B.; Encina, M. V.; Lissi, E. A. The photolysis of 3-pentanone. *Journal of Photochemistry* **1972**, *1*, 387-396.

(256) Yang, M.; Zhang, X. Comparative developmental toxicity of new aromatic halogenated DBPs in a chlorinated saline sewage effluent to the marine polychaete *Platynereis dumerilii*. *Environ Sci Technol* **2013**, *47*, 10868-10876.

(257) Mosbah, R.; Djerrou, Z.; Mantovani, A. Protective effect of *Nigella sativa* oil against acetamiprid induced reproductive toxicity in male rats. *Drug and chemical toxicology* **2018**, *41*, 206-212.

(258) Muranli, F. D. G.; Rasgele, P. G.; Kekecoglu, M.; Kanev, M.; Ozdemir, K. Potential Genotoxicity of Acetamiprid and Propineb Singly or in Combination in Cultured Human Peripheral Blood Lymphocytes by Using Mn Assay. *Fresenius Environmental Bulletin* **2015**, *24*, 3947-3955.

(259) Chakroun, S.; Ezzi, L.; Grissa, I.; Kerkeni, E.; Neffati, F.; Bhourri, R.; Sallem, A.; Najjar, M. F.; Hassine, M.; Mehdi, M.; Haouas, Z.; Ben Cheikh, H. Hematological, biochemical, and toxicopathic effects of subchronic acetamiprid toxicity in Wistar rats. *Environmental science and pollution research international* **2016**, *23*, 25191-25199.

(260) Shu, W.; Price, G. W.; Jamieson, R.; Lake, C. Biodegradation kinetics of individual and mixture non-steroidal anti-inflammatory drugs in an agricultural soil receiving alkaline treated biosolids. *The Science of the total environment* **2021**, *755*, 142520.

(261) Larsson, E.; Al-Hamimi, S.; Jonsson, J. A. Behaviour of nonsteroidal anti-inflammatory drugs and eight of their metabolites during wastewater treatment studied by hollow fibre liquid phase microextraction and liquid chromatography mass spectrometry. *The Science of the total environment* **2014**, *485-486*, 300-308.

(262) Jelic, A.; Gros, M.; Ginebreda, A.; Cespedes-Sanchez, R.; Ventura, F.; Petrovic, M.; Barcelo, D. Occurrence, partition and removal of pharmaceuticals in sewage water and sludge during wastewater treatment. *Water Res* **2011**, *45*, 1165-1176.

(263) de Souza, S. M.; Vasconcelos, E. C.; Dziedzic, M.; de Oliveira, C. M. Environmental risk assessment of antibiotics: an intensive care unit analysis. *Chemosphere* **2009**, *77*, 962-967.

(264) European Environmental Bureau. The environmental and health impacts caused by emissions of APIs to the environment <http://eeb.org/publications/31/chemicals/92673/the-environmental-and-health-impacts-caused-by-emissions-of-apis-to-the-environment-final-clean-1.pdf> (accessed 2 November 2020).

(265) Pharmaceuticals in the environment: global occurrence and potential cooperative action under the strategic approach to international chemicals management (SAICM).

www.umweltbundesamt.de/sites/default/files/medien/1968/publikationen/iww_abschl_ussbericht_saicm_arzneimittel_final.pdf (accessed 28 July 2020).

(266) Bilal, M.; Adeel, M.; Rasheed, T.; Zhao, Y.; Iqbal, H. M. N. Emerging contaminants of high concern and their enzyme-assisted biodegradation - A review. *Environ Int* **2019**, *124*, 336-353.

(267) Wojcieszynska, D.; Guzik, U. Naproxen in the environment: its occurrence, toxicity to nontarget organisms and biodegradation. *Appl Microbiol Biotechnol* **2020**, *104*, 1849-1857.

(268) Fent, K.; Weston, A. A.; Caminada, D. Ecotoxicology of human pharmaceuticals. *Aquat Toxicol* **2006**, *76*, 122-159.

(269) Norman Network. <https://www.norman-network.net/?q=Home> (accessed 2 November 2020).

(270) Verlicchi, P.; Al Aukidy, M.; Zambello, E. Occurrence of pharmaceutical compounds in urban wastewater: removal, mass load and environmental risk after a secondary treatment--a review. *The Science of the total environment* **2012**, *429*, 123-155.

(271) European Commission. *Document on risk assessment*. Technical Guidance Document on Risk Assessment Part II 337, 2003. DOI: 10.1002/mp.12308. <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/technical-guidance-document-risk-assessment-part-1-part-2>.

(272) Sangion, A.; Gramatica, P. PBT assessment and prioritization of contaminants of emerging concern: Pharmaceuticals. *Environ Res* **2016**, *147*, 297-306.

(273) Cassani, S.; Gramatica, P. Identification of potential PBT behavior of personal care products by structural approaches. *Sustainable Chemistry and Pharmacy* **2015**, *1*, 19-27.

(274) E., C. Commission directive 93/67/EEC of 20 July 1993 laying down the principles for assessment of risks to man and the environment of substances notified in accordance with council directive 67/548/EEC. . *Off J Eur Commun* **1993**, *L227*, 1-3.

(275) GWRC. *Pharmaceuticals and personal care products in the water cycle. An international review*. London, UK: Author, 2004.

(276) Monteiro, S. C.; Boxall, A. B. Occurrence and fate of human pharmaceuticals in the environment. *Rev Environ Contam Toxicol* **2010**, *202*, 53-154.

(277) ISO 11348-3:2007(en). Water quality — determination of the inhibitory effect of water samples on the light emission of *Vibrio Fisheri* (Luminescent bacteria test) — part 3: method using freeze-dried bacteria.

(278) Naproxen Safety Data Sheet. <https://www.caymanchem.com/msdss/70290m.pdf> (accessed 2 November 2020).

(279) Isidori, M.; Lavorgna, M.; Nardelli, A.; Parrella, A.; Previtiera, L.; Rubino, M. Ecotoxicity of naproxen and its phototransformation products. *The Science of the total environment* **2005**, *348*, 93-101.

(280) Cory, W. C.; Welch, A. M.; Ramirez, J. N.; Rein, L. C. Naproxen and Its Phototransformation Products: Persistence and Ecotoxicity to Toad Tadpoles (*Anaxyrus terrestris*), Individually and in Mixtures. *Environ Toxicol Chem* **2019**, *38*, 2008-2019.

(281) Kern, S.; Fenner, K.; Singer, H. P.; Schwarzenbach, R. P.; Hollender, J. Identification of transformation products of organic contaminants in natural waters by computer-aided prediction and high-resolution mass spectrometry. *Environ Sci Technol* **2009**, *43*, 7039-7046.

(282) Ternes, T. A. Occurrence of drugs in German sewage treatment plants and rivers. *Water Research* **1998**, *32*, 3245-3260.

(283) Farré, M. I.; Ferrer, I.; Ginebreda, A.; Figueras, M.; Olivella, L.; Tirapu, L.; Vilanova, M.; Barceló, D. Determination of drugs in surface water and wastewater samples by liquid chromatography–mass spectrometry: methods and preliminary results including toxicity studies with *Vibrio fischeri*. *Journal of Chromatography A* **2001**, *938*, 187-197.

(284) Hall, L. H.; Mohny, B.; Kier, L. B. The Electrotopological State - Structure Information at the Atomic Level for Molecular Graphs. *Journal of Chemical Information and Computer Sciences* **1991**, *31*, 76-82.

Chapitre IV - DEVELOPPEMENT D'UN LOGICIEL
D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE
SPECTRES DE MASSE — APPLICATION AU SUIVI
REACTIONNEL DE MELANGES HAUTEMENT
COMPLEXES

Ce chapitre présente le développement d'un logiciel visant à détecter des produits et intermédiaires réactionnels au sein de mélanges complexes analysés par spectrométrie de masse. Ce logiciel baptisé SPIX a été développé en étroite collaboration entre l'INRIA, le Centre de Mathématiques Appliquées (CMAP) de l'Ecole polytechnique et notre laboratoire.

Ce chapitre se divise en 2 parties. L'une traite de la subjectivité et de la variabilité induites par l'opérateur lors de l'analyse des résultats. La seconde montre comment le logiciel SPIX traite cette variabilité et permet de réduire la subjectivité de l'opérateur lors de l'interprétation de résultats ; elle est articulée autour un article publié dans la revue internationale à comité de lecture *Rapid Communications in Mass Spectrometry*.

IV.1. CONCEPTS DE SUBJECTIVITE ET VARIABILITE EN CHIMIE ANALYTIQUE

La variabilité induite par les analyses en GC-MS ou LC-MS et par les différentes étapes de préparation des échantillons est souvent estimée par l'analyste pour évaluer l'incertitude sur les résultats fournis. La subjectivité due à l'intervention de l'opérateur lors de l'interprétation des résultats est quant à elle peu ou pas étudiée et quasiment jamais discutée dans les publications scientifiques.

Nous allons définir le terme de subjectivité avant de présenter les résultats d'une étude de comparaison d'images, réalisée sur un panel d'individus de formation scientifique, qui illustre les conséquences de la subjectivité sur la lecture des données et l'interprétation des résultats.

IV.1.1. Définitions de la subjectivité et de la variabilité

La subjectivité relève par définition d'un choix personnel. Liée à l'opérateur, elle est présente tout au long de la préparation et de l'analyse des échantillons, depuis le choix du mode de prélèvement, des flaconnages, de la technique de préparation des échantillons, jusqu'au choix des instruments utilisés pour l'analyse. La subjectivité est

aussi présente lors de la gestion des données, qu'elle soit manuelle ou automatisée (par le choix de valeurs « seuil » à considérer lors de l'intégration automatique d'aires de pics chromatographiques par exemple).

Concernant la préparation de l'échantillon, la subjectivité peut être limitée en réduisant le nombre d'étapes, pour minimiser la quantité d'analyte perdue à chacune d'entre elles ou, au contraire, en multipliant les modes d'extraction - menés en parallèle - pour élargir le spectre des analytes susceptibles d'être détectées (polaires, apolaires, petites ou grosses molécules, etc.). Pour ce qui est de l'interprétation des données, l'utilisation de logiciels réduit fortement la subjectivité due à l'utilisateur. En effet, même si les paramètres fournis demeurent subjectifs, le fait qu'ils soient maintenus constants d'une analyse à l'autre permet la comparaison des données.

La variabilité résulte de la somme des variations de rendement inhérentes à chaque étape du protocole analytique. Elle est induite lors du prélèvement (variabilité spatiale et temporelle « naturelle » de l'échantillonnage), lors du transport, par les instruments de préparation et de mesure de l'échantillon. La variabilité induite « post prélèvement » peut être globalement limitée par l'utilisation d'étalons internes.

IV.1.2. Comparaison d'images

Afin d'évaluer les conséquences de la subjectivité de l'opérateur lors de la comparaison de spectres de masse, nous avons organisé deux tests portants sur la comparaison d'images.

La première étude a consisté à présenter 5 jeux de données constitués de spectres de masse simulés en format A4 à un panel de scientifiques. Chaque paneliste devait comparer des spectres deux à deux et repérer les différences par une croix ; 2 minutes étaient allouées à chaque comparaison. Cette étude a permis d'appréhender la variabilité inter-opérateurs mais n'a pas pris en compte la variabilité pour un même opérateur.

La seconde étude a été organisée au cours de l'école d'hiver internationale de spectrométrie de masse à Palaiseau les 4 et 5 mars 2019.²⁸⁵ Les panelistes devaient alors,

en 1 minute, comparer deux images projetées simultanément sur grand écran et noter le nombre de différences observées. Afin de motiver les panelistes à « jouer le jeu » et à s'investir dans le test, un gain (chocolats) était attribué au paneliste ayant obtenu le meilleur taux de bonnes réponses. Cette étude a permis d'étudier la variabilité entre utilisateurs mais aussi pour un même utilisateur, grâce à l'utilisation d'un triplicat introduit à l'insu des panelistes.

Aucune limitation en nombre de différences n'était indiquée pour chacune des études.

IV.1.2.1) Description des panels

Les études susmentionnées ont été réalisées sur un panel de 11 scientifiques pour la première et 37 pour la seconde. La première étude comportait un panel de chercheurs, doctorants et ingénieurs en chimie. La seconde étude utilisait un panel plus vaste comprenant des chercheurs et ingénieurs en chimie de l'eau, mais aussi des industriels, des doctorants en chimie analytique et des étudiants en master de chimie environnementale. Tous possédaient un solide bagage en chimie analytique et étaient habitués à comparer des spectres (de masse, de RMN, de spectrophotométrie, etc.)

IV.1.2.2) Choix des images

Le jeu de données de la première étude comprenait 5 lots de 2 spectres de masses simulés. Les variations d'intensité de pics entre les deux spectres allaient de moins de 30% à 100% (apparition ou disparition totale d'un pic). Le nombre de différences allait de 15 à 39 entre les deux spectres. Pour la seconde étude, le jeu de données comprenait 10 lots de 2 images à comparer dont 5 comparaisons d'images plus ou moins complexes, et 5 comparaisons de spectres simulés dont un triplicat, les types de comparaisons (images ou spectres simulés) étaient présentés en alternance (voir Figure IV-1). Le choix d'utiliser des comparaisons d'images en plus de comparaisons de spectres avait pour but de rendre le test plus ludique mais surtout de masquer efficacement l'utilisation du triplicat.

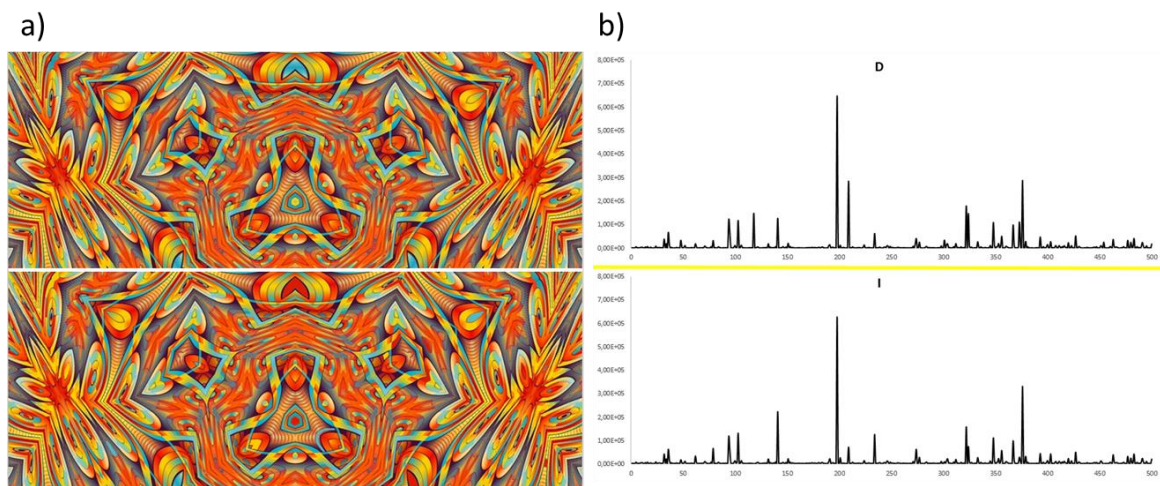


Figure IV-1 : Exemple d'images projetées lors de l'école d'hiver de spectrométrie de masse.²⁸⁵ a) comparaison de deux mandalas comprenant 22 différences. b) comparaison de deux spectres simulés comprenant 15 différences.

Les modifications au sein des comparaisons d'images étaient de natures diverses : changements de couleur, de formes, suppression, diminution ou agrandissement de motifs, etc. Certaines des images étaient simples (plusieurs formes sur un fond blanc, référencées A dans le Tableau IV-3) tandis que d'autres étaient très complexes et incluaient de très petites variations (fractales ou mandalas, référencées B dans le Tableau IV-3). Le nombre de différences allait de 5 à 22 pour les comparaisons d'images. Pour les modifications au sein des spectres de masses simulés (au nombre de 15), les variations allaient de 100% (disparition/apparition totale d'un pic) à 15% de la hauteur du pic. Les variations étaient effectuées sur des pics de différentes intensités afin d'évaluer les effets de masquage et les effets de seuil. Les 5 comparaisons de spectres simulés incluaient 3 comparaisons de jeux identiques (triplicats) sans que les panelistes n'en soient informés, ceci afin d'évaluer la variabilité intra-opérateur en plus de la variabilité inter-opérateurs (subjectivité).

IV.1.2.3) Résultats des tests

Les Tableau IV-1 et Tableau IV-2 résument les différences observées par les panelistes de la première étude en fonction des pourcentages de variation de l'intensité des pics. 51% des variations ont été identifiées en moyenne : 29% de variations faibles, 64% de variation

**CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES – APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES**

moyennes et 72% de variations fortes (Tableau IV-2). Le faible taux d'identification des pics variant faiblement peut s'expliquer par un effet de masquage de ces variations par d'autres, plus importantes, le paneliste se concentrant sur les fortes variations avant de chercher des différences plus faibles. Les valeurs des moyennes et des médianes des résultats sont relativement proches pour chacune des variations, ce qui démontre une distribution symétrique des résultats. Les coefficients de variation sont très élevés (de 38 à 92% de variation entre utilisateurs) démontrant une forte disparité des résultats. Par exemple, le paneliste n°7 identifie près de 95% des différences entre les spectres tandis que le paneliste 11 n'en identifie que 21% (voir Tableau IV-1). Ceci illustre à quel point les comparaisons « manuelles » de spectres peuvent être dépendantes de l'opérateur. Il est intéressant de mettre ce résultat en perspective : alors que de très nombreuses études sont consacrées à la variabilité inhérente à tout protocole analytique (prélèvement, préparation et analyse de l'échantillon), la variabilité imputable à l'opérateur n'est quasiment jamais questionnée alors qu'elle peut induire à elle seule plus de variabilité que l'ensemble du protocole physico-chimique. Il peut être aussi vertigineux de constater – après tous les efforts de normalisation des protocoles – que 2 analystes présenteront des conclusions différentes à partir de données identiques !

Tableau IV-1 : Différences observées pour la première étude (x : variabilité de l'intensité du pic entre les deux spectres comparés)

Paneliste	1	2	3	4	5	6	7	8	9	10	11
N.D.O. total	59 (49%)	76 (63%)	84 (70%)	50 (42%)	90 (75%)	66 (55%)	114 (95%)	43 (36%)	41 (34%)	31 (26%)	25 (21%)
N.D.O pour 30% < x < 50%	21 (39%)	18 (33%)	24 (44%)	8 (15%)	28 (52%)	11 (20%)	49 (91%)	5 (9%)	7 (13%)	1 (2%)	2 (4%)
N.D.O pour 50% < x < 75%	8 (42%)	14 (74%)	15 (79%)	11 (58%)	17 (89%)	16 (84%)	19 (100%)	11 (58%)	10 (53%)	8 (42%)	4 (21%)
N.D.O pour x. ≥ 75%	30 (64%)	44 (94%)	45 (96%)	31 (66%)	45 (96%)	39 (83%)	46 (98%)	27 (57%)	24 (51%)	22 (47%)	19 (40%)

N.D.O. : nombre de différences observées

**CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES**

Tableau IV-2 : Tableau récapitulatif des différences observées pour la première étude (x : variabilité de l'intensité du pic entre les deux spectres comparés)

	Nombre de différences réelles	Moyenne	Médiane	Ecart type	Coefficient de variation
N.D.O. total	120	62 (51%)	59 (49%)	31	47%
N.D.O pour 30% < x < 50%	54	16 (29%)	11 (20%)	18	92%
N.D.O pour 50% < x < 75%	19	12 (64%)	11 (58%)	5	38%
N.D.O pour x. ≥ 75%	47	34 (72%)	31 (66%)	11	30%

N.D.O. : nombre de différences observées

Le Tableau IV-3 présente les résultats de la seconde étude. La moyenne et la médiane sont déterminées pour chacune des comparaisons, afin de déterminer si les données sont symétriques au sein du panel. L'écart-type et le coefficient de variation permettent d'avoir une notion de la dispersion des résultats entre les différents opérateurs. L'étendue des résultats illustre la disparité au sein du panel.

Les moyennes des différences observées sont proches des médianes, ce qui démontre une certaine symétrie des données. Cela indique aussi l'absence de valeurs aberrantes au sein du panel. Les moyennes s'approchent du nombre de différences réellement présentes dans les comparaisons sans pour autant l'atteindre, indiquant de manière générale que les panelistes n'ont pas pu identifier la totalité des différences entre les jeux de données. Cela peut être dû à des variations trop faibles pour être aisément perçues à l'œil nu dans les conditions données. Les coefficients de variation au sein du panel sont tous assez élevés avec une variabilité sur les comparaisons de spectres comprise entre 34 et 47% avec une moyenne à 31%. Les coefficients de variation relatifs aux comparaisons d'images varient de 24 à 88% avec une moyenne à 23% montrant que - même si les résultats montrent plus de disparité pour la comparaison d'images complexes - les réponses sont globalement moins disparates pour les comparaisons d'images que de spectres. Il est intéressant de noter que certains panelistes observent plus de différences qu'il n'y en a réellement, majoritairement lors de la comparaison de spectres.

**CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES**

Tableau IV-3 : Comparaison des résultats obtenus lors du test visuel

<i>Comparaison</i>	<i>Type</i>	<i>Nombre de différences réelles</i>	<i>Moyenne de différences observées</i>	<i>Médiane des différences observées</i>	<i>Ecart-type</i>	<i>Coefficient de variation</i>	<i>Etendue</i>
1	Spectres	15	8.5	8	2.9	34%	2-15
2	Images (A)	13	9.3	10	2.4	26%	3-14
3	Spectres	15	9.6	10	3.8	40%	0-19
4	Images (A)	5	3.6	4	1.1	31%	1-6
5	Spectres	15	9.1	9	4.3	47%	0-19
6	Images (A)	20	17.2	18	4.1	24%	3-23
7	Spectres	15	10.6	11	3.9	37%	2-17
8	Images (B)	13	2.6	2	2.3	88%	0-8
9	Spectres	15	9.3	10	3.6	39%	0-15
10	Images (B)	22	5.6	5	2.7	48%	0-11
Total		148	85.6	90	20.7	24%	31-122
Images		73	38.4	38	8.7	23%	10-53
spectres		75	47.2	51	14.8	31%	0-19

(A) Images simples, (B) images complexes

Les résultats montrent une grande variabilité entre panelistes avec un écart type de plus de 20 différences sur l'ensemble des données et des valeurs moyennes et médianes de respectivement 85,6 et 90 pour un total de 148 différences à identifier. Comme dans la première étude, la disparité des résultats entre panelistes est très grande : 31 à 122 différences observées selon le paneliste considéré.

Les résultats obtenus pour les triplicats montrent autant de variabilité que les résultats globaux. La moyenne et la médiane des différences observées sont respectivement de 9,6 et 9,7, avec un écart-type de 2,3 et un coefficient de variation de 24%. L'étendue des résultats est très vaste : de 0 à 19 différences observées selon le paneliste considéré.

Ces résultats montrent que les variabilités intra- et inter-opérateurs peuvent être très importantes : un opérateur peut conclure que deux spectres sont quasiment identiques alors qu'un autre les considère comme significativement différents. Il arrive qu'un

opérateur identifie plus de différences qu'il n'y en a réellement. Il est également possible qu'un même opérateur n'arrive pas aux mêmes conclusions en comparant deux fois les mêmes images à quelques minutes d'intervalle.

Il est très perturbant d'observer une telle variabilité sur le rendu des résultats dans un domaine tel que l'identification de polluants potentiellement toxiques pour l'être humain et les écosystèmes. C'est la raison pour laquelle nous avons entrepris de développer un logiciel assistant la comparaison de jeux de spectres de masse, en particulier en prenant en compte le critère de variabilité pour extraire les informations pertinentes de données hautement complexes.

IV.1.3. Comparaison d'échantillons naturels

Lors de la comparaison d'échantillons prélevés avant et après un traitement physique ou chimique, il est important de prendre en compte la variabilité intrinsèque à l'échantillon. La comparaison ne se fait donc pas seulement entre deux données, mais sur - au minimum - 2 jeux de données (lots de $n * m$ échantillons de populations à comparer où n correspond au nombre d'échantillons de la population prétraitement et m au nombre d'échantillons de la population post traitement). Lors d'une comparaison visuelle de spectres de HRMS, l'opérateur doit donc non seulement comparer deux spectres « moyens » comprenant des milliers de données, mais également prendre en compte la variabilité relative à chacun des pics au sein de l'ensemble des spectres. L'opérateur peut donc facilement passer à côté d'une donnée importante ou même extraire une donnée qui n'est pas significativement différente entre les deux populations. L'utilisation d'un logiciel d'analyse statistique devient donc non seulement un atout pour limiter la subjectivité mais surtout un outil indispensable à l'analyse de données *de facto* impossible à l'œil nu.

IV.2. SPIX : UN LOGICIEL INNOVANT POUR CARACTERISER DES REACTIONS CHIMIQUES EN MILIEUX HAUTEMENT COMPLEXES

De nombreux logiciels existent pour la comparaison de spectres, beaucoup sont onéreux (quelques milliers d'euros), d'autres nécessitent de solides compétences informatiques, mais de manière générale ces logiciels sont aptes à retraiter des données complexes et à comparer deux jeux de données par le biais d'un traitement univarié (voir paragraphe I.3.2.1). La plupart sont également capables de réaliser d'autres traitements statistiques multivariés (voir paragraphe I.3.2.2) mais, à ce jour, aucun ne permet de prendre en compte des évolutions au sein d'un ensemble de populations et non plus uniquement deux populations. Or, dans le cas de suivi de réactions de dégradation, il peut être nécessaire d'identifier les intermédiaires réactionnels afin d'élucider les structures chimiques et mécanismes de formation des produits finaux. Nous avons donc, en collaboration avec le Centre de Mathématiques Appliquées de l'École polytechnique (CMAP) et l'INRIA, développé un logiciel d'analyse statistique baptisé SPIX. Celui-ci permet la comparaison de populations d'échantillons deux à deux par un traitement univarié de type test de Student ou de Wilcoxon ; il est également capable d'extraire les données traduisant une évolution significative de l'abondance d'un ion au sein des différents jeux de données en fonction d'un paramètre donné : temps de réaction, évolution du pH, ajout progressif d'un réactif etc. Dans ce dernier cas, l'intérêt n'est pas uniquement d'identifier les ions dont l'intensité varie significativement, il s'agit également de décrire cette évolution en utilisant des modèles - prédéfinis mais évolutifs - paramétrés dans le logiciel (voir *Supplementary Information IV-2*).

Le logiciel SPIX est gratuitement mis à disposition de la communauté scientifique par téléchargement sur le site web dédié : <http://spix.webpopix.org>. Il s'agit d'un logiciel possédant une interface simple d'utilisation pour un utilisateur novice en informatique - basé sur de la programmation en Matlab 2018a - dont le principe est détaillé dans un article publié en 2020 dans la revue internationale *Rapid Communications in Mass Spectrometry* (voir IV.2.3).²⁸⁶

Afin de tester la pertinence des résultats fournis par le logiciel, deux types d'échantillons ont été testés. Tout d'abord une solution d'acétamipride a été irradiée dans de l'eau ultra pure afin de voir si SPIX était capable de retrouver par des tests statistiques les photoproduits précédemment caractérisés par notre équipe suite à l'analyse visuelle de spectres de masse.²³⁰ Puis l'acétamipride a été dilué dans une solution d'acide fulvique à une concentration proche de celles retrouvées en eaux de rivière dans le but de tester la capacité du logiciel à identifier et suivre la dégradation d'un contaminant et l'évolution de ces produits de dégradation - même à l'état de trace - au sein d'un mélange complexe. Ces deux expériences sont détaillées ci-après.

IV.2.1. Test de SPIX sur un modèle simple : photolyse de l'acétamipride dans de l'eau ultrapure

Une solution d'acétamipride a été diluée à 10 ppm dans de l'eau ultra pure puis irradiée par une lampe à vapeur de mercure pendant 90 minutes. Afin de pouvoir réaliser une analyse cinétique, des volumes de 1 mL de solution ont été prélevés en triplicat aux temps 0, 1, 2, 5, 10, 30, 60 et 90 min.

Les échantillons ont ensuite été analysés par spectrométrie de masse FT-ICR de la même façon que les échantillons d'acétamipride dilué dans l'acide fulvique (voir paragraphe IV.2.3), puis analysés par le logiciel SPIX en comparant la totalité des points (cinétique) ou seulement les échantillons de référence (t0) et les échantillons prélevés après 30 minutes d'irradiation (t30) par un test de Student en posant comme hypothèse H_0 que les intensités moyennes des populations t0 et t30 étaient identiques. Le temps t30 a été choisi sur la base des résultats obtenus lors de l'étude précédente qui a montré que la majorité des photoproduits étaient présents à 30 minutes (à l'exception de PP3 et PP4a voir *Supplementary Information III-28*).

La Figure IV-2 illustre les ions ayant subi une variation d'intensité significative : (i) lors de la comparaison des deux jeux d'échantillons (Fig. IV-2 a), (ii) sur l'ensemble de la cinétique (Fig. IV-2 b). Pour les deux comparaisons, l'axe des x correspond aux rapports m/z . Dans le cas de la comparaison deux à deux, l'axe y correspond à la différence entre les

moyennes des intensités à t_0 et à t_{30} pour chaque m/z , et la couleur des pics correspond à la p -value obtenue par le biais d'un test de Student ou de Wilcoxon, suivant l'option choisie par l'utilisateur. Plus la valeur absolue de y d'un pic est grande, plus sa variation est importante, plus la couleur du pic tend vers le rouge, plus la p -value est faible et donc plus la variation est significative. Sans même procéder à l'extraction des résultats, il est facile d'identifier la diminution des pics liés à l'acétamipride ionisé (m/z 223 ; 245 ; 445 et 467 représentant respectivement $[\text{Acétamipride} + \text{H}]^+$; $[\text{Acétamipride} + \text{Na}]^+$; $[2 \times \text{Acétamipride} + \text{H}]^+$; $[2 \times \text{Acétamipride} + \text{Na}]^+$) ; l'extraction des données sélectionnées par SPIX permet de repérer d'autres ions ayant subi des variations d'abondance significatives (voir Tableau IV-4). Dans le cas de l'étude de la cinétique, l'axe y correspond au r^2 (fraction de la variance des données expliquée par le modèle) du modèle choisi par le logiciel (choix du modèle expliqué au paragraphe IV.2.3) et la couleur à la p -value correspondante. Plus un pic est grand et rouge, meilleure est la correspondance avec le modèle proposé par le logiciel.

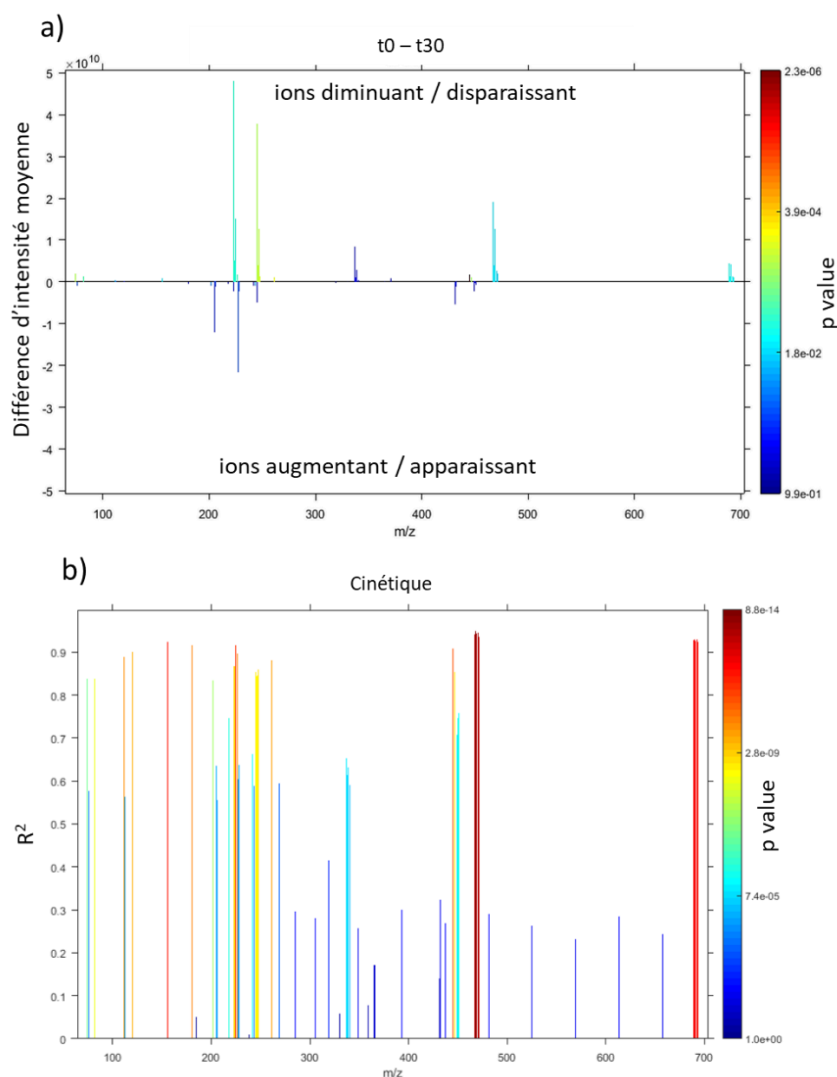


Figure IV-2 : Graphiques obtenus par le logiciel SPIX pour les données issues de l'irradiation de l'acétamipride dilué dans l'eau ultrapure à 10 ppm. a) comparaison avant (t_0) et après 30 minutes d'irradiation (t_{30}) par un test de Student. b) comparaison sur l'ensemble de la cinétique

Il est possible de faire un agrandissement sur un ion d'intérêt afin d'observer les différents spectres simulés pour cet ion (voir Figure IV-3). La visualisation des spectres pour la comparaison deux à deux (a) ou pour la cinétique (c) permet d'obtenir une vue des données et de l'évolution des intensités en fonction des points demandés. Il est ainsi possible de vérifier rapidement qu'il n'y a pas d'erreur d'attribution de pic ou de dédoublement de ce dernier. La visualisation sous forme de « boîte à moustache » (b) permet d'observer la variabilité des résultats ainsi que les valeurs moyennes d'intensité de l'ion concerné, tandis que la visualisation de la variation des intensités au cours du

temps (d) trace le modèle proposé par le logiciel ainsi que le temps nécessaire à la disparition/apparition d'un ion, ou bien le point de cinétique pour lequel cet ion est à son maximum.

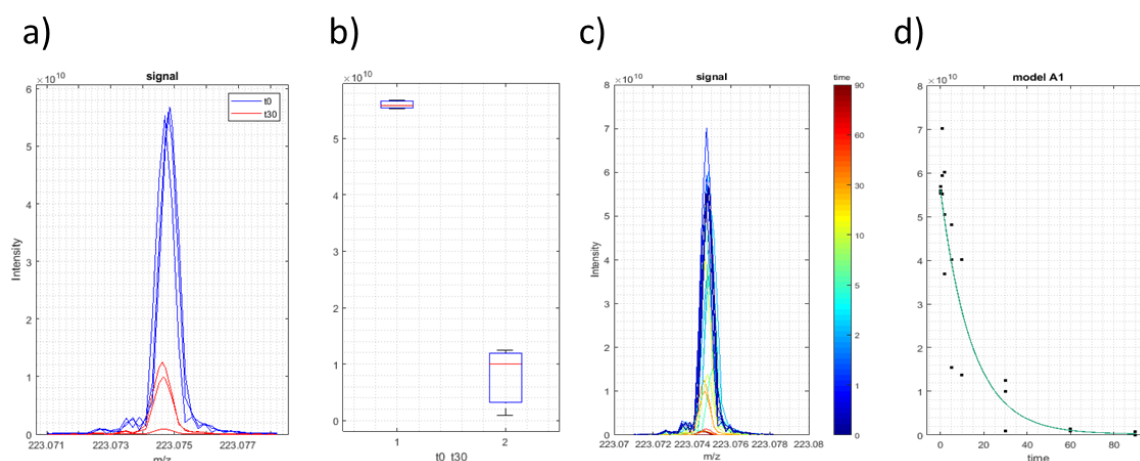


Figure IV-3 : Zoom sur le pic de l'ion de l'acétamipride protoné (m/z 223.0747) a) spectres simulés lors de la comparaison t_0 / t_{30} minutes, avant irradiation en bleu, après 30 minutes d'irradiation en rouge. b) boîte à moustache des intensités moyennes pour les deux conditions étudiées (condition 1 : t_0 ; condition 2 : t_{30}). c) spectres simulés pour la totalité de la cinétique. La couleur représente le temps d'irradiation. d) courbe de dégradation déterminée à partir des intensités des points de cinétique pour le même ion

Les données extraites fournissent les résultats indiqués dans le Tableau IV-4 après identification des formules brutes via le logiciel DataAnalysis de Bruker, regroupement des espèces isomères et suppression des harmoniques. Que ce soit pour l'analyse de la cinétique complète ou des points t_0 et t_{30} minutes, les ions subissant une variation d'abondance significative sont des ions liés à l'acétamipride (monomères, dimères et trimères) sous diverses formes cationisées (H^+ , Na^+ et K^+).

CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES

DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES

Tableau IV-4 : Extraction par le logiciel SPIX des rapports m/z ayant une p -value < 0,05 pour la cinétique et pour la comparaison $t_0 - t_{30}$. Identification des formules brutes par le logiciel DataAnalysis, ions associés à ces formules, regroupement des massifs isotopiques et suppression des harmoniques

m/z	Formules brutes	Identification (basée sur la formule brute des ions)	Cinétique			Comparaison $t_0 - t_{30}$	
			r^2	p -value	modèle	Différence $\bar{I}_0 - \bar{I}_{30}$	p -value
120.05335	$C_4H_7N_3Na^+$	Fragment de $PP2 + Na]^+$ et $[PP4 + Na]^+$	0.90	9.04E-10	B2		
180.11339	$C_9H_{14}N_3O^+$	$[PP4 + H]^+$	0.92	1.98E-10	B2		
202.09537	$C_9H_{13}N_3NaO^+$	$[PP4 + Na]^+$	0.84	1.22E-07	B2	-9.54E+08	2.84E-02
205.10863	$C_{10}H_{13}N_4O^+$	$[PP2 + H]^+$	0.64	1.99E-04	C1		
223.07475	$C_{10}H_{12}ClN_4^+$	[Acétamipride + H] ⁺	0.87	1.65E-09	A1	4.82E+10	4.70E-03
223.11932	$C_{10}H_{15}N_4O_2^+$	$[PPx + H]^+ *$	0.75	7.29E-06	B2		
227.09057	$C_{10}H_{12}N_4NaO^+$	$[PP2 + Na]^+$	0.61	4.19E-04	C1	-2.17E+10	4.85E-02
242.11391	$C_{10}H_{16}N_3O_4^+$	$[PP1 + H]^+$	0.66	1.93E-05	B1		
243.06461	$C_{10}H_{12}N_4KO^+$	$[PP2 + K]^+$	0.59	6.28E-04	C1	-9.66E+08	2.32E-02
245.05668	$C_{10}H_{11}ClN_4Na^+$	[Acétamipride + Na] ⁺	0.85	4.27E-09	A1	3.79E+10	8.47E-04
261.03074	$C_{10}H_{11}ClKN_4^+$	[Acétamipride + K] ⁺	0.88	5.60E-10	A1	1.03E+09	3.88E-04
337.10420	$C_{13}H_{19}ClN_4NaO_3^+$	$[PPx + Na]^+ *$	0.65	2.43E-05	A1		
445.14170	$C_{20}H_{23}Cl_2N_8^+$	[2 x Acétamipride + H] ⁺	0.91	3.96E-11	A1	1.83E+09	2.31E-06
449.15770	$C_{20}H_{23}ClN_8NaO^+$	$[PPx + Na]^+ *$	0.71	2.49E-05	C1		
449.20232	$C_{20}H_{26}N_8NaO_3^+$	$[PPx + Na]^+ *$	0.40	5.61E-03	B1	-1.62E+09	4.77E-02
467.12399	$C_{20}H_{22}Cl_2N_8Na^+$	[2 x Acétamipride + Na] ⁺	0.94	4.69E-13	A1	1.92E+10	1.10E-02
689.19184	$C_{30}H_{33}Cl_3N_{12}Na^+$	[3 x Acétamipride + Na] ⁺	0.93	3.73E-12	A1	4.59E+09	7.16E-03

*photoproduits, liés à l'acétamipride, non identifiés lors de l'étude précédente

A l'exception de PP3, les photoproduits identifiés lors de l'étude précédente ont tous été « retrouvés » par SPIX sous diverses formes ionisées, ainsi qu'un fragment (m/z 120.05335) ayant pu être formé lors de l'ionisation en source ou lors du stockage des ions dans la cellule de collision. Des ions dont les intensités varient significativement, non identifiés lors de l'étude précédente, ont pu être caractérisés ici : m/z 223.11932 ; m/z 337.10420 ; m/z 449.15770 ; m/z 449.20232. Leurs formules brutes suggèrent qu'ils peuvent être produits lors de l'ionisation de photoproduits connus ou liés à des dimères de l'acétamipride. Leur extraction par le logiciel montre l'efficacité de ce dernier pour repérer rapidement de nouveaux produits avant de les caractériser par LC-HRMS/MS.

L'analyse de la cinétique permet une identification de produits intermédiaires (modèles cinétiques de type C1 et C2) qui peuvent avoir été manqués lors de la comparaison t0 / t30, ainsi que de produits finaux (modèles de type B1 et B2) pouvant être apparus après 30 minutes.

IV.2.2. Test de SPIX sur des données complexes : photolyse de l'acétamipride à très faible concentration dans une solution aqueuse d'acide fulvique

Les précédents résultats ont permis de valider que le logiciel permettait de mettre en évidence l'apparition ou la disparition de molécules suite à un évènement (photolyse ou autre) mais aussi l'évolution des molécules au cours de cet évènement. Une autre expérience, plus proche de conditions environnementales réelles a permis de tester les capacités du logiciel SPIX lorsque les variations surviennent à l'état de trace dans un milieu chimique très complexe. L'acétamipride a donc été dilué à 0,4 ppm dans un mélange d'acide fulvique aqueux à 200 ppm afin de simuler l'eau d'une rivière. 6 aliquots de cette solution ont été irradiés pendant 2 heures avec une lampe à vapeur de mercure. Des prélèvements à 0, 2, 5, 10, 15, 30, 60 et 120 minutes ont été effectués pour chaque tube. Tous les points ont été analysés en infusion directe, en mode positif, par spectrométrie de masse FT-ICR. Les résultats obtenus pour la cinétique ainsi que pour la comparaison des points t0 et t30 minutes sont illustrés Figure IV-4 et détaillés dans le Tableau IV-5.

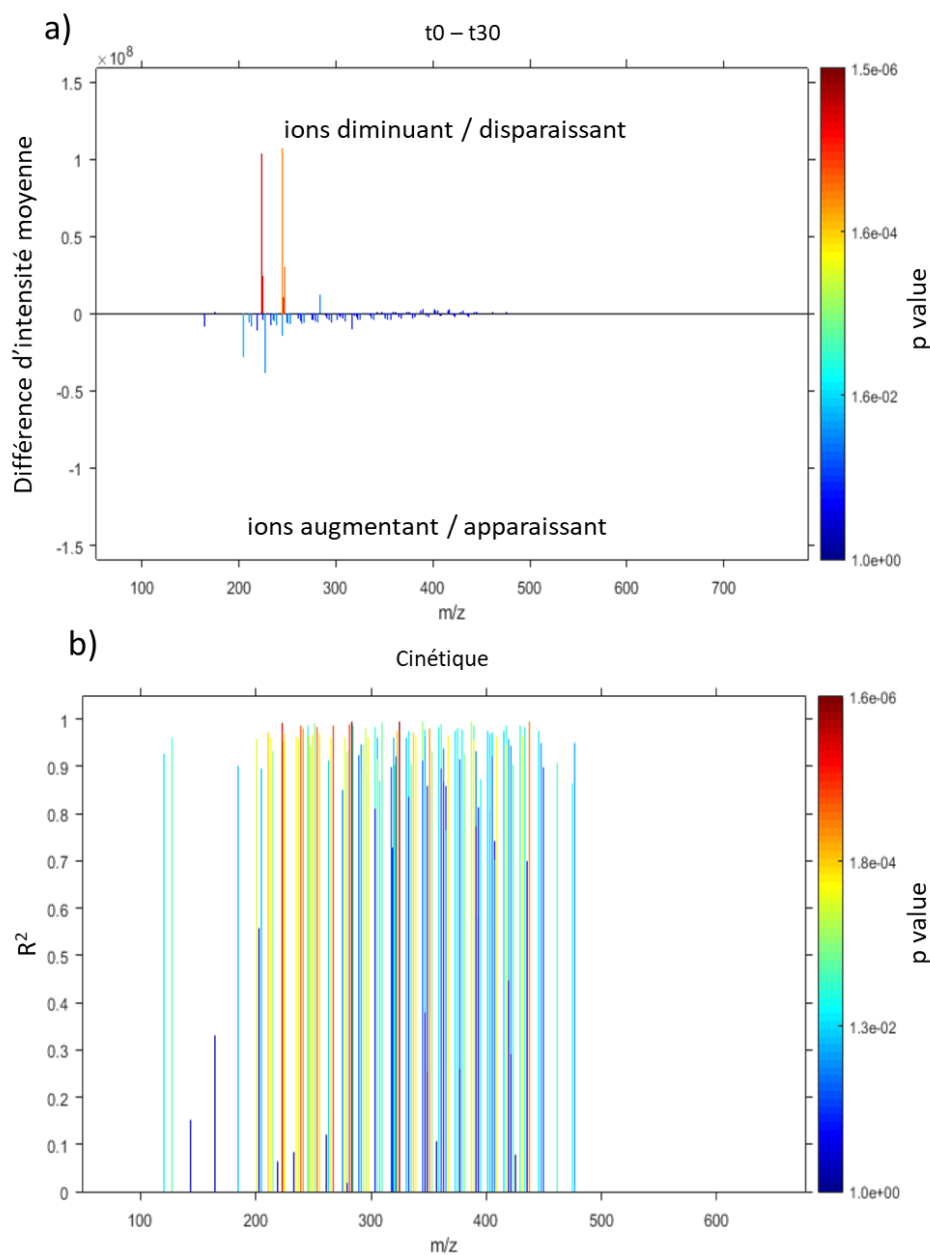


Figure IV-4 : Graphiques obtenus par le logiciel SPIX pour les données issues de l'irradiation de l'acétamipride dilué à 0,4 ppm dans l'acide fulvique à 200 ppm. a) comparaison avant (t_0) et après 30 minutes d'irradiation (t_{30}) par un test de Student après soustraction des « blancs » b) comparaison pour les données de la cinétique après soustraction des « blancs »

Sur cette figure, il est facile de voir que la principale variation concerne les ions de l'acétamipride ionisé (m/z 223 et m/z 245), en rouge et orange pour la comparaison deux à deux. En effet, comme le montre le Tableau IV-5, il s'agit des ions possédant la variabilité la plus significative (p -value la plus faible). Pourtant ces ions représentent moins de 10% du pic de base dans le spectre de référence qui est très « chargé ». Ceci illustre la

**CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES**

puissance de SPIX pour identifier des variations significatives, même sur des ions de faible intensité au sein d'un mélange complexe et donc l'intérêt d'un tel outil pour l'analyse d'échantillons environnementaux pour lesquels les polluants les plus toxiques sont rarement en forte abondance. Dans le cas de l'étude cinétique, le grand nombre d'ions dont les variations d'abondance sont considérées comme significatives (Figure IV-4) nécessite une extraction des données, en particulier pour établir les formules brutes des ions d'intérêt.

Tableau IV-5 : Extraction par le logiciel SPIX des rapports m/z auxquels est associée une p-value < 0.05 pour la cinétique, et la comparaison t0 – t30. Identification des formules brutes par le logiciel DataAnalysis et regroupement des massifs isotopiques

m/z	Formules brutes	Identification (basée sur la formule brute des ions)	Comparaison t0 – t30		Cinétique		
			Différence $\bar{I}_0 - \bar{I}_{30}$	p-value	r ²	p-value	modèle
120.05333	C ₄ H ₇ N ₃ Na ⁺	fragment			0.93	9.5E-03	B2
127.00027	C ₃ H ₄ NaO ₄ ⁺	DOM			0.96	2.7E-03	B2
185.00576	C ₅ H ₆ NaO ₆ ⁺	DOM			0.90	1.8E-02	B2
201.03705	C ₆ H ₁₀ NaO ₆ ⁺	DOM			0.96	3.8E-04	B1
205.10847	C ₁₀ H ₁₃ N ₄ O ⁺	[PP2 + H] ⁺ *	-2.8E+07	2.1E-02	0.90	1.9E-02	C1
211.02138	C ₇ H ₈ NaO ₆ ⁺	DOM	-5.8E+06	3.7E-02	0.97	1.4E-04	B1
213.03705	C ₇ H ₁₀ NaO ₆ ⁺	DOM	-8.7E+06	4.0E-02	0.96	3.3E-04	B1
215.05269	C ₇ H ₁₂ NaO ₆ ⁺	DOM			0.93	1.2E-03	B1
223.07459	C ₁₀ H ₁₂ ClN ₄ ⁺	[Acétamipride + H] ⁺ *	1.0E+08	3.2E-06	0.99	5.5E-06	A1
225.03465	C ₈ H ₁₀ NaO ₆ ⁺	DOM			0.97	1.7E-04	B1
227.09036	C ₁₀ H ₁₂ N ₄ NaO ⁺	[PP2 + Na] ⁺ *	-3.9E+07	3.7E-02			
235.05778	C ₁₀ H ₁₂ NaO ₅ ⁺	DOM			0.96	2.6E-04	B1
237.03704	C ₉ H ₁₀ NaO ₆ ⁺	DOM	-4.9E+06	3.3E-02	0.96	4.0E-04	B1
239.0527	C ₉ H ₁₂ NaO ₆ ⁺	DOM	-7.6E+06	3.8E-02	0.99	2.3E-05	B1
241.03195	C ₈ H ₁₀ NaO ₇ ⁺	DOM			0.98	5.8E-05	B1
245.05653	C ₁₀ H ₁₁ ClN ₄ Na ⁺	[Acétamipride + Na] ⁺ *	1.1E+08	4.5E-05	0.99	3.7E-03	C2
245.10095	C ₁₀ H ₁₄ N ₄ NaO ₂ ⁺	*	-1.5E+07	3.6E-02			
249.07344	C ₁₁ H ₁₄ NaO ₅ ⁺	DOM			0.97	2.0E-04	B1
251.05269	C ₁₀ H ₁₂ NaO ₆ ⁺	DOM	-7.1E+06	2.7E-02	0.99	1.5E-03	D1
253.03194	C ₉ H ₁₀ NaO ₇ ⁺	DOM			0.98	3.3E-05	B1
253.06834	C ₁₀ H ₁₄ NaO ₆ ⁺	DOM	-7.1E+06	4.6E-02	0.98	4.3E-05	B1
255.084	C ₁₀ H ₁₆ NaO ₆ ⁺	DOM			0.97	1.8E-04	B1
263.08907	C ₁₂ H ₁₆ NaO ₅ ⁺	DOM			0.91	1.4E-02	C1
265.06833	C ₁₁ H ₁₄ NaO ₆ ⁺	DOM			0.96	2.9E-04	B1
267.0476	C ₁₁ H ₁₆ NaO ₆ ⁺	DOM			0.99	2.0E-05	B1
267.084	C ₁₁ H ₁₆ NaO ₆ ⁺	DOM			0.96	3.4E-04	B1
277.06833	C ₁₂ H ₁₄ NaO ₆ ⁺	DOM			0.96	3.4E-04	B1

CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES

DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES

<i>m/z</i>	Formules brutes	Identification (basée sur la formule brute des ions)	Comparaison t0 – t30		Cinétique		
			Différence $\bar{I}_0 - \bar{I}_{30}$	<i>p-value</i>	<i>r</i> ²	<i>p-value</i>	modèle
279.04761	C ₁₁ H ₁₂ NaO ₇ ⁺	DOM			0.93	1.2E-03	B1
281.06326	C ₁₁ H ₁₄ NaO ₇ ⁺	DOM			0.99	1.5E-05	B1
281.09963	C ₁₂ H ₁₈ NaO ₆ ⁺	DOM			0.92	1.6E-03	B1
283.04253	C ₁₀ H ₁₂ NaO ₈ ⁺	DOM			1.00	1.6E-06	B1
284.12725	C ₁₂ H ₁₉ ClN ₅ O ⁺	*	1.3E+07	4.4E-02	0.99	3.8E-03	C2
289.10472	C ₁₄ H ₁₈ NaO ₅ ⁺	DOM			0.92	4.9E-02	C2
291.12039	C ₁₄ H ₂₀ NaO ₅ ⁺	DOM			0.95	3.0E-02	C2
293.06327	C ₁₂ H ₁₄ NaO ₇ ⁺	DOM			0.95	6.9E-04	B1
295.0789	C ₁₂ H ₁₆ NaO ₇ ⁺	DOM			0.98	6.2E-04	B2
295.1153	C ₁₃ H ₂₀ NaO ₆ ⁺	DOM			0.94	6.1E-03	B2
297.05819	C ₁₁ H ₁₄ NaO ₈ ⁺	DOM			0.96	3.4E-04	B1
303.12039	C ₁₅ H ₂₀ NaO ₅ ⁺	DOM			0.98	4.9E-03	C2
305.06326	C ₁₃ H ₁₄ NaO ₇ ⁺	DOM			0.92	2.0E-03	B1
305.13603	C ₁₅ H ₂₂ NaO ₅ ⁺	DOM			0.96	2.0E-02	C2
309.05818	C ₁₂ H ₁₄ NaO ₈ ⁺	DOM			0.99	2.1E-03	D1
309.09456	C ₁₃ H ₁₈ NaO ₇ ⁺	DOM			0.95	7.0E-04	B1
319.11531	C ₁₅ H ₂₀ NaO ₆ ⁺	DOM			0.96	1.9E-02	C2
321.05817	C ₁₃ H ₁₄ NaO ₈ ⁺	DOM			0.90	2.8E-03	B1
323.07384	C ₁₃ H ₁₆ NaO ₈ ⁺	DOM			0.97	1.2E-04	B1
325.05308	C ₁₂ H ₁₄ NaO ₉ ⁺	DOM			0.99	2.1E-06	B1
325.08945	C ₁₃ H ₁₈ NaO ₈ ⁺	DOM			0.98	5.5E-03	D1
331.11528	C ₁₆ H ₂₀ NaO ₆ ⁺	DOM			0.96	1.9E-02	C2
333.13095	C ₁₆ H ₂₂ NaO ₆ ⁺	DOM			0.98	9.5E-03	C2
335.07383	C ₁₄ H ₁₆ NaO ₈ ⁺	DOM			0.90	2.8E-03	B1
337.05309	C ₁₃ H ₁₄ NaO ₉ ⁺	DOM			0.97	1.3E-04	B1
339.10515	C ₁₄ H ₂₀ NaO ₈ ⁺	DOM			0.96	2.5E-04	B1
345.13097	C ₁₇ H ₂₂ NaO ₆ ⁺	DOM			0.99	1.1E-03	C2
347.11021	C ₁₆ H ₂₀ NaO ₇ ⁺	DOM			0.96	1.7E-02	C2
347.14661	C ₁₇ H ₂₄ NaO ₆ ⁺	DOM			0.98	8.1E-03	C2
351.06874	C ₁₄ H ₁₆ NaO ₉ ⁺	DOM			0.98	5.8E-05	B1
351.10514	C ₁₅ H ₂₀ NaO ₈ ⁺	DOM			0.94	7.5E-03	B2
353.08442	C ₁₄ H ₁₈ NaO ₉ ⁺	DOM			0.93	1.2E-03	B1
359.11019	C ₁₇ H ₂₀ NaO ₇ ⁺	DOM			0.98	8.8E-03	C2
359.14663	C ₁₈ H ₂₄ NaO ₆ ⁺	DOM			0.98	5.3E-03	C2
361.12588	C ₁₇ H ₂₂ NaO ₇ ⁺	DOM			0.99	3.2E-03	C2
363.14156	C ₁₇ H ₂₄ NaO ₇ ⁺	DOM			0.94	3.6E-02	C2
367.06368	C ₁₄ H ₁₆ NaO ₁₀ ⁺	DOM			0.96	2.7E-04	B1
367.10008	C ₁₅ H ₂₀ NaO ₉ ⁺	DOM			0.94	3.9E-02	D1
373.12588	C ₁₈ H ₂₂ NaO ₇ ⁺	DOM			0.98	9.1E-03	C2
375.10513	C ₁₇ H ₂₀ NaO ₈ ⁺	DOM			0.98	7.4E-03	C2
375.14154	C ₁₈ H ₂₄ NaO ₇ ⁺	DOM			0.98	8.1E-03	C2
379.06372	C ₁₅ H ₁₆ NaO ₁₀ ⁺	DOM			0.98	7.9E-03	D1
379.13656	C ₁₇ H ₂₄ NaO ₈ ⁺	DOM			0.97	1.6E-02	C2

**CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES**

<i>m/z</i>	Formules brutes	Identification (basée sur la formule brute des ions)	Comparaison t0 – t30		Cinétique		
			Différence $\bar{I}_0 - \bar{I}_{30}$	<i>p-value</i>	<i>r</i> ²	<i>p-value</i>	modèle
381.07937	C ₁₅ H ₁₈ NaO ₁₀ ⁺	DOM			0.93	1.4E-03	B1
387.10504	C ₁₈ H ₂₀ NaO ₈ ⁺	DOM			0.95	4.8E-04	A1
387.14153	C ₁₉ H ₂₄ NaO ₇ ⁺	DOM			1.00	8.2E-04	C2
389.12079	C ₁₈ H ₂₂ NaO ₈ ⁺	DOM			0.96	4.0E-04	A1
389.1572	C ₁₉ H ₂₆ NaO ₇ ⁺	DOM			0.99	4.4E-03	C2
391.13647	C ₁₈ H ₂₄ NaO ₈ ⁺	DOM			0.93	4.3E-02	C2
401.12075	C ₁₉ H ₂₂ NaO ₈ ⁺	DOM			0.97	1.1E-02	C2
401.15721	C ₂₀ H ₂₆ NaO ₇ ⁺	DOM			0.97	1.7E-04	A1
403.13645	C ₁₉ H ₂₄ NaO ₈ ⁺	DOM			0.97	1.3E-02	C2
405.15215	C ₁₉ H ₂₆ NaO ₈ ⁺	DOM			0.97	1.2E-02	C2
409.07429	C ₁₆ H ₁₈ NaO ₁₁ ⁺	DOM			0.96	2.4E-04	B1
415.13642	C ₂₀ H ₂₄ NaO ₈ ⁺	DOM			0.97	1.0E-02	C2
417.11571	C ₁₉ H ₂₂ NaO ₉ ⁺	DOM			0.95	6.3E-04	A1
417.15212	C ₂₀ H ₂₆ NaO ₈ ⁺	DOM			0.99	4.2E-03	C2
419.13137	C ₁₉ H ₂₄ NaO ₉ ⁺	DOM			0.96	2.1E-02	C2
421.14709	C ₁₉ H ₂₆ NaO ₉ ⁺	DOM			0.94	3.3E-02	C2
423.08996	C ₁₇ H ₂₀ NaO ₁₁ ⁺	DOM			0.90	3.0E-03	B1
429.15211	C ₂₁ H ₂₆ NaO ₈ ⁺	DOM			0.99	4.0E-03	C2
431.13137	C ₂₀ H ₂₄ NaO ₉ ⁺	DOM			0.96	2.4E-04	A1
433.14704	C ₂₀ H ₂₆ NaO ₉ ⁺	DOM			0.98	5.0E-03	C2
437.10555	C ₁₈ H ₂₂ NaO ₁₁ ⁺	DOM			1.00	4.5E-05	B2
445.14704	C ₂₁ H ₂₆ NaO ₉ ⁺	DOM			0.97	9.7E-03	C2
447.16269	C ₂₁ H ₂₈ NaO ₉ ⁺	DOM			0.95	2.8E-02	C2
461.14193	C ₂₁ H ₂₆ NaO ₁₀ ⁺	DOM			0.91	2.7E-03	A1
477.13686	C ₂₁ H ₂₆ NaO ₁₁ ⁺	DOM			0.95	2.9E-02	C2

*Produits issus de l'ionisation de l'acétamipride ou de ses photoproduits

DOM : matière organique dissoute

Dans le Tableau IV-5, les ions liés à l'acétamipride ionisé et aux isomères du photoproduit PP2 (photoproduit majoritaire) ont pu être identifiés par le logiciel SPIX malgré leurs très faibles intensités au sein du mélange. Au vu de l'abondance de ces ions, il est difficile de conclure que les autres photoproduits ne sont pas formés dans ces conditions, leur absence pouvant aussi être justifiée par un manque de sensibilité de l'instrument ou par une suppression de signal lors de l'ionisation. Une étude en couplage LC-MS aurait ici permis d'améliorer la sensibilité de l'instrument en permettant l'ionisation de plus de molécules, mais aussi en séparant les isomères de PP2 afin de savoir lesquels sont formés lors de la photolyse en présence d'acide fulvique. Ceci étant dit, une analyse en infusion

directe couplée à l'utilisation du logiciel SPIX permet une identification rapide des principales différences entre deux échantillons ou au cours du suivi d'une évolution d'un échantillon. Lors de cette étude, SPIX a aussi permis de mettre en évidence l'impact du traitement UV sur la matière organique dissoute en extrayant de nombreux ions résultants de l'oxydation de cette dernière. Ces ions montrent une variation significative d'intensité après irradiation UV-Vis. SPIX a aussi permis, par exemple, d'identifier deux ions, m/z 284.12725 ($C_{12}H_{19}ClN_5O^+$) et m/z 245.10095 ($C_{10}H_{14}N_4NaO_2^+$), résultant sans aucun doute (au vu de leurs formules brutes) de l'ionisation de photoproduits de l'acétamipride. Ces ions n'ont pas été détectés dans l'eau pure ; ils présentent une intensité si faible (< 1%) qu'ils n'auraient probablement pas été repérés sans l'utilisation de SPIX.

Cet exemple démontre la capacité du logiciel SPIX à mettre en évidence des changements significatifs entre deux conditions, à partir de changements d'intensité d'ions de faible abondance au sein d'une matrice complexe. Ceci ouvre la voie à de futures études concernant les conséquences globales d'un traitement physico-chimique sur le milieu aquatique. Les études actuelles suivent l'abattement des polluants d'intérêt mais ne considèrent quasiment jamais la portée du traitement sur le biotope.

IV.2.3. SPIX : un nouveau logiciel pour l'identification de réactions chimiques au sein de mélanges complexes, à partir de données de spectrométrie de masse à ultra haute résolution

Les principes de fonctionnements du logiciel SPIX, les résultats d'une comparaison avant/après irradiation UV-Visible de l'acétamipride à l'état de trace dans de l'acide fulvique ainsi qu'une étude cinétique de la dégradation de la maprotiline sont détaillés dans un article publié dans la revue *Rapid Communication in Mass Spectrometry* qui est présenté ci-après.²⁸⁶

SPIX: A NEW SOFTWARE PACKAGE TO REVEAL CHEMICAL REACTIONS AT TRACE AMOUNTS IN VERY COMPLEX MIXTURES FROM HIGH- RESOLUTION MASS SPECTRA DATA SETS

Edith Nicol^{a*}, Yao. Xu^{b,c}, Zsuzsanna Varga^a, Said Kinani^d, Stéphane Bouchonnet^a, Marc Lavielle^{b,c}

^a Laboratoire de Chimie Moléculaire, CNRS - IP Paris, Ecole polytechnique, Route de Saclay, 91128 Palaiseau, France

^b Centre de Mathématiques Appliquées, CNRS - IP Paris, Ecole polytechnique, Route de Saclay, 91128 Palaiseau, France

^c Inria, École polytechnique, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

^d Laboratoire National d'Hydraulique et Environnement (LNHE), Division Recherche et Développement, Electricité de France (EDF), 6 Quai de Watier, 78401 Chatou Cedex 01, France

Rapid Commun. Mass Spectrom. 2021;35:e9015

<https://doi.org/10.1002/rcm.9015>

RATIONALE:

High-resolution mass spectrometry based non-targeted screening has a huge potential for applications in environmental sciences, engineering and regulation. However, it produces large datasets for which full appropriate processing is a real challenge; the development of processing software is the last building-block to enable large-scale use of this approach.

METHODS:

A new software application, SPIX, has been developed to extract relevant information from high-resolution mass spectral datasets. Dealing with intrinsic sample variability and reducing operator subjectivity, it opens up opportunities and promising prospects in many areas of analytical chemistry. SPIX is freely available at: <http://spix.webpopix.org>.

RESULTS:

Two features of the software are presented in the field of environmental analysis. An example illustrates how SPIX reveals photodegradation reactions in wastewater by fitting kinetic models to significant changes in ion abundance over time. A second example shows the ability of SPIX to detect photoproducts at trace amounts in river water, through comparison of datasets from samples taken before and after irradiation.

CONCLUSIONS:

SPIX has shown its ability to reveal relevant modifications between two series of large datasets, allowing, for instance, the study of the consequences of a given event on a complex substrate. Most of all – and it is to our knowledge the only software currently available allowing this – it can reveal and monitor any kind of reaction in all types of mixture.

1. INTRODUCTION

High-resolution mass spectrometry (HRMS) is now experiencing unprecedented growth. It appeared in the early 1970s with dual-focus devices combining magnetic and electrostatic fields, and continued its development with the introduction of time-of-flight (TOF), Orbitrap and Fourier transform ion cyclotron resonance (FT-ICR) analyzers. While FT-ICR mass spectrometers remain the most accurate today, high-end QTOFs and QEx Orbitraps provide accuracies below 3 ppm. The resolution of an analyzer reflects its ability to separate ions with close m/z values. High-resolution analyzers can thus differentiate isobaric ions: i.e. ions with the same nominal mass but different exact masses, and therefore different chemical formulae, such as N_2^{+*} (m/z 28.0056) and CO^{+*} (m/z 27.9944). High resolution is a very valuable asset: it not only greatly improves the selectivity and specificity of “traditional” detection and quantification methods (in comparison with low-resolution analyzers), but also greatly facilitates structural elucidation by assigning chemical formulae to the detected ions.²⁸⁷

More recent use of high resolution takes advantage of its ability to separate isobaric ions, in an attempt to break free from separation methods – mainly gas (GC) or liquid chromatography (LC), or more rarely capillary electrophoresis or ion mobility – so as to expand the range of molecules detectable in a single analysis. This is particularly interesting in the context of non-targeted analyses – in which the operator does not know which molecules are likely to be present in a sample – because the choice of a chromatographic system focuses the analysis of certain classes of compounds based on their properties (volatile or not, polar or apolar, large or small, etc.) and thereby introduces a bias attributable to the subjectivity of the analyst. Direct introduction into the ion source without prior pretreatment or chromatographic separation was shown to be a useful alternative for rapid and comprehensive diagnosis of environmental samples, but this approach remains very challenging due to the extreme complexity of environmental matrices and the large number of contaminants likely to be present.²⁸⁸ On direct introduction of a mixture, various molecules are simultaneously ionized, resulting in mass spectra yielded by the overlapping of spectra of the detectable species. Thus, complex mixture analyses provide mass spectra that can contain tens or hundreds of

thousands of ions, even with soft ionization techniques such as electrospray ionization, atmospheric pressure ionization or atmospheric pressure photoionization; these spectra are of no possible use to the operator without the help of adapted software. Finding a molecule showing significant change between two conditions (upstream/ downstream or after treatment, for instance) in its trace amounts in an environmental sample is like looking for a needle in a haystack. Being able to quickly evaluate all the chemical consequences of an industrial accident on the biotope can be crucial to decision-making. In these situations, non-targeted HRMS-based screening is one of the last resorts for identifying unexpected or unknown contaminants.^{287,289-293} This approach has recently been evaluated in a comprehensive collaborative study organized by the NORMAN association, in which a total of 18 institutes from 12 European countries analyzed an extract of the same water sample collected from the Danube River. The results revealed that non-targeted analytical techniques were already widespread and that practices were substantially harmonized between the participants, but that data processing remained complicated and time-consuming.²⁹⁴ Among the main recommendations formulated to improve the non-targeted approach is the development of robust user friendly processing software. Likewise, AQUAREF – the French national reference laboratory for aquatic environment monitoring, which works in close concert with other European reference laboratories – published guidelines for HRMS untargeted analysis, for which SPIX could be a powerful tool.²⁹⁵

The first part of this article discusses the notions of uncertainty and subjectivity related to untargeted analysis. The second part presents the general working principle of SPIX software. The third part is dedicated to the presentation of results obtained on real samples. It discusses the strengths and limitations of the software and its specificities compared with the few programs currently commercially available. A brief overview of current computational and statistical approaches to extract relevant information from the big data of mass spectrometry analyses is provided in Supplementary Information IV-1; it describes the approaches of Kendrick^{61,75} and van Krevelen,^{83,96,296} as well as multivariate statistical analysis.^{106,123-125,297-299} Multivariate analysis tools enable global understanding of many concomitant variables and of their inter-correlations. Metabolomics processing

pipelines often include univariate and multivariate statistical approaches. Univariate analysis is usually used as a preprocessing step, while multivariate analysis is used for classification of samples or features. For example, principal component analysis is used to characterize differences of two groups of metabolomics GC/MS data for the diagnosis of gastric cancer. The Wilcoxon rank sum test showed the marker metabolites specific to the tumor group. Multivariate analysis, specifically principal component analysis, successfully divided the two groups of samples of normal and malignant gastric tissue.³⁰⁰ A comprehensive workflow for univariate analysis of LC/HRMS data was developed to follow human adult urinary metabolome variations. Univariate analysis was used as a preprocessing step: nonparametric hypothesis testing was used to assess correlations with covariables and the Wilcoxon test was used to calculate the median differences between genders. The univariate *p-value* results together with multivariate importance in projection evidenced that there were 108 urine metabolites whose concentrations varied with either age, body mass index or gender.³⁰¹ Concerning direct infusion mass spectrometry, a comprehensive workflow for data processing and quality control was developed for metabolomics analysis of cardiac tissue extract. It can be used for different metabolomics analyses as it focuses on the correction of intra- and inter-batch variations and offers best-practice workflows and rigorous quality assessment. The data processing steps include the Wilcoxon test and multivariate analysis.³⁰² These applications could be extended to environmental samples; however, no approach has been reported using univariate or multivariate analysis which focuses on the kinetics of compounds in HRMS datasets. The concept behind multivariate analysis is different from that of the SPIX software: the latter aims at observing all statistically relevant variables individually. Examples of SPIX applications are given below.

2. NOTIONS OF UNCERTAINTY AND SUBJECTIVITY IN MODERN UNTARGETED ANALYTICAL APPROACHES, AND INTRODUCTION TO SPIX

To illustrate the functionality of the SPIX software, it is necessary to address the notions of uncertainty and subjectivity that are fundamental in analytical chemistry. We propose to take an example in environmental chemistry. Consider a plant located on the bank of a river; it may be a treatment plant or, on the contrary, a source of pollution. The question is whether its presence significantly alters the composition of the water. The question seems simple enough, but providing a relevant answer is much less so. The conventional approach is to take water samples upstream and downstream of the plant, analyze them chemically and compare the results. This approach, while scientifically reasonable, nevertheless raises many questions at each step of the process. How many samples are needed to take account of the spatial and temporal variability of upstream and downstream water composition? Where, when and how to sample? What sample preparation to adopt, given that each choice of solvent, filter, solid-phase extraction column, chromatographic protocol and mass spectrometry ionization mode conditions the results of the analysis by favoring detection of certain molecules based on their size or polarity? Every single step in the analytical process introduces metrological uncertainties related to the measuring instruments used (balances, pipettes, etc.), but also to the so-called “matrix effect”: i.e. the matrix of the reference used to validate the method is generally not rigorously identical to the matrix being analyzed. Stochastic biases and uncertainties are also caused by adsorption, evaporation, etc. The proliferation of sources of error obliges analysts to use internal standards to reduce the overall uncertainty of the results and try to conform to industry-specific standards. Limiting the subjectivity in a method needs to make no assumptions at all, which is in contradiction with the use of an internal standard; thus, the analytical scientist is left with choosing between limiting subjectivity or limiting uncertainties. To the problem of uncertainties must be added that of operator subjectivity, at two main levels. As mentioned above, this subjectivity comes into play before measurement: when the operator establishes the

analytical protocol, choices are made, conditioned by assumptions – the operator's own or those of third parties – as to what might have contaminated the water of the river. Even if the method is not “targeted” (i.e. specifically designed for the selective detection of given analytes), it cannot be considered totally “non-targeted” as there is no effective protocol capable of extracting and detecting everything simultaneously (e.g. both polar and apolar molecules) and any selected protocol effectively excludes some potential analytes. This will lead the analyst to try to minimize sample preparation, with the dual objective of limiting uncertainties and of reducing operator-induced subjectivity; an immediate consequence of this simplification is to increase the complexity of the data. For example, mass spectra recorded from environmental samples will be much more complex if the sample is introduced directly into the mass spectrometer without prior purification and separation. A point that is generally much less considered is operator subjectivity in interpreting results, especially when the data are complex and voluminous, when it comes to manually integrating a peak or comparing two chromatograms or two mass spectra, for example.

In 2019, a visual trial devoted to subjectivity evaluation was carried out during a European winter school on mass spectrometry, on a panel of 37 people with a strong scientific background in analytical chemistry. It consisted of a series of one-minute projections of two images with 5 to 22 differences; panelists were asked to note the number of differences that they were able to spot. Some images were quite simple (pictures with modified areas) while others were very complex (fractals containing very small differences within complex areas). A set of simulated mass spectra containing 15 differences (variations in peak intensity, addition and removal of peaks) was presented to the panel – in triplicate and not consecutively – without prior notice. The variability between the results of these triplicates gave an average standard deviation of 2.3 observed differences per individual, with mean and median values of 9.6 and 9.7, respectively, and a range of 0–19. Considering the variability between panelists, a standard deviation of 20.6 differences was determined over the whole dataset, with mean and median values of 85.6 and 90, respectively, for a total 148 differences to be identified. The number of observed differences ranged from 31 to 122. The number of differences identified varied to the

point that one operator would conclude that two spectra were almost identical while another would consider them significantly different!²⁸⁵

The problem is substantially more complicated when comparing not only spectra but series of spectra corresponding, for example, to samples taken upstream and downstream of a treatment plant. A large variation in an ion count between upstream and downstream spectra may not be significant if the magnitude of variation is equal within and between the downstream and upstream populations; changes in the abundance of this ion reflect only the intrinsic chemical variability of the environment and are not a relevant marker of the impact of the plant. On the other hand, a slight change in the abundance of an ion between “upstream and downstream spectra” may be significant if the abundance is almost constant within each population; it then reflects a real effect of the plant on water quality. The SPIX software was created to remedy the observed fact that it is impossible for an operator to determine what makes sense based on simple observation of complex datasets, especially since the data are subject to intrinsic variability. The aim is to extract relevant information from numerous complex datasets. As explained below, the software can identify significant differences between mass spectra series and track the kinetic evolution of reagents, unknown reaction intermediates and reaction products at low concentrations in complex mixtures.

3. MATERIALS AND METHODS

3.1. The SPIX software

SPIX was developed in MATLAB 2018a (MathWorks, Natick, MA, USA). A stand-alone version is freely available on the website (<http://spix.webpopix.org>). The source code can be made available on request. Prior to performing any statistical analysis of the data, pre-processing is required to identify and align significant peaks in the data. The method used for detection and alignment actually depends on the type of data available:

- When the device provides data in xml format, these data have already been filtered and contain only the most significant peaks. These peaks are then aligned by using the `malign` function of the Bioinformatics Toolbox (MATLAB) with the “shortest-path” option.
- When the data obtained are raw data (e.g. xy Bruker format in the present study), i.e. intensities measured on a fine and regular grid, the following algorithm is used: considering K series to analyze, the approximate positions of the significant peaks are first roughly determined by building a single series, consisting of the maximum intensities of the K series at all data points, and by thresholding this series. This procedure is used to determine disjoint segments in which the peaks of each of the K series are located.

The position and intensity of each of these peaks are then estimated for each spectrum by fitting a model of the form $A \exp(-\alpha(x - m))$ for which the maximum value A is reached for $x - m$.

SPIX can be used in essentially two situations. The first one allows evidencing some modifications in the composition of a complex mixture over time. The focus here is on how the abundance of certain species varies as a function of a given parameter (time, pH, reagent, etc.). The objective is twofold: to detect ions with significantly varying abundance (in terms of statistical relevance), and to describe how the abundance varies by kinetic modeling. After aligning the peaks as previously described, different kinetic models, including various patterns associated with compound degradation and formation and reaction intermediates, are fitted to the data. A library including seven typical kinetics profiles is currently available; examples of graphical representations are provided in Supplementary Information IV-2. The selected model minimizes the Bayesian information criterion. The coefficient of determination r^2 is calculated to quantify the part of the variability of the data explained by the model and an ANOVA assesses whether this part of explained variability is statistically significant. The p -value of the F -test and the r^2 value are represented graphically so as to easily visualize ions with abundance accurately fitting a kinetic model.

SPIX also permits two series of samples collected under two experimental conditions to be compared. The objective is to identify the ions with significant differences in intensity and to quantify these differences. The algorithm first consists of identifying the peaks considered significant: i.e. present and above a given threshold in at least one of the two conditions. For an ion detected in this way, the procedure is as follows. First, the series are locally shifted so that all the peaks are aligned. The maximum intensity at the peak is estimated for each spectrum by fitting a model of the form $A \exp(-\alpha(x - m))$ for which the maximum value A is reached when $x = m$. This provides two series of values that can be compared on statistical tests. A t-test detects differences in the mean while a nonparametric Wilcoxon test more generally detects whether the peak intensity tends to be higher in one condition than in the other. A graphical representation of the p -values obtained for all the peaks detected, as well as of the size effects (i.e. differences in mean values between the two conditions), provides quick visualization of the chemically significant differences and the statistical relevance of the differences.

Blank correction can be done as follows. The user chooses as a threshold, a ratio and a percentile. By default, the median of intensities is used for the calculations ($p = 0.5$). For the given percentile, the ratio is defined as:

$$r_p(m/z) = \frac{B_p(m/z)}{S_p(m/z)}$$

with $B_p(m/z)$ being the percentile of order p of the blank intensities' maximum and $S_p(m/z)$ the percentile of order p of the experimental data intensities' maximum. If the peak intensity is higher than $S_p(m/z)$ (as a threshold value) in at least one of the experiment spectra, it will be kept as a peak; if not it will be ignored.

With .mat or .xml files, SPIX occupies about 500 MB (it is the MATLAB runtime that takes up all the space). With .xy formats the .mat conversion stage has to be added (sequentially): if a sub-repository (time_0 for example) is 250 MB, then SPIX occupies about 750 MB of memory. It does not represent the total volume of all sub-repositories because SPIX loads them and converts them one by one. In all cases, it works very well on a standard PC.

3.2. Chemicals, reagents, irradiation processes and analysis

The ability of SPIX to extract relevant information from sets of complex high-resolution mass spectra is illustrated in two experiments. The first concerned peroxide-photocatalyzed degradation of maprotiline (an antidepressant drug) in a wastewater treatment pilot plant. In this case, the comparison aimed at revealing reagents, intermediates and products using kinetic models, from mass spectra recorded at different irradiation times. The second experiment concerned UV irradiation of acetamiprid (a neonicotinoid insecticide) in a complex mixture of aqueous fulvic acid to simulate river water; it aimed at revealing acetamiprid photoproducts at trace levels and evaluating the impact of UV treatment on dissolved organic matter. The comparison covers two datasets, for spectra recorded before and after irradiation. The chemical structures of maprotiline and acetamiprid are shown in Figure IV-5. File Supplementary Information IV-3 describes the chemicals, sampling and irradiation processes used for the two experiments.

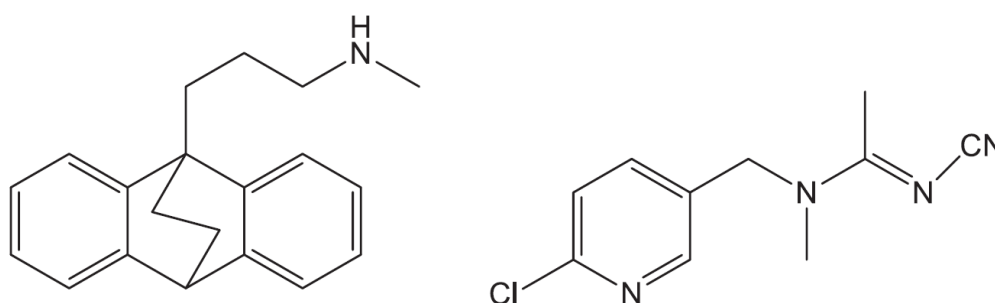


Figure IV-5: Chemical structures of maprotiline (left) and acetamiprid (right)

3.3. HRMS analysis

An ultrahigh-resolution mass spectrometer (FT-ICR SolarixXR 9.4 T; Bruker Daltonics, Bremen, Germany) was used for direct infusion mass spectrometry analysis. The electrospray ionization (ESI) source was set in positive mode and solutions were injected using an automated Acquity HPLC system (Waters, Saint Quentin en Yvelines, France). The injection volume was 10 μL . Elution was carried out using a 0.002 mL/min flow of water/acetonitrile/formic acid ($\text{H}_2\text{O}/\text{ACN}/\text{FA}$; 50/50/0.1). Nitrogen was used as nebulizer and as drying gas, set at 1 bar pressure and 4 L/min flow rate, respectively. The drying gas

temperature was set at 180 °C. The capillary voltage and endplate offset potential were set at -4500 and -500 V, respectively. Ions were accumulated for 0.2 s in the collision cell, and 50 scans were summed. The resolution was set at 4 Mpt (million data points) on a scan range from m/z 57 to m/z 1000 in order to obtain a resolution greater than 400 000 at m/z 200. A tune mix (Agilent Technologies, Les Ulis, France) was used for mass calibration. Exact formulae were assigned with error < 1 ppm.

4. RESULTS AND DISCUSSION

4.1. Highlighting chemical reactions in complex mixtures

Degradation of maprotiline in wastewater under an advanced oxidation process (peroxide/UV) was carried out in a pilot plant, with the aim of testing the ability of SPIX to follow the degradation of contaminants and the evolution of their transformation products. This pilot plant was set up by FACSA, a Spanish company operating water treatment plants, to design, optimize and compare novel water treatment processes; the operational parameters and analytical conditions are given in Supplementary Information IV-3. Considering that the abundances of reagents, intermediates and products of a chemical reaction are not expected to evolve stochastically, an original way to extract relevant information from untargeted analysis consists of filtering the results based on ion abundance trends. Briefly, A-type models are selected to detect molecules with decreasing abundance during the reaction while B-type and C-type models detect the products and intermediates, respectively. The software detects all significant changes over time and provides a kinetic model; statistical data can be exported for further analysis in the table format described in Supplementary Information IV-4. As a first example, from one set of samples using the software default threshold (1.2×10^8), SPIX automatically extracted the m/z 278.19056 signal (protonated maprotiline) for each irradiation time, and associated the fitting model referred to as A1 in SPIX with $r^2 > 0.99$ (Figure IV-6).

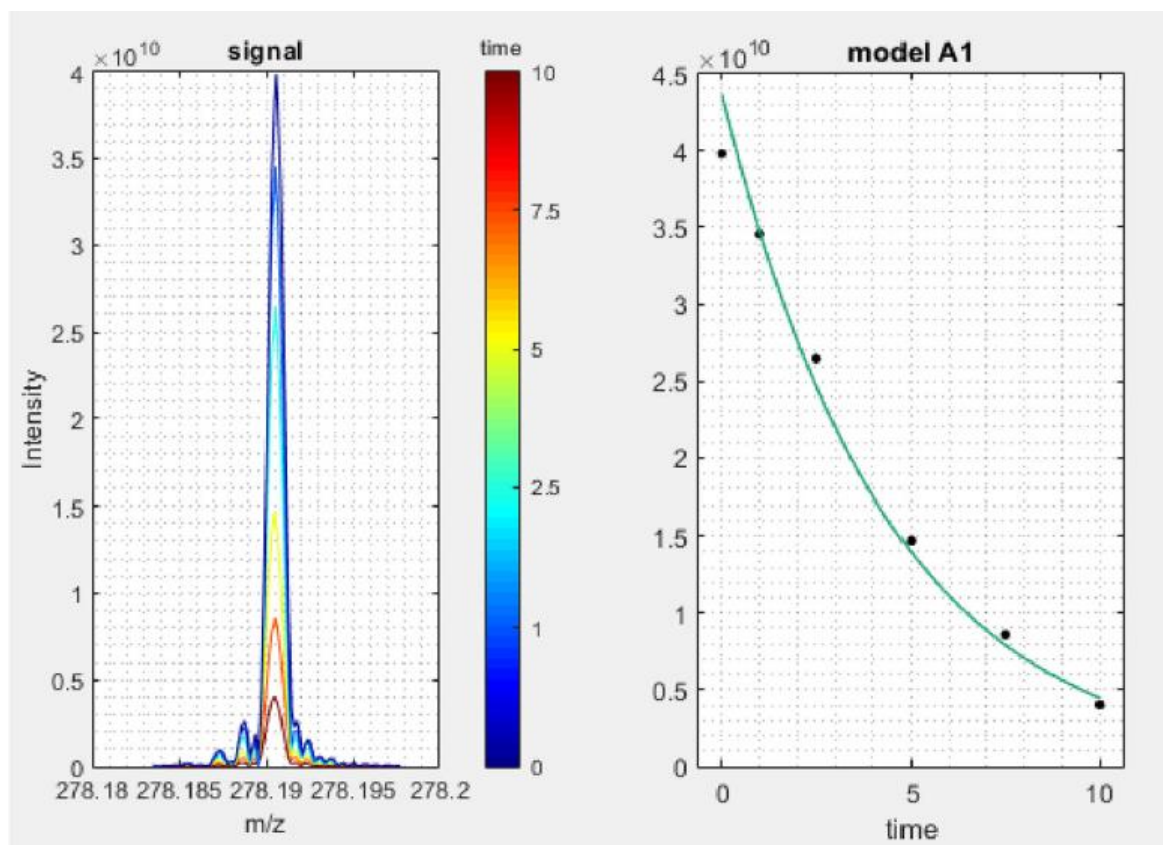


Figure IV-6: m/z 278.19056 signal (protonated maprotiline) extracted for each irradiation time and fitting model associated with degradation kinetics: $8.19 \times 10^{-5} + 4.37 \times 1010 \times \exp(-0.23x)$; $r^2 = 0.991$

A first attempt to extract transformation products from the background noise, conducted on the basis of one set of measurements, gave a few maprotiline-related peaks, but a thorough study of the raw data showed that using data from duplicate measurements yielded a kinetic model better fitting ion intensity evolution. The aim was to maximize the relevant data obtained while minimizing the number of parallel measurements, to gain valuable analysis time. Using two parallel measurements for each irradiation time, 88 ions fitting one of the SPIX kinetic models were extracted with default threshold intensity 1.2×10^8 ; this list was reduced to 12 ions keeping m/z fitting a kinetic model with $r^2 > 0.9$ (Table IV-6), these fittings being also those corresponding to the lowest p values. The formulas were assigned using Bruker software based on accurate mass measurements (sub-*ppm* accuracy) and isotopic pattern-matching. All the extracted m/z values were related to maprotiline or its photoproducts (oxidized compounds); they corresponded to singly charged ions with ^{12}C and ^{13}C isotopic contributions. One signal (m/z 92.73055)

CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES

corresponded to an artefact related to the harmonics of the m/z 278.19054 signal, resulting from signal digitization and Fourier transformation, a phenomenon previously described and explained by Mathur and O'Connor.³⁰³ To study the threshold effect and determine whether additional photoproducts would be found if more peaks were considered, the threshold was halved (6×10^7) and the same methodology was applied. Thus 197 peaks were selected by SPIX and the data were ordered according to statistical relevance. Then 23 peaks were selected on the criterion $r^2 > 0.9$ (see Table IV-6). Here again, the m/z values were all related to maprotiline and its photoproducts; a second artifact (m/z 93.06499) was found and attributed to the harmonics of m/z 279.19386.

Table IV-6: Ions extracted and associated kinetic models related to the photodegradation of Maprotiline in wastewater with $r^2 > 0.9$ (data ordered by decreasing intensity)

m/z	r^2	p value	Model ^a	Maprotiline-related	Ion formula	Intensity	Relative intensity (%)
With intensity $> 1.2 \times 10^8$							
278.19054	0.93	5.25×10^{-6}	A1	Yes	$C_{20}H_{24}N^+$	8 350 669 911	100.0
279.19386	0.94	3.82×10^{-6}	A1	Yes	$C_{19}H_{24}N^{13}C^+$	1 643 233 281	19.7
294.18552	0.98	3.91×10^{-7}	C1	Yes	$C_{20}H_{24}NO^+$	1 126 960 171	13.5
276.17486	0.92	7.58×10^{-5}	C1	Yes	$C_{20}H_{22}N^+$	451,170,002	5.4
292.16992	0.94	2.73×10^{-5}	B2	Yes	$C_{20}H_{22}NO^+$	352,390,361	4.2
92.73055	0.92	1.48×10^{-5}	A1	Yes	^b	339,591,819	4.1
310.18038	0.94	4.09×10^{-6}	B1	Yes	$C_{20}H_{24}NO_2^+$	228 987 461	2.7
295.18881	0.97	2.30×10^{-6}	C1	Yes	$C_{19}H_{24}NO^{13}C^+$	218 499 045	2.6
280.19727	0.94	3.44×10^{-6}	A1	Yes	$C_{18}H_{24}N^{13}C_2^+$	157 067 516	1.9
308.16472	0.91	1.69×10^{-4}	B2	Yes	$C_{20}H_{22}NO_2^+$	140 798 543	1.7
344.18594	0.95	1.52×10^{-5}	B2	Yes	$C_{20}H_{26}NO_4^+$	135 672 030	1.6
342.17029	0.95	1.50×10^{-5}	B2	Yes	$C_{20}H_{24}NO_4^+$	112 723 518	1.3
With $1.2 \times 10^8 > \text{intensity} > 6 \times 10^7$							
360.18088	0.96	8.69×10^{-6}	B2	Yes	$C_{20}H_{26}NO_5^+$	95 678 296	1.1
312.19604	0.94	2.71×10^{-5}	C1	Yes	$C_{20}H_{26}NO_2^+$	84 376 971	1.0
328.19096	0.94	2.74×10^{-5}	C1	Yes	$C_{20}H_{26}NO_3^+$	79 243 551	0.9
318.17025	0.94	2.46×10^{-5}	B2	Yes	$C_{18}H_{24}NO_4^+$	75 240 290	0.9
302.1753	0.98	6.44×10^{-7}	C1	Yes	$C_{18}H_{24}NO_3^+$	66 900 122	0.8
93.06499	0.92	9.63×10^{-6}	A1	Yes	^b	66 023 322	0.8
242.15416	0.95	1.34×10^{-5}	B2	Yes	$C_{16}H_{20}NO^+$	63 982 463	0.8
326.17543	0.92	8.04×10^{-5}	C1	Yes	$C_{20}H_{24}NO_3^+$	50 496 456	0.6
300.17259	0.95	1.27×10^{-6}	A1	Yes	$C_{20}H_{23}NNa^+$	20 928 662	0.2
139.09569	0.91	9.43×10^{-4}	C2	Yes	$[C_{20}H_{24}N]^{2+}$	7 061 477	0.1
336.14927	0.91	1.76×10^{-5}	A1	No	$C_{23}H_{18}N_3^+$	104 081 459	1.2

^a Currently available SPIX kinetic models are given in Supplementary Information IV-2.

^b m/z 93.06499 and m/z 92.73055 signals correspond to artifact peaks related to the harmonics of m/z 278.19054 and m/z 279.19386 ions, respectively; they resulted from signal digitization.³⁰³

Using the selection parameters referred to above, one protonated species was detected at m/z 336.14927. Considering the “seven golden rules” of Kind and Fiehn and selecting

atoms C, H, N, O, P, S, F, Cl, Br, Si, Na and K, the only matching formula was $C_{23}H_{18}N_3^+$.³⁰⁴ This species is logically assumed not to be related to maprotiline; it could correspond to a contaminant in high concentration in wastewater, degrading under UV radiation. According to the kinetics revealed by SPIX, some photoproducts were present in detectable abundance after 2.5 min irradiation. To estimate the relevance of the results provided by SPIX, one of the spectra recorded at this reaction time was selected. After blank subtraction (the blank consisting of wastewater matrix without maprotiline), the spectrum was exported in .csv format (Bruker's FTICR-MS file format). The spectra were recorded using 8 Mpt, and as the experiments were carried out using secondary treated wastewater, 4479 peaks were exported by the Bruker software which were above the signal-to-noise ratio threshold of 4. Out of these 4479 peaks, the 11 most abundant ions in the selected spectrum corresponded to the 11 most statistically relevantly changing m/z values extracted by SPIX. The two m/z signals corresponding to harmonics of major ions were also extracted with good fit, and ranked 18th and 115th in the original file (overall blank-subtracted spectrum). The photoproducts showing the highest significance (lowest p values) were those corresponding to m/z 294.18552 and m/z 302.17530 (protonated molecules); the fitting is presented in Figure IV-7 for the former. One of the maprotiline-related peaks that was not in the list of the highest intensities was not found in the spectrum recorded at 2.5 min of irradiation: it was removed by the blank subtraction process, since the wastewater matrix was very complex. It is thus noteworthy that SPIX does not require blank subtraction to provide valuable information, allowing relevant peaks that coincidentally overlap with some of the matrix peaks not to be removed. The experiment conducted on photodegradation of maprotiline showed that the SPIX software efficiently revealed the most relevant changes in the composition of the irradiated mixture on the basis of only 12 mass spectra. It was able to automatically detect reagents, intermediates and products at trace levels. Most extracted ions were related either to maprotiline or to its photoproducts. Only one compound was found which was assumed not to be related to maprotiline on the basis of its molecular formula ($C_{23}H_{17}N_3$); no significant change in the composition of the dissolved organic matter was found, although of course only electrospray-ionized species were considered. The

photodegradation pathways of maprotiline have been reported in a study more oriented toward structural elucidation.³⁰⁵

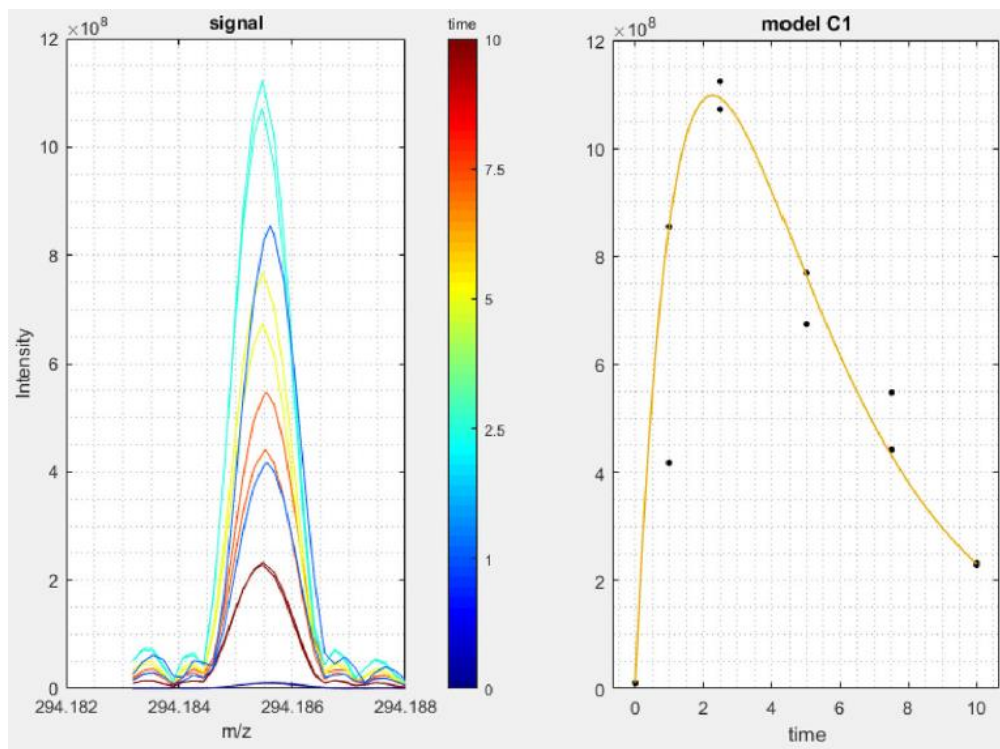


Figure IV-7: m/z 294.1852 signal extracted for each irradiation time and associated kinetic model

4.2. Comparison of two conditions: example of photodegradation of acetamiprid in aqueous solution of fulvic acid

This example demonstrates the ability of the SPIX software to point out relevant changes between two conditions from changes in low-abundance ions within a complex matrix. We studied the effect of UV radiation on acetamiprid in a complex mixture, simulating river water. A prior photolysis study of acetamiprid in ultrapure water identified acetamiprid degradation products in ultrapure water and demonstrated that the presence of other substances in the matrix leads to the formation of different degradation products.²³⁰ These results led us to study the effect of dissolved organic matter on the photodegradation of acetamiprid as may happen under real environmental conditions. The peak intensity of acetamiprid represented 9.2% of the base peak of the mass

spectrum before photolysis and only 1.9% after 30 min irradiation. A sodium adduct, originating from the use of glassware during sample preparation, impurity in solvents or electrospray ionization needle for instance, was detected with a relative intensity of 10.4% in the spectrum recorded before and 2.4% after photolysis. These differences, not detectable looking at the whole spectrum, are obvious when zooming on the region from m/z 223.00 to 223.20 (Figure IV-8).

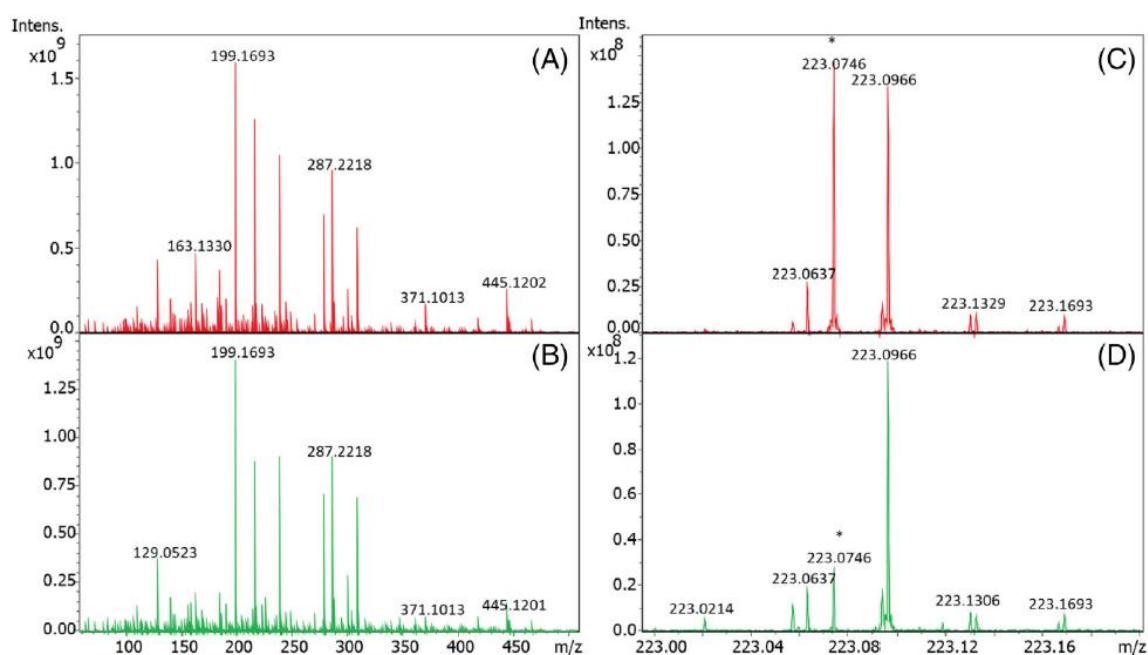


Figure IV-8: Mass spectra of acetamiprid in mixture with fulvic acid. (A) Mass spectrum before photolysis; (B) mass spectrum after 30 min photolysis; (C) zoom on the protonated acetamiprid peak (m/z 223.0746) in (A); (D) zoom on the protonated acetamiprid peak in (B)

Sets of spectra recorded before and after photolysis were compared using the SPIX software. Given the intensity of acetamiprid within the mixture, a peak detection limit was set at only three times the average intensity of spectral noise (average noise at 1×10^6 , detection threshold set at 3×10^6). This threshold was set as low as possible so as to identify acetamiprid degradation products in small amounts. Blank spectra were subtracted to eliminate any interference from solvents or instruments. Exported from SPIX, Table IV-7 lists the ions the abundance of which underwent significant change after irradiation. Here, only ions with a probability of $\geq 95\%$ ($p \leq 0.05$) were retained, i.e. with significant difference in intensity between the two conditions. In this example, results are

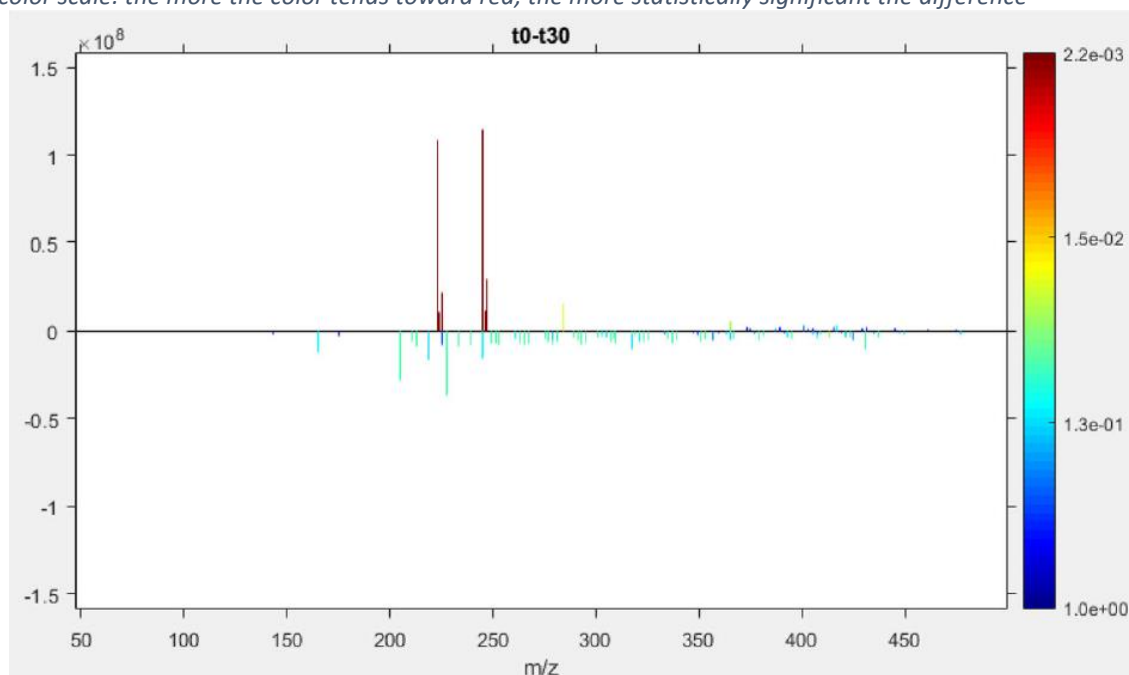
organized by increasing p value, but any parameter can be chosen for presentation of the results. A negative value in the “Difference” column indicates that the ion abundance increased after photolysis. Visual comparison (Figure IV-9) confirms the results displayed in Table IV-7: the greatest change in intensity – and with the highest significance – was the concomitant decrease of the m/z 223 (MH^+) and m/z 245 (MNa^+) ions. Many other peaks decreased or increased after photolysis, but with lower p values for the difference in intensity; some were related to acetamiprid photodegradation, others to changes in dissolved organic matter. The ions at m/z 205.10847 ($C_{10}H_{13}N_4O^+$) and m/z 227.09036 ($C_{10}H_{12}N_4NaO^+$) shown here correspond to the protonated and cationized forms of a photoproduct previously described, the structure of which was elucidated in a study of the UV irradiation of acetamiprid in pure water.³⁰⁰ This photoproduct was the major one. Thus, it is not possible to say whether others which were previously described were not detected here due to too high detection thresholds or because they are not formed in the presence of fulvic acid. It is interesting to note that SPIX revealed two ions, m/z 284.12725 ($C_{12}H_{19}ClN_5O^+$) and m/z 245.10095 ($C_{10}H_{14}N_4NaO_2^+$) – resulting from ionization of acetamiprid photoproducts based on their formulae – that were not detected by LC/MS in pure water. With a relative intensity below 1% in infusion mode, these two molecules could not have been revealed without SPIX. Fifteen ions with abundance significantly varying before and after UV irradiation were indicated. Based on their chemical formulae, they were assumed not to be related to acetamiprid or to its photoproducts; they all included a large number of oxygen atoms (≥ 6) and probably resulted from oxidation of dissolved organic matter. This is of great interest because it opens a way to investigate the global consequences of a depollution treatment, evaluating the treatment – apart from its ability to efficiently degrade pollutants – in terms of biotope preservation.

Table IV-7: *m/z* values for which the intensity significantly varied between series of spectra recorded before and after 30 min of irradiation (*n* = 6). Ions are listed by increasing *p* value

<i>m/z</i>	Difference in intensity	<i>p</i> value	Ion formula	Related to acetamidrid ²
224.07792	10 523 197	1.47×10^{-6}	$C_9H_{12}ClN_4^{13}C^+$	Yes
223.07459	103 518 217	3.20×10^{-6}	$C_{10}H_{12}ClN_4^+$	Yes
225.07161	24 820 606	1.08×10^{-5}	$C_{10}H_{12}N_4^{37}Cl^+$	Yes
246.05988	10 966 361	1.30×10^{-5}	$C_9H_{11}ClN_4Na^{13}C^+$	Yes
247.05358	30 754 403	2.86×10^{-5}	$C_{10}H_{11}N_4Na^{37}Cl^+$	Yes
245.05653	107 329 585	4.50×10^{-5}	$C_{10}H_{11}ClN_4Na^+$	Yes
201.03705	-3 542 730	1.84×10^{-2}	$C_6H_{10}NaO_6^+$	No
205.10847	-28 497 851	2.11×10^{-2}	$C_{10}H_{13}N_4O^+$	Yes
251.05269	-7 062 764	2.71×10^{-2}	$C_{12}H_{11}O_6^+$	No
297.05819	-5 696 496	2.72×10^{-2}	$C_{11}H_{14}NaO_8^+$	No
267.04760	-5 586 613	2.99×10^{-2}	$C_{10}H_{12}NaO_7^+$	No
283.04253	-4 479 881	3.20×10^{-2}	$C_{10}H_{12}NaO_8^+$	No
237.03704	-4 871 105	3.25×10^{-2}	$C_9H_{10}NaO_6^+$	No
253.03194	-4 495 727	3.33×10^{-2}	$C_9H_{10}NaO_7^+$	No
245.10095	-14 590 229	3.65×10^{-2}	$C_{10}H_{14}N_4NaO_2^+$	Yes
211.02138	-5 849 796	3.65×10^{-2}	$C_7H_8NaO_6^+$	No
227.09036	-38 558 966	3.70×10^{-2}	$C_{10}H_{12}N_4NaO^+$	Yes
239.05270	-7 601 354	3.77×10^{-2}	$C_9H_{12}NaO_6^+$	No
213.03705	-8 698 910	4.00×10^{-2}	$C_7H_{10}NaO_6^+$	No
275.05267	-3 822 758	4.02×10^{-2}	$C_{12}H_{12}NaO_6^+$	No
284.12725	12 781 460	4.41×10^{-2}	$C_{12}H_{19}ClN_5O^+$	Yes
253.06834	-7 132 886	4.57×10^{-2}	$C_{10}H_{14}NaO_6^+$	No
255.08400	-3 840 684	4.60×10^{-2}	$C_{10}H_{16}NaO_6^+$	No
325.05308	-3 163 824	4.66×10^{-2}	$C_{12}H_{14}NaO_9^+$	No
241.03195	-4 410 055	4.98×10^{-2}	$C_8H_{10}NaO_7^+$	No

²Assumption based on the ion chemical formula.

Figure IV-9: Visual result provided by the SPIX software after processing of mass spectra series recorded from samples taken before and after 30 min photolysis. The differences in ion intensities are given on the y-axis, positive values corresponding to decreased intensity after irradiation. The associated *p* value is given by the color scale: the more the color tends toward red, the more statistically significant the difference



5. CONCLUSIONS

The SPIX software aims at extracting relevant data from mass spectra datasets. User-friendly and totally free, it is available for all at: <http://spix.webpopix.org>. Two features of SPIX are presented in this article, based on examples taken from the field of environmental analysis. The first example showed how SPIX revealed photodegradation reactions by correlating significant changes in ion abundance over time with kinetic models. Thus, the software revealed the reagents, products and intermediate species in a very complex mixture (wastewater). This functionality can be extended to monitor any kind of reaction – even unknown – in all types of mixtures. Some features of the SPIX software are still under development. One aims at extending compatibility with many more file formats (.wiff, .d, .pkl, .qgd, etc.), in order to be used with most marketed mass spectrometers. Another focus of development – already running but requiring some improvements – concerns the extension of SPIX to three-dimensional datasets from hyphenated techniques such as GC/MS, LC/MS or IM-MS. Some commercial software applications allow comparison of chromatograms but, to our knowledge, an approach consisting of extracting relevant data from hyphenated techniques based on fitting with kinetic models has never been reported. The second example showed the ability of SPIX to detect photoproducts at trace amounts in an aqueous solution containing dissolved organic matter, through comparison of datasets for two conditions (before/after photolysis in the present case). Regarding this example, the ability of SPIX to deal with intrinsic variability and reduce operator subjectivity opens up promising prospects in all areas of analytical chemistry. It could, in particular, be a very useful tool to assess fragrance and flavor counterfeits, where assessment is very challenging due to the normal variations in abundance of natural substances featuring at low levels in their composition. This feature can also be used to estimate the global consequences of a given treatment on the treated medium, for example to monitor the oxidation of dissolved organic matter and its consequences on biotope preservation.

ACKNOWLEDGEMENTS

Financial support from the National FT-ICR network (FR 3624 CNRS) and from Inria (Institut national de recherche en informatique et en automatique) for conducting the research is gratefully acknowledged. This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie Grant Agreement No. 765860 (AQUALity).

SUPPLEMENTARY INFORMATION

Supplementary Information IV-1: State of the art of modern approaches in managing high-resolution mass spectrometry data

In mass spectrometry, the emergence of high-resolution analyzers has enabled analysis of samples of ever-increasing complexity. Whether in direct infusion or with hyphenated techniques (LC-MS and GC-MS couplings), the amount of information issuing from high-resolution analysis of complex mixtures requires computer processing and simplified data representation. Direct infusion of a sample can thus provide a mass spectrum including several thousands of distinct ions. Various representations are commonly used to simplify the visualization and comparison of samples analyzed by mass spectrometry. As the approaches are so diverse, this section does not seek to be exhaustive, but restricts itself to presenting the Kendrick and Van Krevelen diagrams and the multivariate statistical analyses most commonly used by high-resolution mass spectrometry specialists.

Kendrick diagrams

The Kendrick diagram allows easy identification, from a mass spectrum, of series of compounds that include the same number of heteroatoms and unsaturations but differ from each other by the number of $-CH_2-$ groups.⁶¹ The diagram is built by plotting the Kendrick mass defect (KMD) for each ion (eq. 1) as a function of the Kendrick mass (KM) (eq. 2).

$$\text{KMD} = (\text{Nominal Kendrick Mass} - \text{Exact Kendrick Mass}) \quad (\text{eq. 1})$$

$$\text{KM} = \text{IUPAC mass} * (14/14.01565) \quad (\text{eq. 2})$$

Compounds in the same series (i.e., with the same number of heteroatoms and degrees of unsaturation) will have the same KMD. In the diagram, each series is aligned horizontally with a deviation of 14 that reflects a difference of one $-CH_2-$ pattern. A shift of 0.01340 on the vertical axis corresponds to implementation of 1 unsaturation. Originally reported for the investigation of petroleomics-type samples in the early

2000s,^{62,63} the use of the Kendrick diagram has been extended and adapted over the years for complex environmental samples,^{66,67,72} metabolomic studies,^{70,71} and proteomics on phosphopeptides.⁷⁴ As needs differ between environmental chemistry and petroleomics, many studies have focused on modification of the mass defect⁷⁵ in order to characterize reaction products such as oxidation or chlorination.^{76,77} Even with a mass measurement accuracy of 1 ppm, a mass of 200 Da is assigned only 1 chemical formula, while a mass of 500 Da is assigned 21.⁷⁹ Regarding this issue, the Kendrick diagram can significantly increase the number of single chemical formulae that can be assigned from m/z values in a mass spectrum. From the chemical formula of the first compound, the identification of homologous series aids in assigning the chemical formula of other compounds in the series regardless of their mass. This allows a complex spectrum to be recalibrated, to obtain the best possible accuracy and thus assign as many chemical formulae as possible (e.g., prior to a principal component analysis).^{80,81}

Van Krevelen diagrams

The van Krevelen diagram, originally used in petroleomics to control oil and kerosene quality, represents the H/C ratio as a function of the O/C or N/C ratio for each ion of a complex mixture.⁸³ This allows the composition of a sample to be quickly estimated based on constituent molecular families (lipids, proteins, sugars, carbohydrates, lignin, tannins, etc.).²⁹⁶ Today, this representation is commonly used for environmental samples to track their evolution following an event such as treatment or pollution.^{78,88-91} Some studies have extended the van Krevelen diagram over 3 dimensions to achieve better classification of compounds and better differentiation between complex mixtures. In some cases, this approach is associated with other t-test type statistical tests.^{95,96}

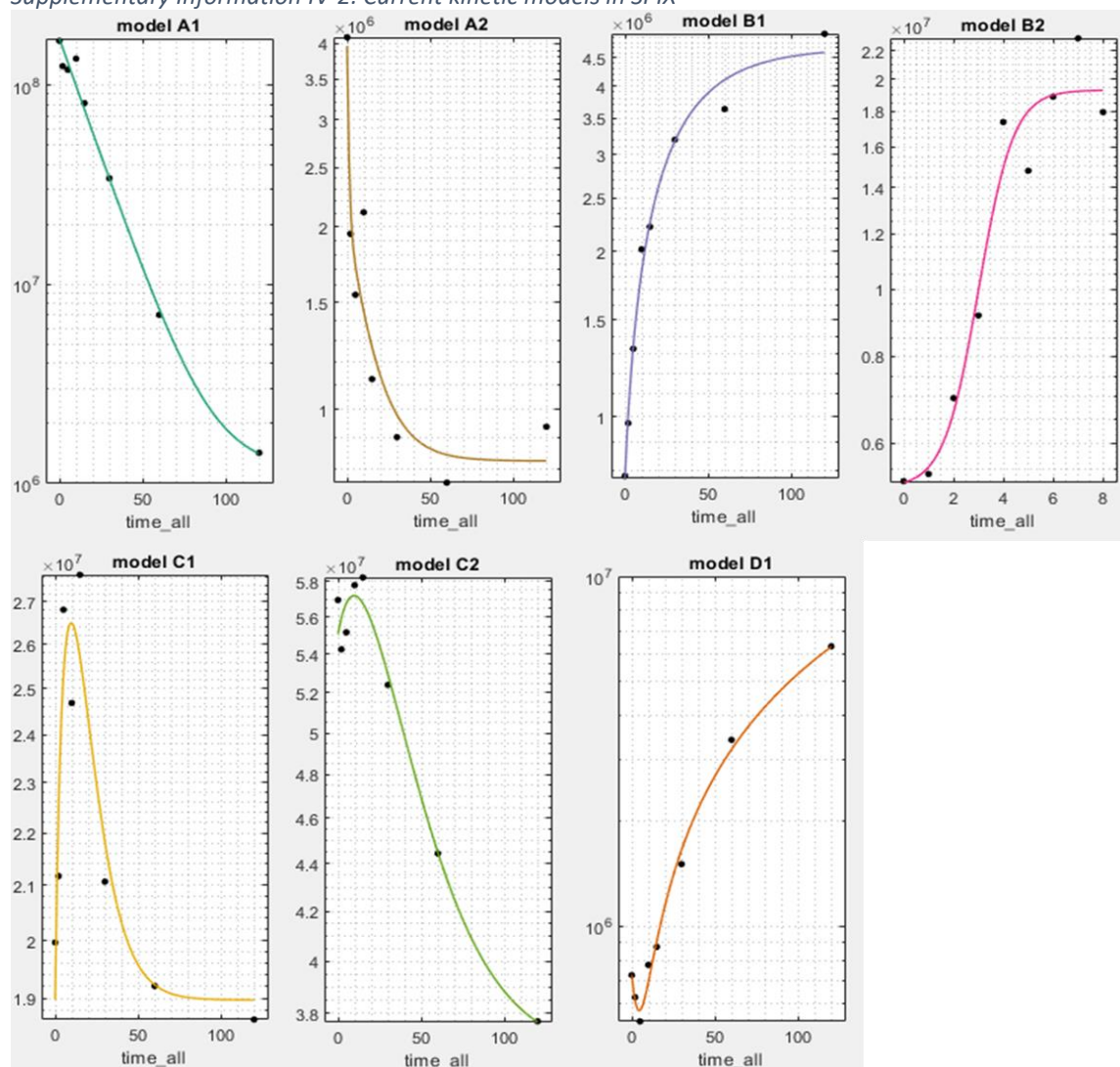
Multivariate statistical analysis

Multivariate statistical analysis is a versatile tool for dealing with high-dimensional datasets, and many methods can be used to extract valuable information, perform data compression, assess subclasses and compare groups of samples assessing relationships

between variables. For quantitative datasets, two categories of model can be distinguished in terms of the relationship between variables and response, according to the parameters: linear and non-linear. An example is UV-Vis absorbance analysis of a complex mixture, where absorbance depends on the concentrations of all the compounds present in the mixture, based on a linear relationship; in this case, multivariate linear regression can describe the correlations.²⁹⁷ On the other hand, supervised and unsupervised methods are applied for qualitative datasets.^{103,298} Principal component analysis, an unsupervised method, is usually implemented as a first approach for visualization, dimensionality reduction, classification, and finding patterns of similarities in the dataset.^{106,107} Supervised methods require *a priori* information, meaning that classes, determined by specific qualitative properties, are known in advance, and this information is used to sharpen the distinction between the given classes. A subclass of these methods, discriminant analysis, studies why the classes are different and which variables drive their separation, bearing the largest discriminatory power (e.g., Partial Least Squares Discriminant Analysis).²⁹⁹ The main areas of application in mass spectrometry data interpretation include food analysis and authentication,^{108-110,112} environmental sample analysis,^{113,114} proteomics,^{116,117} metabolomics in diagnostics and biology,¹¹⁸⁻¹²⁰ and imaging.^{121,122} However, when dealing with high-dimensionality data (e.g., direct infusion HRMS) it is difficult to assess and visualize which variables account for the differences. Usually, variable selection¹²³ or sparse methods¹²⁴ must be applied, which are able to remove or suppress variables that are irrelevant to response prediction or classification.¹²⁵ These methods proved their efficiency but have to be used with caution to avoid losing valuable information, to prevent overfitting, and to handle chance correlations correctly. In summary, multivariate analysis tools enable global understanding of many concomitant variables and of their inter-correlations. The concept behind multivariate analysis is different from that of the SPIX software: the latter aims at observing all statistically relevant variables individually.

CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES — APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES

Supplementary Information IV-2: Current kinetic models in SPIX



Model	Theoretical equation	Experimental equations for the examples shown
A1	$f=f_0+A*\exp(-k*t)$ $k>0$	$1.18E+06 + 1.70E+08 * \exp(-5.50E-02 * t)$
A2	$f=f_0+A1*\exp(-k1*t)+A2*\exp(-k2*t)$ $k1>0 ; k2>0$	$8.22E+05 + 1.89E+06 * \exp(-9.61E-01 * t) + 1.25E+06 * \exp(-6.93E-02 * t)$
B1	$f=f_0+B*(1-\exp(-k*t))$ $k>0$	$4.69E+06 - 3.91E+06 * \exp(-3.17E-02 * t)$
B2	$f=A/B/(1+C*\exp(-k*t))$ $k>0$	$1.21E-12 + 1.47E+07 / (1 + \exp(-4.83E-02 * (t - 65.23)))$
C1	$f=f_0+A*(\exp(-k1*t)-\exp(-k2*t))$ $k2 > k1 > 0$	$1.90E+07 + 3.25E+08 * (\exp(-1.00E-01 * t) - \exp(-1.06E-01 * t))$
C2	$f=f_0+A1*\exp(-k1*t)-A2*\exp(-k2*t)$ $k2 > k1 > 0 ; A1 > A2$	$3.64E+07 + 3.70E+08 * \exp(-3.96E-02 * t) - 3.52E+08 * \exp(-4.29E-02 * t)$
D1	$f=f_0+A1*\exp(-k1*t)-A2*\exp(-k2*t)$ $k1 > k2 > 0 ; A2 > A1$	$1.01E+09 + 6.22E+05 * \exp(-2.14E-01 * t) - 1.01E+09 * \exp(-5.12E-05 * t)$

Supplementary Information IV-3: Chemicals, reagents and sample preparation

Chemicals and reagents

Acetamiprid ((E)-N-(6-chloro-3-pyridylmethyl)-N'-cyano-N-methylacetamidine), maprotiline hydrochloride (N-methyl-3-(1-tetracyclo[6.6.2.0^{2,7}.0^{9,14}]hexadeca-2,4,6,9,11,13-hexaenyl) propan-1-amine;hydrochloride), acetonitrile (ACN) and formic acid (FA) (chromatographic grade purity > 99.99% for both) were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). Suwannee River Fulvic Acid Standard was purchased from International Humic Substances Society (Denver, Colorado, USA). Ultrapure water (specific resistance, 18 MΩ cm⁻¹ at 25 °C) was produced by a Purelab Chorus 1 water purification system purchased from Veolia Water Technologies (Wissous, France).

Sample preparation

1. Peroxide/UV photodegradation of maprotiline in wastewater

The UV photocatalyzed degradation of maprotiline was carried out in a 45-liter pilot plant with continuous flow at a wastewater treatment plant operated by the FACSA company in Alhama de Murcia (Spain). The molecule was submitted to peroxide/UV advanced oxidation process. The pilot plant included a tank for mixing with a stirrer, a reactor with UV lamp (model UVLA-325-4, controlled by Synergy 3 control panel - ATG UV Technology), a pump for water circulation, a rotameter to assess the water flow, and a compressor (Metabo Basic 250-50 W) to provide airflow in the rotameter equipped-reactor. The wastewater, secondary treated water originating from municipal and industrial sources, was transferred into the pilot plant after undergoing preliminary treatment (screening, sand and grease removal): decantation, biological treatment and sand filtration, before the chlorination step. At this point there were still bacteria in the mixture as well as micropollutants, which the traditional wastewater treatment methods are not able to remove. The wastewater was spiked with Maprotiline hydrochloride at 5 ppm, the total

organic carbon content of the mixture was measured as 37.2 mg/L. 4 mL of hydrogen peroxide of technical grade (33 v/w%, VWR chemicals, Llinars del Vallès, Spain) were added to the pilot plant; it corresponds to the stoichiometric amount needed for Maprotiline mineralization. Since the reaction with peroxide radical is usually fast, the reaction time was 10 minutes and the sampling for HRMS (2 x 1 mL) was done at the following times: t₀ (3 samples), 1 min, 2.5 min, 5 min, 7.5 min and 10 min. The only sample preparation before direct infusion mass spectrometry analysis was the addition of 0.1 mL of acetonitrile and 0.1% of formic acid to the 1 mL samples, to achieve better ionization and solubility. Using direct infusion HRMS aimed at suppressing many sample preparation steps, to gain a considerable amount of time and avoid too much variability and operator subjectivity.

2. UV irradiation of acetamiprid in an aqueous solution of humic acid

Acetamiprid has been detected in agriculture water samples at concentrations of up to 44 µg/L.¹³² Therefore, a 40 µg/L acetamiprid solution was prepared using an aqueous solution of fulvic acid at 20 mg/L, a mean value regarding the amount usually found in natural waters.³⁰⁶ Six glass tubes of the solution were simultaneously irradiated for 30 minutes in a laboratory-made reactor equipped with a UV-Vis high-pressure mercury lamp HPL-N125W/542 E27 SC (Philips, Ivry-sur-Seine, France) emitting light on wavelengths ranging from 200 nm to 650 nm, with a maximum irradiation wavelength at 254 nm and a radiation flux of 6200 lm. Each tube contained 50 mL of solution to ensure good surface irradiation. 1 mL of solution was taken twice from each tube before and after irradiation. Samples were analyzed in ESI-MS using automated direct infusion with a solvent made of 50% H₂O/AF (0.1%) and 50% ACN/AF (0.1%) at a flow of 0.002 mL/min. The six replicates and blanks (H₂O/ACN 50/50 v/v) were randomly analyzed and data were extracted in a «.xy» text format so that they can be treated with the SPIX software.

**CHAPITRE IV - DEVELOPPEMENT D'UN LOGICIEL D'ASSISTANCE A L'ANALYSE DE POPULATIONS DE SPECTRES
DE MASSES – APPLICATION AU SUIVI REACTIONNEL DE MELANGES HAUTEMENT COMPLEXES**

Supplementary Information IV-4: Exported data from the SPIX software after assignation of a kinetic model

Segment	<i>m/z</i>	Intensity	Time	File
65	278.19055	3.53E+10	0.0	Maprotiline WW H2O2 0min B.xy
65	278.19057	3.39E+10	1.0	Maprotiline WW H2O2 1min B.xy
65	278.19050	8,34E+09	2.5	Maprotiline WW H2O2 2min B.xy
65	278.19052	3,74E+09	5.0	Maprotiline WW H2O2 5min B.xy
65	278.19057	1,95E+09	7.5	Maprotiline WW H2O2 7min B.xy
65	278.19053	8,84E+08	10.0	Maprotiline WW H2O2 10min B.xy
<i>m/z</i>	<i>r</i> ²	p-value	model	
278.19053	0.99	0.000776	A1	

IV.2.4. Conclusion sur les travaux publiés

Ces travaux ont permis de présenter l'outil SPIX à la communauté scientifique et de montrer la puissance d'un tel outil pour l'extraction de données pertinentes au sein d'un jeu de données issues d'analyses par spectrométrie de masse. D'autres logiciels existent pour des comparaisons de jeux de données deux à deux, mais ils sont payants ou nécessitent des compétences en différents langages (R, Matlab etc.), tandis que SPIX est totalement gratuit et possède une interface conviviale même pour un utilisateur novice en informatique.

Dans cet article, deux échantillons liés à des problématiques environnementales ont été étudiés.

Le suivi cinétique de la dégradation de la maprotiline au sein d'eaux usées a permis la mise en évidence des produits mais aussi d'espèces intermédiaires. Les différents modèles cinétiques proposés par le logiciel ont été corrélés avec les changements d'abondance des ions au cours du temps. Ces modèles peuvent aussi être appliqués à tout type de paramètre réactionnel tel que l'évolution de la température, l'ajout d'un réactif, le changement du pH, etc.

Un échantillon d'acide fulvique contenant un pesticide, l'acétamipride, à l'état de trace a été analysé par la comparaison de deux jeux de données, avant et après irradiation UV-Vis. Après photolyse les variations les plus significatives s'avèrent correspondre aux ions des produits d'ionisation du pesticide et de ses photoproduits. Les autres ions subissant des variations significatives d'intensité sont issus de la fraction de matière organique dissoute modifiée suite au traitement UV-Vis. Cela nous amène à envisager de futures études sur les conséquences des différents processus de traitement des eaux sur le milieu, et non plus seulement sur les micropolluants habituellement ciblés dans ce type d'étude.

IV.3. CONCLUSION SUR LE DEVELOPPEMENT DE L'OUTIL SPIX

Les résultats présentés dans ce chapitre démontrent que SPIX est un outil puissant pour la comparaison d'échantillons deux à deux mais aussi pour le suivi de l'évolution d'un échantillon au cours du temps ou en fonction d'un paramètre déterminé.

La capacité de SPIX à déceler de très faibles différences, mais significatives, au sein de mélanges complexes, même sur des ions de faibles intensité, en fait un outil puissant pour les analyses d'échantillons environnementaux, mais aussi pour toutes analyses pour lesquelles la variabilité intrinsèque de l'échantillon est élevée et les composés d'intérêt en faible abondance. SPIX peut aussi être utilisé dans le cadre de l'estimation des conséquences globales d'un traitement sur un milieu donné comme nous avons pu le voir avec les effet d'un traitement UV-Vis sur la matière organique dissoute.

Le suivi d'une cinétique en DI-MS avec un outil aussi puissant que le logiciel SPIX permet d'extraire rapidement les ions présentant une variation d'abondance significative lors d'un traitement, de déterminer le temps auquel l'abondance d'une espèce est maximale. Il est alors possible de mener une analyse par LC-MS/MS pour en obtenir la structure et une étude *in silico* pour estimer sa toxicité.

SPIX est le premier logiciel « clé en main » permettant l'extraction de données pertinentes sur une cinétique. Suite à la présentation de SPIX au public, cette fonctionnalité a été très récemment implémentée sur d'autres logiciels commerciaux payants. Le logiciel SPIX reste gracieusement ouvert à la communauté scientifique. Les principaux développements en cours concernent l'analyse de données bidimensionnelles (LC-MS, GC-MS ou IM-MS) ; ils sont actuellement en phase de tests.

RÉFÉRENCES BIBLIOGRAPHIQUES

- (61) Kendrick, E. A Mass Scale Based on $CH_2 = 14.0000$ for High Resolution Mass Spectrometry of Organic Compounds. *Analytical chemistry* **1963**, *35*, 2146-2154.
- (62) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Analytical chemistry* **2001**, *73*, 4676-4681.
- (63) Marshall, A. G.; Rodgers, R. P. Petroleomics: the next grand challenge for chemical analysis. *Accounts of chemical research* **2004**, *37*, 53-59.
- (66) Kramer, R. W.; Kujawinski, E. B.; Hatcher, P. G. Identification of black carbon derived structures in a volcanic ash soil humic acid by Fourier transform ion cyclotron resonance mass spectrometry. *Environ Sci Technol* **2004**, *38*, 3387-3395.
- (67) Sleighter, R. L.; Hatcher, P. G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *Journal of mass spectrometry : JMS* **2007**, *42*, 559-574.
- (70) Haiying, Z.; Donglu, Z.; Ray, K.; Mingshe, Z. H. U. Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry. *Journal of mass spectrometry* **2009**, *44*, 999-1016.
- (71) Shumao, N.; Dawei, Q.; Duan, J.-A.; Jianming, G.; Shang, E.-X.; Yan, S.; Caifu, X. UPLC-QTOF/MS-based screening and identification of the constituents and their metabolites in rat plasma and urine after oral administration of Glechoma longituba extract. *Journal of chromatography. B* **2010**, *878*, 2741-2750.
- (72) Chu, F. L.; Pirastru, L.; Popovic, R.; Sleno, L. Carotenogenesis up-regulation in *Scenedesmus* sp. using a targeted metabolomics approach by liquid chromatography-high-resolution mass spectrometry. *Journal of agricultural and food chemistry* **2011**, *59*, 3004-3013.
- (74) Bruce, C.; Shifman, M. A.; Miller, P.; Gulcicek, E. E. Probabilistic enrichment of phosphopeptides by their mass defect. *Analytical chemistry* **2006**, *78*, 4374-4382.
- (75) Sleno, L. The use of mass defect in modern mass spectrometry. *Journal of Mass Spectrometry* **2012**, *47*, 226-236.
- (76) Jobst, K. J.; Shen, L.; Reiner, E. J.; Taguchi, V. Y.; Helm, P. A.; McCrindle, R.; Backus, S. The use of mass defect plots for the identification of (novel) halogenated contaminants in the environment. *Analytical and bioanalytical chemistry* **2013**, *405*, 3289-3297.
- (77) Taguchi, V. Y.; Nieckarz, R. J.; Clement, R. E.; Krolik, S.; Williams, R. Dioxin analysis by gas chromatography-Fourier transform ion cyclotron resonance mass spectrometry (GC-FTICRMS). *Journal of the American Society for Mass Spectrometry* **2010**, *21*, 1918-1921.
- (78) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Analytical chemistry* **2003**, *75*, 5336-5344.
- (79) Kind, T.; Fiehn, O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC bioinformatics* **2006**, *7*, 234.

(80) Ajaero, C.; McMartin, D. W.; Peru, K. M.; Bailey, J.; Haakensen, M.; Friesen, V.; Martz, R.; Hughes, S. A.; Brown, C.; Chen, H.; McKenna, A. M.; Corilo, Y. E.; Headley, J. V. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Characterization of Athabasca Oil Sand Process-Affected Waters Incubated in the Presence of Wetland Plants. *Energy and fuels* **2017**, *31*, 1731-1740.

(81) Ajaero, C.; Peru, K. M.; Hughes, S. A.; Chen, H.; McKenna, A. M.; Corilo, Y. E.; McMartin, D. W.; Headley, J. V. Atmospheric Pressure Photoionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Characterization of Oil Sand Process-Affected Water in Constructed Wetland Treatment. *Energy and fuels* **2019**, *33*, 4420-4431.

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

(88) Hertkorn, N.; Benner, R.; Frommberger, M.; Schmitt-Kopplin, P.; Witt, M.; Kaiser, K.; Kettrup, A.; Hedges, J. I. Characterization of a major refractory component of marine dissolved organic matter. *Geochimica Et Cosmochimica Acta* **2006**, *70*, 2990-3010.

(89) Wozniak, A. S.; Bauer, J. E.; Sleighter, R. L.; Dickhut, R. M.; Hatcher, P. G. Technical Note: Molecular characterization of aerosol-derived water soluble organic carbon using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Atmospheric Chemistry and Physics* **2008**, *8*, 5099-5111.

(90) D'Andrilli, J.; Foreman, C. M.; Marshall, A. G.; McKnight, D. M. Characterization of IHSS Pony Lake fulvic acid dissolved organic matter by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and fluorescence spectroscopy. *Organic Geochemistry* **2013**, *65*, 19-28.

(91) Minor, E. C.; Swenson, M. M.; Mattson, B. M.; Oyler, A. R. Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environmental science. Processes & impacts* **2014**, *16*, 2064-2079.

(95) Martins, N.; Jimenez-Morillo, N. T.; Freitas, F.; Garcia, R.; Gomes da Silva, M.; Cabrita, M. J. Revisiting 3D van Krevelen diagrams as a tool for the visualization of volatile profile of varietal olive oils from Alentejo region, Portugal. *Talanta* **2020**, *207*, 120276.

(96) Wu, Z.; Rodgers, R. P.; Marshall, A. G. Two- and three-dimensional van krevelen diagrams: a graphical analysis complementary to the kendrick mass plot for sorting elemental compositions of complex organic mixtures based on ultrahigh-resolution broadband fourier transform ion cyclotron resonance mass measurements. *Analytical chemistry* **2004**, *76*, 2511-2516.

(103) Kemsley, E. K. Discriminant analysis of high-dimensional data: A comparison of principal components analysis and partial least squares data reduction methods. *Chemometrics and Intelligent Laboratory Systems* **1996**, *33*, 47-61.

(106) Abdi, H.; Williams, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* **2010**, *2*, 433-459.

(107) Ringner, M. What is principal component analysis? *Nat Biotechnol* **2008**, *26*, 303-304.

(108) Callao, M. P.; Ruisanchez, I. An overview of multivariate qualitative methods for food fraud detection. *Food Control* **2018**, *86*, 283-293.

(109) Marti, M. P.; Busto, O.; Guasch, J. Application of a headspace mass spectrometry system to the differentiation and classification of wines according to their origin, variety and ageing. *Journal of Chromatography A* **2004**, *1057*, 211-217.

(110) Rubert, J.; Lacina, O.; Zachariasova, M.; Hajslova, J. Saffron authentication based on liquid chromatography high resolution tandem mass spectrometry and multivariate data analysis. *Food Chem* **2016**, *204*, 201-209.

(112) Kenar, A.; Çiçek, B.; Arslan, F. N.; Akin, G.; Elmas, Ş. N. K.; Yilmaz, I. Electron Impact–Mass Spectrometry Fingerprinting and Chemometrics for Rapid Assessment of Authenticity of Edible Oils Based on Fatty Acid Profiling. *Food Analytical Methods* **2019**, *12*, 1369-1381.

(113) Corilo, Y. E.; Podgorski, D. C.; McKenna, A. M.; Lemkau, K. L.; Reddy, C. M.; Marshall, A. G.; Rodgers, R. P. Oil spill source identification by principal component analysis of electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Analytical chemistry* **2013**, *85*, 9064-9069.

(114) Karpuzcu, M. E.; Fairbairn, D.; Arnold, W. A.; Barber, B. L.; Kaufenberg, E.; Koskinen, W. C.; Novak, P. J.; Rice, P. J.; Swackhamer, D. L. Identifying sources of emerging organic contaminants in a mixed use watershed using principal components analysis. *Environmental science. Processes & impacts* **2014**, *16*, 2390-2399.

(116) Gaspari, M.; Verhoeckx, K. C.; Verheij, E. R.; van der Greef, J. Integration of two-dimensional LC-MS with multivariate statistics for comparative analysis of proteomic samples. *Analytical chemistry* **2006**, *78*, 2286-2296.

(117) Wang, X.; Chambers, M. C.; Vega-Montoto, L. J.; Bunk, D. M.; Stein, S. E.; Tabb, D. L. QC metrics from CPTAC raw LC-MS/MS data interpreted through multivariate statistics. *Analytical chemistry* **2014**, *86*, 2497-2509.

(118) Wang, C.; Kong, H.; Guan, Y.; Yang, J.; Gu, J.; Yang, S.; Xu, G. Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. *Analytical chemistry* **2005**, *77*, 4108-4116.

(119) Tsugawa, H.; Tsujimoto, Y.; Arita, M.; Bamba, T.; Fukusaki, E. GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC bioinformatics* **2011**, *12*, 131.

(120) Kiss, A.; Lucio, M.; Fildier, A.; Buisson, C.; Schmitt-Kopplin, P.; Cren-Olive, C. Doping control using high and ultra-high resolution mass spectrometry based non-targeted metabolomics—a case study of salbutamol and budesonide abuse. *PLoS One* **2013**, *8*, e74584.

(121) Dill, A. L.; Eberlin, L. S.; Zheng, C.; Costa, A. B.; Ifa, D. R.; Cheng, L.; Masterson, T. A.; Koch, M. O.; Vitek, O.; Cooks, R. G. Multivariate statistical differentiation of renal cell carcinomas based on lipidomic analysis by ambient ionization imaging mass spectrometry. *Analytical and bioanalytical chemistry* **2010**, *398*, 2969-2978.

(122) Alexandrov, T. MALDI imaging mass spectrometry: statistical data analysis and current computational challenges. *BMC bioinformatics* **2012**, *13*, S11.

(123) Peres, F. A. P.; Fogliatto, F. S. Variable selection methods in multivariate statistical process control: A systematic literature review. *Computers & Industrial Engineering* **2018**, *115*, 603-619.

(124) Zou, H.; Hastie, T.; Tibshirani, R. Sparse principal component analysis. *Journal of Computational and Graphical Statistics* **2006**, *15*, 265-286.

(125) Filzmoser, P.; Gschwandtner, M.; Todorov, V. Review of sparse methods in regression and classification with application to chemometrics. *Journal of Chemometrics* **2012**, *26*, 42-51.

(132) Anderson, T. A.; Salice, C. J.; Erickson, R. A.; McMurry, S. T.; Cox, S. B.; Smith, L. M. Effects of landuse and precipitation on pesticides and water quality in playa lakes of the southern high plains. *Chemosphere* **2013**, *92*, 84-90.

(230) Nicol, E.; Varga, Z.; Vujovic, S.; Bouchonnet, S. Laboratory scale UV-visible degradation of acetamiprid in aqueous marketed mixtures - Structural elucidation of photoproducts and toxicological consequences. *Chemosphere* **2020**, *248*, 126040.

(285) Nicol E; Xu Y; Varga Z; Grosshans R; Lavielle M; S., B. Variability and subjectivity in analytical chemistry. *Results presented at the International Winter School on Mass Spectrometry* **2019**.

(286) Nicol, E.; Xu, Y.; Varga, Z.; Kinani, S.; Bouchonnet, S.; Lavielle, M. SPIX: A new software package to reveal chemical reactions at trace amounts in very complex mixtures from high-resolution mass spectra dataset. *Rapid communications in mass spectrometry : RCM* **2021**, *35*, e9015.

(287) Hollender, J.; van Bavel, B.; Dulio, V.; Farmen, E.; Furtmann, K.; Koschorreck, J.; Kunkel, U.; Krauss, M.; Munthe, J.; Schlabach, M.; Slobodnik, J.; Stroomberg, G.; Ternes, T.; Thomaidis, N. S.; Togola, A.; Tornero, V. High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management. *Environmental Sciences Europe* **2019**, *31*.

(288) Giorio, C.; Bortolini, C.; Kourtchev, I.; Tapparo, A.; Bogialli, S.; Kalberer, M. Direct target and non-target analysis of urban aerosol sample extracts using atmospheric pressure photoionisation high-resolution mass spectrometry. *Chemosphere* **2019**, *224*, 786-795.

(289) Hoh, E.; Dodder, N. G.; Lehotay, S. J.; Pangallo, K. C.; Reddy, C. M.; Maruya, K. A. Nontargeted comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry method and software for inventorying persistent and bioaccumulative contaminants in marine environments. *Environ Sci Technol* **2012**, *46*, 8001-8008.

(290) Guo, J.; Chen, D.; Potter, D.; Rockne, K. J.; Sturchio, N. C.; Giesy, J. P.; Li, A. Polyhalogenated carbazoles in sediments of Lake Michigan: a new discovery. *Environ Sci Technol* **2014**, *48*, 12807-12815.

(291) Schymanski, E. L.; Singer, H. P.; Longree, P.; Loos, M.; Ruff, M.; Stravs, M. A.; Ripolles Vidal, C.; Hollender, J. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. *Environ Sci Technol* **2014**, *48*, 1811-1818.

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

(293) Kinani, A.; Kinani, S.; Bouchonnet, S. Formation and determination of organohalogen by-products in water. Part III. Characterization and quantitative approaches. *Trac-Trend Anal Chem* **2016**, *85*, 295-305.

(294) Schymanski, E. L.; Singer, H. P.; Slobodnik, J.; Ipolyi, I. M.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund, P.; Letzel, T.; Grosse, S.; Thomaidis, N. S.; Bletsou, A.; Zwiener, C.; Ibanez, M.; Portoles, T.; de Boer, R.; Reid, M. J.; Onghena, M.; Kunkel, U.; Schulz, W.; Guillon, A.; Noyon, N.; Leroy, G.; Bados, P.; Bogialli, S.; Stipanicev, D.; Rostkowski, P.; Hollender, J. Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. *Analytical and bioanalytical chemistry* **2015**, *407*, 6237-6255.

(295)

https://www.ineris.fr/sites/ineris.fr/files/contribution/Documents/AQUAREF_To_gola_NTS.pdf.

(296) Kew, W.; Blackburn, J. W.; Clarke, D. J.; Uhrin, D. Interactive van Krevelen diagrams - Advanced visualisation of mass spectrometry data of complex mixtures. *Rapid communications in mass spectrometry : RCM* **2017**, *31*, 658-662.

(297) Massart, D. L.; Vandeginste, B. G. M.; Deming, S. M.; Michotte, Y.; Kaufman, L.: *Regression Methods in Chemometrics : a textbook*, 1988. pp. 165-182.

(298) Jurs, P. C. Pattern recognition used to investigate multivariate data in analytical chemistry. *Science* **1986**, *232*, 1219-1224.

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

(300) Wu, H.; Xue, R.; Tang, Z.; Deng, C.; Liu, T.; Zeng, H.; Sun, Y.; Shen, X. Metabolomic investigation of gastric cancer tissue using gas chromatography/mass spectrometry. *Analytical and bioanalytical chemistry* **2010**, *396*, 1385-1395.

(301) Thevenot, E. A.; Roux, A.; Xu, Y.; Ezan, E.; Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J Proteome Res* **2015**, *14*, 3322-3335.

(302) Kirwan, J. A.; Weber, R. J.; Broadhurst, D. I.; Viant, M. R. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Sci Data* **2014**, *1*, 140012.

(303) Mathur, R.; O'Connor, P. B. Artifacts in Fourier transform mass spectrometry. *Rapid communications in mass spectrometry : RCM* **2009**, *23*, 523-529.

(304) Kind, T.; Fiehn, O. Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC bioinformatics* **2007**, *8*, 105.

(305) Goncalves, N. P. F.; Varga, Z.; Bouchonnet, S.; Dulio, V.; Alygizakis, N.; Dal Bello, F.; Medana, C.; Calza, P. Study of the photoinduced transformations of maprotiline in river water using liquid chromatography high-resolution mass spectrometry. *The Science of the total environment* **2021**, *755*, 143556.

(306) Thurman, E. M.: Amount of Organic Carbon in Natural Waters. In *Organic Geochemistry of Natural Waters*; Springer Netherlands: Dordrecht, 1985; pp 7-65.

CONCLUSION GENERALE

L'étude des contaminants d'intérêt émergent (CECs) est un sujet d'actualité mondiale. A l'heure actuelle, seuls quelques produits de dégradation des CECs sont recherchés dans les échantillons environnementaux. La plupart des produits de dégradation n'ayant pas encore été identifiés, il est impossible de les inclure dans des analyses de routine qui sont le plus souvent ciblées. Le travail présenté dans ce manuscrit a majoritairement porté sur l'identification de produits de dégradation de CECs par spectrométrie de masse à haute résolution. Pour cela, deux approches ont été utilisées et comparées : la modélisation en laboratoire et l'analyse directe de mélanges complexes.

La première partie de ces travaux, présentée au chapitre III, a consisté à modéliser la photodégradation en laboratoire de 4 polluants d'intérêt émergent : le cyprodinil, la benzisothiazolinone, l'acétamipride et le naproxène. La modélisation consiste à étudier la dégradation du polluant en eau ultra pure, puis à augmenter progressivement la complexité du milieu pour approcher les conditions environnementales et étudier l'effet matrice sur la dégradation. Cette approche, longue et fastidieuse, a néanmoins fait ses preuves au laboratoire depuis de nombreuses années ; elle a permis d'identifier des produits de dégradation dans des échantillons réels et d'élucider les mécanismes de formation de ces composés. Pour chacun des 4 composés, les produits de dégradation ont été caractérisés par spectrométrie de masse en tandem à haute résolution, avec introduction directe ou en couplage avec différentes techniques chromatographiques. Les résultats ont été interprétés « manuellement », c'est-à-dire par l'opérateur, sur la base de ses connaissances de la chimie des ions en phase gazeuse.

Ces études ont permis de caractériser plus d'une dizaine de nouveaux photoproduits pour chacun des composés étudiés et de répertorier des mécanismes récurrents, débutant pour la plupart par la perte de l'halogène lorsque la molécule est chlorée - rupture homolytique - suivie de réactions dont les plus courantes sont l'hydroxylation, la cyclisation et l'isomérisation.

Des études de toxicité *in silico* ont été effectuées par le biais du logiciel T.E.S.T. afin d'estimer la toxicité individuelle des photoproduits. Des études d'écotoxicité des solutions avant et pendant la photolyse ont été conduites au moyen de tests *in vitro* d'extinction de bioluminescence de bactéries marines *Vibrio fischeri*. Ces études *in silico* et *in vitro* ont démontré que la plupart des photoproduits présentaient une toxicité comparable ou

supérieure à la molécule mère, démontrant la pertinence de mener ce type d'étude sur les CECs, leurs métabolites et produits de dégradation. Ceci dit, les tests *in silico* utilisés lors de l'étude de la dégradation du naproxène ont montré leur limite avec, pour l'un des composés, une toxicité estimée deux fois supérieure à celle obtenue expérimentalement. Il est donc important d'être vigilant et de garder à l'esprit que les toxicités obtenues *via* les méthodes *in silico* sont des toxicités estimées et non réelles et que les paramétrages du logiciel utilisés peuvent biaiser les résultats.

Lors de la photodégradation de formulations commerciales contenant de l'acétamipride, les photoproduits identifiés à l'issue de l'irradiation du principe actif en solution aqueuse ont été retrouvés ; un nouveau photoproduit a pu être mis en évidence, caractérisant un effet matrice significatif - même au sein d'un mélange relativement simple - et démontrant la nécessité de compléter les études de modélisation par des études sur des échantillons naturels.

La photolyse UV-Visible n'est pas la seule voie de dégradation des polluants en milieu naturel. Il conviendrait d'étudier également l'effet d'autres mécanismes de dégradation (hydrolyse, dégradation thermique, réaction catalysées par les métaux, etc.) et ce dans différents environnements. Il n'est pas envisageable d'appliquer systématiquement une approche modélisée pour l'étude exhaustive de chaque polluant d'intérêt émergent. Il est donc indispensable de disposer d'outils pour le *screening* du devenir de polluants dans l'environnement afin de limiter les risques de faux négatifs vis-à-vis de produits de dégradation potentiellement toxiques pour l'être humain ou pour l'environnement.

Avec cet objectif, nous avons développé, en collaboration avec l'INRIA et le CMAP, un logiciel innovant de comparaison d'ensembles de spectres de masses complexes enregistrés en HRMS avec introduction directe. Ce logiciel, baptisé SPIX, utilise des outils statistiques tels que les tests de Student et de Wilcoxon ; il est présenté au chapitre IV de ce mémoire. SPIX permet l'identification de réactifs, d'intermédiaires réactionnels et de produits à l'état de traces au sein de mélanges complexes. A notre connaissance, il s'agit du premier logiciel permettant le suivi de l'évolution de la composition d'un échantillon avec des propositions de modèles cinétiques. Cette capacité de SPIX à identifier des variations significatives du milieu à partir de jeux de spectres HRMS le rend

particulièrement adapté à l'étude du devenir des polluants dans des échantillons environnementaux.

La modélisation en laboratoire et l'analyse directe avec gestion statistique des résultats constituent deux approches complémentaires. Une analyse directe d'échantillons complexes peut être réalisée dans un premier temps, afin d'établir si les échantillons ont évolué dans des proportions qui méritent que l'on s'y intéresse plus en détail. Le cas échéant, la modélisation avec suivi en LC-HRMS/MS pourra permettre de caractériser les produits de dégradation en travaillant à des concentrations compatibles à la réalisation d'expériences de MS/MS.

En parallèle à ces travaux, des études de comparaisons visuelles ont été réalisées, qui démontrent le rôle déterminant de la subjectivité de l'opérateur lors de l'interprétation de résultats issus de données MS, à tel point que cette subjectivité peut conduire à des interprétations complètement différentes, non seulement entre deux utilisateurs, mais aussi pour un même utilisateur à qui l'on confie les mêmes données à deux instants différents. Le logiciel SPIX, en plus des fonctionnalités précédemment décrites, contribue à réduire très significativement la subjectivité de l'opérateur lors de l'interprétation des résultats expérimentaux.

La suite de ces études consistera à analyser des échantillons de plus en plus complexes en associant les deux approches susmentionnées. Une campagne de tests avec diverses solutions d'acétamipride et de formulations commerciales à différentes concentrations est en cours sur des tomates cerises plantées au laboratoire et placées sous lampe UV jusqu'à maturation. L'objectif est d'étudier la cinétique de dégradation de l'acétamipride sur la peau des fruits, pour voir si de nouveaux produits de dégradation sont engendrés en comparaison de ceux formés lors de la modélisation en milieux aqueux. Cette étude permettra également d'établir si le rinçage à l'eau du fruit permet l'élimination des produits de dégradation les plus toxiques ou si la peau des tomates est perméable aux

différents photoproduits ; nous avons précédemment montré que c'était le cas pour les photoproduits du thiophanate de méthyle.²⁷

SPIX sera également utilisé pour diverses analyses en cours, au laboratoire et en collaboration avec d'autres centres de recherche, pour des analyses de mélanges complexes et le suivi de leur évolution au cours du temps. Ce logiciel est un outil puissant qui a déjà montré ses atouts, mais qui vise à être constamment amélioré grâce aux retours de la communauté scientifique. Les développements en cours portent sur l'ajout de la dimension chromatographique : les comparaisons des abondances des ions seront alors réalisées sur des jeux de données à deux dimensions (temps de rétention ; rapports m/z)

²⁷ Chayata, H.; Lassalle, Y.; Nicol, E.; Manolikakes, S.; Souissi, Y.; Bourcier, S.; Gosmini, C.; Bouchonnet, S. Characterization of the ultraviolet-visible photoproduits of thiophanate-methyl using high performance liquid chromatography coupled with high resolution tandem mass spectrometry-Detection in grapes and tomatoes. *Journal of chromatography. A* **2016**, *1441*, 75-82.

Titre : Etude de la dégradation photochimique de contaminants en matrices environnementales - Modélisation en laboratoire versus analyses directes.

Mots clés : Contaminants, photodégradation, spectrométrie de masse à très haute résolution, matrices biologiques, écotoxicologie

Résumé : L'étude des contaminants émergents (CECs) est un sujet d'actualité mondiale. Actuellement seuls quelques produits de dégradation des CECs sont recherchés dans les échantillons environnementaux, la majorité d'entre eux n'ayant pas encore été identifiés. Le travail présenté dans ce manuscrit porte sur l'identification de produits de dégradation de CECs par spectrométrie de masse à haute résolution par deux approches : la modélisation en laboratoire et l'analyse directe de mélanges complexes.

Ce manuscrit présente la modélisation en laboratoire de la photodégradation de 4 CECs : le cyprodinil, la benzisothiazolinone, l'acétamipride et le naproxène. Cette approche a permis pour chacun des 4 composés de caractériser plus d'une dizaine de produits de dégradation et de répertorier des mécanismes récurrents. Les résultats des études de toxicité *in silico* et *in vitro* ont démontré que la plupart des photoproduits ont une toxicité comparable ou supérieure à la molécule mère. L'étude des formulations commerciales de l'acétamipride démontre un effet matrice significatif, témoignant de la nécessité de compléter les études de modélisation par des études en

milieu réel. Cependant, il n'est pas envisageable d'appliquer systématiquement une approche modélisée pour l'étude exhaustive de chaque polluant d'intérêt émergent. Une approche plus rapide, de type screening, doit être envisagée.

La deuxième partie de ce manuscrit présente un logiciel innovant, développé en collaboration avec l'INRIA et le CMAP, de comparaison d'ensembles de spectres de masses complexes. Ce logiciel, baptisé SPIX, permet l'identification de réactifs, d'intermédiaires réactionnels et de produits à l'état de traces au sein de mélanges complexes. A notre connaissance, il s'agit du premier logiciel permettant le suivi de l'évolution de la composition d'un échantillon avec des propositions de modèles cinétiques. Cette capacité de SPIX à identifier des variations significatives du milieu à partir de jeux de spectres HRMS le rend particulièrement adapté à l'étude du devenir des polluants dans des échantillons environnementaux.

Ces deux approches sont décrites et comparées pour l'identification de produits de dégradation de CECs.

Title : Photoinduced degradation of contaminants in environmental matrices - Laboratory modelisation versus direct analysis.

Keywords : Contaminants, photodegradation, ultra high resolution mass spectrometry, biological matrices, ecotoxicology

Abstract: The study of emerging contaminants (CECs) is a worldwide concern. Currently, only a few CEC degradation products are detected in environmental samples, most of them have not been identified yet. The work presented in this manuscript focuses on the identification of CEC degradation products by high resolution mass spectrometry using two approaches: laboratory modelling and direct analysis of complex mixtures.

This manuscript presents the laboratory modelling of the photodegradation of 4 CECs: cyprodinil, benzisothiazolinone, acetamiprid and naproxen. This approach led to the characterization of more than ten degradation products for each of the 4 compounds and to identify recurrent mechanisms. The results of the *in silico* and *in vitro* toxicity studies showed that most of the photoproducts have a toxicity similar or superior to the parent compound. The study of commercial formulations of acetamiprid shows a significant matrix effect, highlighting the need to complement modelling studies

with real-life studies. However, it is not an option to systematically apply a model-based approach to the comprehensive study of each emerging pollutant of interest. A faster approach, such as screening, should be considered.

The second part of this manuscript presents an innovative software, developed in collaboration with INRIA and CMAP, for the comparison of sets of complex mass spectra. This software, named SPIX, allows the identification of reactants, reaction intermediates and trace products in complex mixtures. To our knowledge, it is the first software allowing the monitoring of the evolution of the composition of a sample with kinetic models' proposals. This ability of SPIX to identify significant variations in the environment from sets of HRMS spectra makes it particularly adapted to the study of the fate of pollutants in environmental samples.

These two approaches are described and compared for the identification of CEC degradation products.