

ÉCOLE DOCTORALE
SANTÉ, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT

Institut de Chimie Organique et Analytique

THÈSE présentée par :
Laëtitia FOUGERE

soutenue le : **7 Mai 2019**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : Chimie Analytique

Signature moléculaire de milieux complexes :
Stratégies de couplage à la spectrométrie de
masse et d'interprétation des données

THÈSE dirigée par :

Emilie DESTANDAU
Claire ELFAKIR

Professeur, Université d'Orléans
Professeur, Université d'Orléans

RAPPORTEURS :

Carlos AFONSO
Sophie TOMASI

Professeur, Université de Rouen
Professeur, Université Rennes 1

JURY :

Carlos AFONSO
Arnaud LANOUE
Nicolas SOMMERER
Sophie TOMASI
Emilie DESTANDAU
Claire ELFAKIR

Professeur, Université de Rouen
Maître de Conférences, Université de Tours
Ingénieur de recherche, INRA Montpellier
Présidente, Professeur, Université Rennes 1
Professeur, Université d'Orléans
Professeur, Université d'Orléans

Remerciements

Je tiens tout d'abord à remercier Pr. Sophie Tomasi de l'Université de Rennes et Pr. Carlos Afonso de l'Université de Rouen qui m'ont fait l'honneur de juger ce travail en qualité de rapporteurs. Je remercie également Dr. Nicolas Sommerer et Dr. Arnaud Lanoue pour avoir accepté d'être examinateur de ce travail. Leurs interventions lors de congrès et certains de nos échanges m'ont permis d'aller plus loin en me donnant de nouvelles idées.

Je suis reconnaissante envers les deux directeurs de l'école doctorale Luigi Agrofoglio et Agnès Delmas d'avoir accepté que je réalise cette thèse en parallèle de mon travail d'ingénieur d'études en posant des conditions, qui ont permis que ces travaux de thèse soient équivalents à tous travaux de doctorants. Je remercie également Véronique Thénot de son aide en tant que secrétaire de l'école doctorale.

Mes remerciements vont également vers M. Olivier Martin et Pascal Bonnet, directeurs de l'Institut de Chimie Organique et Analytique, de m'avoir permis de réaliser cette thèse dans le cadre de mon travail, en espérant apporter une plus-value à l'institut.

J'exprime toute ma gratitude et ma plus profonde reconnaissance à mes deux directrices de thèse Claire Elfakir et Emilie Destandau, d'avoir accepté de me suivre dans cette folle aventure et d'avoir fait preuve de tant de compréhension dans cette rédaction. Claire, merci d'avoir toujours cru en moi et dès le début de ma carrière, de m'avoir en partie lancée dans cette grande aventure, merci de m'avoir aidée à écrire cette belle histoire. Votre clairvoyance et votre amour pour la science donnent envie de discuter et de travailler à vos côtés. Emilie, ta générosité de cœur et d'esprit font de toi une directrice exceptionnelle. Il est simple de travailler avec toi sur des sujets qui nous animent toutes les deux.

Je tiens à remercier tout particulièrement David Da Silva pour toutes les longues, voire très longues, discussions. Ton amour irréprochable pour le MALDI/ToF MS n'a pas été contagieux, mais bien utile pour réaliser ces couplages entre la TLC et la spectrométrie de masse.

Un grand merci aussi à Benoit Maunit, pour m'avoir insufflé l'idée que faire une thèse était également possible pour moi. Son amour pour la spectrométrie de masse est aussi très grand.

Un grand merci à tous mes collègues de l'ICDA de m'avoir aidée tout au long de cette thèse : Florian pour ces discussions de chantier et ta gentillesse (Un grand merci pour Totoro), Thomas, bien que loin, j'adore toujours travailler, discuter et boire des verres avec toi, Caroline, Pascal et Stéphane pour le temps passé, vos conseils et les cours en traitements de données, Richard pour ton aide dans mes dossiers, Mathieu notre rugbyman en synthèse, M. Martin et Chrystel pour leur aide dans une structure de sucre, Cyril pour son expertise en haute résolution. Mes pensées vont également aux étudiants avec qui j'ai partagé des joies, des doutes, des attentes, des envies (dont de Kebab) lors de nos parcours de doctorants : Elise, Synthia, Justine, Adrien, Nhi, Thibault, Souhila, Jérémy et aussi, même si elles ne sont plus étudiantes : Claire, Gaëlle, Angéline et Rouba. Un grand merci plus particulièrement, à Sandrine, ma "mamie" du bureau, mais où est donc passé ta tête ? Tu as su faire preuve de patience quand j'avais mes moments de stress et de concentration et tu m'as toujours proposé ton aide. Merci pour tes magnifiques dosages de CD, ton énorme saisie de données pour pouvoir construire le diagramme de Van Krevelen et d'avoir dompter le logiciel rTLC. La patience est de rigueur avec le TR.

A tous ceux qui ont croisé mon chemin, vous m'avez tous permis de m'améliorer d'une façon ou d'une autre, vous aussi vous avez laissé une empreinte :

Béatrice Rhino pour ce beau sujet de co-culture ;

Michel Pinault pour avoir travaillé ensemble dans ton domaine, heureuse de t'avoir aidé dans l'analyse des stéroïdes par TLC/MS ;

Ainsi que tous les stagiaires, dont leurs participations ont permis de réaliser ces beaux projets : Laura, Vanille, Gaëlle, Cécile, ...

A toutes les personnes avec qui je collabore en parallèle de ma thèse :

- Thomas, Jérémy, Tiago et Lauriane dans le cadre d'analyse de drogues et de médicaments dans les eaux et les carottes sédimentaires ;
- Marius et Benoit dans la dégradation de principes actifs ;
- Christoph Poincilié dans votre combat pour lutter contre la maladie orpheline de Niemann pick et Fabrizio Montecucco et Serge Rudaz pour l'aider
- Célia et Franck dans l'analyse et la caractérisation de cyclodextrines ;
- Thomas, Nicolas, Jean-Paul, et tous ceux du projet chêne et vin ;
- Sylvain et Manhattan dans l'analyse des exsudats racinaires ;
- Patrice et toutes les personnes de l'entreprise TIC dans les empreintes de plantes finlandaises.

Ces remerciements vont également auprès de mon équipe de choc d'amis, ensemble depuis des années, et notre groupe s'agrandit avec nos enfants, merci à vous : Céline qui adore les enfants, Julien le bricoleur pâtissier, Coralie l'amoureuse des chevaux, Julien notre grand rêveur, Pierre notre cuisinier écrivain, Cindy et sa fibre sociale, Emmanuelle la grande maman, Fabrice notre super informaticien.

Et enfin, mes dernières pensées vont vers ma famille, mes parents, ma mamie, à Frédérique ma coach, mais surtout à mes trois hommes : Fabrice, Antonin & Maxence, qui ont marqué ma vie personnelle de vos magnifiques empreintes et qui ont su m'aider à tout moment avec vos sublimes sourires.

Publications dans les journaux internationaux à comités de lectures

(1) L. Fougère, D. Da Silva, E. Destandau, C. Elfakir

TLC-MALDI-TOF-MS based identification of flavonoid compounds using an inorganic matrix
Phytochemical analysis, 2019, 30 (2), 218-225

Received: 29 April 2018 | Revised: 15 October 2018 | Accepted: 23 October 2018
 DOI: 10.1002/pca.2807

RESEARCH ARTICLE

WILEY **Phytochemical
Analysis**

TLC-MALDI-TOF-MS-based identification of flavonoid compounds using an inorganic matrix

Laëtitia Fougère | David Da Silva | Emilie Destandau  | Claire Elfakir

University of Orléans, CNRS, ICOA, UMR
 7311, F-45067 Orléans, France

Correspondence
 Emilie Destandau, University of Orléans,
 CNRS, ICOA, UMR 7311, F-45067 Orléans,
 France.
 Email: emilie.destandau@univ-orleans.fr

Abstract

Introduction: Thin layer chromatography (TLC) is frequently used to obtain the fingerprint of a plant extract. Although the retardation factor and the response to visualisation give primary information about compound identification, the direct TLC-mass spectrometry (MS) coupling allows a more detailed characterisation of samples.

Objectives: To demonstrate the potential for the flavonoid dereplication using an inorganic matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) method with and without TLC separation.

Material and Methods: Samples derived from wine, apple or rose were deposited on an aluminium-backed silica gel TLC sheet compatible with the MS adapter. Unlike the wine sample, for apple and rose samples compound derivatisation was necessary. These two samples were deposited twice and the plate was cut in two parts. One half was oversprayed with Neu-Peg reagent to visualise flavonoids while the inorganic matrix was deposited on each flavonoid zone on the second half for MS ionisation.

Results: Mass spectra obtained for samples without plate development showed numerous ions corresponding to glycosylated flavonoids. The lower m/z observed could be due either to aglycone flavonoids or to in-source fragment ions. After plate development, a separation of many spots was observed and each spot was analysed separately leading to a deeper identification of the present flavonoids. Moreover, isobaric flavonoids with different hR_f values could be differentiated.

Conclusion: TLC-MALDI-TOF-MS using an inorganic matrix enabled the analysis of anthocyanins in positive mode and of flavonols, flavanols, dihydrochalcones and phenolic acids in negative mode, reducing adduct, aggregate forms giving thus simple and reliable spectra for the dereplication approach of flavonoids in complex samples.

KEYWORDS

dereplication, flavonoids, Inorganic matrix, MALDI-TOF MS, TLC

1 | INTRODUCTION

Flavonoids are the most common group of plant polyphenols and provide much of the flavour and colour of fruits and vegetables. Many flavonoids exhibit antioxidant, antibacterial, antifungal, and antiviral activities.¹ The increasing use of plant extracts in the pharmaceutical, food and cosmetic industries has stimulated the study of these

metabolites. Flavonoids (the aglycones and their various glycosides) are divided into several subgroups: anthocyanins (e.g. malvidin), flavonols (e.g. quercetin), dihydrochalcones (e.g. phloretin), flavanols (e.g. catechin), flavanones (e.g. naringenin), flavones (e.g. apigenin) and chalcones (e.g. butein).^{2,3}

Recognition of the presence of previously undescribed flavonoids in a sample is facilitated by efficient dereplication techniques that

(2) L. Fougère, B. Rhino, E. Destandau, C. Elfakir

UHPLC/MS targeted metabolic study of flavonoids in corn silks as a part of varietal selection for trap plant
En cours

Communications Orales

Congrès internationaux

- (1) **L. Fougère**, S. Messaili, C. Colas, B. Rhino, Y. Qu, N. Desneux, A.-V. Lavoit, T. Michel, E. Destandau, C. Elfakir
Non-targeted and targeted metabolomics approaches in the tomato protection from pests
3^{ème} symposium international AFERP-STOLON, SympAS3 2018, Rennes

Congrès nationaux

- (1) **L. Fougère**, B. Rhino, E. Destandau, C. Elfakir
Sélection variétale de plants de maïs par leur signature moléculaire
1^{er} Ecole printanière du RFMF junior 2017, Saint Lambert des Bois
- (2) **L. Fougère**, E. Destandau, B. Rhino, C. Elfakir
Evaluation de la teneur en flavonoïdes-C-Glycosylés exprimé dans les soies de maïs par couplage UHPLC/MS
Club Jeunes JFSM 2016, Bordeaux
- (3) **L. Fougère**, D. Da Silva, E. Destandau, B. Maunit, C. Elfakir
Les Flavonoïdes: Caractérisation par TLC@MALDI-MS
16^{ème} journées utilisateurs BRUKER 2014, Paris

Flash Poster

- (1) **L. Fougère**, B. Rhino, E. Destandau, C. Elfakir
Molecular signature: Help in selection of variety for biocontrol
5^{èmes} Journées internationales de l'AFERP, 2017, Angers

Communications par affiches

Congrès internationaux

- (1) D. Da Silva, **L. Fougère**, C. Guimaraes, C. Goupille, B. Maunit, S. Chevalier, M. Pinault
Analysis of sterols and steroids using HPTLC-MS: influence of ionization parameters
International Symposium for High-Performance Thin-Layer Chromatography, HPTLC 2017, Berlin (Allemagne)
- (2) **L. Fougère**, D. Da Silva, E. Destandau, C. Elfakir
HPTLC-Maldi/TOF/MS vs HPLC-ESI/MS² for identification to flavonoids in plants extracts
9th Joint Natural Products Conference 2016, Copenhagen
- (3) **L. Fougère**, A. Chartier, B. Rhino, E. Destandau, C. Elfakir
Analysis of elicitors in different varieties of corn silk
9th Joint Natural Products Conference 2016, Copenhagen

Congrès nationaux

- (1) **L. Fougère**, D. Da Silva, E. Destandau, C. Elfakir
Profil métabolique et analyse ciblée : tout est permis avec la Cghromatographie sur Couche Mince
13^{ème} Congrès francophone sur les Sciences Séparatives et les Couplages de l'AFSEP, 2019, Paris
- (2) **L. Fougère**, D. Da Silva, C. Guimaraes, C. Goupille, B. Maunit, S. Chevalier, M. Pinault
Analysis of sterols and steroids using HPTLC-MS: influence of ionization parameters
Spectrométrie de masse, Métabolomique et Fluxomique & Electrophorèse et Analyse Protéomique, SMMAP 2017, Marne-la-Vallée
- (3) **L. Fougère**, E. Destandau, B. Rhino, C. Elfakir
Métabolomique ciblée par UHPLC/MS pour l'aide à la sélection variétale de plante piège
Spectrométrie de masse, Métabolomique et Fluxomique & Electrophorèse et Analyse Protéomique, SMMAP 2017, Marne-la-Vallée
- (4) **L. Fougère**, B. Rhino, E. Destandau, C. Elfakir
Sélection variétale de plants de maïs par leur signature moléculaire
1^{ere} Ecole printanière du RFMF junior 2017, Saint Lambert des Bois
- (5) **L. Fougère**, B. Rhino, E. Destandau, C. Elfakir
Contribution of UHPLC/TQ for the characterization and quantification of flavonoids in different varieties of corn silks
Journées Françaises de Spectrométrie de masse, JFSM 2016, Bordeaux
- (6) **L. Fougère**, B. Rhino, E. Destandau, C. Elfakir
Flavones-C-glycosylées indicateurs de l'inhibition du développement des larves de la noctuelle Helicoverpa zea
Ecole thématique Ecologie Chimique, ETEC 2016, Roscoff

Sommaire

Liste des abréviations	11
Liste des figures.....	14
Liste des tableaux.....	21
 INTRODUCTION GENERALE	 23
 Chapitre 1 : Méthodes analytiques et outils de traitements de données utiles pour caractériser et comparer des milieux complexes	 27
1. Techniques analytiques à disposition	29
1.1. Méthodes d'extraction	30
1.2. Méthodes séparatives	31
1.3. Modes de détection	32
1.4. Techniques pour l'identification des métabolites	32
2. Stratégies actuelles de déréplication et de métabolomique	35
2.1. Domaines d'applications.....	35
2.2. Définitions des termes : déréplication et métabolomique	36
2.3. Etapes clés d'une analyse métabolomique	37
3. Outils d'aide à l'identification des métabolites.....	39
3.1. Niveaux d'identification	39
3.2. Bases de données en spectrométrie de masse.....	39
3.3. Bases de données spectrales	41
3.4. Prédiction in Silico de la fragmentation en MS	41
4. Outils d'aide à la comparaison des données MS.....	455
4.1. Analyse statistique.....	455
4.1.1. Les analyses non supervisées.....	466
4.1.2. Les analyses supervisées.....	477
4.2. Représentation graphique des échantillons en fonction d'une propriété de leurs molécules .	488
4.2.1. Représentation de Kendrick.....	488
4.2.2. Diagramme de Van Krevelen.....	500
4.2.3. Réseaux de similarité spectrale (Réseaux moléculaires).....	522
5. Etat de l'art sur la fragmentation des flavonoïdes glycosylés MS.....	544
5.1. Structure chimique des flavonoïdes	555
5.2. Nomenclature générale liée à la fragmentation des flavonoïdes en MS	588

5.3. Fragmentation des flavonoïdes <i>O</i> -glycosylés.....	600
5.3.1. Fragmentation des flavonoïdes mono- <i>O</i> -glycosylés	611
5.3.2. Fragmentation des flavonoïdes mono- <i>O</i> -glycosylés- <i>O</i> -glycosylés . Erreur ! Signet non défini.	2
5.3.3. Fragmentation des flavonoïdes di- <i>O</i> -glycosylés.....	633
5.4. Fragmentation des flavonoïdes <i>C</i> -glycosylés	644
5.4.1. Fragmentation des flavonoïdes mono- <i>C</i> -glycosylés.....	644
5.4.2. Fragmentation des flavonoïdes mono- <i>C</i> -glycosylés- <i>O</i> -glycosylés.....	667
5.4.3. Fragmentation des flavonoïdes di- <i>C</i> -glycosylés	69
5.5. Fragmentation des flavonoïdes di- <i>O</i> , <i>C</i> -glycosylés.....	71
5.6. Bilan de la fragmentation des flavonoïdes glycosylés en MS.....	74
6. Conclusion du chapitre	76
Références bibliographiques	77

Chapitre 2 :Identification moléculaire par TLC : Apports du traitement statistique de données et du couplage à la spectrométrie de masse.....	877
1. Description de la technique séparative TLC.....	89
1.1. Considération générale sur la technique	899
1.2. Mise en œuvre expérimentale de la TLC	90
2. Analyse non ciblée de différents organes du Colza (<i>Brassica napus L.</i>) par TLC	922
2.1. Contexte	93
2.2. Mise en évidence par TLC des différentes familles moléculaires présentes	944
2.2.1. Recherche des polyphénols	944
2.2.2. Recherche des lipides et terpènes	95
2.3. Analyses non supervisées sur différents organes du colza (feuille, racine, tige, graine).....	96
3. Etat de l'art de l'analyse par TLC/MS	102
3.1. Différentes méthodes de couplage.....	102
3.1.1. Principe du couplage TLC/API-MS.....	1022
3.1.2. Principe du couplage TLC/MALDI-Tof MS	1055
3.2. Applications des couplages en phytochimie. Focus sur les flavonoïdes.....	1099
4. Développement du couplage TLC/MALDI-Tof MS pour la caractérisation de flavonoïdes	110
4.1. Optimisation des paramètres MALDI-Tof-MS	111
4.1.1. Besoin ou non d'une matrice pour l'ionisation ?	111
4.1.2. Choix de la matrice	112
4.1.3. Mode de dépôt de la matrice	114
4.2. Mise au point du couplage TLC/MALDI-Tof MS à partir de l'analyse de standards.....	115

4.2.1. Le système TLC pour la séparation des flavonoïdes.....	1155
4.2.2. Couplage du système TLC au MALDI-Tof-MS	1166
4.3. Application à l'analyse d'anthocyanes dans le vin rouge	118
4.4. Intérêt du couplage TLC/MALDI-Tof-MS pour caractériser les flavonoïdes présents dans un extrait de baies d'argousier	122
4.5. Application à l'analyse d'un jus de pomme enrichi en dérivés de la quercétine	124
4.5.1. Caractérisation de dérivés de quercétine	1244
4.5.2. Détermination de la position du sucre sur la quercétine et cas d'isoméries	1266
4.5.3. Détection de composés minoritaires	1266
4.6. Transposition de la méthode TLC/MALDI-Tof MS à l'analyse d'une famille moléculaire de structure proche de celle des flavonoïdes : les xanthones prénylés.....	128
4.6.1 Analyse directe de l'extrait par MALDI-Tof MS avant l'élution	1288
4.6.2 Analyse de l'extrait par couplage TLC/MALDI-Tof-MS	129
5. Etude exploratoire du couplage TLC/API MS	132
5.1. Réponse en ESI-MS de molécules de masse inférieure à 200 Da déposées sur gel de silice	133
5.2. Analyse de molécules peu polaires : les oxystérols et les stéroïdes	134
5.2.1. Evaluation des trois sources d'ionisation	134
5.2.2. Etude du solvant de désorption.....	137
5.2.3. Analyse HPTLC/APCI-TQ MS d'un mélange de 23 standards.....	138
6. Conclusions et perspectives.....	141
Références bibliographiques	144

Chapitre 3 : Approche métabolomique : Comparaison ciblée de la composition en flavonoïdes glycosylés de différentes variétés de soies de maïs par UHPLC/MS.....

1. Etat de l'art	1522
1.1. Les plantes pièges	1522
1.2. Le maïs : Plante piège de l' <i>Helicoverpa zea</i>	1522
1.3. La soie de maïs.....	153
1.4. Composition moléculaire des composés non volatils du maïs et de la soie.....	154
1.4.1. Composés azotés	1544
1.4.2. Composés terpéniques	155
1.4.3. Composés phénylpropanoïdes.....	1555
1.5. Activités des flavonoïdes.....	157
2. Etude métabolomique ciblée des soies de maïs par UHPLC/MS	158
2.1. Sélection et récolte du matériel végétal	1588
2.2. Préparation des échantillons	159
2.2.1. Méthode d'extraction	1599

2.2.2. Etude de répétabilité de la méthode d'extraction	160
2.3. Optimisation des conditions chromatographiques en vue de l'analyse UHPLC/MS/MS des flavonoïdes dans les extraits de soies de maïs.....	162
2.3.1. Choix de la phase mobile	162
2.3.2. Choix de la colonne.....	162
2.3.3. Choix du diamètre de particules du support chromatographique.....	164
2.3.4. Choix de la température de la colonne.....	166
2.3.5. Optimisation du débit de la phase mobile	167
2.4. Sélection des composés d'intérêt à suivre pour l'approche métabolomique ciblée des soies de maïs	168
2.4.1. Analyse totale en UHPLC/HRMS en vue de la détection d'un maximum de composés	168
2.4.2. Etude de fragmentation de quelques flavonoïdes glycosylés standards	1699
2.4.3. Démarche appliquée à l'identification d'une molécule dans l'extrait de soies de maïs.....	172
2.4.4. Construction du diagramme de Van Krevelen	1788
2.4.5. Identification des flavonoïdes dans les extraits de soies de maïs	1800
2.5. Mise en place de la méthode de quantification.....	1866
2.5.1. Optimisation de la méthode MRM	1877
2.5.2. Quantification relative par rapport à un standard	1899
2.6. Traitement de données	190
2.6.1. Prétraitement des données	1911
2.6.2. Classification Ascendante Hiérarchique : CAH	193
2.6.3. Analyse en Composantes Principales (ACP)	1944
2.6.4. sPLS-DA	1955
2.6.5. Interprétation des résultats et comparaison avec les résultats terrain	1966
2.7. Identification par réseaux moléculaires.....	200
2.7.1. Réseaux moléculaires dans les deux modes d'ionisation.....	200
2.7.2. Modification du cosinus score	203
3. Conclusion et perspectives	2066
Références bibliographiques	2077
Conclusions – Perspectives.....	211
 ANNEXES Chapitre 2	 2199
ANNEXES Chapitre 3	2299

Liste des abréviations

AcAm	Acétate d'ammonium
AcOEt	Acétate d'éthyle
AcOH	Acide acétique
ACN	Acétonitrile
ACP	Analyse en Composantes Principales
AF	Acide formique
Ag	Aglycone
ANS	Anthocyanidine synthase
APCI	Ionisation chimique à pression atmosphérique
API	Ionisation sous pression atmosphérique
APPI	Photoionisation à pression atmosphérique
Bdd	Bases de données
BPC	Base peak chromatogram
CAH	Classification Ascendante Hiérarchique
CAS	Chemical abstracts service
CASMI	Critical assessment of small molecule identification
CCM	Chromatographie sur couche mince
CCS	Collision cross section
CFM-ID	Competitive Fragmentation Modeling for Metabolite Identification
ChEBI	Chemical Entities of Biological Interest
CHCl₃	Chloroforme
CHI	Chalcone isomérase
CHR	Chalcone réductase
CHS	Chalcone synthase
CIRAD	Centre de Coopération Internationale en Recherche Agronomique pour le Développement
CPG	Chromatographie en phase gazeuse
CO₂	Dioxyde de carbone
CoA	Coenzyme A
CSI-FingerID	Compound Structure Identification : FingerID
Cv	Coefficient de variation
Da	Dalton
DAD	Détecteur à barrettes de diodes
DART	Analyse directe à temps réel
DBE	Nombre équivalent de doubles liaisons
DDA	Acquisition dépendante des données
DEDL	Détecteur évaporatif à diffusion de lumière
Deox	Déoxyhexose
DESI	Ionisation électrospray par désorption
DF	Détecteur à fluorescence
DFR	Dihydroflavonol 4-reductase
DHB	Acide 2,5-dihydroxybenzoïque
DIBOA-Glc	2-β-D-glucopyranosyloxy-4-hydroxy-1,4-benzooxazin-3-one
DIMBOA-Glc	2-β-D-glucopyranosyloxy-4-hydroxy-7-methoxy-1,4-benzooxazin-3-one
DIM₂BOA-Glc	2-β-D-glucopyranosyloxy-4-hydroxy-7,8-dimethoxy-1,4-benzooxazin-3-one

Liste des abréviations

DNP	Dictionary of Natural Products
EC	Electrophorèse capillaire
ECD	Détecteur électrochimique
ESI	Ionisation par électrospray
EtOH	Ethanol
eV	Electron-Volt
FB	Formule Brute
FLS	Flavonol synhase
FNS	Flavone synthase
FTBLAST	Fragmentation Tree Basic Logic Alignment Search Tool
FTICR	Résonance cyclotronique ionique à transformée de Fourier
F3H	Flavanone 3-hydroxylase
ISIS	In silico identification software
Jav	Variété de maïs Java
Gar	Variété de maïs Garrison
Gb	Variété de maïs Golden bantam
Glc	Glucoside
GNPS	Global natural products social molecular networking
H₂O	Eau
HDMBOA-Glc	2-β-D-glucopyranosyloxy-4,7-diméthoxy-1,4-benzooxazin-3-one
Hex	Hexose
HMBD	The human metabolome database
HPLC	Chromatographie en phase liquide
HPTLC	Chromatographie sur couche mince haute performance
HRMS	Spectrométrie de Masse Haute Résolution
<i>H. zea</i>	<i>Helicoverpa zea</i>
H/C	Rapport atomique entre le nombre d'hydrogène et le nombre de carbone
ICOA	Institut de Chimie Organique et Analytique
IFS	Isoflavone synthase
IPP	Isopentenyl-pyrophosphate
KEGG	Kyoto Encyclopedia of Genes and Genomes
KMD	Défaut de masse de Kendrick
KMN	Masse nominale de Kendrick
LCR	Leucoanthocyanidine reductase
LDI	Désorption/ionisation laser
MALDI	Ionisation sous vide par désorption laser assistée par matrice
MeOH	Méthanol
MoNA	MassBank of North America
MRM	<i>Multiple Reaction Monitoring</i>
MS	Spectrométrie de masse
MSⁿ	<i>Multiple-stage mass spectrometry</i>

Liste des abréviations

MSI	« Metabolomic Standards Initiative »
m/z	Rapport masse/charge
NALDI	Désorption/ionisation laser assistée par nanoparticules
NAPRALERT	Natural Products Alert
NEU	Amino-2-éthyle diphénylborate
NH₄OH	Ammoniaque
Nov	Variété de maïs Nova
N/C	Rapport atomique entre le nombre d'azote et le nombre de carbone
N2C	Laboratoire Nutrition, Croissance et Cancer
OPLS-DA	Régression des Moindres Carrés Partiels Orthogonale en Analyse Discriminante
O/C	Rapport atomique entre le nombre d'oxygène et le nombre de carbone
PEG	Polyéthylène glycol
Pent	Pentose
PLS	Régression des Moindres Carrés partiels
PLS-DA	Régression des Moindres Carrés Partiels en Analyse Discriminante
QC	Echantillon de contrôle
QCEIMS	Quantum Chemistry Electron Ionization Mass Spectrometry
QTof	Analyseur quadripôle-temps de vol
RDA	Retro Diels-Alder
ReSpect	RIKEN MSn spectral database
R_f	Rapport frontal
RMN	Résonance magnétique nucléaire
SALDI	Désorption/ionisation laser assistée en surface
SFC	Chromatographie en phase supercritique
Sj	Variété de maïs Sugar Jean
SM	Spectrométrie de Masse
SMILES	<i>Simplified molecular input line entry specification</i>
SPE	Extraction en phase solide
sPLS-DA	Régression des Moindres Carrés Partiels en Analyse Discriminante en méthode parcimonieuse
SRM	Selected reaction monitoring
T°	Température
TIC	Courant ionique total
TLC	Chromatographie sur couche mince
TOF	Temps de vol
TQ	Triple quadripôle
Tr	Temps de rétention
UHPLC	Chromatographie en phase Liquide Ultra Haute pression
UNPD	Universal natural product database
UV	Détecteur ultraviolet mono longueur d'onde
VIP	Projection des variables importantes

Liste des figures

Chapitre I :

Figure I-1: Schéma de fonctionnement d'un triple quadripôle en mode MRM avec un exemple de transition suivie.

Figure I-2: Schéma des différentes étapes clés lors d'une analyse métabolomique.

Figure I-3: Méthodes métabolomiques non ciblées et ciblées en relation avec le nombre de métabolites détectés et la fiabilité des résultats quantitatifs (Cajka 2016).

Figure I-4: Arbre de fragmentation hypothétique de la molécule (-)-épicatéchine calculé à partir des données du spectre de masse obtenu sur un orbitrap (Thermo Scientific). Les nœuds en bleu sont les pics de spectres de masse et leur formule brute, les liens en rouge correspondent à des pertes de neutres hypothétiques, d'après Rasche 2011.

Figure I-5: L'arbre spectral (c) avec ses formules brutes est obtenu à partir des relations entre les pics précurseurs et fragments (b) obtenus des différents spectres de fragmentation en MS^n (a). Un chemin de formules brutes identifiées (d) sans ambiguïté est proposé, les pics, nœuds et formules correspondants sont en gras, d'après Kasper (2012).

Figure I-6: Représentation de Kendrick basée sur les masses exactes de molécules avec une fonction acide carboxylique extraites de la base de données KEGG, d'après la publication de Werner 2008.

Figure I-7: Représentation de Kendrick de trois fractions d'hétérocycles de soufre aromatiques polycycliques séparées sur une phase stationnaire de cyclodextrine, d'après Panda (2006). Représentation des défauts de masse de Kendrick (KMD) et du nombre équivalent de doubles liaisons (DBE) en fonction de la masse nominale (KMN).

Figure I-8: Représentation de Van Krevelen basée sur 10 000 formules brutes extraites de la base de données KEGG, d'après la publication de Werner 2008.

Figure I-9: Représentation des modifications structurales suite à des réactions enzymatiques extrait de la publication d'Ohta 2010.

Figure I-10: Représentation de Van Krevelen de produits naturels obtenus par des recherches de formules brutes dans des bases de données avec le positionnement des familles moléculaires, d'après Kuhnert (2013).

Figure I-11: Illustration d'un cluster de réseaux moléculaires issus de spectres LC-MS/MS des dérivés du tryptophane et de la biotine, d'après Quinn 2017.

Figure I-12: Structure générique des flavonoïdes.

Figure I-13: Voies de biosynthèse des principaux flavonoïdes.

Figure I-14: Différents types de glycosylation sur un flavonoïde.

Figure I-15: Schéma de fragmentation et nomenclature des ions fragments des flavonoïdes aglycones.

Figure I-16: Schéma de fragmentation et nomenclature des ions fragments des flavonoïdes *O*-glycosylés et *C*-glycosylés, *O*-glycosylés.

Figure I-17: Spectres de fragmentation des ions $[M-H]^-$ de différents flavonoïdes *O*-glycosylés, d'après Ablajan (2006 et 2013).

Figure I-18: Spectres de fragmentation des molécules naringine (1->2) et narirutine (1->6) en mode d'ionisation négatif et positif d'après Cuyckens (2000).

Figure I-19: Spectres de fragmentation de flavonoïdes di-*O*-glycosylés d'après Vukics and Guttman (2010).

Figure I-20: Spectres de fragmentation en mode d'ionisation négatif (**A**) et en mode d'ionisation positif (**B**) de l'apigénine 8-*C*-glucoside, d'après Vukics and Guttman (2010).

Figure I-21: Schéma de fragmentation de l'orientine en mode d'ionisation positif.

Figure I-22: Spectres de fragmentation de deux flavonoïdes *C*-glycosylés, isomères de position : apigénine 8-*C*-glycosylé (Vitexine) et apigénine 6-*C*-glycosylé (Isovitexine) acquis avec un analyseur de type Q-Tof (a), avec un analyseur de type trappe d'ions (b) d'après Waridel (2001).

Figure I-23: Spectre de fragmentation de l'isovitexine-2''-*O*-glucoside (ou apigénine 6-*C*-glucoside-2''-*O*-glucoside) en mode d'ionisation négatif, d'après Cuyckens and Claeys (2005).

Figure I-24: Spectres de fragmentation de l'isovitexine-2''-*O*-glucoside en mode d'ionisation positif, d'après Piasecka (2015).

Figure I-25: Comparaison des spectres de masse de fragmentation de l'ion m/z 609 des molécules flavonoïde 2''-*O*-glycosylé-6-*C*-glycosylé et 6''-*O*-glycosylé-6-*C*-glycosylé, d'après Ferreres (2007).

Figure I-26: Spectres de fragmentation de la molécule apigénine 6,8-di-*C*-glucoside en mode d'ionisation positif (A) et en mode d'ionisation négatif (B), d'après Cao (2014).

Figure I-27: Spectres de fragmentation des molécules isoschaftoside (Apigénine-6-*C*-arabinose-8-*C*-glucose) (A) et schaftoside (Apigénine-6-*C*-glucose-8-*C*-arabinose) (B) en mode d'ionisation négatif, d'après Picariello (2017).

Figure I-28: Spectres de fragmentation de la molécule Lutéoline 4'-*O*-glucoside-8-*C*-glucoside en mode d'ionisation négatif d'après Ferreres (2007).

Figure I-29: Spectres de fragmentation de l'apigénine 8-*C*-glucoside-4'-*O*-rhamnoside à une énergie de collision de 20 eV en mode d'ionisation positif, d'après Abad-Garcia (2009).

Figure I-30: Spectres de fragmentation de l'isovitexine-7-*O*-glucoside en mode d'ionisation positif, d'après Piasecka (2015).

Figure I-31: Schéma d'aide à l'identification des flavonoïdes glycosylés basé sur les ions fragments spécifiques et leurs intensités en mode d'ionisation négatif.

Figure I-32: Schéma d'aide à l'identification des flavonoïdes glycosylés basé sur les ions fragments spécifiques et leurs intensités en mode d'ionisation positif.

Chapitre II :

Figure II-1 : Principe de la chromatographie sur couche mince et visualisation des composés d'un échantillon et d'un mélange de standards.

Figure II-2 : Classification des solvants selon leurs propriétés d'accepteur ou de donneur de protons ou de force dipolaire. Représentation des 8 classes de solvant selon le triangle de Snyder.

Figure II-3 : Différentes parties de la plante de colza.

Figure II-4 : Empreintes moléculaires des différents organes du colza (tiges, feuilles, fleurs, racines, graines,) par TLC.

Figure II-5 : Influence des corrections sur l'analyse en composante principale du système d'élution pour les composés lipidiques et système de révélation avec la primuline.

Figure II-6 : Analyse en composantes principales des différentes parties (Feuilles, tiges, graines et racines) de la plante de colza. Correction de la ligne de base et données centrées et réduites. Systèmes d'élution et de révélation associés à la famille des polyphénols.

Figure II-7 : Analyse en composantes principales des différentes parties (Feuilles, tiges, graines et racines) de la plante de colza. Correction de la ligne de base et données réduites. Système d'élution pour les composés lipidiques et système de révélation avec la primuline.

Figure II-8 : Analyse en composantes principales des différentes parties (Feuilles, tiges, graines et racines) de la plante de colza avec indication de la plaque d'analyse 1 ou 2 pour les échantillons et une correction de la ligne de base. Système d'élution pour les composés moins polaires et révélation avec l'anisaldéhyde.

Figure II-9 : Schéma du couplage TLC/API-MS avec l'interface de CAMAG.

Figure II-10 : Comparaison des spectres de masse de la molécule estriol en TLC/APCI TQ-MS. (A) Solubilisation dans l'acétonitrile acidifié. (B) Solubilisation dans du méthanol acidifié.

Figure II-11 : Comparaison des spectres de masse de la molécule testostérone en TLC/API-MS avec de l'acétonitrile comme solvant d'élution. (A) En mode d'ionisation négatif avec une source ESI. (B) En mode d'ionisation positif avec une source ESI. (C) En mode d'ionisation positif avec une source APCI. (D) En mode d'ionisation négatif avec une source APCI.

Figure II-12 : Schéma du couplage TLC/MALDI-Tof MS (Support ; Source ; Analyseur ; Détecteur et enregistreur).

Figure II-13 : Spectre de masse en MALDI-Tof MS de la matrice DHB en mode d'ionisation négatif.

Figure II-14 : Spectre de masse en mode d'ionisation négatif de la matrice inorganique par MALDI-Tof MS.

Figure II-15 : Spectre de masse par MALDI-Tof-MS de 2 standards : l'acide ellagique (A) et l'isorhamnétine 3-glucoside (B) en mode d'ionisation négatif en l'absence de matrice (LDI) ou en présence d'une matrice inorganique.

Figure II-16 : Analyse sans élution de 1 μL de standards de flavonoïdes en mode d'ionisation positif et négatif en présence de 2 μL d'une matrice DHB ou inorganique appliquées en spray.

Figure II-17 : Comparaison du type de dépôt de la matrice (pulvérisation ciblée ou immersion) sur une plaque TLC silice 60 F₂₅₄ (5 x 7,5 cm).

Figure II-18 : Elution en TLC de standards de différentes sous-familles de flavonoïdes.

Figure II-19 : Séparation en TLC/MALDI-ToF MS en mode d'ionisation négatif d'un mélange de standards. (A) Révélation NEU-PEG avec détection à 366 nm. (B) Spectres de masse des spots après élution et dépôt de la matrice inorganique sur une plaque non révélée en mode d'ionisation négatif.

Figure II-20 : Spectre de masse en mode d'ionisation positif obtenu pour 10 μL de vin rouge déposés sur plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm) non éluee, avec 2 μL de la matrice inorganique déposée au-dessus.

Figure II-21 : Analyse par TLC/MALDI-ToF MS d'un échantillon de vin rouge avec un volume de 2 μL de matrice inorganique déposée sur chacun des spots. A) Plaque TLC dans le visible. B) Spectres de masse de chaque spot élué en mode d'ionisation positif.

Figure II-22 : Analyse d'un extrait éthanolique de baies d'argousier (*Hippophae rhamnoides*) par A) MALDI-ToF MS en mode d'ionisation négatif et par B) TLC/MALDI-ToF MS sur plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm). C) Spectres de masse associés à quelques spots de la plaque TLC.

Figure II-23 : Analyse en TLC/MALDI-ToF MS en mode d'ionisation négatif d'un jus de pomme enrichi en dérivés de quercétine avec 2 μL de matrice inorganique déposé sur chaque spot. (A) Plaque TLC après dérivation. (B) Spectres de masse en TLC/MALDI-ToF MS en mode d'ionisation négatif associé à chaque spot, obtenus à partir d'une plaque TLC non dérivée. (C) Identification des dérivés de quercétine.

Figure II-24 : Analyse TLC/MALDI-ToF-MS du jus de pomme enrichi en dérivés de quercétine. Représentation des valeurs de rapport m/z en fonction des zones ciblées et des rapports frontaux R_f . La taille des bulles correspond à l'abondance des ions. La couleur des bulles est associée à la structure aglycone des composés détectés et associée à la couleur des taches observées après dérivation au NEU.

Figure II-25 : Spectre MALDI-ToF MS sans élution de l'extrait de *Garcinia mangostana* en mode d'ionisation négatif.

Figure II-26 : Analyse par TLC/MALDI-ToF MS en mode d'ionisation négatif de l'extrait de *Garcinia mangostana* avec la matrice DHB déposée par immersion à 200 mg.mL^{-1} . (A) Plaque TLC révélée après élution et dérivation. (B) Spectres de masse associés à plusieurs spots.

Figure II-27 : Spectres de masse de quatre acides organiques analysés en TLC/ESI TQ-MS en mode d'ionisation négatif après élution avec une phase éluante $\text{NH}_4\text{OH}/\text{EtOH}/\text{H}_2\text{O}$ 12,5 : 75,5 : 12.

Figure II-28 : Spectres de masse du standard de 24-OH-Cholestérol obtenus avec les sources APCI, MALDI et ESI en mode d'ionisation négatif et positif.

Figure II-29 : Spectres de masse du standard de testostérone obtenus avec les sources APCI, MALDI et ESI en mode d'ionisation négatif et positif.

Figure II-30 : Comparaison de la réponse de trois standards en couplage TLC/API MS, avec deux sources (ESI et APCI) en mode d'ionisation négatif et positif.

Figure II-31 : Analyse HPTLC/APCI-MS en mode d'ionisation positif pour l'identification d'un mélange de 23 standards d'oxystérols et de stéroïdes. (A) Plaque du mélange élué à 366 nm. (B) Plaque après analyse en HPTLC/APCI-MS. (C) Spectres de masses en mode d'ionisation positif avec le couplage HPTLC/APCI MS.

Figure II-32 : Analyse HPTLC/APCI MS d'un mélange de standards d'oxystérols et de stéroïdes Graphique à bulles corrélant l'empreinte spectrale massique et la migration (*R_f*) des standards. La taille des bulles indique l'intensité des ions.

Chapitre III :

Figure III-1 : Schéma d'attraction de *H. zea* (A et D) et des prédateurs (B) sur le maïs (C). Larve de *H. zea* (D) se nourrissant de soie de maïs (E).

Figure III-2 : Squelettes des différents types de flavonoïdes identifiés dans les soies de maïs. (A) Dérivés de la molécule apimaysin. (B) Dérivés de la molécule apigénine 6-C-glycosylés. (C) Dérivés des molécules lutéoline ou chrysoériol 7-O-glycosylé.

Figure III-3 : Schéma des différentes étapes de l'étude métabolomique appliquée aux soies de maïs.

Figure III-4 : Photographie des 5 variétés de soies de maïs après réception à l'ICOA.

Figure III-5 : Répartition des individus (1-05 ; 1-04 ; 1-03 ; 5-05) en fonction de la répétabilité d'extraction (M01, M02, M03) et de variétés (Garrison (Gar) et Java (Jav)).

Figure III-6 : Schéma de la composition de l'échantillonnage.

Figure III-7 : Structures des greffons des deux types de colonnes nucleodur testés.

Figure III-8 : Séparation de 6 flavonoïdes standards sur trois colonnes Nucléodur® différentes.

Figure III-9 : Séparation de flavonoïdes (en extraction d'ions) contenus dans un extrait de soies de maïs sur une colonne Nucleodur Sphinx RP et une colonne Nucleodur Pyramid.

Figure III-10 : Effet de la réduction du diamètre des particules du support chromatographique sur la séparation d'un mélange de six standards de flavonoïdes.

Figure III-11 : Effet de la température de la colonne sur la séparation d'un mélange de six standards de flavonoïdes.

Figure III-12 : Effet du débit de la phase mobile (0,3 et 0,6 mL/min) sur la séparation d'un mélange de six standards de flavonoïdes.

Figure III-13 : Base Peak Chromatogram (BPC) d'un mélange de soies de maïs obtenu en UHPLC/HRMS.

Figure III-14 : Représentation des masses exactes de l'ensemble des composés détectés dans les extraits de soies de maïs en fonction de leur temps de rétention.

Figure III-15 : Spectre de fragmentation de la molécule maysin provenant d'un extrait de soie de maïs possédant un 6-deoxy-xylo-hexos-4-ulopyranosyl en mode d'ionisation négatif, obtenu sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV.

Figure III-16 : Extractions d'ion de rapport de masse m/z 609,1459 u dans les analyses de mélange de standard (violet) et de l'extrait (rouge).

Figure III-17 : Spectre de fragmentation du composé élué au temps de rétention 3,12 min.

Figure III-18 : Spectre de fragmentation du composé élué au temps de rétention 4,80 min en mode d'ionisation négatif.

Figure III-19 : Spectre de masse en full scan associé au temps de rétention 4,79 min.

Figure III-20 : Spectres de fragmentation en mode d'ionisation négatif des quatre ions de rapport m/z 573,1241 ; 559,1445 ; 575,1399 ; 577,1562.

Figure III-21 : Spectres de fragmentation en mode d'ionisation positif des quatre ions de rapport m/z 575,13984 ; 561,16077 ; 577,1538 ; 579,1711.

Figure III-22 : Diagramme de Van Krevelen des 124 molécules extraites des soies de maïs analysées en UHPLC/HRMS en mode d'ionisation négatif.

Figure III-23 : Comparaison des chromatogrammes acquis en mode MRM pour les 5 variétés de soies de maïs.

Figure III-24 : Abondance relative des ions (Parents et fragments) en fonction de l'énergie de collision pour trois standards représentatifs des trois sous-groupes de flavonoïdes présents dans les soies de maïs.

Figure III-25 : Effet du temps d'acquisition des transitions lors de l'analyse UHPLC-MS/MS d'un mélange de six standards de flavonoïdes.

Figure III-26 : Gamme d'étalonnage ($n = 5$) du standard Vitexine 2''-O-Rhamnoside.

Figure III-27 : Suivi des analyses au cours du temps ; somme des concentrations des 67 composés quantifiés sans correction (à gauche) et après correction des concentrations (à droite).

Figure III-28 : Répartition des 35 premières variables (molécules) et des échantillons (variétés) avant et après transformation (normalisation quantile ; range scaling).

Figure III-29 : Classification ascendante hiérarchique des individus (variétés de soies de maïs).

Figure III-30 : Score plot (variétés) et loading plot (molécules) de l'ACP.

Figure III-31 : Score plot (variétés) du sPLS-DA.

Figure III-32 : Poids moyen des stades larvaires de *H. zea*, tiré de la publication Rhino 2016.

Figure III-33 : Loading plot des composantes 1 et 2 avec leur valeur de score, ainsi que leur abondance (rouge = forte ; vert = faible) dans la variété.

Figure III-34 : Boîte à moustache des 4 principaux VIP de la composante 1 (p-value < 0.05).

Figure III-35 : Base Peak Chromatogram (BPC) d'un mélange de soies de maïs de la variété Java obtenu en UHPLC/HRMS, avec les extraits d'ions des molécules VIP de la composante 1.

Figure III-36 : Boîte à moustache des 3 principaux VIP de la composante 2 (p-value < 0.05).

Figure III-37 : Base Peak Chromatogram (BPC) d'un mélange de soies de maïs de la variété Nova obtenu en UHPLC/HRMS, avec les extraits d'ions des molécules VIP de la composante 2.

Figure III-38 : Réseaux moléculaires en mode d'ionisation négatif (A) et en mode d'ionisation positif (B).

Figure III-39 : Représentations graphiques : (A) Nombre équivalent de double liaison en fonction de la masse moléculaire et (B) Van Krevelen, ratio H/C en fonction du ratio O/C.

Figure III-40 : Réseau moléculaire avec cosinus score à 0,6.

Figure III-41 : Spectres de fragmentation en mode d'ionisation négatif en UHPLC/HRMS-MS de trois molécules à une masse de 589 Da. Inconnus 1 et 2 dans le cluster 4 et chrysoériol C-(6-deoxy-xylohexos-4-ulopyranosyl)-2''-O-désoxyhexose dans le cluster 2.

Liste des tableaux

Chapitre I :

Tableau I-1: Base de données de composés chimiques.

Tableau I-2: Bases de données (Bdd) spectrales.

Tableau I-3: Logiciels de fragmentation in silico.

Tableau I-4: Liste de logiciel de pré-traitement et/ou de statistique.

Tableau I-5: Masse (en u) des pertes de neutre dans le cas de deux sucres.

Tableau I-6: Perte de neutre lors de fragmentation de flavonoïdes glycosylés.

Chapitre II :

Tableau II-1 : Caractérisation par MALDI-Tof MS du vin rouge déposé sur la plaque TLC après élution.

Tableau II-2 : Caractérisation des masses de l'extrait *Garcinia mangostana* en TLC/MALDI-Tof MS. Attribution des masses détectées en se basant sur la publication de Destandau (2015).

Tableau II-3 : Liste des oxystérols, phytostérols et stéroïdes (Annexe 5) avec masse et exemple de structure par famille.

Chapitre III :

Tableau III-1 : Flavonoïdes identifiés dans les soies de maïs.

Tableau III-2 : Echantillonnage des soies de maïs.

Tableau III-3 : Nomenclature des différents essais de répétabilité.

Tableau III-4 : Flavones aglycones et les ions diagnostiques pour des flavones di-C-glycosylés et des flavones mono-C-glycosylés en mode d'ionisation négatif.

Tableau III-5 : Proposition de formules brutes pour le fragment et l'ion moléculaire du composé élué au temps de rétention de 3,12 min.

Tableau III-6 : Proposition de formules brutes des fragments de l'ion m/z 573,12485.

Tableau III-7 : Ions spécifiques à la fragmentation des ions moléculaires en mode d'ionisation positif.

Tableau III-8 : Caractérisation des flavonoïdes-O-glycosylés d'un extrait de soies de maïs.

Tableau III-9 : Caractérisation des flavonoïdes mono-C-glycosylés d'un extrait de soies de maïs.

Tableau III-10 : Caractérisation des flavonoïdes mono-C-glycosylés-O-glycosylés d'un extrait de soies de maïs.

Tableau III-11 : Caractérisation des flavonoïdes di-C-glycosylés d'un extrait de soies de maïs.

Annexes :

Tableau A1 : Conditions opératoires d'extraction assistée par micro-ondes.

Tableau A2 : Conditions de dépôts.

INTRODUCTION GENERALE

Dans différents domaines d'activités tels que l'environnement, le médical, le pharmaceutique ou le cosmétique, l'interaction entre le vivant et son environnement peut se traduire par une modification plus ou moins marquée de l'expression moléculaire de l'organisme vivant. Ainsi dans le domaine de l'écologie chimique et de l'écophysiologie des plantes, on étudie les métabolites secondaires, dits maintenant spécialisés, qui permettent aux plantes d'interagir avec leur environnement ainsi que la réponse de ces voies métaboliques suite à des perturbations externes (attaque d'insectes, chimiques, climatiques...). De même dans le domaine médical, on peut étudier l'interaction d'une molécule sur l'homme, mais aussi l'impact des résidus médicamenteux rejetés dans l'environnement.

Dans ce contexte, il est essentiel de développer des méthodes analytiques qui permettent de répondre à ces problématiques et d'aider à la compréhension des différentes interactions. Ainsi pour pouvoir caractériser et quantifier ces métabolites dans des matrices complexes qu'elles soient végétales ou biologiques, il est nécessaire d'optimiser et de maîtriser différentes étapes telles que l'extraction, le traitement d'échantillon, la séparation, la purification, la caractérisation et la quantification des molécules tout en prenant en compte que les molécules ciblées peuvent être éventuellement exprimées à des teneurs très variables.

Les méthodes complémentaires développées donneront accès à l'empreinte moléculaire des milieux complexes représentative d'un état du système à un instant donné et dans un environnement donné. La comparaison des différentes signatures peut dans ces conditions constituer une approche intéressante pour évaluer l'impact de l'environnement et comprendre les interactions avec le vivant.

L'objectif de cette thèse était de développer des méthodes de caractérisation des systèmes moléculaires en milieu complexe par des couplages avec la spectrométrie de masse afin de répondre à différentes problématiques :

- Identifier et reconnaître rapidement les molécules connues pour cibler plus facilement les composés inconnus (Déréplication) ;
- Cartographier les molécules d'un échantillon afin d'aider à leur identification ;
- Comparer des échantillons pour visualiser les différences et cibler l'identification sur les molécules différenciantes (Métabolomique) ;
- Quantifier des molécules d'intérêt dans des milieux complexes et montrer l'interaction de ces métabolites dans ces milieux.

Le premier chapitre de cette thèse est consacré à la présentation des différents outils complémentaires à la disposition de l'analyste pour obtenir et comprendre les signatures moléculaires de milieux complexes. Après un rappel des différentes techniques analytiques employées pour séparer et détecter les molécules présentes dans un échantillon, une présentation des deux stratégies actuelles (déréplication et métabolomique) mises en place pour caractériser les molécules liées au processus des voies de biosynthèse est développée. Après l'étape de séparation et de détection, l'étape d'identification des métabolites peut être longue et fastidieuse au vue de la diversité des structures

moléculaires potentielles. De plus en plus de bases de données moléculaires sont développées pour aider dans cette étape. C'est pourquoi nous avons établi aussi un bref récapitulatif de l'ensemble de ces outils accessibles. Les outils statistiques pour le traitement de données sont maintenant indispensables à l'analyste compte-tenu de la somme de données acquises lors des analyses en couplage avec la spectrométrie de masse et permettent d'accéder aux comparaisons des données et aux représentations visuelles des empreintes moléculaires. Une présentation des outils statistiques plus particulièrement utilisés durant ce travail de thèse est également effectuée. Enfin le chapitre 1 se termine avec la présentation de la famille de métabolites secondaires que nous avons plus particulièrement ciblée à savoir les flavonoïdes et avec un état de l'art sur la fragmentation des flavonoïdes glycosylés.

Les études de déréplication peuvent se faire de façon rationalisée par chromatographie sur couche mince (TLC). A partir de la photographie des plaques CCM obtenues après révélation, il est possible d'appliquer une approche statistique pour analyser les différentes taches de couleur révélées. Ainsi un grand nombre d'échantillons peuvent être regroupés par similarité spectrale, après avoir été élués côte à côte avec des systèmes d'élution et de révélation spécifiques aux familles moléculaires étudiées. Après une présentation des méthodes de couplages TLC/MS, le deuxième chapitre expose le développement d'un couplage TLC/MALDI-ToF MS sur des matrices végétales pour obtenir des analyses ciblées avec une identification des molécules d'intérêt. Cette méthode permet de mettre en évidence rapidement la composition de différents mélanges complexes avec une proposition d'identification des composés. Dans le cas d'une technique complémentaire, les couplages TLC/API MS sont utilisés pour résoudre des problèmes d'analyse avec la première méthode et avoir un panel plus large de familles moléculaire analysable.

Le troisième chapitre est consacré au développement d'une méthode d'analyse ciblée par UHPLC/MS en mode MRM de flavonoïdes glycosylés dans des soies de maïs. Cette étude se situe dans un contexte de recherche par le CIRAD de Martinique d'une plante piège (potentiellement le maïs) pour protéger la culture des plants de tomates contre les attaques de la noctuelle de *Helicoverpa zea* en Martinique. Après une caractérisation des flavonoïdes glycosylés présents dans 5 variétés de soies de maïs différentes, des approches de cartographie moléculaire des extraits de soie de maïs ont été initiées (Van Krevelen, réseau moléculaire), afin d'aider à repérer plus facilement l'appartenance des composés aux familles moléculaires, de lever des indéterminations de caractérisation de molécules et d'aider à l'identification des molécules différenciantes entre variétés. Ces approches peuvent permettre d'aller plus vite dans la caractérisation de composés cibles.

Ces études se finaliseront par une conclusion sur l'obtention de signature moléculaire dans des milieux complexes et ses perspectives.

Méthodes analytiques et outils de traitements de données utiles pour caractériser et comparer des milieux complexes

Le travail de thèse présenté dans ce manuscrit est centré sur l'obtention de signatures moléculaires reflètes à un instant donné de la composition chimique d'un milieu complexe (extraits de plante, milieux réactionnels de synthèse, fluides biologiques...) ou d'un état du système en fonction de conditions internes ou externes données. Les molécules que nous avons cherchées à caractériser sont toutes de petites tailles (< 1500 Da) et sont nommées des métabolites. Tout changement interne (ex : modification du métabolisme) ou externe (ex : application d'un stress abiotique ou biotique) du milieu engendre une modification de l'expression de ces molécules, soit par leur quantité soit par leur nature chimique. La caractérisation de la signature moléculaire du système biologique a donc besoin d'être la plus fidèle possible afin de pouvoir identifier de nouveaux composés et/ou interpréter des modèles de réaction qui engendrent ces modifications métaboliques.

Pour répondre aux différents niveaux de complexité des problématiques rencontrées, différentes stratégies méthodologiques plus ou moins lourdes doivent être mises en place en s'appuyant sur la panoplie d'outils développés ces dernières décennies, dans le but de cibler les molécules d'intérêt, les séparer, les analyser, les visualiser et interpréter les résultats qui génèrent de plus en plus de données. Pour caractériser au mieux un échantillon complexe, des techniques analytiques performantes et notamment chromatographiques sont maintenant disponibles (Dunn and Ellis 2005, Dettmer 2007, Zhang 2012) en couplage ou non entre elles ou avec des techniques spectrométriques pour obtenir dans un domaine donné une empreinte moléculaire. Une fois cette signature obtenue, les bases de données moléculaires et spectrales peuvent aider à l'identification des composés. Dans le cas d'étude où plusieurs échantillons doivent être confrontés, des outils de représentation graphique d'un échantillon ainsi que des outils complémentaires d'aide à la comparaison et à la visualisation des données des différents échantillons peuvent de plus être utilisés.

Ce premier chapitre va ainsi introduire brièvement les techniques d'analyse fondamentales disponibles en se focalisant plus particulièrement sur les couplages de méthodes chromatographiques avec la spectrométrie de masse ainsi que les différents outils d'aide à l'identification des métabolites et d'aide à la comparaison des données acquises en spectrométrie de masse sur lesquels nous nous sommes appuyés lors de ces travaux de thèse. Une attention particulière sera également consacrée à la description de mécanismes de fragmentation qui sont essentiels pour garantir l'identification de métabolites. C'est pourquoi les schémas de fragmentation des flavonoïdes glycosylés, famille moléculaire que nous avons plus particulièrement ciblée dans notre étude de différents milieux complexes, seront décrits en fin de ce chapitre.

1. Techniques analytiques à disposition

Il existe actuellement une diversité de techniques d'extraction et de séparation permettant de fractionner ou d'isoler des molécules naturelles à partir de milieux complexes, de même que plusieurs systèmes de détection dédiés à leur visualisation voire à leur identification ont été développés. Selon le contexte analytique de l'étude, ces techniques peuvent être combinées entre elles pour permettre la caractérisation chimique d'extraits en comparaison à des molécules standards de référence ou établir un profilage haut débit de collections d'échantillons. De nombreuses revues étant disponibles dans la littérature, les principes détaillés de chacune de ces techniques ne seront pas repris mais seront présentées les techniques les plus couramment utilisées, leurs avantages et leurs limites dans un

contexte d'étude de petites molécules organiques pour obtenir les signatures moléculaires de milieux complexes.

1.1. Méthodes d'extraction

L'extraction est un procédé qui permet de séparer des molécules ciblées d'un échantillon et de les faire passer dans un contenu analysable. Dans ce paragraphe seront décrites certaines méthodes d'extraction dites « solide/liquide » permettant d'extraire des composés d'une matrice solide par un solvant liquide. L'extraction est une première étape analytique qui cible les métabolites à analyser en ajustant la polarité du solvant d'extraction au plus proche de la polarité des molécules tout en choisissant une technique adaptée au nombre d'échantillons, aux propriétés physico-chimiques des molécules et en ajustant les paramètres spécifiques de la technique choisie. Cette étape permet ainsi de simplifier la composition de l'échantillon en éliminant les molécules de polarité opposée mais de ce fait, peut conduire aussi à une perte de représentativité en métabolites si la gamme de polarité de ces molécules est très étendue. C'est donc un facteur clé à ne pas négliger qui va selon les conditions utilisées permettre d'obtenir des extraits sélectifs des familles moléculaires ciblées ou au contraire des extraits plus exhaustifs contenant une large gamme de métabolites.

L'extraction par macération est la méthode la plus ancienne et la plus classiquement utilisée pour extraire les composés. Toutefois c'est une méthode demandant un long temps de contact (plusieurs heures) entre la matière végétale et le solvant pour obtenir une extraction optimale. C'est pourquoi d'autres méthodes alternatives d'extraction ont été développées plus récemment en réduisant la consommation de solvant et le temps d'extraction sur des quantités d'échantillons petites ou grandes tout en conservant voire améliorant les rendements d'extraction et la reproductibilité et tout en étant plus automatisables. Ainsi des techniques comme **l'extraction assistée par micro-ondes** ou par **ultrasons**, **l'extraction par solvant pressurisé** et **l'extraction par fluide supercritique** permettant d'extraire davantage d'échantillons et avec de meilleures performances qu'une extraction classique par macération se sont développées pour extraire les composés non-volatils. Le choix de la méthode d'extraction se fera en fonction de la nature de l'échantillon et du type de composés à extraire (plus ou moins polaire ou apolaire). L'optimisation des conditions d'extraction reste toujours en premier lieu le choix du solvant d'extraction le plus approprié possible pour obtenir les composés recherchés, puis selon la technique employée, une optimisation du temps, du nombre de cycles d'extraction, de la température, de la puissance de chauffage ou de la pression pourra être effectuée. A partir d'un échantillon donné et pour être le plus représentatif possible de sa composition, il est possible de développer des procédés d'extraction complet en travaillant avec des solvants de polarité croissante et en collectant les fractions entre chaque étape afin d'éviter la saturation de certains composés. Ce type d'étude plus exhaustive demandera par la suite plus d'analyses et de travail d'interprétation.

1.2. Méthodes séparatives

L'extrait obtenu pouvant être plus ou moins riche en composés, une séparation des différentes molécules présentes, en fonction de leurs propriétés physico-chimiques, est souvent nécessaire afin d'avoir une meilleure visibilité de la composition moléculaire et de la complexité de l'extrait. Dans ce paragraphe, sont rappelées brièvement les différentes techniques de séparation.

La chromatographie sur couche mince (CCM ou TLC) est une première méthode séparative couramment utilisée en première intention en phytochimie ou lors de suivi de synthèse. C'est une méthode qui permet d'analyser plusieurs échantillons et standards en même temps sur une même plaque. Une comparaison rapide des échantillons peut être réalisée visuellement ainsi qu'une identification facile des familles moléculaires sur la base des spots de couleurs développées après application d'agents de dérivation spécifiques aux différentes familles moléculaires. C'est une méthode qui a une bonne résolution, qui peut être automatisable et qui permet de faire de la quantification de molécules sous certaines conditions. Des développements récents en couplage de cette technique séparative avec la spectrométrie de masse sont apparus, apportant une dimension complémentaire d'identification structurale. Une spécificité de la TLC est le développement possible de couplage de la TLC à des bio-essais effectués directement sur la plaque (Hook 1997) pour évaluer la potentielle activité biologique des composés séparés.

La chromatographie en phase gazeuse (CPG) possède une très bonne résolution chromatographique avec une détection en spectrométrie de masse efficace liée à des banques de données nombreuses et précieuses comme aide à l'identification des molécules détectées. Néanmoins, cette méthode se limite à l'analyse des composés volatils et thermostables. Elle n'est pas appropriée à l'analyse directe sans dérivation chimique préalable des composés polaires ni à celle des composés à haut poids moléculaire.

L'électrophorèse capillaire (EC) reste plus appropriée pour la séparation des analytes ioniques ou ionisables selon leur état de charge. Son couplage avec la spectrométrie de masse est moins automatisable et demande un peu plus de développement que celui de la chromatographie en phase gazeuse.

La chromatographie en phase supercritique (SFC), permet d'analyser une large gamme de composés et d'isoler des molécules connues et inconnues pour procéder à leur identification (Nothias 2017). La faible consommation de solvant organique est favorable à son utilisation à l'échelle semi-préparative ou préparative, car les molécules sont collectées dans de faibles volumes rapidement évaporables. Le couplage SFC/MS est également réalisable. Il nécessite une attention particulière pour l'étape de décompression du CO₂ (positionnement du régulateur de pression dans le circuit) afin d'éviter la précipitation de l'analyte (Grand-Guillaume Perrenoud 2014) avant son introduction dans la source d'ionisation à pression atmosphérique.

La chromatographie en phase liquide (HPLC) est devenue une méthode automatisable et très résolutive qui peut être couplée en direct avec la spectrométrie de masse ou la RMN. Le passage de l'HPLC à la chromatographie liquide à ultra haute pression (UHPLC) avec des particules de plus faible diamètre (< 2µm) et des appareils pouvant travailler à de hautes pressions (1000 bars) a permis d'améliorer l'efficacité et la résolution des séparations nécessitant de coupler avec des appareils de spectrométrie de masse pouvant analyser rapidement les pics chromatographiques très fins (Plumb

2004). Ainsi les analyses à haut débit permettent maintenant de cribler plus facilement des séries d'échantillons, afin de vérifier la diversité chimique des extraits (Strege 1999).

1.3. Modes de détection

Afin de visualiser et d'identifier directement ou après séparation les composés extraits, une méthode adéquate de détection doit être sélectionnée. Malheureusement, aucun détecteur ne peut détecter tous les composés d'un extrait donné en une seule analyse. Ils ont tous, une spécificité, des avantages et inconvénients, que cela soit des détecteurs simples pour réaliser des profilages ou de la quantification (détecteur ultraviolet mono longueur d'onde (UV), détecteur évaporatif à diffusion de lumière (DEDL), détecteur électrochimique (ECD), détecteur à fluorescence (DF)) ou des détecteurs qui donnent des informations à deux dimensions (détecteur à barrettes de diodes (DAD), spectromètre de masse (MS), spectromètre de résonance magnétique nucléaire (RMN)). Ces détecteurs à deux dimensions permettent d'accéder, au fur et à mesure de la détection des composés, à leurs spectres d'absorbance (DAD), à leurs spectres de masse (accès aux rapports masse sur charge m/z) (MS) ou à leurs spectres RMN (accès au déplacement chimique) (Wolfender 2009). Ces données spectrales apportent des informations très utiles pour l'identification des composés analysés.

1.4. Techniques pour l'identification des métabolites

En vue de caractériser les métabolites présents dans un échantillon, plusieurs stratégies peuvent être appliquées selon la problématique et l'objectif de départ et peuvent ainsi impliquer différentes techniques analytiques, en réalisant soit une introduction directe de l'échantillon en résonance magnétique nucléaire (RMN) ou en spectrométrie de masse (SM) soit des couplages chromatographiques avec ces analyseurs (Wolfender 2015). L'introduction directe offre l'avantage d'une grande rapidité d'exécution et d'un traitement de l'échantillon limité et permet de s'affranchir de la large diversité de polarité des analytes. Toutefois Wolfender et al. (2010) ont montré les différents avantages à travailler en couplage pour l'analyse d'échantillons complexes. Chaque couplage possède ses avantages et ses inconvénients, ils peuvent également être associés de différentes façons et en complémentarité, ce qui demande un travail de développement méthodologique supplémentaire.

L'analyse directe en RMN permet de faire des analyses à haut débit, non destructives et sans beaucoup de préparation d'échantillon, d'obtenir des empreintes globales de l'échantillon et de réaliser des élucidations structurales fines. Wolfender et al. (2006) ont montré que les différentes techniques de RMN permettent de répondre à des questions biologiques. Ainsi le couplage LC/RMN a permis d'identifier divers isoflavones et isoflavones prénylés avec un potentiel antifongique dans *Erythrina vogelii*. Une étude structurale par RMN a permis de mettre en évidence des molécules naturelles labiles, qui après isolation des composés, peuvent être sous la forme de deux isomères. Toutefois, seuls les métabolites les plus abondants peuvent être caractérisés par analyse directe. Un couplage LC/MS/SPE/RMN permet de mettre en évidence de petites quantités de métabolites dans des milieux complexes sans passer par une étape de purification pour réaliser de l'élucidation structurale (Van der

Hooft 2013). Les analytes d'intérêt sont piégés sur une cartouche d'extraction en fonction de certaines caractéristiques du signal UV ou MS au fur et à mesure de leur élution de la colonne chromatographique. Ce piège permet de concentrer les composés et d'échanger le solvant LC non deutéré par un solvant deutéré adapté à des analyses RMN.

La **spectrométrie de masse** offre une grande versatilité, en réalisant aussi bien des analyses qualitatives via des empreintes globales, que des analyses quantitatives avec une large plage de linéarité (Bedair and Sumner 2008, Lei 2011). Carter a montré les avancées et l'apport de la spectrométrie de masse pour les produits naturels depuis les années 70, en passant de l'introduction directe en impact électronique, à l'ionisation douce et par laser, aux couplages GC/MS, puis LC/MS et le bond réalisé avec la spectrométrie de masse haute résolution (HRMS) (Carter 2014). Cela a permis de déterminer plus efficacement des structures. Les avancées en spectrométrie de masse permettent d'avoir un impact sur la découverte de nouveaux métabolites secondaires (Bouslimani 2014). Lors de l'infusion directe d'échantillons complexes en spectrométrie de masse, l'ionisation peut être sélective pour une famille moléculaire donnée mais des phénomènes de suppression de signal peuvent très souvent survenir. Le couplage avec une technique chromatographique peut permettre alors de minimiser ce problème de suppression d'ions, d'améliorer la sensibilité de détection, la qualité des données et de réaliser de l'identification voire de la quantification de nombreux métabolites. Bien que la spectrométrie de masse permette de caractériser les composés en fonction de leur poids moléculaire, de leur composition élémentaire et des caractéristiques structurales basées sur des modèles de fragmentation, la caractérisation fine de la structure moléculaire ne peut pas néanmoins être atteinte dans le cas d'isomères de position qui produisent des fragments similaires.

L'apport de la **spectrométrie de masse par mobilité ionique** permet d'ajouter une dimension à celle du rapport masse sur charge (m/z) des espèces détectées : c'est le paramètre de « collision cross section » (CSS). Ce paramètre permet d'avoir des informations complémentaires dans la description des molécules. Il décrit la taille, la forme et la charge d'un composé et peut être utilisé dans la caractérisation de composés inconnus (Schrimpe-Rutledge 2018). Ce paramètre supplémentaire peut ainsi permettre de distinguer certains isomères de position, si le CSS est différent, bien qu'ils aient un spectre de fragmentation identique.

Certains analyseurs comme la résonance cyclotronique ionique à transformée de Fourier (FTICR), le temps de vol (TOF) et l'orbitrap permettent d'être plus résolutif et avoir une meilleure précision de mesure, ainsi la **spectrométrie de masse haute résolution** (HRMS) peut répondre à certains besoins. En effet, travailler avec un appareil de spectrométrie à haute résolution permet d'obtenir pour chaque composé une masse exacte avec une précision de mesure allant au quatrième chiffre derrière la virgule, ce qui permet d'obtenir des propositions de formule brute pour l'ion parent, mais aussi pour ses fragments. Suivant le type d'analyseur, il permet également d'avoir une haute résolution allant de 20 000 à 10^6 pour un même rapport m/z . Cet outil est donc souvent utilisé pour réaliser de l'identification de molécules. De plus, la haute résolution des ions permet de distinguer des molécules ayant une même masse nominale, sans avoir une même formule brute (molécules dites isobares). Ainsi, les molécules bézafibrate ($C_{19}H_{20}ClNO_4$) et ofloxacine ($C_{18}H_{20}N_3FO_4$) ont une même masse nominale de 361 Da, mais elles ont respectivement une masse exacte de 361.819 u et de 361.3675 u. Ces deux masses peuvent être distinguées par un appareil à haute résolution. La spectrométrie de masse haute résolution peut permettre de travailler selon différentes approches ciblées ou non ciblées, en même temps ou non, sans amoindrir la qualité des chromatogrammes (Koulman 2009), et

ainsi permettre d'obtenir des informations complémentaires, montrer des différences qualitatives de l'ensemble des molécules composant les extraits ou distinguer des différences d'abondances relatives de composés connus ou inconnus.

Pour quantifier les molécules d'intérêt et être sélectif, le choix d'un spectromètre de masse possédant un analyseur de type **triple quadripôle** est souvent utilisé. Bien qu'il soit dit de basse résolution car affichant seulement une masse nominale avec une précision en masse allant seulement au premier chiffre derrière la virgule, ce type d'appareillage permet de sélectionner des ions parents, ainsi qu'un ou des fragments de l'ion parent sélectionné. Ce mode de travail est nommé « Multiple Reaction Monitoring » (MRM) ou « Selected Reaction monitoring » (SRM) suivant les constructeurs. Le principe (Fig. I-1) est que le premier quadripôle Q_1 sélectionne un ion d'intérêt (●), dit ion parent. Cet ion est fragmenté ensuite dans la cellule de collision (Q_2) avec une certaine énergie. Puis un des ions fragments (●), dit ion fils, est ensuite sélectionné dans le troisième quadripôle Q_3 . Ainsi, on demande au spectromètre de masse de suivre une ou des transitions (ion parent $\rightarrow$ ion fils) caractéristiques d'un composé pendant une plage de temps donnée.

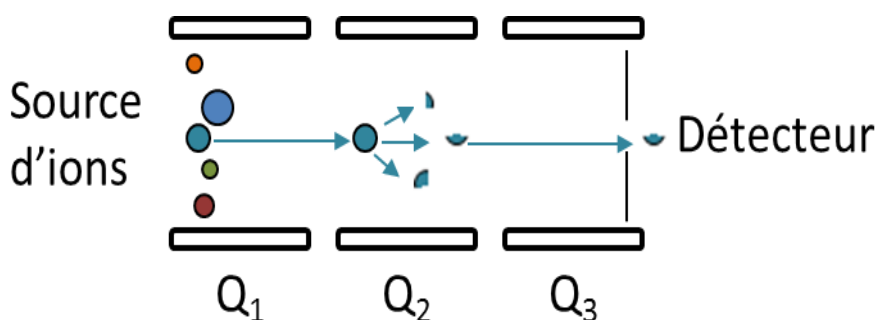


Figure I-1: Schéma de fonctionnement d'un triple quadripôle en mode MRM avec un exemple de transition suivie.

Ce mode MRM est souvent utilisé pour réaliser la quantification des molécules, car il permet d'être sélectif de la molécule et ainsi d'augmenter le rapport signal/bruit. En effet, comme les analyseurs quadripolaires ne laissent passer dans ce mode de fonctionnement que certaines molécules, le détecteur ne compte plus tous les autres ions considérés alors comme interférents, donc le bruit de fond diminue énormément. De plus, ce mode est spécifique à la molécule d'intérêt, car lors de la sélection des transitions, le fragment choisi peut être un fragment caractéristique de la molécule. En revanche, il faut connaître les molécules à quantifier, ainsi que leur fragmentation pour pouvoir convenablement choisir les transitions d'ions à suivre. La méthodologie nécessite souvent deux analyses. Une première analyse est acquise avec un balayage d'une plage de masse (full scan) pour avoir une vision globale de l'extrait et choisir les transitions des molécules d'intérêt. Puis une deuxième analyse est réalisée en mode MRM pour analyser spécifiquement les molécules d'intérêt.

Tous les différents types de couplages permettent de faire des analyses qualitatives et/ou quantitatives, d'obtenir des informations sur la polarité des composés, sur l'appartenance à une

famille moléculaire, sur la structure putative ou fine des composés, d'identifier des biomarqueurs d'intérêt (Seger 2013) et apportent des informations capitales pour répondre à beaucoup de problématiques. Dans les travaux que nous avons développés au cours de cette thèse et qui sont présentés dans les chapitres suivants, nous avons privilégié, l'utilisation de la chromatographie sur couche mince (CCM) et de la chromatographie liquide à ultra haute pression couplée à la spectrométrie de masse (basse et haute résolution) pour obtenir les signatures moléculaires des différents milieux complexes, car les familles moléculaires étudiées correspondent principalement à des molécules polaires et moyennement polaires. La possibilité offerte par la CCM permettra de réaliser des analyses comparatives de différents échantillons (Chapitre 2). Les couplages TLC/MALDI-Tof MS et TLC/ESI et APCI MS seront étudiés (Chapitre 2). Les couplages UHPLC/HRMS et UHPLC/MS en mode MRM permettront l'identification de composés extraits de soies de maïs et des études de quantification en ciblant certaines molécules d'intérêt dans les échantillons (Chapitre 3).

2. Stratégies actuelles de déréplication et de métabolomique

Les extraits de produits naturels sont des milieux complexes qui contiennent de nombreux composés. Beaucoup d'études ont permis d'identifier un certain nombre d'entre eux, mais la connaissance de la composition moléculaire dans son exhaustivité n'est pas encore maîtrisée. Donc la recherche de nouvelles approches se poursuit afin de connaître le plus finement possible la plante et son contenu. Les produits naturels sont une source importante de nouveaux métabolites avec des caractères uniques qui présentent des activités biologiques d'intérêt dans différents domaines. Ils existent donc de nombreux domaines d'applications notamment dans les secteurs de la santé ou de l'environnement qui ont besoin d'obtenir des empreintes moléculaires plus ou moins précises et détaillées suivant les cas.

2.1. Domaines d'applications

La détermination d'empreinte métabolique est nécessaire pour des applications allant de la toxicologie environnementale jusqu'au diagnostic de maladies (Cox 2014).

Dans le domaine **médical**, l'étude métabolomique par GC/MS de souris obèses et maigres a permis de distinguer les groupes sains et malades et de détecter des biomarqueurs comme des acides gras et des stéroïdes, de l'urée, qui peut montrer un dysfonctionnement des reins (Major 2006). Les études métabolomiques peuvent permettre d'aider dans la découverte des voies biologiques liées à des pathologies, comme des maladies cardiovasculaires (Wang 2011). La compréhension de ces mécanismes de réaction peut permettre de développer de nouveaux tests de diagnostic de maladie et de trouver des nouvelles approches thérapeutiques.

Dans le domaine **pharmaceutique**, les chercheurs sont toujours à la recherche de plus de diversités chimiques et criblent des chimiothèques entières pour trouver de nouveaux analogues. Les produits naturels sont également une grande source de composé actifs pour les produits pharmaceutiques. Ils peuvent provenir de sources végétales, fongiques et bactériennes (Krug and Müller 2014). La

découverte de nouveaux composés bioactifs est très difficile et demande beaucoup de temps. Il faut cribler un grand nombre de matières premières avant de sélectionner la plus adéquate et trouver la bonne méthode pour l'étudier efficacement. Dans ce cas, l'identification rapide de composés déjà connus présents dans l'extrait sans avoir besoin de les purifier est un gain de temps considérable pour pouvoir se concentrer sur l'étude des molécules inconnues susceptibles d'être de nouvelle source de molécules thérapeutiques. En **médecine traditionnelle chinoise**, les empreintes métaboliques sont d'une grande aide dans l'amélioration des préparations des plantes médicinales chinoises (Sheridan 2012).

Dans le domaine de **l'agrochimie**, les insecticides, herbicides et fongicides représentent une bonne part des perturbateurs endocriniens, avec une large gamme de structures chimiques et de mécanismes d'action (Mantovani 2008). Pour mettre en évidence les effets **écotoxicologiques** chroniques du fénitrothion, des études métabolomiques ciblées et non-ciblées ont été réalisées à partir d'analyses par RMN, par spectrométrie de masse en infusion directe et en couplage avec la chromatographie liquide. Elles ont permis d'observer les réponses moléculaires dans différents organes du gardon mâle après exposition à ces molécules (Southam 2011). De plus, une étude métabolomique non-ciblée a également permis de montrer les effets toxicologiques d'un polluant environnemental, la cyperméthrine, et les biomarqueurs sensibles à son exposition sur les vers de terre (Ratnasekhar 2015). C'est dans ce contexte que l'optique de remplacer les insecticides par des molécules naturelles s'est ancrée et que le domaine de **l'écologie chimique** s'est développé. Le principe étant d'analyser une grande quantité d'échantillons de plantes ayant subi des stress ou qui seraient entrées en interaction soit avec d'autres plantes, des insectes, des microbes pour interpréter les modifications chimiques ou les signaux engendrés. Ce domaine peut permettre d'exploiter les plantes d'une autre manière en travaillant avec des co-cultures adéquates. Un appui complémentaire pourrait aider pour une meilleure compréhension de la plante et de ces métabolites en réalisant l'imagerie 3D basée sur la LC/MS et MS² des plantes pour visualiser spatialement les métabolites des plantes, comme pour les plants de tomate et de poivre (Floros 2017).

2.2. Définitions des termes : déréplication et métabolomique

La volonté d'améliorer l'efficacité des programmes de criblage des ressources naturelles a conduit à deux concepts très souvent interconnectés : celui de la déréplication et celui de la métabolomique dont les définitions sont rappelées.

La **déréplication** est une approche de profilage pour l'identification rapide de composés présents dans les extraits afin de mettre en évidence les molécules connues et déjà décrites dans la littérature. L'identification peut ensuite se focaliser sur les molécules non reconnues qui peuvent être de potentiels nouveaux composés d'intérêt (Lang 2008) non connus.

Les études par déréplication sont apparues pour permettre d'identifier des molécules en minimisant les étapes d'isolement afin d'accélérer et d'être plus efficace dans la recherche de molécules cibles. En 1997, Constant et al. montraient déjà les prémices de la déréplication pour l'analyse de flavonoïdes aglycones et glycosides en LC/ESI-MS² dans un extrait d'*Eugenia jambos* L. sans passer par l'isolement des composés. Par ailleurs, une approche de déréplication d'un extrait de *Amphimas pterocarpoides* a permis d'identifier rapidement les isoflavonoïdes composant l'extrait (Tchoumtchoua 2013).

D'une espèce végétale à une autre, d'une variété à une autre, d'une partie de la plante à une autre, selon le climat et le lieu de récolte, les extraits peuvent présenter des profils de métabolites différents pouvant conduire jusqu'à des activités biologiques différentes. Un réel intérêt pour étudier et mieux comprendre l'influence de ces différents paramètres sur l'expression moléculaire a conduit à développer un nouveau type d'études, dites métabolomiques, complémentaires à l'approche dite de déréplication.

La **métabolomique** est l'analyse de l'ensemble des molécules de petite taille (inférieure à 1500 Da), appelées métabolites, présents dans un système vivant. Ces molécules sont liées au processus des voies de biosynthèse. La connaissance de l'identité des métabolites permet une meilleure compréhension des systèmes biologiques. Les études métabolomiques permettent de mettre en évidence de nouveaux métabolites secondaires bioactifs, leur voie de biosynthèse et de vérifier la modulation de leur production en réalisant des modifications biologiques (Breitling 2013). L'étude métabolomique permet de comparer des empreintes de métabolites détectables dans un échantillon donné avec une classification ultérieure des échantillons. Elle permet aussi de regrouper les populations d'échantillons, de mettre en évidence des différences de métabolites entre différentes populations (différentes variétés, différentes conditions de culture, de stress, d'interaction ...), de discriminer ou de classer ces métabolites en groupes afin de mettre en lumière les métabolites qui interviennent dans les changements ou dans les différences de groupes. La métabolomique est un outil précieux pour étudier le phénotype et les changements causés par des influences environnementales, des maladies ou autre variation de facteurs biotiques et/ou abiotiques. La distribution des métabolites est soumise à la variabilité temporelle et spatiale. On peut donc décrire la métabolomique comme une approche transversale combinant des systèmes analytiques de haute résolution à des outils d'exploitation de données et des analyses statistiques.

2.3. Etapes clés d'une analyse métabolomique

Une analyse métabolomique se fait en plusieurs étapes clés (Fig. I-2). A partir d'une problématique de départ, on réalise un plan d'échantillonnage. Une fois tous les échantillons récoltés et stockés, le traitement de ces derniers est exécuté pour récupérer les métabolites d'intérêt. Les empreintes de ces extraits sont obtenues via un ou plusieurs appareils analytiques. Quand toutes les analyses sont terminées un traitement des données peut être effectué, afin de procéder aux analyses statistiques univariées et/ou multivariées. Les résultats obtenus sont alors interprétés voire validés selon le niveau d'exigence de la problématique de départ.

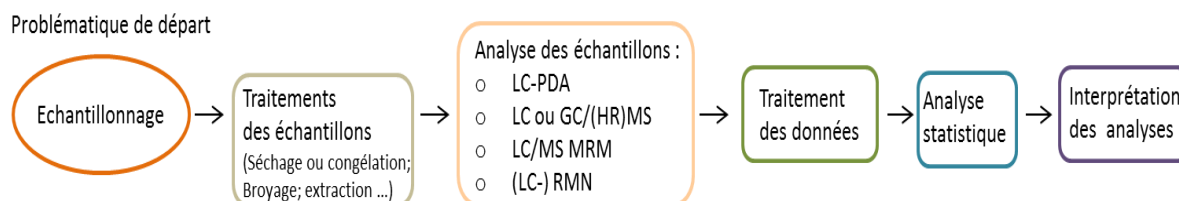


Figure I-2: Schéma des différentes étapes clés lors d'une analyse métabolomique.

La comparaison des populations peut se réaliser de différentes manières. L’empreinte métabolique et le profilage sont deux approches complémentaires. **L’empreinte métabolique** consiste à dévoiler tous les composés qu’il peut y avoir dans un extrait donné, sans identification préalable de ces composés. Une empreinte métabolique peut donc être faite afin d’obtenir une classification rapide des échantillons sans a priori, c’est une **analyse non ciblée**. L’étude métabolomique peut être aussi effectuée en réalisant une **analyse ciblée** sur un groupe de métabolites lié à une voie de biosynthèse ou à une famille moléculaire, c’est le **profilage métabolique** qui s’intéresse alors à des molécules qui ont été préalablement identifiées et pourront être quantifiées. Une analyse ciblée ne peut être menée que sur une liste bien précise de molécules sélectionnées. La quantification peut s’effectuer de deux manières en absolu ou en relatif. La quantification absolue se réfère à tous les étalons externes à l’échantillon ou à des étalons marqués ajoutés aux séquences d’analyse, alors que la quantification relative normalise l’intensité du signal des métabolites par rapport à d’autres standards de la même famille moléculaire dont la zone de rétention est proche de celle des métabolites à doser (Fig. I-3). Bien qu’il puisse y avoir un biais dans ce cas, car les composés peuvent avoir un facteur d’ionisation différent en spectrométrie de masse, la deuxième méthode de quantification est la plus simple à mettre en place quand on ne connaît pas tous les composés d’un mélange complexe -comme c’est le cas pour un extrait de plante - et que la plupart des standards ne sont pas disponibles dans le commerce.

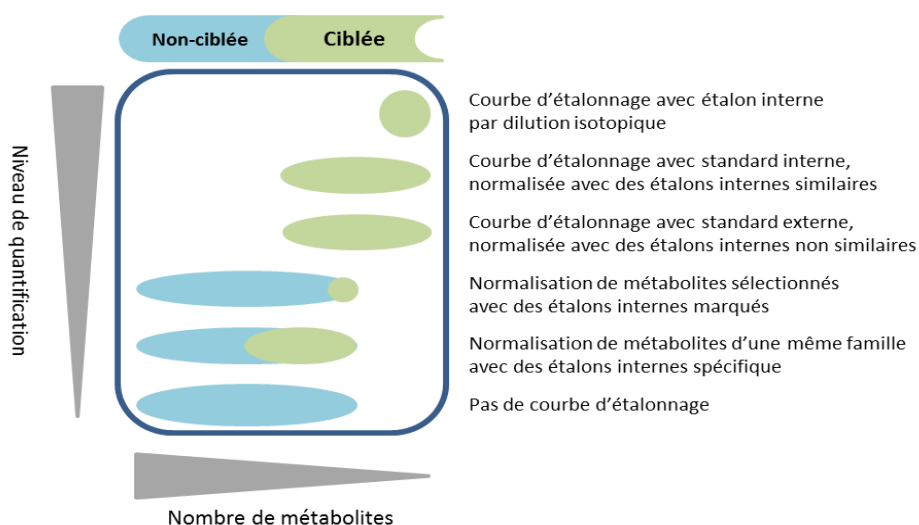


Figure I-3: Méthodes métabolomiques non ciblées et ciblées en relation avec le nombre de métabolites détectés et la fiabilité des résultats quantitatifs (Cajka and Fiehn 2016).

En analyse ciblée et non ciblée, il est nécessaire de pouvoir identifier les métabolites d’intérêt quel que soit le domaine d’application. Il est également nécessaire de connaître à quel point l’identification est fiable. Ainsi, de nombreux outils ont été développés pour répondre à cette demande.

3. Outils d'aide à l'identification des métabolites

Une des difficultés majeures à laquelle on est confronté une fois l'empreinte moléculaire obtenue reste l'identification des métabolites d'intérêt. Pour aider dans l'élucidation de métabolites inconnus, différents outils de traitement des spectres de masse ont été mis en place et sont ouverts à la communauté scientifique (Blaženović 2018).

3.1. Niveaux d'identification

Dans le cadre de l'identification des métabolites, un groupe de recherche, « Metabolomic Standards Initiative » (MSI) a défini quatre seuils de confirmation de l'identification des métabolites pour harmoniser les recherches dans les publications de métabolomique (Sumner 2007).

Le premier niveau, qui présente le plus de confiance dans l'identification du composé, est celui pour lequel le composé est identifié via le standard correspondant, dans le cadre d'une caractérisation en HPLC/HRMS : même temps de rétention, même spectre de masse HRMS et même spectre de fragmentation.

Le niveau deux correspond à un niveau un peu moins élevé, sans référence de standard, la correspondance est basée sur la similitude entre le spectre de fragmentation et une librairie extérieure ou une prédiction de spectre.

Le niveau trois est une similarité spectrale de la fragmentation avec celle de composés connus de la même famille.

Le dernier niveau est celui pour lequel le composé reste inconnu, mais il est quand même différencié ou quantifié.

Comme il est difficile d'obtenir des standards de référence à des prix raisonnables ou que l'on ne peut pas posséder tous les standards constituant un échantillon, le niveau deux est souvent le niveau visé. Le niveau deux de confirmation de métabolite demandant d'avoir des similitudes avec un spectre de fragmentation extérieur ou prédit, un point sur les bases de données web est réalisé dans les paragraphes suivants. Pour identifier les composés, différentes recherches sont possibles via : les bases de données, les bibliothèques spectrales, la prédiction in silico de spectre de masse, la fragmentation combinatoire ou les arbres de fragmentation (Hufsky 2014a).

3.2. Bases de données en spectrométrie de masse

Une base de données répertorie des composés chimiques. La recherche de molécule peut se faire par nom, par formule développée ou parfois par SMILES (Simplified Molecular Input Line Entry Specification). Des informations complémentaires dans certaines bases de données sur les composés peuvent être présentes comme les propriétés physico-chimiques des molécules, des références bibliographiques et parfois leurs activités et leurs rôles biologiques. Une liste des bases de données les plus souvent citées dans les articles et libres d'accès est donnée dans le tableau I-1.

Il est important de restreindre la recherche des métabolites dans les végétaux en incluant des informations de taxonomie. Par exemple, si les métabolites chez l'homme peuvent être obtenus à partir de la base de données HMDB (The Human Metabolome Database), les métabolites des végétaux peuvent être recherchés dans UNPD (Universal Natural Product Database).

Tableau I-1: Base de données de composés chimiques.

Base de données	Description*	Site	Référence
PubChem	96 478 070 composés	https://pubchem.ncbi.nlm.nih.gov/	Kim et al., 2016
ChemSpider	67 millions de structures	http://www.chemspider.com/	
SUPER NATURAL II	325 508 composés naturels	http://bioinf-applied.charite.de/supernatural_new/index.php	Banerjee et al., 2015
HMDB (The Human Metabolome Database)	114 100 métabolites dans le corps humain	http://www.hmdb.ca/	Wishart et al., 2018
ChEBI (Chemical Entities of Biological Interest)	54 881 composés	https://www.ebi.ac.uk/chebi/	Hastings et al., 2016
KNAPsACK	51 179 métabolites	http://kanaya.naist.jp/knapsack_jsp/top.html	
KEGG (Kyoto Encyclopedia of Genes and Genomes)	18 378 métabolites et 591 935 voies	https://www.genome.jp/kegg/	
UNPD	Produits naturels	http://pkuxj.pku.edu.cn/UNPD/	
DNP (Dictionary of Natural Products)	Produits naturels	http://dnp.chemnetbase.com	
Phenol-Explorer	polyphenols dans les aliments et les plantes	http://phenol-explorer.eu/	Rothwell et al., 2013
NAPRALERT (Natural Products Alert)	Produits naturels (Métabolites secondaires, éthnomédicales, pharmacologiques) décrits dans les articles	www.napralert.org	
CAS		https://www.cas.org/support/documentation/chemical-substances	

*Nombres donnés en 2018/08

3.3. Bases de données spectrales

De plus en plus, des bibliothèques spectrales se développent et sont accessibles bien qu'il faille rester prudent lors de leur utilisation. En effet, lors de l'acquisition en couplage LC/MS d'un spectre de masse de métabolite, il faut bien prendre en compte la technologie de l'appareillage, qui peut induire d'un appareil à un autre des modifications de réponses quant à l'abondance relative des ions de fragmentation (Bristow 2004). Le tableau I-2 donne quelques-unes des bases générales en libre accès qui peuvent servir de références.

Tableau I-2: Bases de données (Bdd) spectrales.

Base de données		Description	site	référence
GNPS	MS/MS	Bdd communautaire	https://gnps.ucsd.edu/	Horai 2010 Guijas 2018 Zhu 2013
Massbank	EI, MS/MS,MS ⁿ	Grande bdd communautaire	http://www.massbank.jp/	
Metlin	CIS-MS/MS	Analyseurs Qtof	https://metlin.scripps.edu/	
MoNA	EI, MS/MS,MS ⁿ	collection automatique de spectres	http://mona.fiehnlab.ucdavis.edu/	
mzCloud	MS/MS, MS ⁿ , CID et HCD	Spectre MS ⁿ	https://www.mzcloud.org/	
ReSpect	MS/MS, RT	Bdd de métabolomique des plantes	http://spectra.psc.riken.jp/	

Devant la grande diversité et le nombre de molécules chimiques connues, la diversité des analyseurs d'ions en spectrométrie de masse ainsi que la diversité des valeurs de l'énergie de collision à appliquer pour fragmenter des ions parents, les bases de données spectrales ne peuvent pas regrouper toutes les possibilités de spectres de fragmentation. Des chercheurs se sont donc intéressés à créer des outils pour prédire des spectres de fragmentation et pouvoir identifier au mieux les nouveaux composés.

3.4. Prédiction in Silico de la fragmentation en MS

Le principe de la fragmentation in silico est de calculer les fragmentations les plus probables. Ensuite, les approches prédictives permettent de simuler des spectres de masse de fragmentation à partir de ces calculs. Les calculs sont réalisés à partir de données théoriques issues de standards et de modélisation par des calculs qui suivent les règles de chimie. La génération de spectres de masse informatiquement s'est beaucoup améliorée ces dernières années. Ces générations sont simulées selon quatre méthodes basées sur : la chimie quantique, l'apprentissage automatique, l'heuristique ou les réactions chimiques. Les logiciels permettent en plus de comparer les spectres expérimentaux aux spectres prédits. Quelques-uns des logiciels de fragmentation in-silico sont décrits dans le tableau I-3.

Tableau I-3: Logiciels de fragmentation in silico.

Base de données	Approches prédictives	Bibliothèques	sites	références
QCEIMS	Chimie quantique		qceims@thch.uni-bonn.de	Grimme 2013
ISIS	Apprentissage auto.			Hufsky 2014b
FingerID	Apprentissage auto.		https://www.csi-fingerid.uni-jena.de/	Hufsky 2014b
LipidBlast	Approche heuristique		http://fiehnlab.ucdavis.edu/projects/LipidBlast	Kind 2013
Mass frontier	Réactions chimiques			
MS-Finder	Interprétation chimique		http://prime.psc.riken.jp/Metabolomics_Software/MS-FINDER/	Lai 2018
MetFrag	Fragm. combinatoire	Kegg, PubChem, Chempider,...	https://msbi.ipb-halle.de/MetFragBeta/	Wolf 2010
MetFusion	Fragm. combinatoire	MassBank Kegg, PubChem, Chempider	https://msbi.ipb-halle.de/MetFusion/	Gerlich and Neumann 2013
CFM-ID	Apprentissage auto.	Kegg, HMDB	http://cfmid.wishartlab.com/	Allen 2014
FTBLAST	Arbre de fragmentation			Hufsky 2014b
CSI:FingerID	Apprentissage auto. Arbre de fragmentation	PubChem	https://www.csi-fingerid.uni-jena.de/	Dührkop 2015
MIDAS-G	Fragm. combinatoire		http://facultyweb.mga.edu/yingfeng.wang/Assets/midasg/midas-g.htm	Wang 2017
ChemDistiller	Apprentissage auto.		https://bitbucket.org/iAnalytica/chemdistillerpython	Laponogov 2018
MAGMA	Arbre hiérarchique	PubChem	https://github.com/NLeSC/MAGMA	Ridder 2014

La **chimie quantique** utilise les notions de physique et de chimie pour générer des spectres de masse. L'**apprentissage automatique** peut se faire par réseaux de neurones artificiels ou en utilisant les machines à vecteur support.

Les **approches heuristiques** trouvent des motifs similaires entre des motifs connus et définis et les fragments. Cette approche est utilisée pour des classes de molécules ayant les mêmes types de fragmentation telles que les lipides ou les peptides.

La **fragmentation combinatoire** tente d'expliquer les pics dans un spectre en utilisant des sous-molécules d'une structure moléculaire candidate générée par la déconnexion des liens. Certains inconvénients des méthodes basées sur la fragmentation combinatoire sont les difficultés à expliquer les fragments issus de réarrangements structuraux, plus il y a de fragments, plus on peut avoir de « faux fragments » et diminuer le score de bonne réponse.

Dans la littérature, deux types d'**arbres de fragmentation** sont présentés (arbre de fragmentation et arbres spectraux).

Les arbres de fragmentation sont construits par prédiction des schémas de fragmentation in silico en se basant sur les pertes de neutre connues (Rasche 2011). Chaque nœud peut expliquer un pic de masse dans le spectre de fragmentation (Fig. I-4). Il n'y a pas besoin de bibliothèque spectrale ni de base de données.

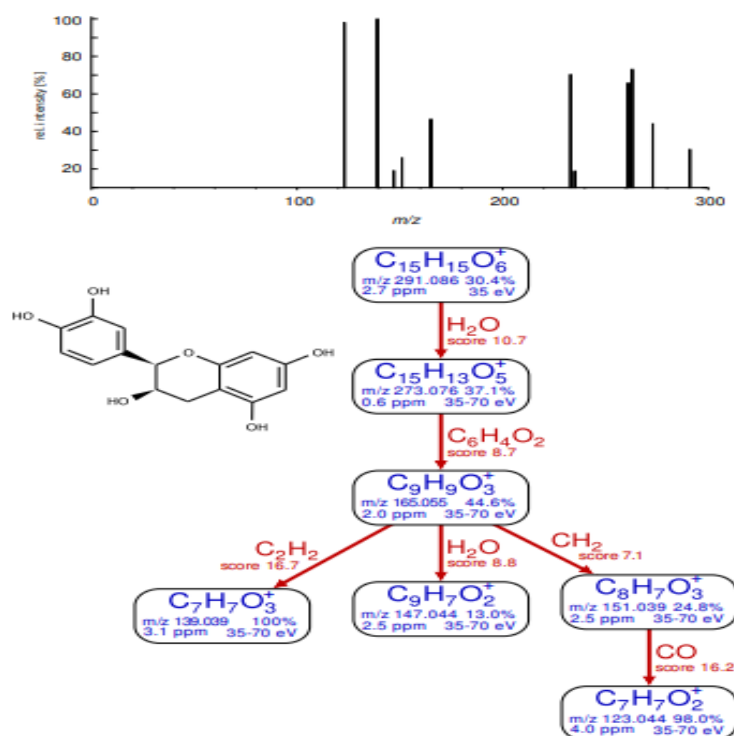


Figure I-4: Arbre de fragmentation hypothétique de la molécule (-)-épicatéchine calculé à partir des données du spectre de masse obtenu sur un orbitrap (Thermo Scientific). Les nœuds en bleu sont les pics de spectres de masse et leur formule brute (m/z ; erreur ; énergie de collision), les liens en rouge correspondent à des pertes de neutres hypothétiques, d'après Rasche 2011.

Les arbres spectraux sont basés sur l'observation des ions dans les spectres de masse MS^n obtenus et leur hiérarchisation (fig. I-5). Ils sont caractéristiques de la structure moléculaire (Kasper 2012). Ces arbres peuvent être transformés en empreintes digitales et peuvent être comparés entre eux. Des arbres similaires sont dus à des composés ayant des structures proches, alors que l'inverse n'est pas forcément vrai, seule une partie de l'arbre peut se ressembler. Ainsi par cet outil, une bibliothèque de composés peut être créée et peut être interrogée pour avoir un score de similarité ou des informations de structure pour permettre d'élucider les composés inconnus.

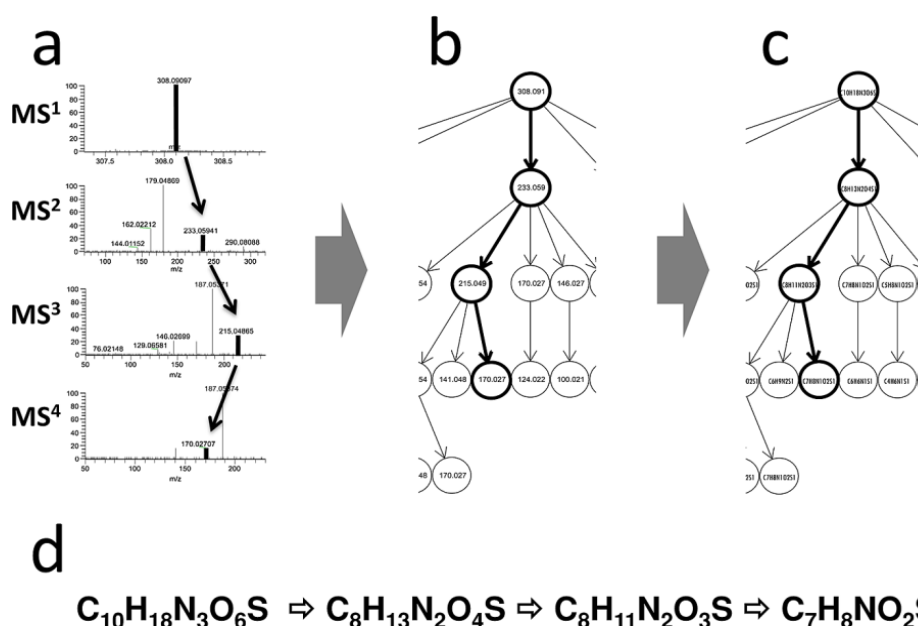


Figure I-5: L'arbre spectral (c) avec ses formules brutes est obtenu à partir des relations entre les pics précurseurs et fragments (b) obtenus des différents spectres de fragmentation en MSⁿ (a). Un chemin de formules brutes identifiées (d) sans ambiguïté est proposé, les pics, nœuds et formules correspondants sont en gras, d'après Kasper (2012).

Afin d'améliorer ces outils, des chercheurs ont créé en 2012 un concours CASMI (Critical Assessment of Small Molecule Identification) ouvert à la communauté sur l'identification de petites molécules à partir de spectres de masse, afin de tester et d'améliorer leurs différentes approches d'identification. Les logiciels tels que MS-Finder, MetFrag, CFM-ID et CSI : FingerID ont été testés via ce concours d'identification de spectres inconnus (CASMI 2017). Certains logiciels se sont démarqués comme CSI : FingerID, mais sans jamais pouvoir tout prédire correctement. Le concours permet d'évaluer pour ceux qui veulent y participer, les différentes approches prédictives qu'elles soient basées sur l'apprentissage (comme CSI : Finder) ou sur l'interprétation chimique (comme MS-Finder). En 2016, ce concours a permis de mettre en évidence les progrès des différentes méthodes et la complémentarité des deux types d'interprétation (Schymanski 2017). Par contre, les fragments obtenus après réarrangements, comme ceux de McLafferty, restent souvent sous-estimés avec ces approches.

Beaucoup de bases de données et d'outils d'aide à l'annotation des fragments ont donc été développés. Dans cette thèse le recours à ces outils a été nécessaire notamment dans le chapitre 3 pour l'identification des molécules présentes dans des extraits de soies de maïs. Metlin fait partie des bases de données ayant procuré le plus de réponses d'identification sur les molécules standards à partir des spectres de fragmentation en mode d'ionisation négatif obtenus avec un analyseur QToF. Ces standards ont été analysés dans les mêmes conditions et dans la même séquence que les extraits. La recherche d'identification des molécules de l'extrait a été effectuée avec cette base de données.

Afin de visualiser les résultats de métabolomique, d'autres outils existent, comme les traitements statistiques pour comparer un très grand nombre d'échantillons en même temps ou des traitements graphiques pour représenter la diversité moléculaire d'un échantillon.

4. Outils d'aide à la comparaison des données MS

Les analyses en spectrométrie de masse sont souvent utilisées pour étudier la diversité chimique d'extraits complexes lors d'étude métabolomique afin de répondre à une question biologique. De plus en plus, l'étude métabolomique met en jeu l'acquisition d'un très grand nombre de données liées à l'analyse simultanée d'un grand nombre de molécules et d'échantillons. Un traitement des données est donc crucial afin de faciliter l'interprétation et de ne pas donner de conclusions erronées. Différentes étapes sont nécessaires pour passer des données brutes obtenues lors de l'analyse à l'identification de biomarqueurs (Boccard 2010). En premier lieu, les données brutes sont corrigées au niveau de leur ligne de base et de leur bruit de fond. Ensuite une déconvolution du spectre de masse de chaque composé peut être envisagée pour associer les adduits, les isotopes et parfois les fragments aux pics moléculaires. Un alignement des chromatogrammes des échantillons qui sont à comparer est effectué quand il y a un grand nombre d'analyses, et une normalisation des données est générée pour éviter qu'une variable n'ait un poids plus important que les autres. Après ce ou ces pré-traitements, des analyses statistiques peuvent être réalisées afin d'avoir des informations visuelles à partir d'un très grand jeu de données et mettre en évidence les biomarqueurs d'intérêt, qui pourront ensuite être caractérisés par des analyses complémentaires.

4.1. Analyse statistique

Depuis plusieurs années, des logiciels de traitements de données ont été développés. Ils sont facilement accessibles sur internet et pour certains gratuits (Coble 2014). Le tableau I-4 donne quelques sites qui permettent de réaliser des pré-traitements et/ou des analyses statistiques de données sur des données brutes LC/MS. Ils permettent de traiter un grand nombre d'échantillons contenant de nombreuses molécules et permettent une meilleure visualisation de l'ensemble des molécules sur l'ensemble des échantillons analysés pour mettre en évidence des similitudes ou des différences entre échantillons.

Tableau I-4: Liste de logiciel de pré-traitement et/ou de statistique.

Logiciel	Pré-traitement	Statistique	Site	Référence
Mzmine	X		http://mzmine.github.io/	Pluskal 2010
MetAlign	X		https://www.wur.nl/en/show/MetAlign-1.htm	Lommen2009
MeltDB 2,0	X	X	meltdb.cebitec.uni-bielefeld.de	Kessler 2013
XCMS		X	xcmsonline.scripps.edu	Gowda 2014
MetaboAnalyst		X	http://www.metaboanalyst.ca/	Xia and Wishart 2016

Des analyses statistiques univariées et/ou multivariées permettent de représenter visuellement la variabilité des variables (molécules) et la répartition des individus (échantillons) en fonction des variables et d'observer s'ils forment ou non des groupes. Elles permettent également de mettre en évidence les métabolites pertinents qui sont liés à la discrimination de ces groupes. Ces métabolites peuvent être ensuite identifiés ci-cela n'a pas été réalisé avant le traitement des données.

Les analyses de données **univariées** permettent d'étudier les variables séparément par des méthodes descriptives ou prédictives. Ces analyses permettent ainsi de décrire la variable, de représenter sa dispersion, et son intervalle de confiance, que cette variable soit qualitative ou quantitative. Les analyses de données univariées sont aussi bien utilisées dans la représentation de la réponse d'un composé dans différentes analyses que lors de la comparaison d'échantillons ou lors de la validation de méthode.

Les analyses de données **multivariées** sont multidimensionnelles (Liland 2011). Elles permettent une confrontation entre de nombreuses informations, ce qui est souvent plus riche que l'examen séparé des données. Elles permettent également de faire une représentation simplifiée de grands tableaux de données, d'en tirer les tendances les plus marquantes et de réaliser une hiérarchisation. La représentation simplifiée se fait par la projection des données initiales dans un espace de plus faibles dimensions qui permet une visualisation des données plus facile et ainsi facilite leur interprétation. Dans le cadre d'études métabolomiques il est nécessaire de disposer d'outils statistiques adaptés à l'interprétation des données (Worley and Powers 2013). Il existe deux approches de traitement statistique multivarié: les méthodes non supervisées et celles supervisées.

4.1.1. Les analyses non supervisées

Les analyses non supervisées permettent d'observer la variation de toutes les données sans préciser quels sont les individus à regrouper ensemble. Les modèles statistiques que l'on peut retrouver sont la Classification Ascendante Hiérarchique (CAH) et l'Analyse en Composantes Principales (ACP).

La **Classification Ascendante Hiérarchique** est une méthode de classification qui permet de regrouper les individus en minimisant la variance intra-groupe et en maximisant la variance inter-groupe. La CAH permet de classer les individus par mesure de la similarité sans tenir compte de leur appartenance à des groupes similaires ou différents. La représentation des individus se fait suivant un dendrogramme (Gordon 1987).

L'**Analyse en Composantes Principales** est une méthode de changement de variables (Wold et al. 1987, Abdi and Williams 2010). Elle crée un nouveau référentiel à partir des variables initiales pour représenter les données dans des espaces de plus faible dimensionnalité. Elle produit de nouveaux axes (composantes principales) orthogonaux indépendants qui sont créés par des combinaisons linéaires des variables initiales. La première composante principale porte le maximum de la variance. Cette méthode permet de représenter l'ensemble des données et de mettre en évidence la variabilité via les graphiques du « score plot » et du « loading plot ». Le premier graphique permet de représenter les individus (les échantillons), alors que le deuxième est réalisé à partir des variables (les molécules).

Ces représentations permettent de classer et/ou de regrouper, sans à priori, les échantillons en fonction de leur similitude. Elles corrélient également les molécules aux échantillons, en montrant les composés qui influencent la séparation des groupes.

4.1.2. Les analyses supervisées

Les analyses supervisées étudient les individus en indiquant au préalable les groupes dans lesquels ils appartiennent. Elles permettent de mettre en évidence les variables prédictives et dépendantes. Les modèles statistiques que l'on peut retrouver sont la régression des moindres carrés partiels (PLS), en analyse discriminante (PLS-DA) ou en méthode parcimonieuse dite sparse PLS-DA (sPLS-DA) ou Orthogonale PLS-DA (O-PLSDA).

La méthode **PLS** permet de réduire la dimensionnalité du problème pour améliorer la performance des méthodes de création du modèle. La **PLS-DA** est une méthode supervisée basée sur les régressions linéaires. Le modèle maximise la séparation entre les classes auxquelles appartiennent les individus. Il permet de mettre en évidence les métabolites qui contribuent à la séparation des différents groupes. La méthode **O-PLSDA** est une amélioration de la méthode PLS-DA pour discriminer des groupes avec des données multivariées (Trygg and Wold 2002). Cette méthode est un modèle de régression calculé à partir des données multivariées et une variable contenant les informations d'un groupe. Un composant unique est utilisé comme prédicteur d'une classe et les autres décrivent la variation orthogonale au premier composant prédictif (Westerhuis 2010). Cette méthode, peut aussi bien être utilisée quand on a des jeux de données provenant d'analyse en spectrométrie de masse acquis aussi bien en mode d'ionisation positif que négatif, provenant d'analyse bidimensionnelle en RMN ou des données hétérogènes de biologie et de métabolomique (Boccard and Rutledge 2013). La méthode **sPLS-DA** est basée sur la méthode PLS-DA avec une diminution du nombre de variables (Lê Cao 2011). Ainsi une sélection des variables pour un ensemble de données est effectuée et permet d'éliminer le bruit. Comme chaque dimension PLS se concentre sur la différenciation de certaines populations et puisque sPLS sélectionne un sous-ensemble de variables sur ces mêmes dimensions, chacun de ces sous-ensembles permet de mieux différencier les populations.

Tous ces traitements de données peuvent donner des conclusions erronées, comme des faux positifs ou des faux négatifs dus entre autres à des tailles d'échantillon inadéquates, une mauvaise utilisation des outils statistiques ou un sur-apprentissage des jeux de données (Broadhurst 2006).

Les analyses statistiques multivariées sont de bons outils de visualisation qui permettent de mettre en évidence des similitudes ou des différences entre échantillons et de comparer ainsi un grand nombre d'échantillons. Il peut être également intéressant d'associer à ces analyses statistiques des représentations graphiques basées cette fois-ci sur la structure chimique des molécules et ainsi de proposer un tri visuel des molécules de par leur structure chimique afin d'aider à l'identification des molécules d'intérêt. Ainsi des représentations telles que les représentations de Kendrick, de Van Krevelen et des réseaux moléculaires peuvent être utilisées pour cibler plus rapidement des familles d'intérêt ou identifier spécifiquement les biomarqueurs différenciants issus des analyses statistiques.

4.2. Représentation graphique des échantillons en fonction d'une propriété de leurs molécules

Les molécules constituant les milieux complexes possèdent des structures chimiques très diverses et un échantillon tel qu'un extrait de plante peut contenir de nombreuses familles moléculaires représentatives de cette diversité structurale. Les représentations graphiques de Kendrick et Van Krevelen utilisent les formules brutes des molécules, alors que les réseaux moléculaires développés par Dorrestein utilisent les propriétés de la fragmentation des molécules.

4.2.1. Représentation de Kendrick

Kendrick a proposé en 1963 de travailler avec une nouvelle échelle de masse dans laquelle la masse du radical CH_2 serait de 14.000000 u au lieu de 14.01565 u. La représentation graphique trace le défaut de masse de Kendrick (différence entre la masse nominale et la masse de Kendrick), en fonction de la masse de Kendrick. La masse de Kendrick correspond à la transformation de la masse exacte mesurée par la masse du méthylène comme référence (14/14.01565). Les molécules qui ne diffèrent que d'un ou plusieurs groupes alkyles auront le même défaut de masse et ils seront plus facilement repérables quand les analyses des échantillons seront réalisées avec un spectromètre de masse à haute résolution. Sur la base de cette proposition, les molécules peuvent être représentées graphiquement, de sorte que les molécules d'une série homologue ne différant que par leur nombre de motifs alkyle seront séparées sur l'horizontale (CH_2O_2 ; $\text{C}_2\text{H}_4\text{O}_2$; $\text{C}_3\text{H}_6\text{O}_2$) et celles ne différant que par leur degré d'insaturation seront séparées en vertical ($\text{C}_3\text{H}_6\text{O}_2$; $\text{C}_3\text{H}_4\text{O}_2$) (Fig. I-6). Ainsi avec cet exemple, les différentes classes d'acides carboxyliques (saturé, insaturé et hydroxy saturé) sont séparées et alignées suivant le même défaut de masse, tandis que les points intermédiaires visualisés ne font donc pas partie de ces différentes classes.

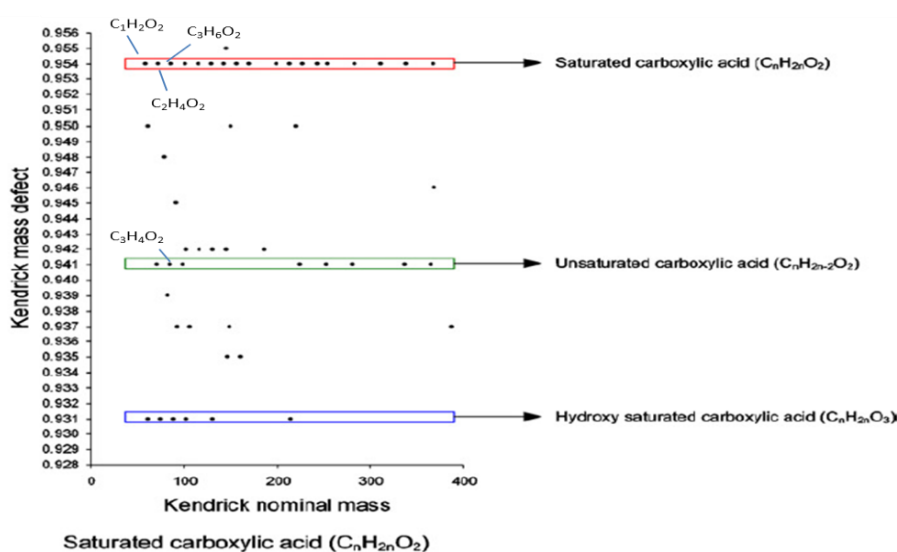


Figure I-6: Représentation de Kendrick basée sur les masses exactes de molécules avec une fonction acide carboxylique extraites de la base de données KEGG, d'après la publication de Werner 2008.

Cette représentation a souvent été utilisée dans la caractérisation des pétroles bruts ou raffinés, qui ont une composition simple en éléments C, H, N, O, S (Marshall and Rodgers 2004). Cette cartographie permet facilement de mettre en évidence les différences entre pétroles bruts de provenances différentes et de faire le suivi des étapes du traitement (Fig. I-7). Ainsi Panda met en évidence qu'après l'analyse par FTICR-MS de trois fractions d'un mélange de composés polycycliques aromatiques sulfurés, la représentation de Kendrick permet de montrer les différences dans la composition de ces trois fractions. Bien que les trois fractions répondent sensiblement dans la même gamme de masse, les défauts en masse de Kendrick (KMD) et les nombres équivalents de doubles liaisons (DBE) sont différents. Cette différence est due à un changement de classe ou de type de composé. Ainsi la fraction 1 serait composée de nombreuses molécules de la série des benzothiophènes (DBE 6) et des naphthénobenzothiophènes (DBE 7). La fraction 2 serait riche de la série des dibenzothiophènes (DBE 9) et des acénaphthénothiophènes ou des benzothiophènes avec un substituant phényle (DBE 10). Tandis que la fraction 3 contiendrait la série des benzonaphthothiophènes (DBE 12) et des dérivés de benzonaphthothiophènes avec un cycle naphthéno ou des dibenzothiophènes avec un groupement phényle (DBE 13).

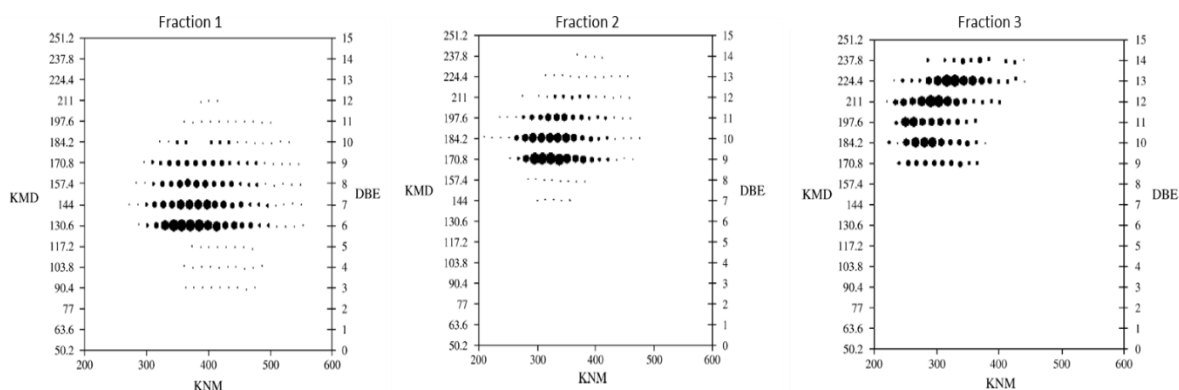


Figure I-7: Représentation de Kendrick de trois fractions d'hétérocycles de soufre aromatiques polycycliques séparées sur une phase stationnaire de cyclodextrine, d'après Panda (2006). Représentation des défauts de masse de Kendrick (KMD) et du nombre équivalent de doubles liaisons (DBE) en fonction de la masse nominale (KNM).

La représentation de Kendrick a également aidé dans l'interprétation des spectres de masse obtenus en haute résolution de matière organique naturelle (Kim 2003). Au lieu de prendre comme nouvelle échelle de masse, la masse du CH_2 à 14 u, cette représentation peut être réalisée avec d'autres échelles de masse, comme celle de H_2 , H_2O ou O pour mettre en évidence d'autres différences. D'autres domaines utilisent également la représentation de Kendrick comme outil complémentaire dans l'analyse des vins (Roullier-Gall 2014) et des thés (Kuhnert 2010). Ces représentations permettent de distinguer les composés de séries homologues.

Kuhnert et al. (2010) ont mis en évidence les limites de la représentation par CH_2 dans le cas d'analyse d'échantillons complexes, tel qu'un extrait de thé noir, lorsque ces échantillons renferment différents types de familles moléculaires et donc plusieurs sortes de différence. Les représentations de Kendrick doivent être établies pour des séries homologues.

4.2.2. Diagramme de Van Krevelen

Un autre type de diagramme a été proposé : le diagramme de Van Krevelen (Kim 2003 et Wu 2004) qui est plus général. Ce diagramme (Fig. I-8) est obtenu en projetant la formule brute ($C_aH_bO_cN_d$) de chaque composé selon deux ou trois axes. Le premier axe classe chaque composé en fonction de la valeur de son rapport atomique entre le nombre d'hydrogène et le nombre de carbone (H/C) présents dans sa formule brute, le deuxième axe classe selon la valeur du rapport atomique entre le nombre d'oxygène et le nombre de carbone (O/C), et le troisième axe classe selon la valeur du rapport atomique entre le nombre d'azote et le nombre de carbone (N/C). D'après la Figure I-8, le rapport H/C permet de séparer les composés en fonction de leur degré d'insaturation (BB'), alors que le rapport O/C distingue les composés en fonction du nombre d'oxygène (DD') et les diagonales permettent de distinguer les molécules par rapport à leur méthylation (AA') et leur hydratation (CC').

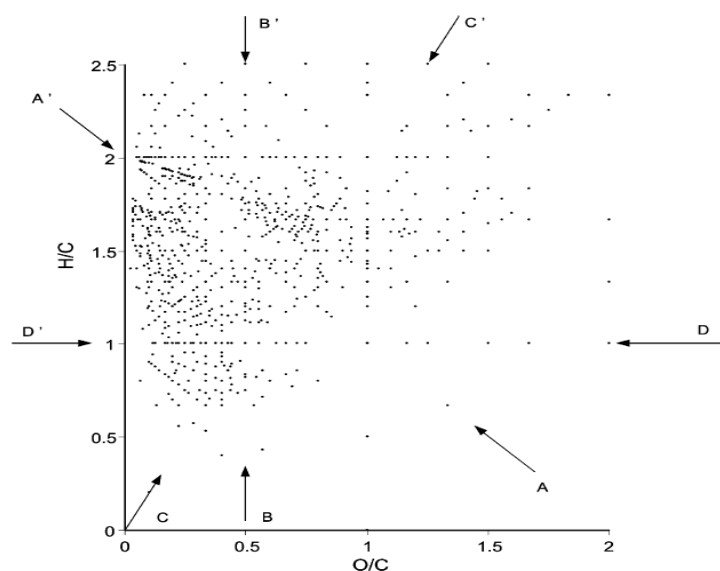


Figure I-8: Représentation de Van Krevelen basée sur 10 000 formules brutes extraites de la base de données KEGG, d'après la publication de Werner 2008.

Ohta et al. ont ainsi montré sur la base de ce type de diagramme, qu'en partant d'une molécule, les autres molécules pouvaient être décrites par les différentes réactions enzymatiques (méthylation, glycosylation ...) et que ces molécules formaient une ou plusieurs voies métaboliques (Fig I-9).

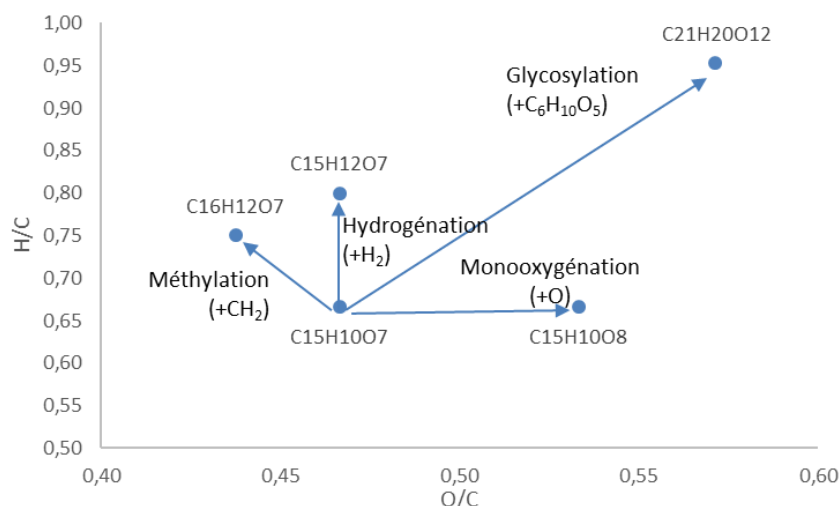


Figure I-9: Représentation des modifications structurales suite à des réactions enzymatiques extrait de la publication d'Ohta 2010.

Souvent utilisé pour représenter des échantillons complexes de pétrole (Wu 2004) et de matière organique naturelle (Kim 2003), le diagramme de Van Krevelen peut également être utilisé dans le domaine des plantes pour avoir une représentation globale des extraits. Une telle approche permet de classer les molécules dans les différentes familles de composés et d'avoir une représentation graphique qualitative sur la distribution des molécules (Fig. I-10). Le diagramme peut être ensuite utilisé pour déterminer rapidement l'appartenance des composés du mélange complexe inconnu aux différentes familles moléculaires en fonction de leur formule brute. Il a été utilisé notamment pour obtenir un regroupement par famille moléculaire des différents espaces chimiques (CHO, CHOS, CHON, CHONS) d'échantillon de vin (Gougeon 2009) ainsi que pour obtenir une meilleure compréhension des molécules modifiées lors de la transformation d'aliment, telle que la fabrication du caramel ou de la torréfaction de café (Kuhnert 2013).

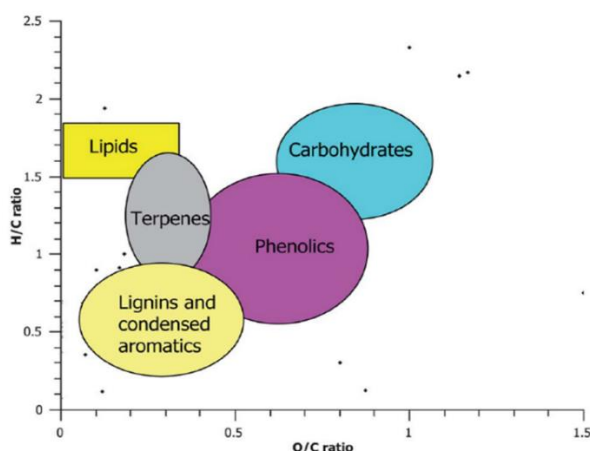


Figure I-10: Représentation de Van Krevelen de produits naturels obtenus par des recherches de formules brutes dans des bases de données avec le positionnement des familles moléculaires, d'après Kuhnert (2013).

Ces deux premiers diagrammes décrits sont assez simples à tracer puisqu'ils sont basés uniquement sur l'utilisation de la formule brute, obtenues en HRMS, des différentes molécules présentes dans l'échantillon. Ils apportent ainsi une première classification intéressante des molécules, qui reste cependant limitée dans le cas d'échantillons complexes. En effet, des molécules de formules brutes identiques peuvent avoir des structures différentes et ils existent des chevauchements entre les différentes familles moléculaires. Un troisième outil complémentaire, le réseau moléculaire, a été développé plus récemment pour représenter les molécules de l'échantillon en fonction de leur similarité structurale en utilisant leur fragmentation en spectrométrie de masse.

4.2.3. Réseaux de similarité spectrale (Réseaux moléculaires)

Beaucoup d'études sur les souches bactériennes ont permis de développer de nouveaux outils comme les réseaux moléculaires qui peuvent maintenant servir dans le domaine des plantes notamment pour compléter la stratégie traditionnelle de dérégulation (Yang 2013). Les réseaux moléculaires peuvent également être un atout dans la découverte de pistes thérapeutiques, du suivi du métabolisme de médicament et des diagnostics cliniques (Quinn 2017). Dorrestein et al. ont créé cet outil puissant de bio-informatique pour faciliter l'identification de molécules. Les réseaux moléculaires permettent ainsi de visualiser sous forme de carte de similarité spectrale l'ensemble des résultats d'analyses MS² et ont été appliqués pour la première fois pour le profilage métabolique en 2012 (Watrous 2012, Nothias-Scaglia 2015). Cet outil s'est développé par la création de la plateforme GNPS (Global Natural Products Social Molecular Networking) (<http://gnps.uscd.edu>) qui est une plate-forme collaborative qui stocke, analyse et diffuse des connaissances sur les spectres de masse bruts et les spectres MS² de produits naturels à l'ensemble de la communauté scientifique avec des incréments en continu de données spectrales déposées (Wang 2016). GNPS permet également d'analyser un grand ensemble de données et de les comparer à toutes les données accessibles au public. Cette approche pourrait devenir une bibliothèque spectrale permettant de répondre aux questions d'identification des composés dans les plantes (Wang 2016) car lors de la soumission des analyses MS² à la plate-forme GNPS, une recherche dans la librairie du GNPS des spectres de fragmentation peut être demandée.

Le réseau fonctionne sur le principe qu'en spectrométrie de masse, des molécules appartenant à une même famille moléculaire ont des pertes de neutres qui se ressemblent. Lors de la fragmentation d'ions parents de molécules proches, il est observé en effet des fragments communs et/ou des voies de fragmentation similaires. Ainsi, la plateforme GNPS compare les spectres de fragmentation deux à deux par alignement spectral et attribue un score. Ce score correspond à un facteur de similarité, appelé **similarité cosinus**, attribué entre chaque paire de spectres. Ce score est proche de 1 quand les spectres de fragmentation se ressemblent et proche de zéro quand ils sont différents. La plateforme GNPS permet d'obtenir un fichier regroupant un ensemble de données (ions précurseurs, clusters, scores de similarité, proposition de molécules de la librairie, nombre de scan, ...), qui peut être visualisé par le logiciel Cytoscape (Shannon 2003) sous forme de réseau. Les nœuds représentent les ions parents et les connexions entre deux nœuds donnent la ressemblance entre les spectres de fragmentation des deux molécules (Fig. I-11). Les connexions se font quand les spectres ont un cosinus supérieur à une valeur donnée (souvent 0,7), ils ont ainsi une forte similarité spectrale. Ce type d'étude et de représentation permet ainsi d'avoir des regroupements de molécules par famille moléculaire et de pouvoir identifier plus facilement les composés via la recherche dans la bibliothèque de spectres

de molécules du GNPS ou en ayant inséré lors de l'analyse de l'échantillon des standards de ces familles qui se retrouveront dans les mêmes groupes et faciliteront l'identification des molécules inconnues.

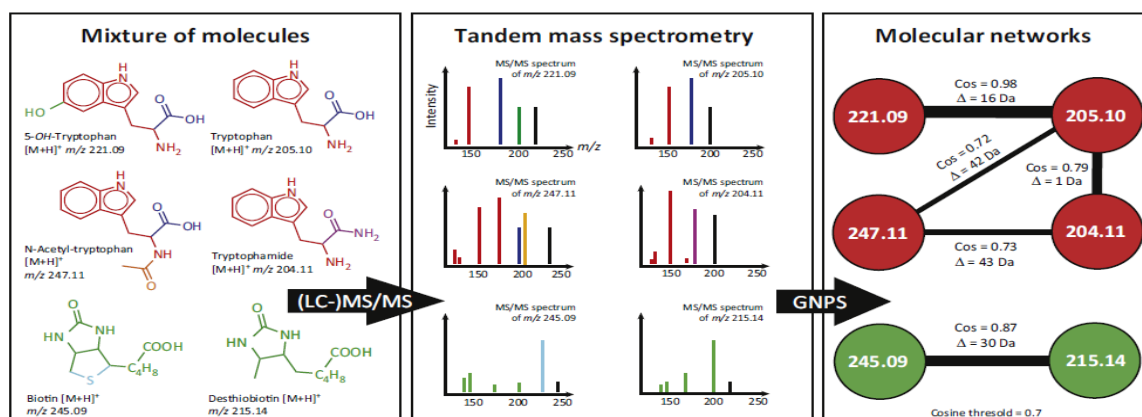


Figure I-11: Illustration d'un cluster de réseaux moléculaires issus de spectres LC/MS-MS des dérivés du tryptophane et de la biotine, d'après Quinn 2017.

Plusieurs échantillons peuvent être comparés par l'utilisation de ces réseaux, mettant ainsi facilement en évidence des clusters communs signes de molécules communes et potentiellement connues ou au contraire des clusters spécifiques signes de molécules potentiellement originales. De nouvelles molécules ou familles moléculaires appartenant spécifiquement à une souche (Duncan 2015), un genre (Purves 2016), une plante (Nothias 2017) ou une partie de plante (Ge 2017), ont ainsi pu être mises en évidence et identifiées plus facilement et rapidement. Les études par les réseaux moléculaires se répandent de plus en plus comme par exemple pour analyser une prolifération de cyanobactéries et trouver de nouveaux analogues de microcystine (Teta 2015), ou de potentiel composés antiviraux dans *Euphorbia semiperfoliata* (Nothias 2017). La combinaison du marquage isotopique des acides aminés et de la représentation en réseaux moléculaires a permis de trouver de nouveaux analogues dans les champignons filamenteux et de fournir en complément des informations sur la biosynthèse de ces composés actifs (Klitgaard 2015). Les activités d'extraits de plantes (Olivon 2017a) ou des fractions d'extraits peuvent y être incorporées et également comparer différentes souches bactériennes de lichens marins (Parrot 2018). En dehors du domaine des plantes, les réseaux moléculaires ont également permis de réaliser un criblage de métabolites de médicaments (Van der Hooft 2016).

Les analyses réalisées sur des appareils UHPLC/HRMS en vue de construire des réseaux moléculaires demandent que la méthode soit bien paramétrée lors de l'acquisition des spectres de fragmentation (nombre d'ions sélectionnés, seuil d'intensité des ions, exclusion dynamique ...). Lors de la soumission des données à la plateforme GNPS des valeurs de paramétrages (score de similarité, taille de cluster minimum ...) doivent être bien ajustées pour obtenir une bonne qualité de réseau moléculaire (Olivon 2017b). En effet, des problèmes de données redondantes, de création de fausses molécules (dû à des molécules dichargées) peuvent être rencontrés. L'un des problèmes rencontrés lors de l'analyse d'extraits de plante est que le réseau ne prend pas en compte les nombreux isomères de position qui peuvent être séparés par chromatographie et avoir des temps de rétention distincts tout en ayant le même spectre de fragmentation. Dans ces conditions, le réseau assimile souvent tous ces isomères à une seule et même molécule et les regroupe sous un même nœud. Pour pallier ce problème, Olivon et

al. (2017c) proposent de réaliser un traitement des données avant de les soumettre au GNPS et de le visualiser avec cytoscape, via des logiciels (MZmine 2 et script sous Python) qui permettront de prendre en compte le temps de rétention afin de distinguer les isomères. Le traitement des données avec le logiciel MZmine 2 permet également d'avoir une annotation de formules brutes et d'ajouter une information de semi-quantification par rapport à l'abondance de l'ion précurseur, que l'on peut visualiser par la suite sur cytoscape. Des informations de quantifications relatives et absolues provenant d'autres expériences peuvent également être ajoutées sur le réseau moléculaire dans Cytoscape (Winnikoff 2014).

Bien qu'il existe des librairies et des bases de données de molécules naturelles non volatiles et de leurs spectres de masse, elles restent encore limitées. Car toutes les molécules n'y sont pas répertoriées. Quand il y a un spectre de fragmentation associé à une molécule de la base de données, ce spectre reste dépendant de l'appareil, de l'énergie de collision ainsi que du mode d'ionisation positif ou négatif utilisés lors de son acquisition.

Comme Werner et al. l'ont montré dans un cas d'étude, il est indispensable de connaître les mécanismes de fragmentation afin de comprendre et prédire certains composés inconnus qui ne seraient pas identifiés par les bases de données et les différents outils mis à notre disposition (Werner 2008). Ainsi lors de la recherche de proposition de molécules pour l'ion de rapport m/z de 201.1133, la base de données HMDB consultée par les auteurs, n'a donné que deux propositions dont une qui ne correspondait pas au spectre de fragmentation. L'acide sébacique a été proposé comme candidat potentiel, mais trois pics chromatographiques dans l'empreinte obtenue correspondaient à la masse de ce composé. Lors de l'injection du standard d'acide sébacique, le temps de rétention obtenu correspondait bien à l'un des trois pics observés, mais pas à celui que l'analyse multivariée avait mis en évidence. Donc, après dépouillement du spectre de fragmentation et en réfléchissant sur les schémas de fragmentation possibles, les auteurs ont mis en évidence des fonctions de la molécule et fait une supposition de structure. Cette structure putative devra être confirmée par des expériences de RMN et de synthèse de la molécule (pour standard) pour une identification formelle.

La connaissance des mécanismes de fragmentation en MS des familles moléculaires ciblées s'avère une aide décisive pour interpréter de manière critique les propositions d'identification moléculaire obtenues au moyen des bases de données. Le paragraphe suivant récapitulera les différents schémas de fragmentation possibles en MS, des flavonoïdes glycosylés, composés que nous avons plus particulièrement ciblés lors des différents développements effectués au cours de cette thèse.

5. Etat de l'art sur la fragmentation des flavonoïdes glycosylés MS

Une grande partie de ce travail de thèse a consisté à développer des stratégies méthodologiques en couplage UHPLC/MS ou TLC/MS pour établir la signature moléculaire de plusieurs extraits naturels et ce, en nous intéressant plus particulièrement à caractériser la présence dans ces extraits, de composés de la famille des flavonoïdes. Ces molécules ont constitué pour nous des composés modèles privilégiés puisque nous avons pu nous appuyer pour leur identification sur l'abondante littérature publiée concernant leur mécanisme de fragmentation en MS. Ainsi dans ce paragraphe, après un bref rappel

de la structure chimique de ces flavonoïdes nous établirons un récapitulatif de l'état des connaissances sur la fragmentation des flavonoïdes glycosylés en MS.

5.1. Structure chimique des flavonoïdes

Les flavonoïdes ont tous comme élément structural de base de 15 carbones (en C6-C3-C6), disposé en deux cycles aromatiques (A et B) liés par un cycle de pyrane hétérocyclique (C) (Fig. I-12).

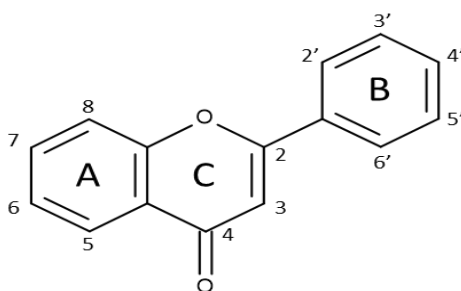


Figure I-12: Structure générique des flavonoïdes.

Les flavonoïdes se situent dans la voie de biosynthèse des phénylpropanoïdes (Fig. I-13). La synthèse des flavonoïdes commence avec la condensation d'une molécule de 4-coumaroyl-Coenzyme A et de trois molécules de malonyl-CoA produisant la naringénine chalcone. La chalcone est ensuite isomérisée par l'enzyme chalcone flavanone isomérase pour obtenir la flavanone. De cet intermédiaire central la voie diverge en plusieurs branches latérales, chacune donnant une classe de flavonoïdes (Corradini 2011), les flavones, les flavanols, les isoflavones et les dihydroflavonols qui donnent les flavonols, les tannins et les anthocyanes (Koes 1994).

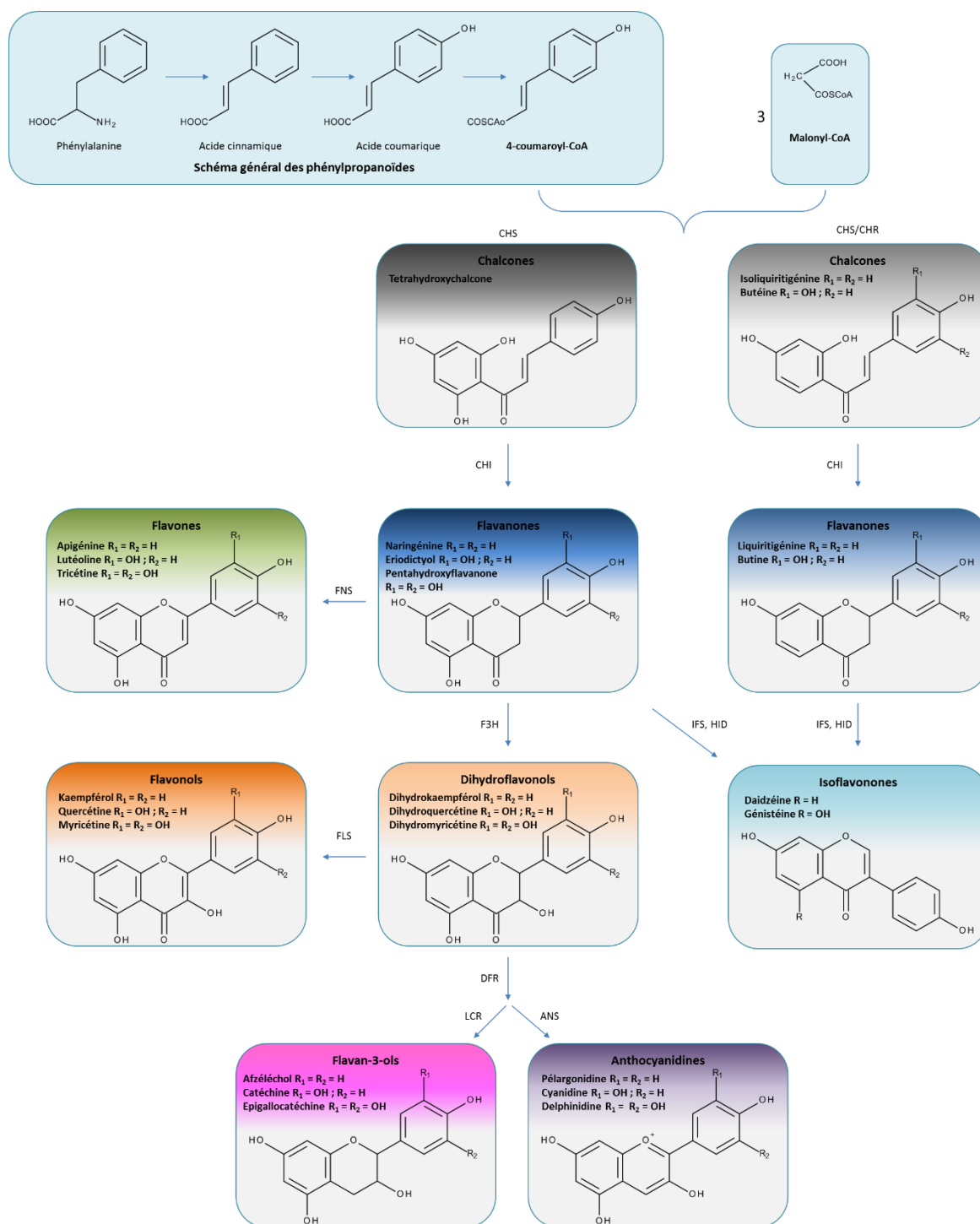


Figure I-13: Shéma partiel illustrant les voies de biosynthèse des principaux flavonoïdes. Les abréviations de nom des enzymes sont les suivantes : Chalcone synthase (CHS), Chalcone réductase (CHR), Chalcone isomérase (CHI), Flavone synthase (FNS), Flavanone 3-hydroxylase (F3H), 2-hydroxyisoflavanone déhydratase (HID), Isoflavone synthase (IFS), Flavonol synthase (FLS), Dihydroflavonol 4-reductase (DFR), Leucoanthocyanidine reductase (LCR), Anthocyanidine synthase (ANS), d'après Winkel-Shirley (2001).

Les flavonoïdes peuvent être ainsi regroupés en sept classes distinctes :

Les **flavones** sont les principaux flavonoïdes présents dans les plantes avec les flavonols et constituent l'un des plus grands groupes, caractérisés par la présence d'une double liaison entre C-2 et C-3 sur le cycle C.

Les **flavonols** sont des dérivés de la 3-hydroxyflavone et possèdent donc une fonction hydroxyle en C-3 sur le cycle C, en plus de la double liaison. Ils sont très présents dans les parties aériennes des plantes.

Les **flavanones** ne possèdent pas la double liaison des flavones. Ce sont des composés sous forme glycosylés caractéristiques des agrumes.

Les **isoflavones** ressemblent beaucoup aux flavones. Ils ne diffèrent que par la position de leur cycle B sur le cycle C. Le phényl (cycle B) se situe en position 3 pour les isoflavones au lieu de la position 2 pour les flavones. Ils sont très présents chez les Leguminosae (ou Fabaceae).

Les **anthocyanidines** sont des pigments naturels qui possèdent en milieu acide une charge positive. C'est la seule classe de flavonoïdes chargés.

Les **chalcones** (et les dihydrochalcones) ne possèdent pas de cycle C, mais une chaîne de carbone. Ils ont une importance dans la voie de biosynthèse des flavonoïdes, car une chalcone (naringénine) est le précurseur immédiat de toutes les classes de flavonoïdes.

Les **flavanols** ne possèdent pas de fonction cétone et ont une fonction hydroxyle en position 3 sur le cycle C. Ils sont présents sous forme monomérique, essentiellement (-)-épicatechine et (+)-catéchine, et sous forme condensée polymérique, les proanthocyanidines. Ces derniers contribuent à l'astringence des aliments et des boissons.

Les flavonoïdes se caractérisent par un degré élevé de diversité chimique fourni par des modifications du squelette, qui comprennent l'hydroxylation, l'*O*-méthylation, l'acylation, la glycosylation, la substitution par des fonctions phénolique ou malonique (Kachlicki 2008) et la prénylation. La glycosylation est une substitution très fréquente, on parle alors d'hétérosides constitués d'une partie phénolique aglycone (ou génine) associée en *O* et/ou *C* à un ou plusieurs sucres.

Les ***O*-glycosylflavonoïdes** (Fig. I-14) sont caractérisés par une liaison entre une fonction hydroxyle de la génine et une fonction hydroxyle du sucre, la condensation entre les deux molécules provoquant la perte d'une molécule d'eau. Les hydroxyles substitués peuvent se situer sur le cycle A en C-7 et C-5, sur le cycle C en C-3 et sur le cycle B en 3', 4' et/ou 5'. Ils sont très répandus dans le règne végétal et dans les différentes parties des plantes. Les *O*-glycosylflavonoïdes ont été retrouvés avec les différents types de squelettes de flavonoïdes existant.

Les ***C*-glycosylflavonoïdes** (Fig. I-14) sont caractérisés par une liaison entre le carbone d'un sucre et la position C-6 ou C-8 sur le cycle A du flavonoïde. Ces composés sont moins répandus et font l'objet d'un intérêt de recherche croissant en raison de leur potentiel biologique élevé. Au cours de la dernière décennie, de nombreux *C*-glycosylflavonoïdes ont été isolés et identifiés à partir d'une grande variété d'espèces végétales. Les *C*-glycosylflavonoïdes peuvent être classés en 3 types selon leur squelette de base : le type I comprend les flavones ; le type II les isoflavones ; et le type III les flavanones.

Les sucres peuvent donc se lier de deux façons différentes (*O* ou *C*-glycosylation) sur l'aglycone (Ag) ou génine (flavonoïdes sans sucres) mais les génines peuvent également être mono ou di-glycosylées,

c'est-à-dire être liées directement à un seul sucre ou à deux sucres respectivement. Quand l'aglycone est **mono-glycosylé**, cela peut être avec un sucre ou un enchaînement de plusieurs sucres liés entre eux, la plupart du temps, par une de leur fonction hydroxyle (Fig. I-14). Dans ce paragraphe, l'exemple sera donné avec deux sucres. Quand l'aglycone est **di-glycosylé**, la glycosylation peut aussi bien être 2 fois en *O* ou 2 fois en *C* mais également une fois sur la position *O* et une fois en position *C* (Fig. I-14).

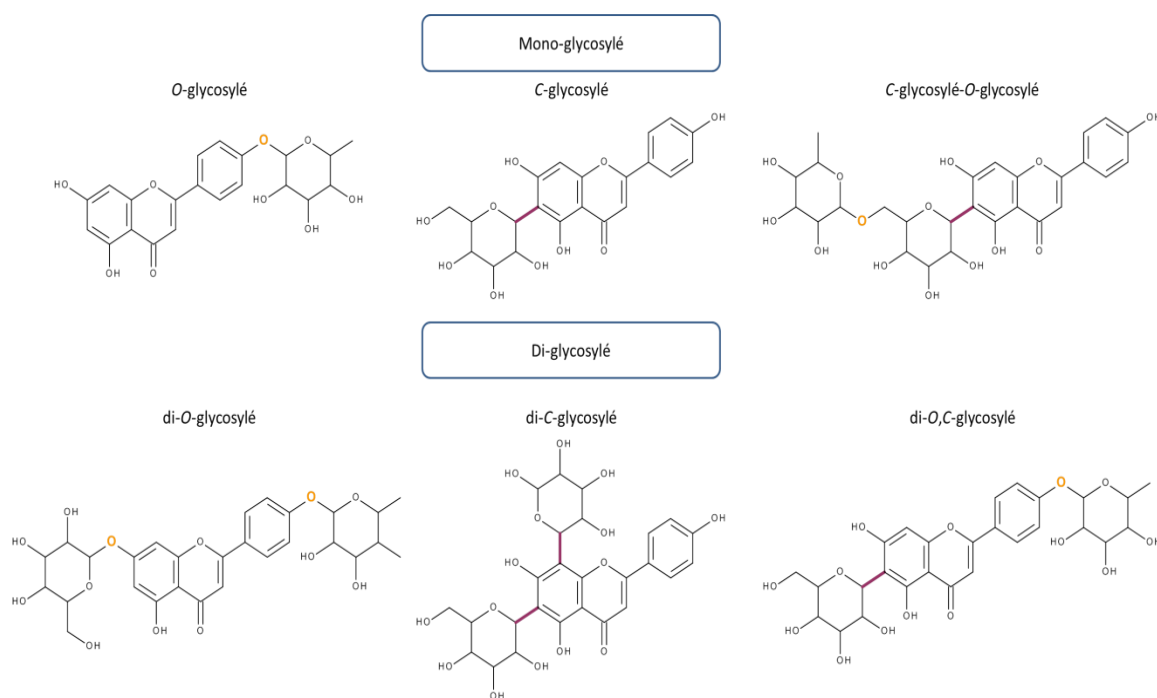


Figure I-14: Différents types de glycosylation sur un flavonoïde.

La fragmentation des flavonoïdes en spectrométrie de masse se fait selon deux schémas bien définis et décrits qui dépendent principalement de la nature de la liaison entre leur partie génine et leur partie glycoside (*O* ou *C*-glycosylation). Ainsi les différents types de substitution peuvent être distingués par une étude des voies de fragmentation.

Dans la suite du paragraphe, il sera développé uniquement la fragmentation des flavonoïdes glycosylés, car ce sont les composés les plus étudiés dans la suite de ces travaux de thèse.

5.2. Nomenclature générale liée à la fragmentation des flavonoïdes en MS

Les flavonoïdes s'ionisent dans les deux modes d'ionisation, avec généralement un meilleur rendement d'ionisation en mode d'ionisation négatif, sauf dans le cas des anthocyanes. En revanche, le mode d'ionisation positif permet d'obtenir plus de fragments que le mode d'ionisation négatif, bien que les fragments essentiels pour l'identification soient généralement présents dans les deux modes.

La fragmentation d'un flavonoïde aglycone peut s'interpréter à partir d'un mécanisme de type Retro Diels Alder (RDA) qui se produit dans le cycle C de la génine par une coupure entre deux liaisons carbone (C-C) donnant préférentiellement deux ions fragments de type $^{1,3}A_0^\pm$ et $^{1,3}B_0^\pm$ (Fig. I-15) (Abad-Garcia 2009, Cuyckens and Claeys 2004, Vukics and Guttman 2010). Ces fragments permettent d'apporter une information sur la nature de la génine et sur le nombre d'hydroxyles présents sur ses cycles A et B.

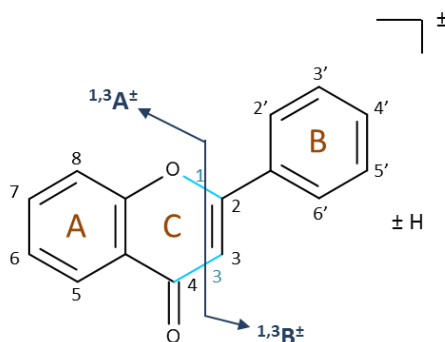


Figure I-15: Schéma de fragmentation et nomenclature des ions fragments des flavonoïdes aglycones.

La fragmentation des flavonoïdes glycosylés suit le même schéma de fragmentation que celui des polysaccharides (Cuyckens 2000). Ainsi, la nomenclature des sucres est reprise pour celle des flavonoïdes glycosylés (Fig. I-16).

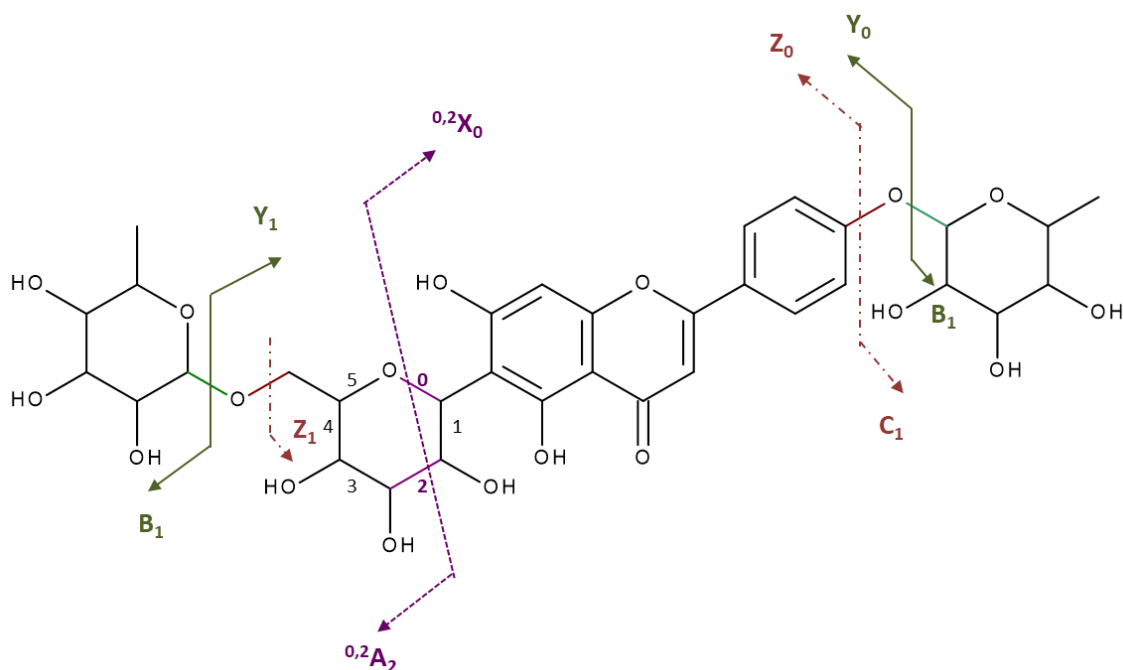


Figure I-16: Schéma de fragmentation et nomenclature des ions fragments des flavonoïdes O-glycosylés et C-glycosylés.

Lors du clivage du flavonoïde entre sa partie génine et sa partie glycosylée, le fragment de la génine est noté Y_n ou Z_n suivant si la coupure se fait respectivement à droite ou à gauche de l'hydroxyle, avec n le nombre de monosaccharides restant liés à la génine (Vukics and Guttman 2010). Les fragments glycosides complémentaires sont notés B_n ou C_n correspondant à n monosaccharides de la partie osidique.

Quand la fragmentation se réalise dans le sucre, les ions sont de type $^{i,j}X_n$, avec l'indice n portant la même valeur d'indice que pour Y et Z . Les indices i et j correspondent aux deux liaisons dans le sucre qui sont coupées lors du clivage (Cuyckens and Claeys 2004, Vukics and Guttman 2010). Ces fragments permettent ainsi d'identifier les sucres liés sur la génine et permettent de décrire les molécules à l'aide de la MS^2 . La différence entre les valeurs des rapports m/z de l'ion parent et des fragments donne l'information sur la perte de neutre et permet donc de déterminer les sucres initialement présents dans la structure moléculaire. Les ions caractéristiques sont ainsi décrits par la suite.

5.3. Fragmentation des flavonoïdes *O*-glycosylés

Lors de la fragmentation de molécules *O*-glycosylées, le fragment de la génine est toujours présent dans le spectre de fragmentation. Deux types de ruptures peuvent avoir lieu pour les flavonoïdes *O*-glycosylés, un clivage hétérolytique favorisé avec une faible énergie de collision, qui s'accompagne d'une migration d'un proton, donnant l'ion aglycone de type $[Y_0 \pm H]^{\pm}$ ou un clivage homolytique favorisé en cas de forte énergie de collision, produisant l'ion aglycone de type $Y_0^{\pm}$ ou un ion aglycone radical de type $[Y_0 \pm H]^{\pm\cdot}$. En plus de déterminer la nature du sucre, la rupture homolytique peut permettre de connaître la position du sucre sur la génine.

Ainsi en mode négatif le rapport d'abondance des ions $[Y_0-H]^{-}/Y_0^{-}$ donne des indications sur le nombre de groupements hydroxyle sur le noyau B, ainsi que sur la position des sucres sur la génine (Ablajan and Tuoheti 2013, Hvattum and Ekeberg 2003). En effet, le rapport d'abondance des ions $[Y_0-H]^{-}/Y_0^{-}$ est plus intense dans le spectre de masse de la myricétine 3-*O*-rhamnose que dans celui de la quercétine 3-*O*-rhamnose qui est lui-même plus intense que dans celui du kaempférol 3-*O*-rhamnose (Fig. I-17A). Ainsi pour une famille moléculaire donnée et un sucre donné, plus le nombre d'hydroxyles augmente sur le cycle B de la génine plus le rapport d'abondance des ions $[Y_0-H]^{-}/Y_0^{-}$ augmente. Le rapport d'abondance des ions $[Y_0-H]^{-}/Y_0^{-}$ dans les spectres de masse s'inverse en passant du kaempférol 3-*O*-glucose au kaempférol 7-*O*-glucose et également de la quercétine 3-*O*-glucose à la quercétine 7-*O*-glucose (Fig. I-17B). Donc pour une même génine et un même sucre, l'ion $[Y_0-H]^{-}$ est plus intense quand le sucre est en position 3 sur la génine, tandis que c'est l'ion Y_0^{-} qui est le plus intense lorsque le sucre est en position 7. Le rapport des ions $[Y_0-H]^{-}/Y_0^{-}$ dans les spectres de masse s'inverse aussi en passant du kaempférol 3-*O*-rhamnose au kaempférol 3-*O*-glucose et également de la quercétine 3-*O*-rhamnose à la quercétine 3-*O*-glucose (Fig. I-17A et B). Attention, pour une génine donnée et une glycosylation dans une position donnée, le rapport entre ces deux ions peut s'inverser en fonction de la nature du motif glycoside.

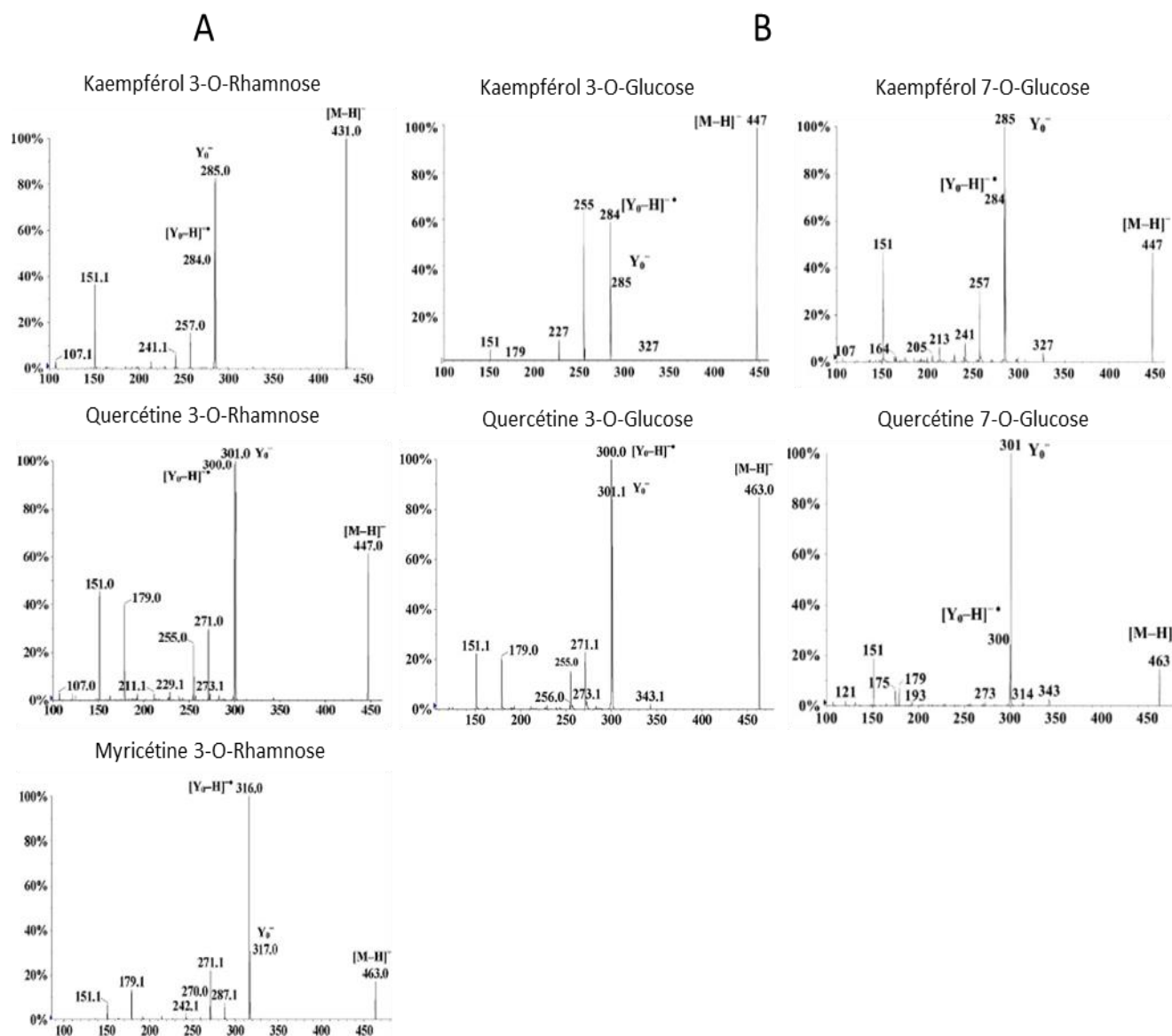


Figure I-17: Spectres de fragmentation des ions $[M-H]^-$ de différents flavonoïdes O-glycosylés, d'après Ablajan (2006 et 2013).

5.3.1. Fragmentation des flavonoïdes mono-O-glycosylés

Quand le sucre est lié à un hydroxyle de la génine, lors de la fragmentation tout le motif sucre se sépare. Le type de sucre peut être facilement déterminé alors, car il existe entre l'ion moléculaire et le fragment de la molécule sans le sucre, des pertes de neutre caractéristiques du sucre. En effet, il y aura respectivement une perte de 162 u pour un hexose, 146 u pour un déoxyhexose, 132 u pour un pentose et 176 u pour un motif acide glucuronique (Cuyckens 2000 et 2004, Ferreres 2007, Figueirinha 2008). Quant à la position du sucre, comme montré précédemment, elle peut être déduite en mode d'ionisation négatif du rapport d'abondance entre les 2 ions $[Y_0-H]^-/Y_0^-$.

5.3.2. Fragmentation des flavonoïdes mono-*O*-glycosylés-*O*-glycosylés

Lorsque la génine est monosubstituée par des polysaccharides, le flavonoïde est nommé mono-*O*-glycosylé-*O*-glycosylé. Comme pour la fragmentation des flavonoïdes mono-*O*-glycosylés, les sucres se séparent de l'aglycone, par une coupure en Y_0 . Les pertes de neutre seront donc une association des pertes de sucres simples (Tableau I-5).

Tableau I-5: Masse (en u) des pertes de neutre dans le cas de deux sucres.

Mono	Pentose	Déoxyhexose	Hexose	Ac. Glucuronique
Pentose	264	278	294	308
Déoxyhexose	278	292	308	322
Hexose	294	308	324	338
Ac. Glucuronique	308	322	338	352

Il peut également y avoir en plus, des fragments intermédiaires correspondant à la perte d'un des sucres (Fig. I-18), notés $Y_1^\pm$. Ces fragments intermédiaires ne permettent néanmoins pas d'indiquer la position du sucre dans la séquence polysaccharidique, car il peut y avoir une réaction de réarrangement qui induit la perte du sucre se trouvant entre la génine et le sucre terminal. En revanche, en mode d'ionisation positif, le rapport d'intensité des ions Y_0^+/Y_1^+ est plus grand pour une liaison de type 1->2 que pour une liaison de type 1->6. Ce rapport permet d'avoir une information sur la liaison interglycosidique (Fig. I-18).

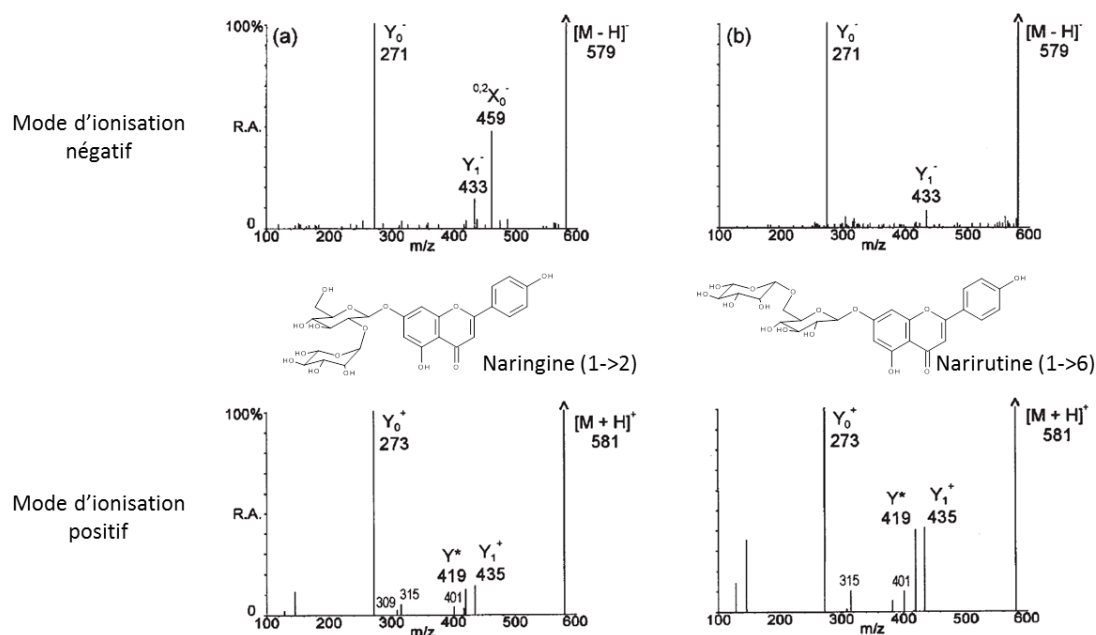


Figure I-18: Spectres de fragmentation des molécules naringine (1->2) et narirutine (1->6) en mode d'ionisation négatif et positif d'après Cuyckens (2000).

5.3.3. Fragmentation des flavonoïdes di-*O*-glycosylés

Les figures I-19a et b illustrent bien que la fragmentation en mode d'ionisation négatif permet de distinguer la mono-substitution *O*-glycosylée-*O*-glycosylée de la di-substitution *O*-glycosylée, notée di-*O*-glycosylée. En effet, comme vu dans le paragraphe précédent, lors de la perte des sucres monosubstitués à la génine, il y a très peu de fragments intermédiaires (Fig. I-19a), alors qu'ils sont bien présents lors d'une di-substitution (Fig. I-19 b-d).

La fragmentation de molécules di-*O*-glycosylées correspond à la même fragmentation que pour des mono-*O*-glycosylées avec les mêmes pertes de neutre. Vukics and Guttman (2010) montrent également que la substitution entre les positions 3 et 7 sur la génine peut être distinguée par la présence ou l'absence de l'ion radical $[Y_0-H]^-$ spécifique qui correspondra à la perte du sucre en position 3 (Fig. I-19c et d).

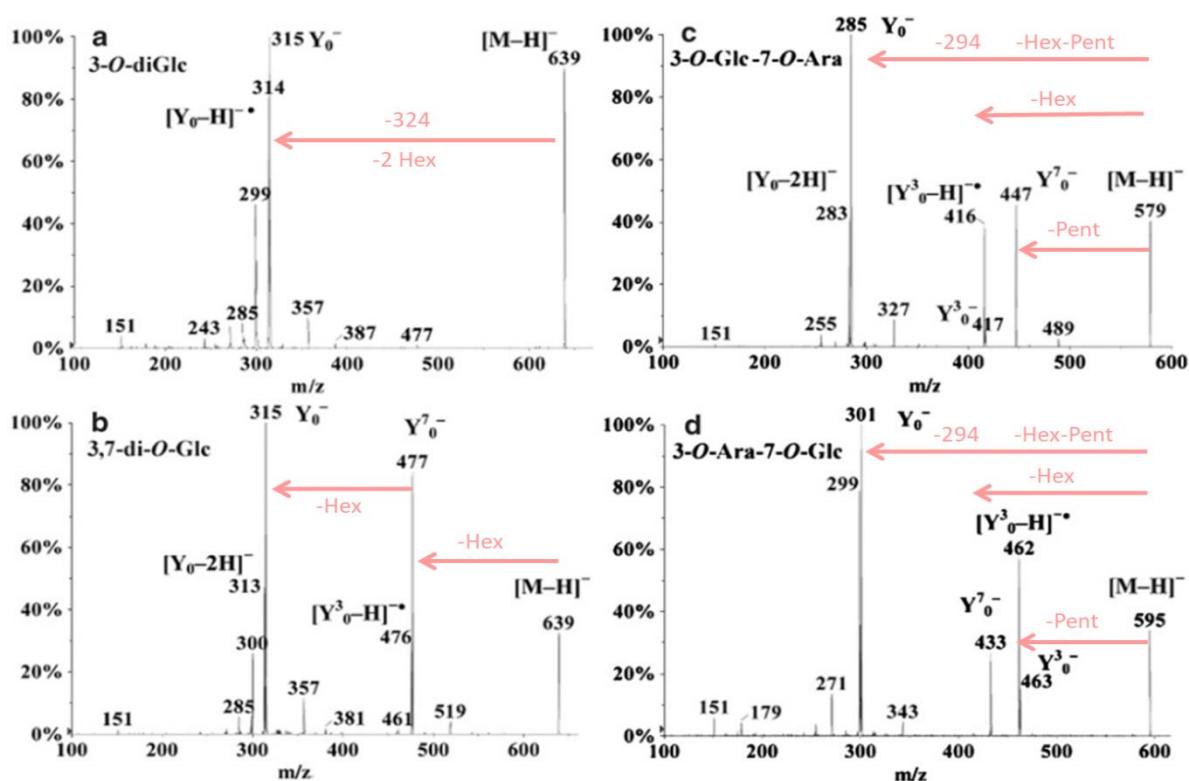


Figure I-19: Spectres de fragmentation de flavonoïdes di-*O*-glycosylés d'après Vukics and Guttman (2010). Les pertes de neutre observées renseignent sur la nature du sucre : 132 u pentose (-Pent), 162 u hexose (-Hex), 294 u hexose et pentose (-Hex -Pent).

Pour résumer, la fragmentation des flavonoïdes *O*-glycosylés en mode d'ionisation positif ou négatif, est basée sur le même schéma de fragmentation que les molécules soient mono ou di-substituées. Le mode positif permettra d'avoir des informations sur la liaison interglycosidique et le mode négatif sur la position (3 ou 7) du sucre sur la génine. La comparaison des rapports des ions n'est valable que s'il y a plusieurs flavonoïdes pour évaluer l'évolution des ratios, or dans un mélange complexe inconnu, les flavonoïdes ne sont pas forcément présents sous toutes leurs formes isomériques.

5.4. Fragmentation des flavonoïdes C-glycosylés

D'après Vukics and Guttman (2010), la C-glycosylation des flavonoïdes est très présente pour une sous-famille moléculaire, les flavones, et se fait principalement en position 6 ou 8 de la génine.

La fragmentation des flavonoïdes C-glycosylés diffère de celle des flavonoïdes O-glycosylés. Davantage d'ions sont observés dans les spectres de masses correspondant aux flavonoïdes C-glycosylés. Au lieu d'avoir une coupure au niveau de Y_n , la coupure se fait dans le sucre (X_n) (Fig. I-16) pour les flavonoïdes C-glycosylés. Contrairement à la substitution O-glycosylée, le fragment correspondant à la masse de l'aglycone n'est donc pas présent dans les spectres de fragmentation que cela soit en mode d'ionisation négatif ou positif. Ci-dessous trois cas différents sont détaillés : le cas de deux types de mono-C-glycosylés et un cas de di-C-glycosylé.

5.4.1. Fragmentation des flavonoïdes mono-C-glycosylés

Comme pour les mono-O-glycosylés, la nature du sucre lié à la génine, peut être déterminée dans le cas des mono-C-glycosylés par la ou les perte(s) de neutre(s) obtenue(s). En effet, en mode d'ionisation négatif, l'observation d'une perte de neutre de 120 u et 90 u correspond à un motif hexose, une perte de 104 u et 74 u correspond à un motif de désoxyhexose et une perte de 90 u et 60 u correspond à celle d'un motif pentose (Cuyckens and Claeys 2004, Ferreres 2007, Figueirinha 2008).

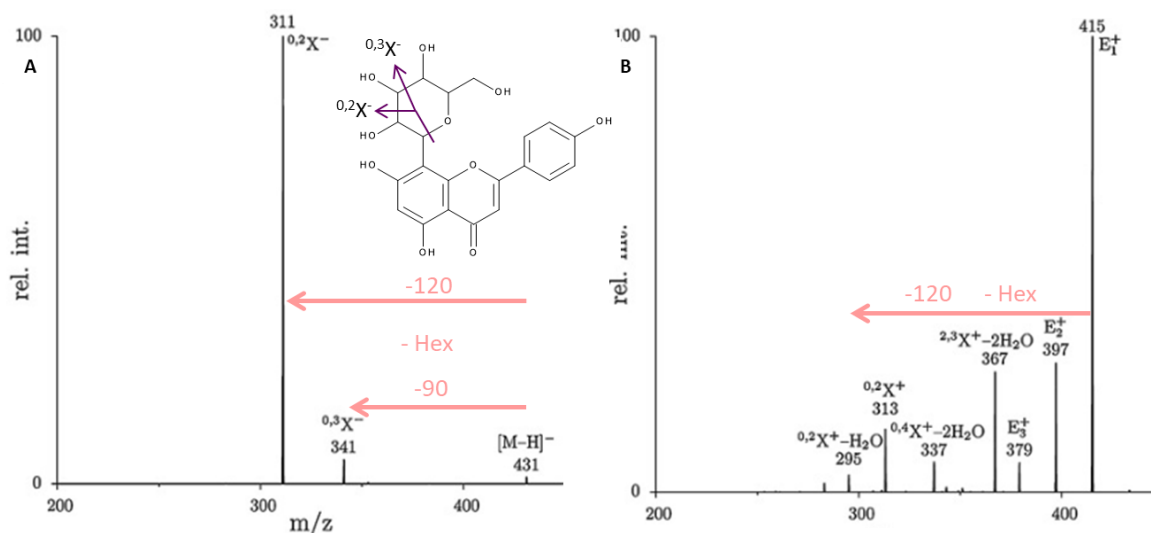


Figure I-20: Spectres de fragmentation en mode d'ionisation négatif (A) et en mode d'ionisation positif (B) de l'apigénine 8-C-glucoside, d'après Vukics and Guttman (2010). Les pertes de neutre observées 120 u et 90 u indiquent une liaison C-hexose (-Hex).

Cette figure I-20 illustre les différences de fragmentation en mode d'ionisation négatif (A) et en mode d'ionisation positif (B). Ainsi en mode d'ionisation négatif, peu de fragments sont observés. Le spectre est composé de fragments de type $^{0,2}X^-$ et $^{0,3}X^-$, qui correspondent à la rupture des liaisons O-2 et O-3

dans la partie glycoside (Fig. I-20A). En mode d'ionisation positif, une fragmentation plus abondante est observée. Ces fragments correspondent à des pertes successives d'eau (E_1^+ à E_3^+), mais on peut observer aussi des coupures du glucose, telles que les ions $^{2,3}X^+$ avec deux pertes d'eau, $^{0,4}X^+$ avec deux pertes d'eau, $^{0,2}X^+$ sans et avec une perte d'eau (Fig. I-20B et 21).

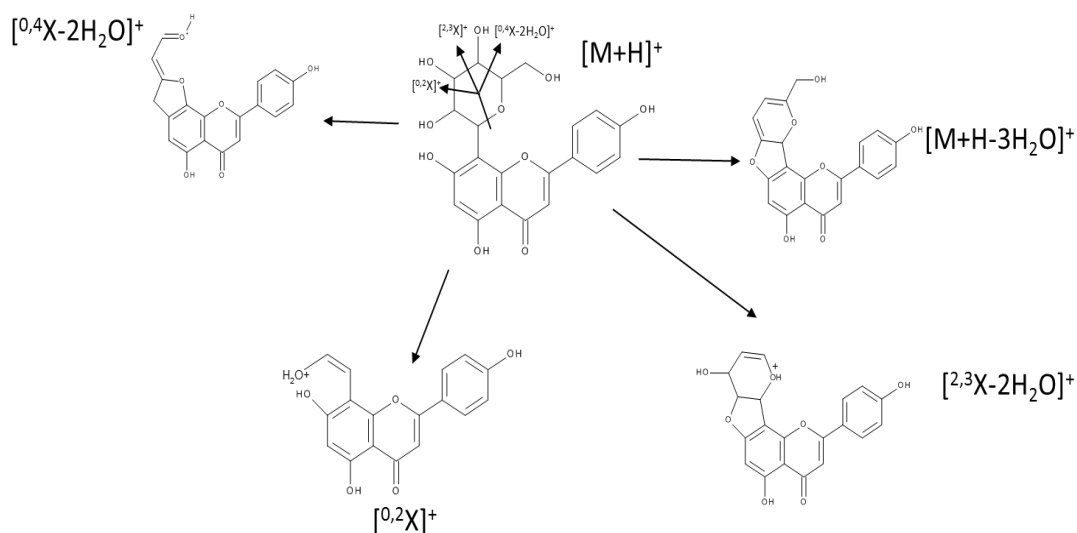


Figure I-21: Schéma de fragmentation de l'orientine en mode d'ionisation positif.

Pour déterminer la nature de la génine en mode d'ionisation négatif, les ions $^{0,2}X^-$ et $^{0,3}X^-$ correspondant à $[Ag+41]^-$ et $[Ag+71]^-$ permettent de déduire la masse de l'aglycone. Dans cet exemple, ces ions sont respectivement de masses m/z 311 u et 341 u, donc l'aglycone a une masse moléculaire de 270 Da, ce qui correspond bien à une apigénine.

Dans le cas d'isomérisation de position, soit en position C-6 soit en position C-8, l'indétermination peut être levée en fonction des fragments obtenus et de leurs intensités (Abad-Garcia 2008, Cuyckens and Claeys 2004, Ferreres 2007, Vukics and Guttman 2010).

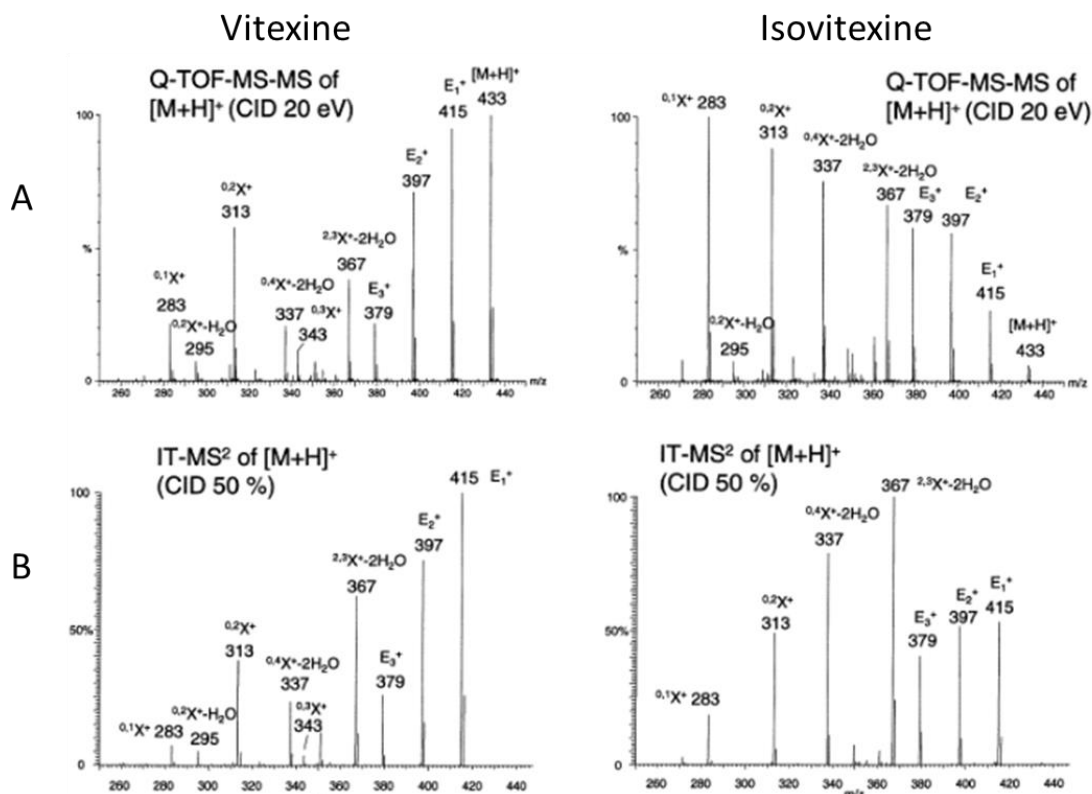


Figure I-22: Spectres de fragmentation de deux flavonoïdes C-glycosylés, isomères de position : apigénine 8-C-glycosylé (Vitexine) et apigénine 6-C-glycosylé (Isovitexine) acquis avec un analyseur de type Q-Tof (a), avec un analyseur de type trappe d'ions (b) d'après Waridel (2001).

En mode d'ionisation positif, la position du sucre C-glycosylé peut être déterminée en observant l'intensité de l'ion correspondant à la première perte d'eau (E_1^+). Si E_1^+ est le pic fragment le plus abondant du spectre MS^2 , alors le sucre est lié au carbone 8 de la génine, sinon il est en position C-6 (Fig. I-22). Waridel et al. (2001) indiquent également qu'en mode d'ionisation négatif, la position du sucre sur la génine peut être déterminée en fonction de l'observation de l'ion E_1^- et de l'ion correspondant à la deuxième perte d'eau (E_2^-). Si ces fragments sont tous les deux présents, alors le sucre est en position C-6, sinon il est en C-8.

D'une manière globale, il sera toujours plus sûr pour déterminer la position d'un sucre sur la génine, de se baser sur l'observation d'ions caractéristiques (comme ici en mode d'ionisation négatif avec les ions E_1^- et E_2^-) plutôt que sur une comparaison de l'abondance relative d'ions. En effet, comme le montre la figure I-22, selon le type d'analyseur avec lequel les expériences de fragmentation sont menées, les abondances relatives de l'ensemble des ions ne seront pas toujours équivalentes. Ainsi, si au préalable, il n'a pas été réalisé sur les appareils de travail, d'études de fragmentation à partir de standards, des interprétations émises sur la seule base des abondances relatives pourront être erronées. Etablir le spectre de fragmentation des flavonoïdes mono-C-glycosylés à partir des 2 modes d'ionisation permet d'obtenir des informations complémentaires.

5.4.2. Fragmentation des flavonoïdes mono-C-glycosylés-O-glycosylés

La fragmentation des flavonoïdes mono-C-glycosylés-O-glycosylés est plus complexe à interpréter du fait de la présence des deux types de liaisons. Il est possible d'observer en mode d'ionisation négatif, la fragmentation au niveau de la liaison mono-C-glycosylée, qui se traduit par une perte de neutre de 120 u, 90 u ou 60 u pour respectivement la perte d'un hexose, un désoxyhexose ou un pentose mais également une fragmentation liée à l'O-glycosylé sur un hydroxyle du premier sucre, et qui peut être suivie d'une perte d'eau. Sur le spectre, ce type de fragmentation O-glycosylée se traduit par une perte de neutre de 162 u, 146 u ou 132 u pour respectivement la perte d'un hexose, un désoxyhexose ou un pentose (coupure Y_1), mais aussi les pertes de masses 180 u, 164 u ou 150 u lorsqu'il y a une perte d'eau supplémentaire (Z_1) (Ferreres 2007 et 2008, Figueirinha 2008).

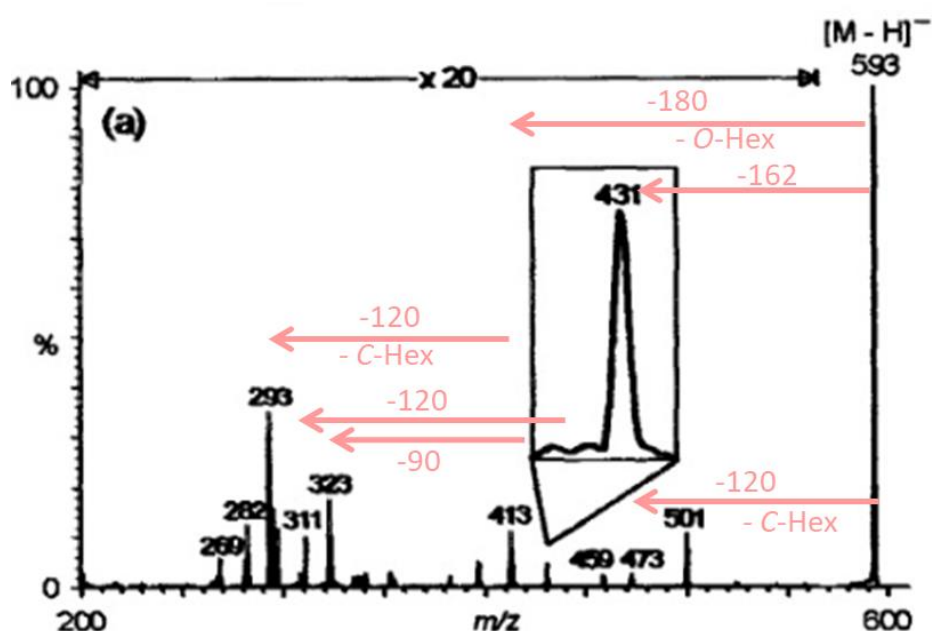


Figure I-23: Spectre de fragmentation de l'isovitexine-2''-O-glucoside (ou apigénine 6-C-glucoside-2''-O-glucoside) en mode d'ionisation négatif, d'après Cuyckens and Claeys (2005). Les pertes de neutre observées 162 u et 180 u correspondent à l'O-hexose (-O-Hex), et les pertes 120 u et 90 u correspondent au C-hexose (-C-Hex).

Ainsi la figure I-23 illustre la fragmentation en mode d'ionisation négatif, de l'isovitexine-2''-O-glucoside, qui est une apigénine 6-C-glucoside-2''-O-glucoside. Les masses m/z 431 (perte de 162 u) et 413 (perte de 180 u) correspondent respectivement aux ions Y_1^- et Z_1^- qui mettent en évidence la présence d'un hexose (dans ce cas, un glucose) en position O terminal. L'ion de masse m/z 473 (perte de 120 u) correspond à l'ion $^{0,2}X^-$ indiquant la présence d'un hexose (dans ce cas, un glucose) lié à un carbone de la génine. Les ions de rapport massique m/z égal à 293, 311, 323 et 331 correspondent respectivement aux ions $[Ag+41-18]^-$, $[Ag+41]^-$, $[Ag+71-18]^-$ et $[Ag+71]^-$ et permettent de calculer la masse de la génine. Dans notre exemple, cela correspond à une masse de 270 Da pour la génine, soit une apigénine, ce qui est bien conforme à la nature de la génine de l'isovitexine.

En mode d'ionisation positive (Fig. I-24), la perte de l'hexose (-162 u) positionné en 2''-O-Glc est d'abord observée et conduit au fragment de rapport m/z égal à 433. Ce fragment correspond à l'ion Y_1^+ . Les fragments provenant du sucre en liaison C-C sur la génine sont obtenus après fragmentation du 2''-O-Glc avec les mêmes types de pertes que celles décrites dans le paragraphe sur la fragmentation de l'isovitexine (6.4.1).

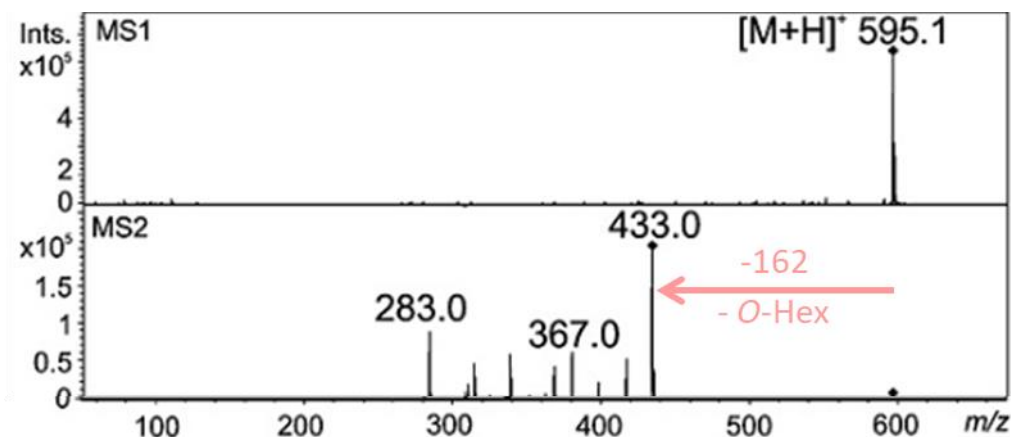
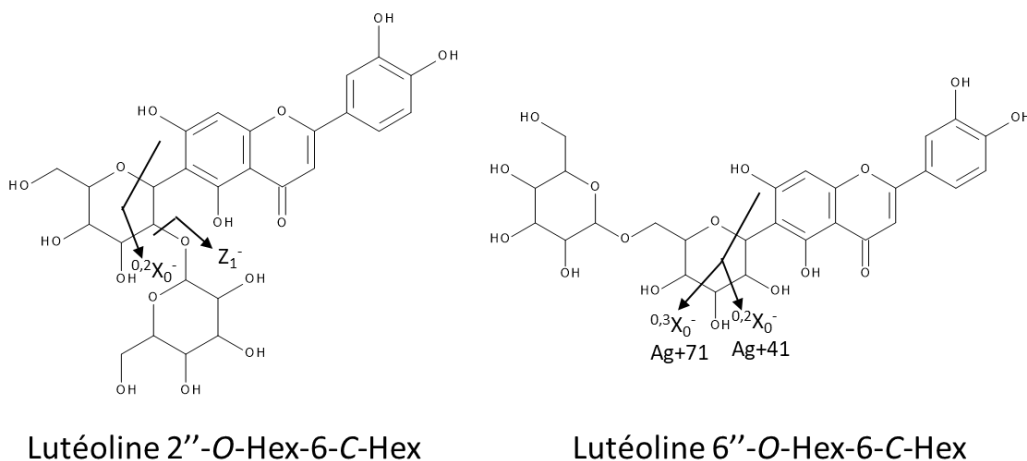


Figure I-24: Spectres de fragmentation de l'isovitexine-2''-O-glucoside en mode d'ionisation positif, d'après Piasecka (2015). Les pertes de neutre observées 162 u indiquent une liaison O-hexose (-O-Hex).

La position en 2''-O- ou en 6''-O- du sucre terminal peut être établie à partir du spectre de fragmentation (Fig. I-25).



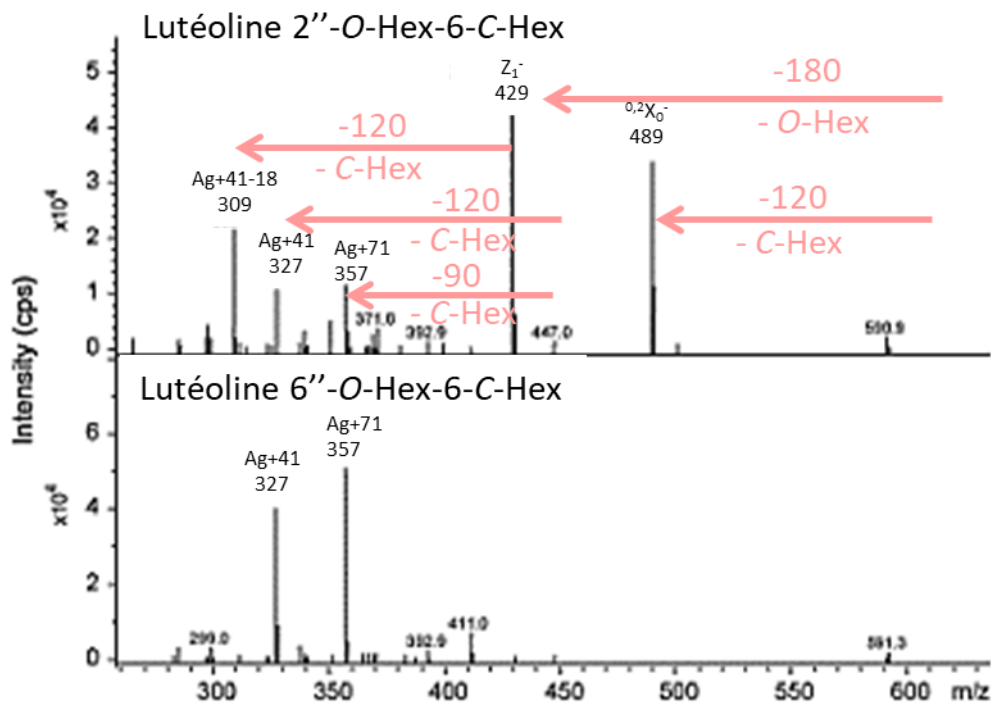


Figure I-25: Comparaison des spectres de masse de fragmentation de l'ion m/z 609 des molécules flavonoïde 2''-O-glycosylé-6-C-glycosylé et 6''-O-glycosylé-6-C-glycosylé, d'après Ferreres (2007). La perte de neutre observée de 180 u correspond à un O-hexose (-O-Hex), et celles de 120 u et 90 u correspondent à un C-hexose (-C-Hex).

Ainsi si le sucre terminal est positionné en 2'' sur le sucre, on observe sur le spectre de masse deux fragments spécifiques : le fragment $[Ag+41-18]^-$ et le fragment Z_1^- qui sont absents du spectre de masse dans le cas d'une glycosylation en position 6''.

5.4.3. Fragmentation des flavonoïdes di-C-glycosylés

La fragmentation des flavonoïdes di-C-glycosylés est identique à celle des mono-C-glycosylés. Il y a la même perte de neutre avec cependant un nombre de fragments deux fois plus importants que dans le cas des mono-glycosylés dû à la présence des deux sucres. Pour connaître le motif de l'aglycone, les deux plus petits fragments en masse observés sur le spectre de fragmentation, correspondent respectivement à $[Ag+83]^-$ et $[Ag+113]^-$ (Ferreres 2007 et 2008, Vukics and Guttman 2010) ce qui indique la perte des deux sucres C-glycosylés.

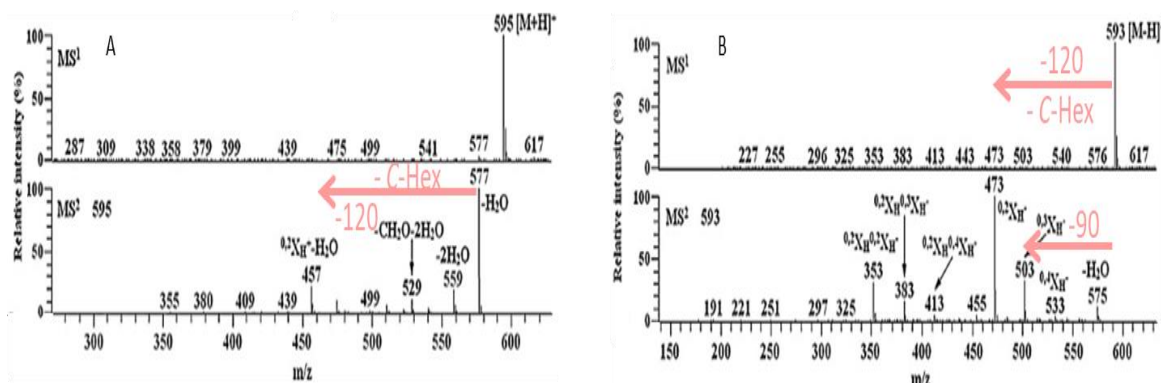


Figure I-26: Spectres de fragmentation de la molécule apigénine 6,8-di-C-glucoside en mode d'ionisation positif (A) et en mode d'ionisation négatif (B), d'après Cao (2014). Les pertes de neutre observées de 120 u et 90 u correspondent à un C-hexose (-C-Hex).

La figure I-26 illustre la fragmentation d'une molécule di-C-glycosylée. En mode d'ionisation positif (Fig I-26A), le spectre possède des ions fragments dus à des pertes de molécules d'eau et d'enchaînements CH_2O . Il possède également l'ion correspondant à une coupure à l'intérieur du motif glycoside ($^{0,2}\text{X}^+ - \text{H}_2\text{O}$). En mode d'ionisation négatif (Fig. I-26B), les ions fragments de type $^{0,2}\text{X}^-$ et $^{0,3}\text{X}^-$ sont bien présents comme dans le cas des mono-C-glycosylés. Les ions de rapport massique m/z 353 u et 383 u permettent de calculer la masse de la génine puisqu'ils correspondent respectivement aux ions $[\text{Ag}+83]^-$ et $[\text{Ag}+113]^-$. Ceci permet de confirmer que la molécule discutée possède bien une base d'apigénine.

Dans le cas présenté figure I-26, la molécule est dite symétrique, car elle possède le même sucre (le glucose) en position C-6 et C-8. D'autres molécules sont dites en revanche asymétriques, car elles possèdent des sucres différents sur les deux carbones. Ainsi Vukics and Guttman (2010) proposent de regarder, quand les deux molécules sont présentes, l'abondance relative des ions correspondant aux pertes respectives des deux sucres (perte de 90 u dans le cas d'un pentose et 120 u dans le cas d'un hexose dans le cas de l'exemple présenté figure I-27) pour définir la position respective de chacun d'eux sur la génine. La perte la plus abondante correspondra au sucre positionné en C-6. L'analyse des spectres de fragmentation des 2 molécules (apigénine-6-C-arabinose-8-C-glucose et apigénine-6-C-glucose-8-C-arabinose), isomères de position, présentés figure I-27 nous permet d'illustrer ces propos. Pour ces 2 molécules, l'ion de rapport m/z 473 correspond à une perte de 90 u par rapport à la molécule initiale, soit à la perte du pentose et l'ion de rapport m/z 443 correspond à la perte d'un hexose (perte de 120 u). La figure I-27A illustre que l'ion de rapport massique m/z 473 est le plus abondant ce qui est conforme avec la structure d'une molécule 6-C-pentose-8-C-hexose tandis que l'ion de rapport massique m/z 443 plus abondant dans la figure I-27B, caractérise bien le spectre de masse d'une molécule de type 6-C-hexose-8-C-pentose.

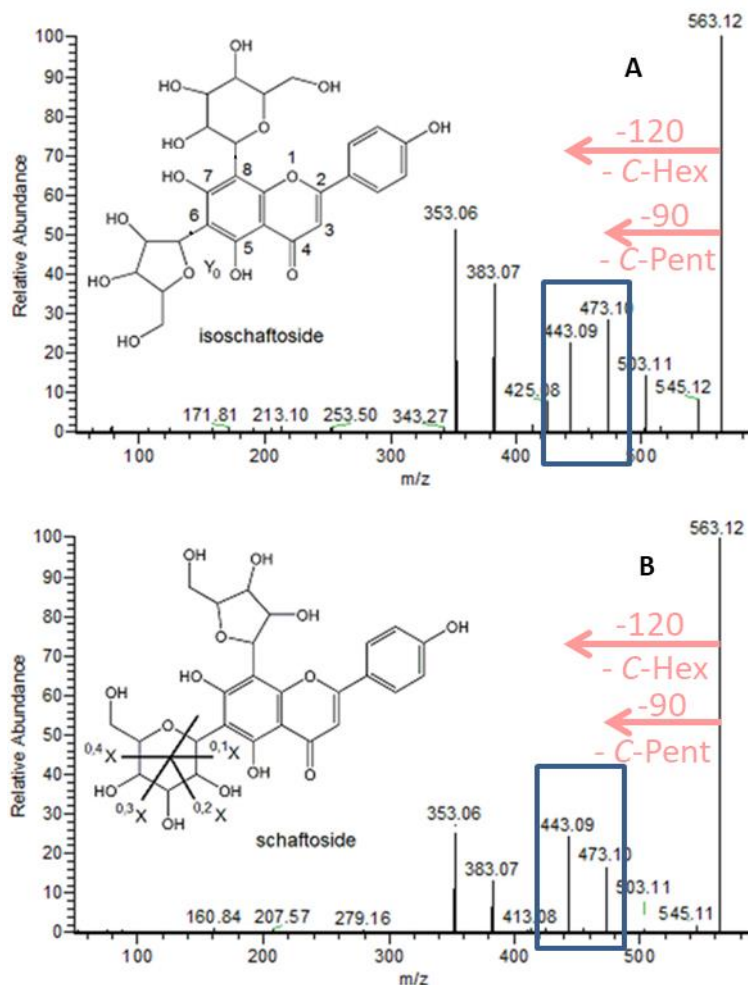
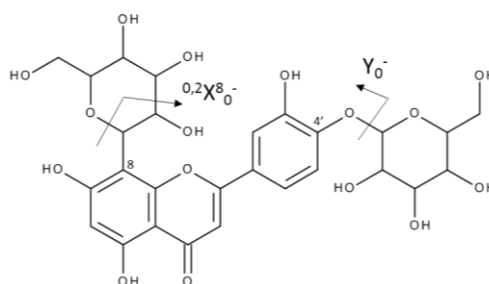


Figure I-27: Spectres de fragmentation des molécules isoschaftoside (Apigénine-6-C-arabinose-8-C-glucose) (A) et schaftoside (Apigénine-6-C-glucose-8-C-arabinose) (B) en mode d'ionisation négatif, d'après Picariello (2017). La perte de neutre de 120 u correspond à un C-hexose (-C-Hex) et celle de 90 u correspond à un C-pentose (-C-Pent).

5.5. Fragmentation des flavonoïdes di-*O,C*-glycosylés

Les flavonoïdes di-glycosylés peuvent également être sous une forme, où les deux sucres sont liés directement à la partie génine du flavonoïde, l'un par une liaison C-C et l'autre sur un des hydroxyles de la génine. On parle alors de flavonoïdes di-*O,C*-glycosylés. Leur fragmentation est un peu moins décrite dans la littérature, mais Ferreres (2007) en montre un exemple (Fig. I-28).



Lutéoline 4'-O-glucoside-8-C-glucoside

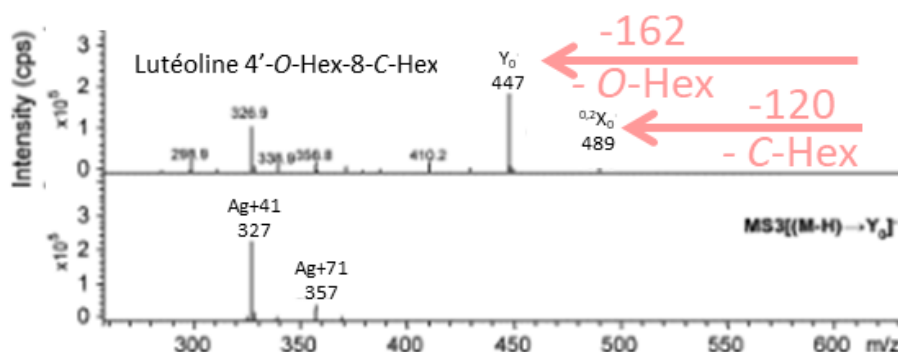


Figure I-28: Spectres de fragmentation de la molécule Lutéoline 4'-O-glucoside-8-C-glucoside en mode d'ionisation négatif d'après Ferreres (2007). La perte de neutre de 162 u correspond à un O-hexose (-O-Hex) et celle de 120 u à un C-hexose (-C-Hex).

Comme pour les flavonoïdes mono-C-glycosylés-O-glycosylés, les ions $[Ag+41]^-$ et $[Ag+71]^-$ sont observés dans les spectres de masse des flavonoïdes di-O,C-glycosylés, ainsi que la perte de 120 correspondant à la coupure de type $^{0,2}X_0^-$ à partir du sucre C-glycosylé. En revanche, ce qui est spécifique au spectre de masse d'un flavonoïde di-O,C glycosylé est la présence d'un ion de type Y_0^- , caractérisé par une perte de 162 u dû à la perte du O-hexose lié à l'aglycone et l'absence de l'ion Z_1^- correspondant à la perte de 180 u pour un O-hexose lié à un hydroxyle C-glycosylé.

En mode d'ionisation positif (Fig. I-29), le spectre de fragmentation de la molécule apigénine 8-C-glucoside-4'-O-rhamnoside indique la présence de l'ion fragment $[Y_0^{4'}]^+$ (m/z 433) correspondant ici à la perte de la fonction rhamnose en position O-glycosylée ainsi que celle de l'ion fragment $[^{0,2}X]^+$ (m/z 313) correspondant à la coupure dans le glucose en position C-glycosylée.

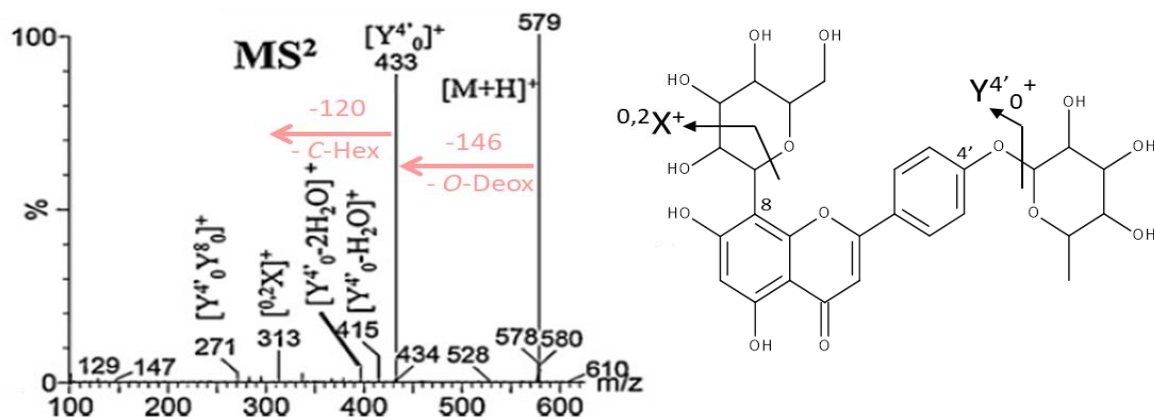


Figure I-29: Spectres de fragmentation de l'apigénine 8-C-glucoside-4'-O-rhamnoside à une énergie de collision de 20 eV en mode d'ionisation positif, d'après Abad-Garcia (2009). La perte de neutre de 146 u correspond à un O-deoxyhexose (-O-Deox) et celle de 120 u correspond à un C-hexose (-C-Hex).

Dans le cas, du spectre de fragmentation de la molécule apigénine 6-C-glucoside-7-O-glucoside (isovitexine 7-O-glucoside), l'ion fragment de masse m/z 529 (Fig. I-30), correspondant à la coupure $^{2,3}X^+ - 2H_2O$ est observé, en plus du fragment de masse m/z 433, correspondant à la coupure de la fonction O-glycosylée.

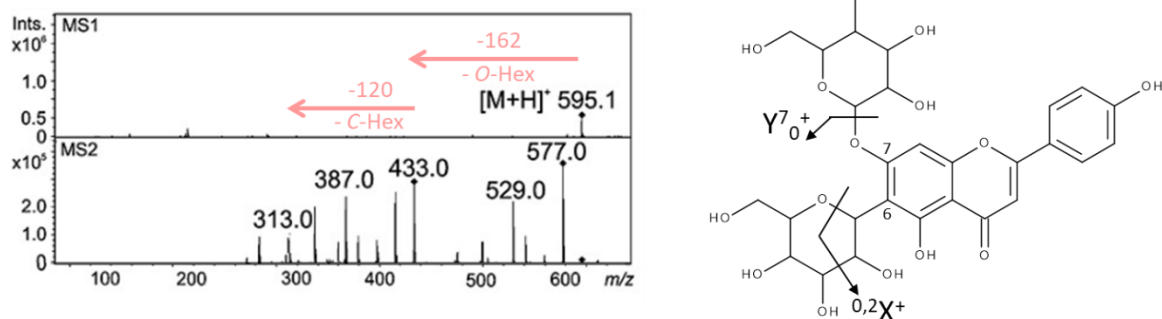


Figure I-30: Spectres de fragmentation de l'isovitexine-7-O-glucoside en mode d'ionisation positif, d'après Piasecka (2015). La perte de neutre de 120 u correspond à un C-hexose (-C-Hex) et celle de 162 u correspond à un O-hexose (-O-Hex).

La différence de fragmentation entre la molécule apigénine 8-C-glucoside-4'-O-rhamnoside (Fig. I-29) et la molécule apigénine 6-C-glucoside-7-O-glucoside (Fig. I-30) est liée au type de coupure dans le sucre en position C-glycosylée. Le fait que la fonction O-glycosylée se situe en position 7 favorise peut-être la coupure en 2,3 au lieu de la coupure habituelle en 0,2.

5.6. Bilan de la fragmentation des flavonoïdes glycosylés en MS

Par l'étude de la fragmentation en SM des flavonoïdes glycosylés, un certain nombre d'informations sur la structure de cette famille moléculaire très présente dans les extraits de plante peut être obtenu. Il est ainsi possible d'avoir accès à une information sur la structure de la génine, sur la nature du sucre lié à la partie aglycone (Tableau I-6) ainsi que sur la position du sucre et du type de liaison interglycosidique (Fig. I-31 et I-32). Dans le cas d'isomères de position, l'analyse de la fragmentation et de l'abondance relative des fragments peut conduire à faire des propositions de structure. Le tableau I-6 résume les pertes de neutre qui peuvent être observées lors de la fragmentation de flavonoïdes glycosylés. Les spectres de fragmentation correspondant aux flavonoïdes di-substitués en C,O-glycosylés et les flavonoïdes mono-substitués en C-glycosylés-6''-O-glycosylés peuvent révéler la présence des mêmes fragments caractéristiques ($[Ag+41]^-$ et $[Ag+71]^-$) (Fig. I-31), donc ces deux types de flavonoïdes ne peuvent pas être différenciés par cette méthode.

Tableau I-6: Perte de neutre lors de fragmentation de flavonoïdes glycosylés.

Fragments	$^{0,1}X$ ($^{0,1}X - H_2O$)	$^{0,2}X$ ($^{0,2}X - H_2O$)	$^{0,3}X$	$^{2,3}X - 2H_2O$	$^{0,4}X - 2H_2O$	Y_i	Z_i
Ac glucuronique	164 (146)	134 (116)	104		110	176	194
Hexose	150 ^a (132)	120 ^a (138)	90 ^a	66 ^a	96 ^a	162 ^a	180 ^a
Deoxyhexose	134 ^a (116)	104 ^a (122)	74 ^a	66 ^a	80 ^a	146 ^a	164 ^a
Oxo Deoxyhexose	132 (114)	102 (120)	72		80	144	162
Pentose	120 ^a (102)	90 ^a (108)	60 ^a	-	66 ^a	132 ^a	150 ^a

^aRef. Vukics and Guttman, 2010

Ce récapitulatif des connaissances (Fig. I-31 et I-32) basé sur les ions fragments spécifiques des flavonoïdes glycosylés et leurs intensités relatives dans les deux modes d'ionisation a été une aide précieuse pour interpréter les spectres de fragmentation obtenus lors des analyses d'extraits de plantes dans les chapitres 2 et 3.

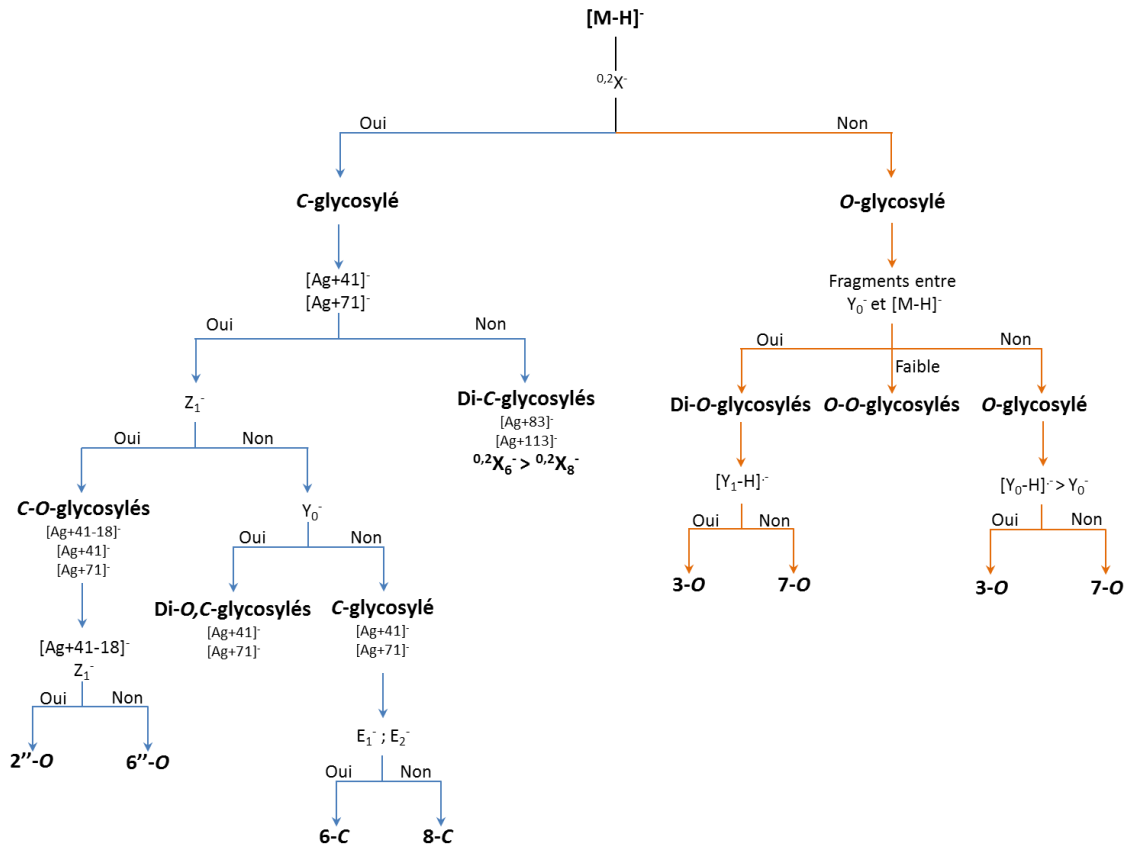


Figure I-31: Schéma d'aide à l'identification des flavonoïdes glycosylés basé sur les ions fragments spécifiques et leurs intensités en mode d'ionisation négatif.

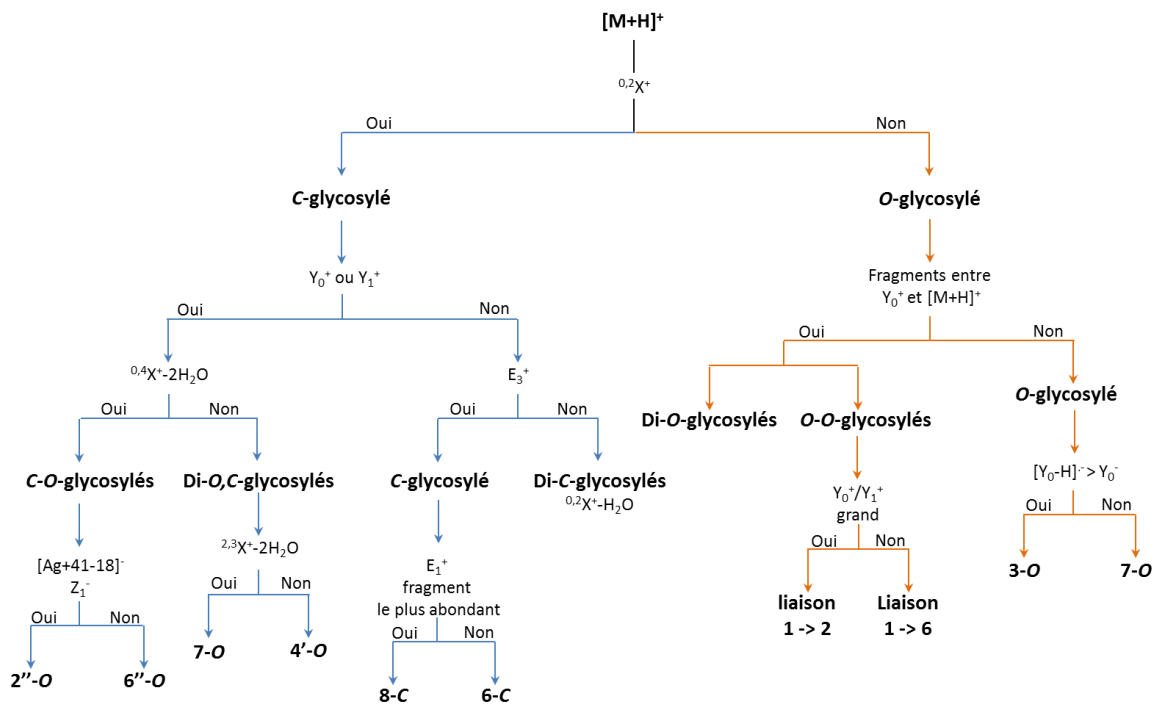


Figure I-32: Schéma d'aide à l'identification des flavonoïdes glycosylés basé sur les ions fragments spécifiques et leurs intensités en mode d'ionisation positif.

Une fois les molécules d'intérêt identifiées dans un extrait donné, il peut être nécessaire de quantifier ces composés afin de connaître leur abondance dans les milieux étudiés. Pour ce faire et avoir une grande fiabilité des résultats obtenus, une démarche de validation de la quantification doit être mise en place. Cette dernière étape essentielle pour aller jusqu'au bout de l'analyse quantitative ne sera pas abordée dans ce manuscrit car nous nous sommes surtout focalisés en première intention sur l'aspect qualitatif des analyses en vue de l'identification des métabolites et de la différenciation des échantillons.

6. Conclusion du chapitre

L'objectif de ce chapitre était de présenter les outils et les stratégies analytiques à disposition pour pouvoir cribler les métabolites non volatils présents en milieu complexe et de comprendre la signature moléculaire d'un système biologique à un moment donné. Les stratégies sélectionnées relèvent d'approches interdisciplinaires combinant des outils de séparation (essentiellement chromatographiques tels que la TLC et l'UHPLC) et de détection (spectroscopie UV et spectrométrie de masse) à des outils d'analyse de données (informatique, chimiométrie) pour aller encore plus loin dans l'interprétation des résultats.

Un large panel de techniques analytiques permettant d'extraire, de séparer et de détecter les molécules présentes dans un échantillon est disponible. Ces techniques permettent d'obtenir des empreintes moléculaires plus ou moins riches selon la manière d'aborder l'étude de façon ciblée ou non. L'analyse des milieux complexes génère de grands ensembles de données spectrales, qui peuvent être traités via l'aide de la bio-informatique pour faciliter l'analyse de ces données et leur rendu visuel. Des méthodes statistiques multivariées peuvent être appliquées pour identifier les métabolites statistiquement significatifs ou différents entre individus. L'identification des métabolites peut se faire à l'aide de différents types de base de données en libre accès. Mais il reste important de connaître les mécanismes de base de fragmentation des familles moléculaires étudiées afin d'avoir un regard critique sur les résultats donnés. L'empreinte moléculaire d'un extrait peut également être visualisée par différents outils complémentaires (Diagramme de Kendricks, Van Krevelen ou réseau moléculaire) qui peuvent être d'une aide utile lors de l'identification des composés.

Les deux chapitres suivants présenteront les stratégies que nous avons mises en place pour obtenir les signatures moléculaires de milieux complexes dans deux contextes d'étude différents.

Le chapitre 2 présentera les développements de méthodes de déréplication réalisés par TLC/MS pour identifier rapidement des molécules connues et évaluer la composition de milieux d'intérêt (échantillons de vin, de jus de pomme) et le chapitre 3 présentera la comparaison que nous avons réalisée par des approches de métabolomique des profils UHPLC/MS d'échantillons de soies de maïs de variétés différentes pour visualiser les différences et cibler l'identification des molécules discriminantes.

Références bibliographiques

- B. Abad-Garcia, S. Garmón-Labato, L. A. Berrueta, B. Gallo, F. Vicente**, New features on the fragmentation and differentiation of C-glycosidic flavone isomers by positive electrospray ionization and triple quadrupole mass spectrometry, *Rapid communication mass spectrometry*, 2008, 22, 1834-1842
- B. Abad-Garcia, L. A. Berrueta, S. Garmón-Lobato, F. Vicente**, A general strategy for the characterization of phenolic compounds in fruit juices by high-performance liquid chromatography with diode array detection coupled to electrospray ionization and triple quadrupole mass spectrometry, *Journal of chromatography A*, 2009, 1216, 5398-5415
- H. Abdi and L. J. Williams**, Principal component analysis, *Wiley interdisciplinary reviews: Computational statistics*, 2010, 2(4), 433-459
- K. Ablajan, Z. Abliz, X.-Y. Shang, J.-M. He, R.-P. Zhang, J.-G. Shi**, Structural characterization of flavonols 3,7-di-O-glycosides and determination of the glycosylation position by using negative ion electrospray ionization tandem mass spectrometry, *Journal of mass spectrometry*, 2006, 41, 352-360
- K. Ablajan and A. Tuoheti**, Fragmentation characteristics and isomeric differentiation of flavonols O-rhamnosides using negative ion electrospray ionization tandem mass spectrometry, *Rapid communication mass spectrometry*, 2013, 27, 451-460
- F. Allen, A. Pon, M. Wilson, R. Greiner, D. Wishart**, CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra, *Nucleic acids research*, 2014, 42, W94-W99
- P. Banerjee, J. Erehrman, B.-O. Gohlke, T. Wilhem, R. Preissner, M. Dunkel**, Super Natural II – a database of natural products, *Nucleic acids research*, 2015, 43, 935-939
- M. Bedair and L. W. Sumner**, Current and emerging mass spectrometry technologies for metabolomics, *Trends in Analytical chemistry*, 2008, 27 (3), 238-250
- J. Blaženović, T. Kind, J. Ji, O. Fiehn**, Software tools and approaches for compounds identification of LC-MS/MS data in metabolomics, *Metabolites*, 2018, 8(2), 31, 1-23
- J. Boccard, J.-L. Veuthey, S. Rudaz**, Knowledge discovery in metabolomics: an overview of MD data handling, *Journal of separation science*, 2010, 33, 290-304
- J. Boccard and DN. Rutledge**, A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion, *Analytica chimica acta*, 2013, 769, 30-39
- A. Bouslimani, L. M. Sanchez, N. Garg, P. C. Dorrestein**, Mass spectrometry of natural products: current, emerging and future technologies, *Natural product reports*, 2014, 31, 718-729
- R. Breitling, A. Cenicerros, A. Jankevics, E. Takano**, Metabolomics for secondary metabolite research, *Metabolite*, 2013, 3, 1076-1083
- AW. Bristow, KS. Webb, AT. Lubben, J. Halket**, Reproducible product-ion tandem mass spectra on various liquid chromatography/mass spectrometry instruments for the development of spectral libraries, *Rapid communications in mass spectrometry*, 2004, 18, 1147-1454
- D. I. Broadhurst and D.B. Kell**, Statistical strategies for avoiding false discoveries in metabolomics and related experiments, *Metabolomics*, 2006, 2(4), 171-196
- T. Cajka and O. Fiehn**, Toward merging untargeted and targeted methods in mass spectrometry based metabolomics and lipidomics, *Analytical chemistry*, 2016, 88(1), 524-545
- GT. Carter**, NP/MS since 1970: from the basement to the bench top, *Natural product reports*, 2014, 31(6), 711-717

- J. Cao, C. Yin, Y. Qin, Z. Cheng, D. Chen**, Approach to the study of flavone di-C-glycosides by high performance liquid chromatography-tandem ion trap mass spectrometry and its application to characterization of flavonoid composition in *Viola yedoensis*, *Journal of mass spectrometry*, 2014, 49, 1010-1024
- JB. Coble and CG. Fraga**, Comparative evaluation of preprocessing freeware on chromatography/mass spectrometry data for signature discovery, *Journal of chromatography A*, 2014, 1358, 155-164
- H. L. Constant, K. Slowing, J. G. Graham, J. M. Pezzuto, G. A. Cordell, C. W. W. Beecher**, A general method for the dereplication of flavonoid glycosides utilizing high performance liquid chromatography/mass spectrometric analysis, *Phytochemical analysis*, 1997, 8, 176-180
- E. Corradini, P. Foglia, P. Giansanti, R. Gubbiotti, R. Samperi, A. Lagana**, Flavonoids: chemical properties and analytical methodologies of identification and quantitation in foods and plants, *Natural product research*, 2011, 25(5), 469-495
- DG. Cox, J. Oh, A. Keasling, KL. Colson, MT Hamann**, The utility of metabolomics in natural product and biomarker characterization, *biochimica et biophysica acta*, 2014, 1840, 3460-3474
- F. Cuyckens, Y. L. Ma, G. Pocsfalvi, M. Claeys**, Tandem mass spectral strategies for the structural characterization of flavonoid glycosides, *Analisis*, 2000, 28, 888-895
- F. Cuyckens and M. Claeys**, Mass spectrometry in the structural analysis of flavonoids, *Journal of mass spectrometry*, 2004, 39, 1-15
- F. Cuyckens and M. Claeys**, Determination of the glycosylation site in flavonoid mono-O-glycosides by collision-induced dissociation of electrospray-generated deprotonated and sodiated molecules, *Journal of mass spectrometry*, 2005, 40, 364-372
- K. Dettmer, PA. Aronov, BD. Hammock**, Mass spectrometry-based metabolomics, *Mass spectrometry reviews*, 2007, 26, 51-78
- KR. Duncan, M. Crüsemann, A. Lechner, A. Sarkar, J. Li, N. Ziemert, M. Wang, N. Bandeira, BS. Moore, PC. Dorrestein, PR. Jensen**, Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora species*, *Chemistry & biology*, 2015, 22, 460-471
- K. Dührkop, H. Shen, M. Meusel, J. Rousu, S. Böcker**, Searching molecular structure databases with tandem mass spectra using CSI:FingerID, *Proceedings of the national academy of sciences of the united states of america*, 2015, 112 (41), 12580-12585
- W. B. Dunn and D. I. Ellis**, Metabolomics: Current analytical platforms and methodologies, *Trends in Analytical Chemistry*, 2005, 24 (4), 285-294
- F. Ferreres, A. Gil-Izquierdo, P. B. Andrade, P. Valentão, F. A. Tomás-Barberán**, Characterization of C-glycosyl flavones O-glycosylated by liquid chromatography tandem mass spectrometry, *Journal of Chromatography A*, 2007, 1161, 214-223
- F. Ferreres, PB. Andrade, P. Valentão, A. Gil-Izquierdo**, Further knowledge on barley leaves O-glycosyl-C-glycosyl flavones by liquid chromatography-UV diode-array detection-electrospray ionization mass spectrometry, *Journal of Chromatography A*, 2008, 1182, 56-64
- A. Figueirinha, A. Paranhos, J. J. Pérez-Alonso, C. Santos-Buelga, M. T. Batista**, *Cymbopogon citratus* leaves: characterization of flavonoids by HPLC-PDA-ESI/MS/MS and an approach to their potential as source of bioactive polyphenols, *Food chemistry*, 2008, 110, 718-728
- D. J. Floros, D. Petras, C. A. Kapon, A. V. Melnik, T. J. Ling, R. Knight, P. C. Dorrestein**, Mass spectrometry based molecular 3D-cartography of plant metabolites, *Frontiers in plant science*, 2017, 8, 429

- YW. Ge, S. Zhu, K. Yoshimatsu, K. Komatsu**, MS/MS similarity networking accelerated target profiling of triterpene saponins in *Eleutherococcus senticosus* leaves, *Food chemistry*, 2017, 227, 444-452
- M. Gerlich and S. Neumann**, MetFusion: integration of compound identification strategies, *Journal of mass spectrometry*, 2013, 48, 291-298
- A. D. Gordon**, A review of hierarchical classification, *Journal of the royal statistical society A*, 1987, 150, 119-137
- R. D. Gougeon, M. Lucio, M. Frommberger, D. Peyron, D. Chassagne, H. Alexandre, F. Feuillat, A. Voilley, P. Cayot, I. Gebefügi, N. Hertkom, P. Schmitt-Kopplin**, The chemodiversity of wines can reveal a metaboledgeography expression of cooperage oak wood, *Proceeding of the national academy of sciences of the united states of America*, 2009, 106 (23), 9174-9179
- H. Gowda, J. Ivanisevic, C. H. Johnson, M.E. Kurczyk, H. P. Benton, D. Rinehart, T. Nguyen, J. Ray, J. Kuehl, B. Arevalo, P. D. Westenskow, J. Wang, A. P. Arkin, A. M. Deutschbauer, G. J. Patti, G. Siuzdak**, Interactive XCMS online: simplifying advanced metabolomic data processing and subsequent statistical analyses, *Analytical chemistry*, 2014, 86 (14), 6931-6939
- A. Grand-Guillaume Perrenoud, J.L. Veuthey, D. Guillaume**, Coupling state-of-the-art supercritical fluid chromatography and mass spectrometry: from hyphenation interface optimization to high-sensitivity analysis of pharmaceutical compounds, *Journal of chromatography A*, 2014, 1339, 174-184
- S. Grimme**, Towards first principles calculation of electron impact mass spectra of molecules, *Angewandte communications*, 2013, S2, 6306-6312
- C. Guigas, JR. Montenegro-Burke, X. Domingo-Almenara, A. Palermo, B. Warth, G. Hermann, G. Koellensperger, T. Huan, W. Uritboonthai, AE. Aisporna, DW. Wolan, ME. Spiker, HP. Benton, G. Siuzdak**, METLIN: a technology platform for identifying knowns and unknowns, *Analytical chemistry*, 2018, 90, 3156-3164
- J. Hastings, G. Owen, A. Dekker, M. Ennis, N. Kale, V. Muthukrishnan, S. Turner, N. Swainston, P. Mendes, C. Steinbeck**, ChEBI in 2016: Improved services and an expanding collection of metabolites, *Nucleic acids research*, 2016, 44, 1214-1219
- D. J. Hook, E. J. Pack, J. J. Yacobucci, J. Guss**, Approaches to automating the dereplication of bioactive natural products- The key step in high throughput screening of bioactive materials from natural sources, *Journal of biomolecular screening*, 1997, 2 (3), 145-152
- H. Horai, M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa, Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, Y. Oda, M. Kakazu, T. Tohge, F. Matsuda, Y. Sawada, MY. Hirai, H. Nakanishi, K. Ikeda, N. Akimoto, T. Maoka, H. Takahashi, T. Ara, N. Sakurai, H. Suzuki, D. Shibata, S. Neumann, T. Lida, K. Tanaka, K. Funatsu, F. Matsuura, T. Soga, R. Taguchi, K. Saito, T. Nishioka**, MassBank: a public repository for sharing mass spectral data for life sciences, *Journal of mass spectrometry*, 2010, 45, 703-714
- F. Hufsky, K. Scheubert, S. Böcker**, Computational mass spectrometry for small-molecule fragmentation, *Trends in analytical chemistry*, 2014, 53, 41-48 (a)
- F. Hufsky, K. Scheubert, S. Böcker**, New kids on the block: novel informatics methods for natural product discovery, *Natural product reports*, 2014, 31 (6), 807-817 (b)
- E. Hvattum and D. Ekeberg**, Study of the collision-induced radical cleavage of flavonoid glycosides using negative electrospray ionization tandem quadrupole mass spectrometry, *Journal of mass spectrometry*, 2003, 38, 43-49
- P. Kachlicki, J. Einhorn, D. Muth, L. Kerhoas, M. Stobiecki**, Evaluation of glycosylation and malonylation patterns in flavonoid glycosides during LC/MS/MS metabolite profiling, *Journal of mass spectrometry*, 2008, 43, 572-586

- P. T. Kasper, M. Rojas-Chertó, R. Mistrik, T. Reijmers, T. Hankemeier, R. J. Vreeken**, Fragmentation trees for the structural characterization of metabolites, *Rapid communications in mass spectrometry*, 2012, 26, 2275-2286
- E. Kendrick**, A mass scale based on CH₂ = 14.0000 for high resolution mass spectrometry of organic compounds, *Analytical chemistry*, 1963, 35 (13), 2146-2154
- N. Kessler, H. Neuweber, A. Bonte, G. Langenkämper, K. Niehaus, T. W. Nattkemper, A. Goesmann**, MeltDB 2.0-advances of the metabolomics software system, *Bioinformatics*, 2013, 29 (19), 2452-2459
- S. Kim, R. W. Kramer, P. G. Hatcher**, Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van krevelen diagram, *Analytical Chemistry*, 2003, 75, 5336-5344
- S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. FU, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, S. H. Bryant**, PubChem substance and compound databases, *Nucleic acids research*, 2016, 44, 1202-1213
- T. Kind, K.-H. Liu, D. Y. Lee, B. DeFelice, J. K. Meissen, O. Fiehn**, LipidBlast in silico tandem mass spectrometry database for lipid identification, *Nature methods*, 2013, 10(8), 1-17
- A. Klitgaard, J. B. Nielsen, R. J. N. Frandsen, M. R. Andersen, K. F. Nielsen**, Combining stable isotope labeling and molecular networking for biosynthetic pathway characterization, *Analytical chemistry*, 2015, 87, 6520-6526
- R. E. Koes, F. Quattrocchio, J. N. M. Mol**, The flavonoid biosynthetic pathway in plants: Function and evolution, *BioEssays*, 1994, 16, 123-132
- A. Koulman, G. Woffendin, V. K. Narayana, H. Welchman, C. Crone, D. A. Volmer**, High-resolution extracted ion chromatography, a new tool for metabolomics and lipidomics using a second-generation orbitrap mass spectrometer, *Rapid communications in mass spectrometry*, 2009, 23, 1411-1418
- D. Krug and R. Müller**, Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products, *Natural product reports*, 2014, 31, 768-783
- N. Kuhnert, J. W. Drynan, J. Obuchowicz, M. N. Clifford, M. Witt**, Mass spectrometric characterization of black tea thearubigins leading to an oxidative cascade hypothesis for thearubigin formation, *Rapid communications in mass spectrometry*, 2010, 24, 3387-3404
- N. Kuhnert, F. Dairpoosh, G. Yassin, A. Golon, R. Jaiswal**, What is under the hump? Mass spectrometry based analysis of complex mixtures in processed food- lessons from the characterization of black tea thearubigins, coffee melanoidines and caramel, *Food & function*, 2013, 4, 1130-1147
- Z. Lai, H. Tsugawa, G. Wohlgemuth, S. Mehta, M. Mueller, Y. Zheng, A. Ogiwara, J. Meissen, M. Showalter, K. Takeuchi, T. Kind, P. Beal, M. Arita, O. Fiehn**, Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics, *Nature methods*, 2018, 15, 53-56
- G. Lang, N. A. Mayhudin, M. I. Mitova, L. Sun, S. Van der Sar, J. W. Blunt, A. L. J. Cole, G. Ellis, H. Laatsch, M. H. G. Munro**, Evolving trends in the dereplication of natural products extracts: new methodology for rapid, small-scale investigation of natural product extracts, *Journal of natural products*, 2008, 71 (9), 1595-1599
- I. Laponogov, N. Sadawi, D. Galea, R. Mirnezami, K. A. Veselkov**, ChemDistiller: an engine for metabolite annotation in mass spectrometry, *Bioinformatics*, 2018, 34(12), 2096-2102
- K.-A. Lê Cao, S. Boitard, P. Besse**, Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems, *BMC Bioinformatics*, 2011, 12:253, 1-16
- Z. Lei, D. V. Huhman, L. W. Sumner**, Mass spectrometry strategies in metabolomics, *The journal of biological chemistry*, 2011, 286 (29), 25435-25442

- K. H. Liland**, Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis, *Trends in analytical chemistry*, 2011, 30 (6), 827-841
- A. Lommen**, MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing, *Analytical chemistry*, 2009, 81, 3079-3086
- H. J. Major, R. Williams, A. J. Wilson, I. D. Wilson**, A metabolomic analysis of plasma from Zucker rat strains using gas chromatography/mass spectrometry and pattern recognition, *Rapid communications in mass spectrometry*, 2006, 20, 3295-3302
- A. Mantovani, F. Maranghi, C. La Rocca, GM. Tiboni, M. Clementi**, The role of toxicology to characterize biomarkers for agrochemicals with potential endocrine activities, *Reproductive toxicology*, 2008, 26, 1-7
- A. G. Marshall and R. P. Rodgers**, Petroleomics: the next grand challenge for chemical analysis, *Accounts of chemical research*, 2004, 37 (1), 53-59
- L.-F. Nothias-Scaglia, M. Esposito, J. Costa, J. Paolini, D. Touboul, M. Litaudon**, Les réseaux moléculaires, une approche bio-informatique globale pour interpréter les données de spectrométrie de masse tandem, *Spectra analyse*, 2015, 307, 73-78
- L.-F. Nothias, S. Boutet-Mercey, X. Cachet, E. De la Torre, L. Laboureur, J.-F. Gallard, P. Retailleau, A. Brunelle, P. C. Dorrestein, J. Costa, L. M. Bedoya, F. Roussi, P. Leyssen, J. Alami, J. Paolini, M. Litaudon, D. Touboul**, Environmentally friendly procedure based on supercritical fluid chromatography and tandem mass spectrometry molecular networking for the discovery of potent antiviral compounds from *Euphorbia semiperfoliata*, *Journal of natural products*, 2017, 80, 2620-2629
- D. Ohta, S. Kanaya, H. Suzuki**, Application of Fourier-transform ion cyclotron resonance mass spectrometry to metabolic profiling and metabolite identification, *Current opinion in biotechnology*, 2010, 21, 35-44
- F. Olivon, P.-M. Allard, A. Koval, D. Righi, G. Genta-Jouve, J. Neyts, C. Apel, C. Pannecouque, L.-F. Nothias, X. Cachet, L. Marcourt, F. Roussi, V. L. Katanaev, D. Touboul, J.-L. Wolfender, M. Litaudon**, Bioactive natural products prioritization using massive multiinformational molecular networks, *ACS chemical biology*, 2017, 12, 2644-2651 (a)
- F. Olivon, F. Roussi, M. Litaudon, D. Touboul**, Optimized experimental workflow for tandem mass spectrometry molecular networking in metabolomics, *Analytical and bioanalytical chemistry*, 2017, 409, 5767-5778 (b)
- F. Olivon, G. Grelier, F. Roussi, M. Litaudon, D. Touboul**, MZmine 2 data-preprocessing to enhance molecular networking reliability, *Analytical chemistry*, 2017, 89, 7836-7840 (c)
- S. K. Panda, W. Schrader, J. T. Andersson**, β -Cyclodextrin as a stationary phase for the group separation of polycyclic aromatic compounds in normal-phase liquid chromatography, *Journal of chromatography A*, 2006, 1122, 88-96
- D. Parrot, L. Intertaglia, P. Jehan, M. Grube, M. T. Suzuki, S. Tomasi**, Chemical analysis of the Alphaproteobacterium strain MOLA 1416 associated with the marine lichen *Lichina pygmaea*, *Phytochemistry*, 2018, 145, 57-67
- A. Piasecka, A. Sawikowska, P. Krajewski, P. Kachlicki**, Combined mass spectrometry and chromatographic methods for in-depth analysis of phenolic secondary metabolites in barley leaves, *Journal of mass spectrometry*, 2015, 50, 513-532
- G. Picariello, L. Sciammaro, F. Siano, M. G. Volpe, M. C. Puppo, G. Mamone**, Comparative analysis of C-glycosidic flavonoids from *Prosopis* spp. and *Ceratonia siliqua* seed germ flour, *Food research international*, 2017, 99, 730-738

- R. Plumb, J. Castro-Perez, J. Granger, I. Beattie, K. Joncour, A. Wright**, Ultra-performance liquid chromatography coupled to quadrupole-orthogonal time-of-flight mass spectrometry, *Rapid communications in mass spectrometry*, 2004, 18, 2331-2337
- T. Pluskal, S. Castillo, A. Villar-Briones, M. Orešič**, MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics*, 2010, 11(1):395
- K. Purves, L. Macintyre, D. Brennan, G. Ó. Hreggviðsson, E. Kuttner, M. E. Ásgeirsdóttir, L. C. Young, D. H. Green, R. Edrada-Edel, K. R. Duncan et al.**, Using molecular networking for microbial secondary metabolite bioprospecting, *Metabolites*, 2016, 6(1), 2
- R. A. Quinn, L.-F. Nothias, O. Vining, M. Meehan, E. Esquenazi, P. C. Dorrestein**, Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy, *Trends in pharmacological sciences*, 2017, 38 (2), 143-154
- F. Rasche, A. Svatoš, R. K. Maddula, C. Böttcher, S. Böcker**, Computing fragmentation trees from tandem mass spectrometry data, *Analytical chemistry*, 2011, 83, 1243-1251
- C. Ratnasekhar, AK. Singh, P. Pandey, PN. Saxena, MKR. Mudiam**, Identifying the metabolic perturbations in earthworm induced by cypermethrin using gas chromatography-mass spectrometry based metabolomics, *Scientific reports*, 2015, 5 (15674), 1-13
- L. Ridder, J. J. J. Van der Hooft, S. Verhoeven**, Automatic compound annotation from mass spectrometry data using MAGMa, *The mass spectrometry society of Japan*, 2014, 3(2), 1-7
- J. A. Rothwell, J. Perez-Jimenez, V. Neveu, A. Medina-Remón, N. M'Hiri, P. García-Lobato, C. Manach, C. Knox, R. Eisner, D. S. Wishart, A. Scalbert**, Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content, *Database*, 2013, bat070
- C. Roullier-Gall, M. Witting, R. D. Gougeon, P. Schmitt-Kopplin**, High precision mass measurements for wine metabolomics, *Frontiers in chemistry*, 2014, 2 (102), 1-9
- E. L. Schymanski, C. Ruttkies, M. Krauss, C. Brouard, T. Kind, K. Dührkop, F. Allen, A. Vaniya, D. Verdegem, S. Böcker, J. Rousu, H. Shen, H. Tsugawa, T. Sajed, O. Fiehn, B. Ghesquière, S. Neumann**, Critical assessment of small molecule identification 2016: automated methods, *Journal of cheminformatics*, 2017, 9:22
- C. Seger, S. Sturm, H. Stuppner**, Mass spectrometry and NMR spectroscopy: modern high-end detectors for high resolution separation techniques – state of the art in natural product HPLC-MS, HPLC-NMR, and CE-MS hyphenations, *Natural product reports*, 2013, 30, 970-987
- H. Sheridan, L. Krenn, R. Jiang, I. Sutherland, S. Ignatova, A. Marmann, X. Liang, J. Sendker**, The potential of metabolic fingerprinting as a tool for the modernization of TCM preparations, *Journal of ethnopharmacology*, 2012, 140, 482-491
- M. A. Strege**, High-performance liquid chromatographic-electrospray ionization mass spectrometric analyses for the integration of natural products with modern high-throughput screening, *Journal of chromatography B*, 1999, 725, 67-78
- P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker**, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome research*, 2003, 13, 2498-2504
- A. C. Schrimpe-Rutledge, S. D. Sherrod, J. A. McLean**, Improving the discovery of secondary metabolite natural products using ion mobility-mass spectrometry, *Current opinion in chemical biology*, 2018, 42, 160-166

- A. D. Southam, A. Lange, A. Hines, E. M. Hill, Y. Katsu, T. Iquchi, C. R. Tyler, M. R. Viant**, Metabolomics reveals target and off-target toxicities of a model organophosphate pesticide to roach (*Rutilus rutilus*): Implications for biomonitoring, *Environmental science & technology*, 2011, 45, 3759-3767
- L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W.-M. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden, M. R. Viant**, Proposed minimum reporting standards for chemical analysis, *Metabolomics*, 2007, 3, 211-221
- J. Tchoumtchoua, D. Njamen, J. C. Mbanya, A. L. Skaltsounis, M. Halabalaki**, Structure-oriented UHPLC-LTQ orbitrap-based approach as a dereplication strategy for the identification of isoflavonoids from *Amphimas pterocarpoides* crude extract, *Journal of mass spectrometry*, 2013, 48, 561-575
- R. Teta, G. D. Sala, E. Glukhov, L. Gerwick, W. H. Gerwick, A. Mangoni, V. Costantino**, Combined LC-MS/MS and molecular networking approach reveals new cyanotoxins from the 2014 cyanobacterial bloom in Green Lake, Seattle, *Environmental science & technology*, 2015, 49, 14301-14310
- J. Trygg and S. Wold**, Orthogonal projections to latent structures (O-PLS), *Journal of chemometrics*, 2002, 16, 119-128
- J. J. J. Van der Hooft, R. C. H. de Vos, L. Ridder, J. Vervoort, R. J. Bino**, Structural elucidation of low abundant metabolites in complex sample matrices, *Metabolomics*, 2013, 9, 1009-1018
- J. J. J. Van der Hooft, S. Padmanabhan, K. E. V. Burgess, M. P. Barrett**, Urinary antihypertensive drug metabolite screening using molecular networking coupled to high-resolution mass spectrometry fragmentation, *Metabolomics*, 2016, 12(125), 1-15
- V. Vukics and A. Guttman**, Structural characterization of flavonoid glycosides by multi stage mass spectrometry, *Mass spectrometry reviews*, 2010, 29(1), 1-16
- Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. DuGar, A. E. Feldstein, E. B. Britt, X. Fu, Y.-M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. W. Tang, J. A. DiDonato, A. J. Lusis, S. L. Hazen**, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease, *Nature*, 2011, 472, 57-63
- M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. KaApono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrew, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya P, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielson, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. liu, L. Zhang, R. Knight, P. R. Jensen, B. Ø Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein, N. Bandeira**, Sharing and community curation of mass spectrometry data with global natural products social molecular networking, *Nature biotechnology*, 2016, 34 (8), 828-837

- Y. Wang, X. Wang, X. Zeng**, MIDAS-G: a computational platform for investigating fragmentation rules of tandem mass spectrometry in metabolomics, *Metabolomics*, 2017, 13: 116, 1-9
- P. Waridel, J.-L. Wolfender, K. Ndjoko, K. R. Hobby, H. J. Major, K. Hostettmann**, Evaluation of quadrupole time-of-flight tandem mass spectrometry and ion-trap multiple-stage mass spectrometry for the differentiation of C-glycosidic flavonoid isomers, *Journal of chromatography A*, 2001, 926, 29-41
- J. Watrous, P. Roach, T. Alexandrov, B. S. Heath, J. Y. Yang, R. D. Kersten, M. Van der Voort, K. Pogliano, H. Gross, J. M. Raaijmakers, B. S. Moore, J. Laskin, N. Bandeira, P. C. Dorrestein**, Mass spectral molecular networking of living microbial colonies, *Proceedings of the national academy of sciences of the united states America*, 2012, 109 (26), E1743-1752
- E. Werner, J.-F. Heilier, C. Ducruix, E. Erzan, C. Junot, J.-C. Tabet**, Mass spectrometry for the identification of the discriminating signals from metabolomics: Current status and future trends, *Journal of chromatography B*, 2008, 871, 143-163
- J. A. Westerhuis, E. J. J. Van Velzen, H. C. J. Hoefsloot, A. K. Smilde**, Multivariate paired data analysis: multilevel PLS-DA versus OPLS-DA, *Metabolomics*, 2010, 6, 119-128
- B. Winkel-Shirley**, Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology and biotechnology, *Plant physiology*, 2001, 126(2), 485-493
- J. R. Winnikoff, E. Glukhov, J. Watrous, P. C. Dorrestein, W. H. Gerwick**, Quantitative molecular networking to profile marine cyanobacterial metabolomes, *the journal of antibiotic*, 2014, 67, 105-112
- D. S. Wishart, Y. D. Feunang, A. Marcu, A. C. Guo, K. Liang, R. Vázquez-Fresno, T. Sajed, D. Johnson, C. Li, N. K. Z. Sayeeda, E. Lo, N. Assempour, M. Berjanskii, S. Singhal, D. Arndt, Y. Liang, H. Bardran, J. Grant, A. Serra-Cayuela, Y. Liu, R. Mandal, V. Neveu, A. Pon, C. Knox, M. Wilson, C. Manach, A. Scalbert**, HMDB 4.0: the human metabolome database for 2018, *Nucleic acids research*, 2018, 46, 608-617
- S. Wold, K. H. Esbensen, P. Geladi**, Principal component analysis, *Chemometrics and intelligent laboratory systems*, 1987, 2, 37-52
- S. Wolf, S. Schmidt, M. Müller-Hannemann, S. Neumann**, In silico fragmentation for computer assisted identification of metabolite mass spectra, *BMC bioinformatics*, 2010, 11, 148
- J.-L. Wolfender, EF. Queiroz, K. Hostettmann**, The importance of hyphenated techniques in the discovery of new lead compounds from nature, *Expert Opinion Drug Discovery*, 2006, 1(3), 237-260
- J.-L. Wolfender**, HPLC in natural product analysis: The detection issue, *Planta Med*, 2009, 75, 719-734
- J.-L. Wolfender, G. Marti, E. F. Queiroz**, Advances in techniques for profiling crude extracts and for the rapid identification of natural products: Dereplication, Quality Control and Metabolomics, *Current Organic Chemistry*, 2010, 14, 1808-1832
- J.-L. Wolfender, G. Marti, A. Thomas, S. Bertrand**, Current approaches and challenges for the metabolite profiling of complex natural extracts, *Journal of chromatography A*, 2015, 1382, 136-164
- B. Worley and R. Powers**, Multivariate analysis in metabolomics, *Current metabolomics*, 2013, 1(1), 92-107
- Z. Wu, RP. Rodgers, AG. Marshall**, 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
- J. Xia and DS. Wishart**, Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis, *Current protocols in bioinformatics*, 2016, 55, 14.10.1-14.10.91

JY. Yang, LM. Sanchez, CM. Rath, X. Liu, PD. Boudreau, N. Bruns, E. Glukhov, A. Wodtke, R. de Felicio, A. Fenner, WR. Wong, RG. Linington, L. Zhang, HM. Debonisi, WH. Gerwick, PC. Dorrestein, Molecular networking as a dereplication strategy, *Journal of natural products*, 2013, 76, 1686-1699

A. Zhang, H. Sun, Y. Han, X. Wang, Modern analytical techniques in metabolomics analysis, *Analyst*, 2012, 137, 293-300

ZJ. Zhu, AW. Schultz, J. Wang, CH. Johnson, SM. Yannone, GJ. Patti, G. Siuzdak, Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database, *Nature protocols*, 2013, 8 (3), 451-460

Identification moléculaire par TLC : Apports du traitement statistique de données et du couplage à la spectrométrie de masse

Le chapitre précédent a permis d'indiquer les différents outils permettant d'obtenir des empreintes moléculaires de mélanges complexes et de les interpréter par spectrométrie de masse.

Les approches analytiques présentées dans ce chapitre seront centrées plus particulièrement sur des développements méthodologiques que nous avons menés en chromatographie en couche mince (CCM ou TLC) et en couplage de cette technique séparative avec la spectrométrie de masse. Deux approches de couplage ont été évaluées pour analyser les plaques TLC : l'une via l'utilisation d'une source d'ionisation assistée par désorption laser et d'un analyseur à temps de vol (MALDI-Tof MS) et l'autre, via l'utilisation d'une source à ionisation à pression atmosphérique (API) et d'un analyseur de type quadripôle. La TLC est régulièrement utilisée pour effectuer un premier screening moléculaire des différentes familles contenues dans des extraits de plante. Les différents profils d'organes, de variétés sont souvent comparés visuellement (Michel 2012 ; Riffault 2014) ou avec les profils densitométriques. L'objectif de ce chapitre est de proposer de nouvelles approches à l'utilisation classique au laboratoire, en apportant des informations supplémentaires et complémentaires. Ainsi aux analyses TLC, il peut être associé des traitements statistiques qui permettent de réaliser des comparaisons rationalisées. De plus, le couplage avec la spectrométrie de masse permet d'obtenir un début de caractérisation des différentes taches de la séparation TLC et donc de la composition moléculaire de l'extrait.

Le premier travail que nous présenterons concerne une comparaison de différents organes de la plante *Brassica napus* L.. Cette comparaison a été effectuée après séparation des différents extraits sur plaques TLC et révélation des familles moléculaires présentes au moyen de révélateurs spécifiques puis traitement des images obtenues à l'aide d'une première approche statistique multivariée.

Puis dans un second temps, après avoir expliqué le principe des couplages TLC/MALDI-Tof MS et TLC/API MS, une étude sur le contenu en flavonoïdes de différents produits issus de la transformation agro-alimentaire, tels que le vin ou le jus de pomme sera détaillée. En nous basant sur la connaissance déjà bien décrite dans la littérature des flavonoïdes contenus dans ces types d'échantillons, cette étude nous a permis d'évaluer toutes les potentialités du couplage TLC/MALDI-Tof MS et ses limites dans le cadre de l'identification de constituants polaires appartenant à une même famille moléculaire.

La troisième étude qui sera présentée concerne l'évaluation des potentialités du couplage TLC/API MS. Une collaboration avec Michelle Pinault du laboratoire Nutrition, Croissance et Cancer (N2C) nous a permis de développer des méthodes d'analyse de stéroïdes et d'oxystérols sans dérivation par TLC/MS. La stratégie analytique entreprise pour les détecter par un couplage direct entre la TLC et la spectrométrie de masse sera discutée en incluant une comparaison des performances de différentes sources d'ionisation disponibles au laboratoire (sources MALDI, ESI et APCI en mode d'ionisation positif et négatif) pour ioniser ces composés. Enfin, une analyse de composés de bas poids moléculaires, les acides organiques par couplage TLC/ESI MS, sera également présentée.

1. Description de la technique séparative TLC

1.1. Considération générale sur la technique

Bien que la chromatographie sur couche mince (CCM ou TLC) offre une efficacité de séparation inférieure à celle de l'UHPLC, elle reste une technique de première intention dans de nombreux domaines d'analyse (Kumar et Sharma 2018) dont celui de la phytochimie. Parmi les différents

avantages de cette technique d'analyse citons le fait que de nombreux échantillons et standards peuvent être appliqués sur une même plaque TLC et donc être simultanément analysés tout en conservant l'intégralité des échantillons sur la plaque après séparation ; de plus, les étapes de purification des échantillons avant dépôt peuvent être minimisées. La différenciation visuelle des constituants d'un extrait peut être obtenue soit directement grâce à l'absorption lumineuse ou à la fluorescence naturelle de certaines familles moléculaires soit après application d'un réactif de dérivation directement sur la plaque TLC. Les taches colorées résultantes peuvent alors être identifiées par comparaison avec des standards de référence. En l'absence de composés standard l'identification est plus difficile, seule une information de polarité et d'appartenance des composés à une famille moléculaire peut être avancée.

Pour surmonter cette limite, plusieurs approches de couplages TLC/MS ont été développées (Morlock et Schwack 2010, Cheng 2011) et parmi elles, les plus utilisés sont les couplages avec l'ionisation sous vide par désorption laser assistée par matrice (MALDI) et l'ionisation sous pression atmosphérique (API). Ainsi, la spectrométrie de masse (MS) apporte une dimensionnalité supplémentaire en permettant l'accès aux valeurs de rapport m/z des composés. Cette information facilite l'élucidation structurale de composés inconnus via leur fragmentation (Silva and Lopes 2015) et permet de mettre en évidence les phénomènes éventuels de co-migration de composés sur la plaque TLC. Ces développements récents apportent de nouveaux outils complémentaires pour mettre en œuvre une stratégie de dérégulation.

1.2. Mise en œuvre expérimentale de la TLC

Le **principe de la chromatographie sur couche mince** consiste à utiliser une phase stationnaire solide qui a été au préalable déposée sur un support plan en aluminium ou en verre. Cette phase stationnaire est un matériau adsorbant constitué de silice (matériau polaire), et elle peut être greffée de groupements plus ou moins polaires (NH_2 , CN, C2, C8 ou C18 principalement) pour modifier sa polarité. Les échantillons sont déposés par nébulisation à environ un centimètre en bas de la plaque (Fig.II-1). Sur une plaque de dimension 10 cm par 20 cm de long, 18 échantillons peuvent ainsi être déposés l'un à côté de l'autre en bande de 8 mm de largeur sur toute la longueur d'une plaque de 20 cm. Une fois le dépôt effectué et séché, les constituants de l'échantillon sont séparés par une phase mobile (appelée aussi éluant). Cette phase mobile constituée d'un mélange de solvants permettant d'affiner au mieux la séparation, migre par capillarité le long de la plaque en entraînant les molécules (Fig. II-1). Selon sa polarité l'éluant interagit plus ou moins avec la phase stationnaire et les solutés ce qui induit une migration plus ou moins importante des composés. Le choix des solvants se fera donc en fonction de la force éluante et de la sélectivité des solvants par rapport aux analytes à séparer. Snyder a classé les solvants en huit groupes de sélectivité répartis dans un triangle en fonction de leur aptitude à être donneur ou accepteur de protons et de leurs interactions dipolaires (Fig.II-2) (Kowalska 2003). L'optimisation du solvant peut se faire en quatre étapes pour obtenir une séparation adéquate des composés sur plaque de silice:

- 1) Sélectionner et tester la force éluante d'un solvant par groupe de sélectivité (à l'exception du groupe IV),
- 2) Réduire ou augmenter la force éluante de ce solvant en ajoutant respectivement de l'hexane ou de l'acide formique,

- 3) Réaliser des mélanges de groupe de solvants et évaluer les sélectivités de ces mélanges,
- 4) Affiner les proportions de solvants dans le mélange offrant les meilleures conditions de sélectivité pour améliorer la résolution.

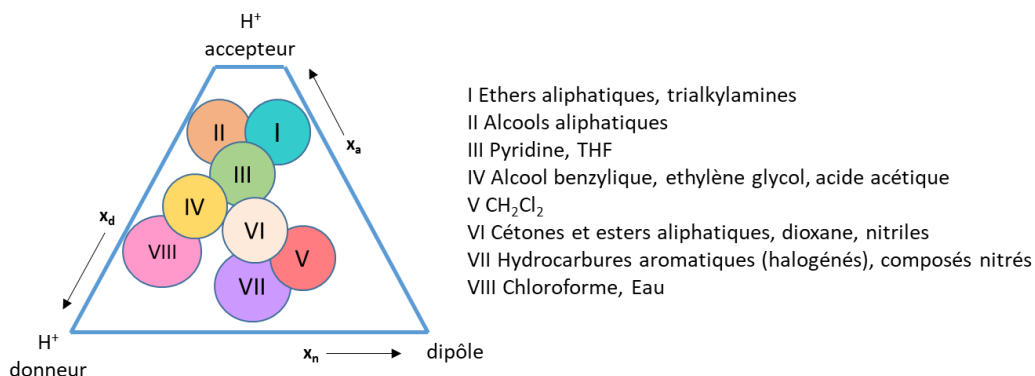


Figure II-2 : Classification des solvants selon leurs propriétés d'accepteur ou de donneur de protons ou de force dipolaire. Représentation des 8 classes de solvant selon le triangle de Snyder.

La séparation ainsi effectuée, les composés sont détectés soit directement dans le visible (molécules colorées naturellement), soit sous une lampe optique ou après révélation par application sur la plaque d'un agent chimique ou biologique (Fig. II-1).

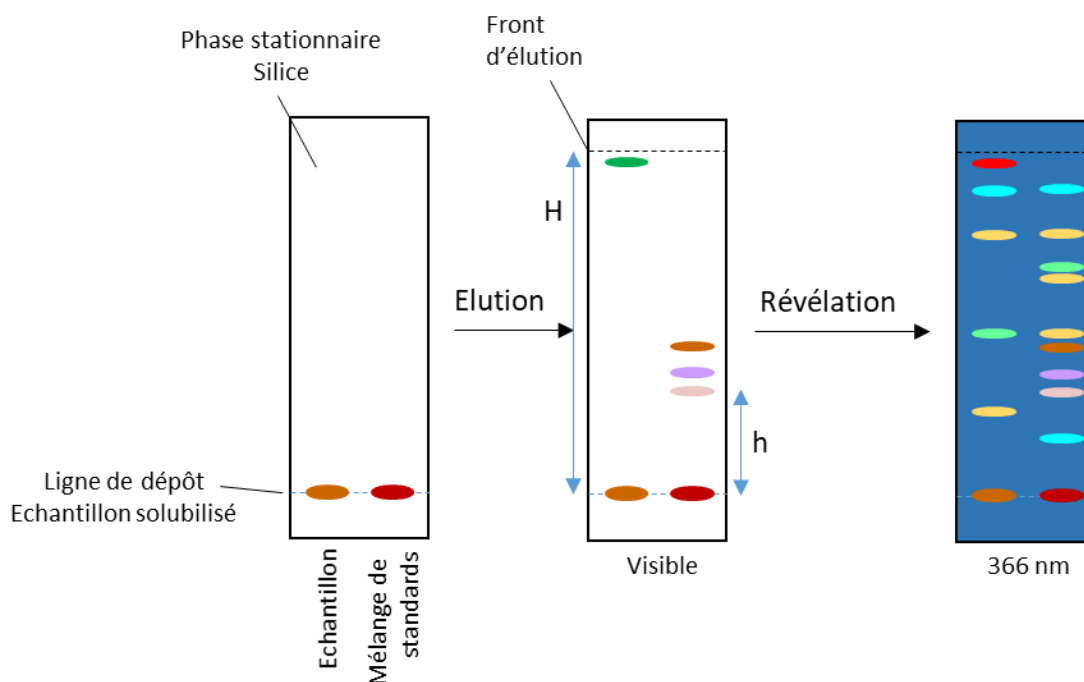


Figure II-1 : Principe de la chromatographie sur couche mince et visualisation (directement dans le visible ou sous lumière UV à 366 nm) des composés d'un échantillon et d'un mélange de standards.

Chaque spot révélé est défini par son rapport frontal (R_f) qui correspond à sa migration relative par rapport au solvant d'élution. Le rapport frontal (R_f) se calcule par la distance parcourue par le soluté sur la distance parcourue par le solvant ($R_f = \frac{h}{H}$).

Dans l'exemple de la figure II-1, la révélation permet de mettre en évidence des composés qui n'étaient pas colorés dans le visible. La comparaison d'un extrait et d'un mélange de standards positionnés sur les pistes côte à côte peut se faire de façon visuelle, en comparant la présence ou l'absence de tache à un R_f donné ainsi que l'intensité et la couleur des taches communes. Ainsi la figure II-1 illustre ces propos : en haut de la plaque deux standards ont le même R_f que des constituants de l'échantillon et sont révélés par des spots de même couleur : bleue ou jaune respectivement. Cette information permet d'indiquer qu'au niveau de ces deux spots, les composés présents dans l'échantillon ont la même polarité que les 2 standards déposés (même R_f) et qu'ils appartiennent à la même famille moléculaire (même couleur de spot après révélation) mais sans certitude que cela soit vraiment les deux mêmes composés. Pour les autres molécules n'ayant pas de correspondance de migration avec un standard déposé, la couleur donne une information d'une appartenance possible à une famille moléculaire. Et la position des taches sur la plaque donne une indication de polarité, par rapport aux autres standards, des composés inconnus.

Dans la majorité des cas, une analyse en TLC permet de limiter la préparation d'échantillon à la seule étape de solubilisation de l'extrait dans le solvant de dépôt et permet de séparer plusieurs échantillons en parallèle, tout en n'ayant pas d'effet de mémoire du support comme on peut parfois le constater en chromatographie liquide (changement de plaque TLC après chaque analyse tandis que de nombreuses analyses successives sont effectuées sur une même colonne HPLC). L'absence d'effet mémoire évite d'induire un biais lors de l'interprétation des analyses. Cette technique permet de réaliser un screening rapide de plusieurs échantillons par famille moléculaire.

Afin d'avoir plus de précision dans les analyses qu'une comparaison visuelle, la plaque peut être lue par un appareil densitométrique. Ceci permet de visualiser le résultat sous forme d'un chromatogramme qui est établi à une longueur d'onde donnée et qui reporte en fonction de la distance de migration, l'absorbance des taches tout le long de la plaque et pour chaque dépôt élué. Pour aller plus loin dans la comparaison des empreintes phytochimiques de différents extraits analysés, ces chromatogrammes peuvent être alors comparés et leur comparaison rationalisée en utilisant un traitement d'image photo de la plaque TLC suivi d'analyses statistiques.

2. Analyse non ciblée de différents organes du Colza (*Brassica napus* L.) par TLC

La TLC utilisée pour un premier criblage rapide d'extraits permet de pouvoir comparer la signature moléculaire des plantes qui peuvent être soit totalement différentes ou de différentes variétés ou des organes différents ou des préparations différentes...

2.1. Contexte

Le **colza** (*Brassica napus* L.) est une plante appartenant à la famille des Brassicacées (Fig. II-3). Une représentation photographique des différents organes de cette plante : graines, racines, tiges, feuilles et fleurs est présentée ci-dessous (figure II-3) :

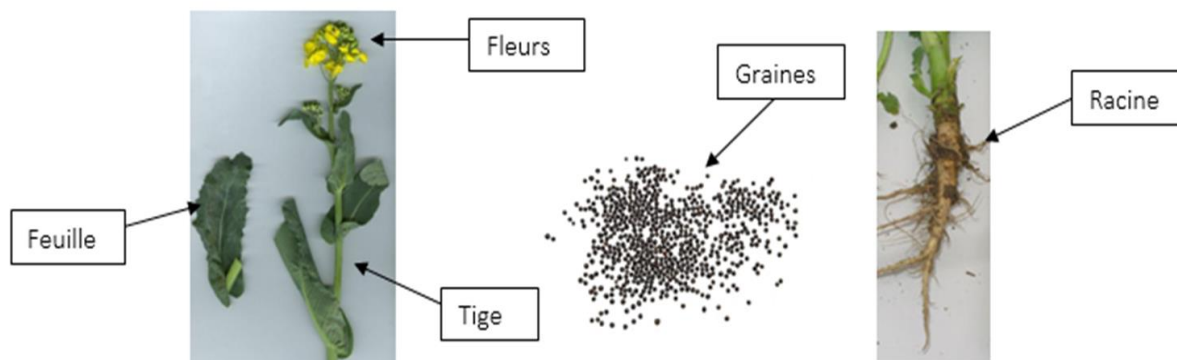


Figure II-3 : Différentes parties de la plante de colza.

Le colza est cultivé actuellement pour la production d'huile alimentaire et d'agrocarburant. Ce sont ainsi les graines qui sont principalement exploitées, les autres parties de la plante restant sous-exploitées. Dans le cadre d'une approche globale de valorisation de la biomasse et des co-produits de cette culture, une meilleure connaissance du contenu moléculaire de cette plante ainsi que des teneurs de ses constituants, s'avère indispensable.

Des polyphénols ont déjà été décrits comme présents dans les graines de colza, tels que des dérivés de kaempférol, d'isorhamnétine (Jing 2015) et de quercétine, des flavanols, des tannins ainsi que des acides hydroxycinnamiques et des esters choline de dérivés d'acide sinapique (Shao 2014). Sang et al. (1984) ont mis en évidence que les organes feuilles, racines et graines du colza contenaient tous des glucosinolates, mais à des proportions différentes. El Beltagi et Amin Mohamed (2010) ont également décrit des glucosinolates de type aliphatique, indolique et aromatique dans les graines, mais également des acides gras, comme majoritairement les acides oléique, linoléique, stéarique et linolénique.

Le travail que nous avons réalisé et que nous présentons dans ce paragraphe, s'inscrit ainsi dans ce contexte. Pour acquérir une meilleure connaissance phytochimique de la plante, une analyse non ciblée a été menée sur la comparaison des différentes parties du colza pour ainsi évaluer la répartition des différentes familles moléculaires connues dans les différents organes et établir des similitudes ou différences de profils moléculaires des différentes parties de la plante.

2.2. Mise en évidence par TLC des différentes familles moléculaires présentes

Après avoir récolté le colza (Annexe 1), les différents organes ont été séparés, puis conservés par congélation. Tous les organes ont été ensuite extraits, par micro-ondes avec de l'éthanol afin de solubiliser sélectivement les composés polaires à moyennement polaires contenus dans ces organes, et pouvoir comparer les compositions moléculaires dans les mêmes conditions (Annexe1). Afin de connaître la composition de ces extraits en polyphénols et en composés lipidiques, moyennement polaires, les différents extraits ont ensuite été analysés par TLC sur plaque de silice selon deux conditions d'élution. Une première analyse ciblant plus spécifiquement les composés phénoliques a été réalisée avec un système d'élution composé du mélange de solvants suivant : acétate d'éthyle (AcOEt), acide formique (AF), acide acétique (CH_3COOH) et eau avec des proportions en volume respectivement de 100 :11 :11 :27. Et la deuxième analyse ciblant les composés lipidiques a été réalisée à l'aide du système d'élution constitué du mélange de solvants suivant : chloroforme (CHCl_3), n-heptane et méthanol (MeOH) avec des proportions en volume respectivement de 2 :2 :0,5. Après la séparation, pour détecter les différents constituants en présence, une révélation adaptée à chaque famille moléculaire a été appliquée. Les réactifs Neu puis PEG ont été appliqués sur la première plaque afin de mettre en évidence les polyphénols à 366 nm. La primuline, puis l'anisaldéhyde ont été appliqués successivement sur la 2^{ème} plaque pour visualiser les composés lipidiques.

2.2.1. Recherche des polyphénols

L'empreinte en polyphénols et notamment celle en flavonoïdes, a donc été réalisée pour les différentes parties de la plante (Fig.II-4A). Pour révéler les polyphénols, le réactif d'amino-2-éthyle diphénylborate dit Neu est d'abord appliqué sur la plaque après élution, puis séché. Ce réactif permet de dériver les flavonoïdes par complexation de l'atome de bore (Kartnig and Göbel 1996). Ensuite une solution de polyéthylène glycol (PEG) 4000 est appliquée afin d'intensifier les couleurs. Les polyphénols sont visualisés à une longueur d'onde de 366 nm sous forme de taches fluorescentes dont la couleur dépend de leur structure moléculaire. Ainsi, les acides phénoliques sont associés à une couleur bleue, les dérivés de quercétine à la couleur jaune orangé et les dérivés de kaempférol et d'isorhamnétine à la couleur verte.

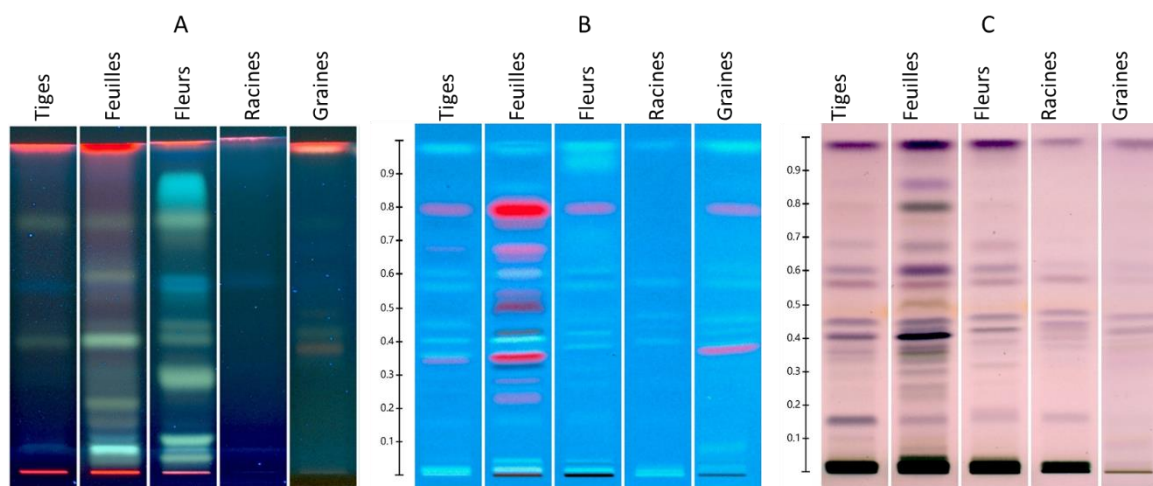


Figure II-4 : Empreintes moléculaires des différents organes du colza (tiges, feuilles, fleurs, racines, graines) par TLC. Phase stationnaire: plaque TLC verre silice 60 F₂₅₄ (20x10 cm). Dépôt 5 µL.

A) Recherche des polyphénols. Solvant de migration: AcOEt/AF/CH₃COOH/Eau (100 :11 :11 :27, V/V/V/V). Détection 366 nm. Agent dérivant Neu-Peg.

B) et C) : Recherche de lipides et terpènes. Solvant de migration: CHCl₃/n-heptane/MeOH (2 :2 :0.5, V/V/V). **B)** : Détection à 366 nm après une 1^{ère} application de primuline **C)** Détection à la lumière blanche après une 2^{ème} dérivatation à l'anisaldéhyde.

Dans les conditions d'élution utilisées (Fig. II-4A) le standard aglycone (quercétine) est élué en front de solvant, alors que la molécule monoglycosylée (quercitrine) est éluée avec un *R_f* d'environ 0,8, et la molécule diglycosylée (rutine) avec un *R_f* voisin de 0,5. Ainsi dans ces conditions d'élution. Les composés ayant une même génine sont séparés en fonction de leur nombre de sucres. Plus la molécule est glycosylée, moins elle migre ce qui est attendu lors d'une séparation TLC utilisant une phase stationnaire polaire (gel de silice).

Visuellement les parties aériennes (fleur et feuille) de la plante possèdent beaucoup de flavonoïdes (Fig. II-4A) contrairement aux racines qui ne semblent pas en contenir. Les tiges et les graines quant à elles, en contiendraient légèrement. D'après la littérature et les couleurs des spots, les différents organes contiendraient des dérivés de kaempférol et/ou d'isorhamnétine, ainsi que des acides phénoliques. Dans les feuilles et les fleurs, les composés présents couvrent une large gamme de polarité, allant du plus polaire (di ou triglycosylés), avec des composés peu élués restant en bas de plaque au moins polaire, (génines), élués en haut de la plaque, en passant par des composés de polarité intermédiaire (monoglycosylés).

2.2.2. Recherche des lipides et terpènes

Les empreintes en composés lipidiques ont été faites sur une plaque en silice comme pour les flavonoïdes, avec un système d'élution moins polaire que pour l'analyse des flavonoïdes, constitué de chloroforme, du n-heptane et de méthanol en des proportions respectives 2 ; 2 et 0,5. Une fois la plaque élue, ne disposant pas de révélateurs spécifiques aux lipides et terpènes, des révélations plus génériques sont appliquées pour mettre en évidence un maximum de composés. Les composés possédant une ou des insaturation(s) sont révélés avec la primuline comme agent dérivant non-

destructif et visualisés à 366 nm (Fig.II-4B). Après avoir pris en photo la plaque, l'anisaldéhyde est pulvérisé pour révéler à la lumière blanche des fonctions alcools allyliques, phénoliques, amines, aldéhydes, cétones, glucides et esters (Fig. II-4C). Avec cette révélation, les alcènes, les alcynes et les composés aromatiques ne seront pas visibles.

Etant donné que l'extraction a été réalisée à l'éthanol, ce sont les lipides les plus polaires qui ont pu être extraits. La deuxième dérivation à l'anisaldéhyde (Fig II-4C) révèle très peu de nouveaux spots par rapport à ceux précédemment détectés avec la 1^{ère} révélation à la primuline (Fig II-B).

Parmi tous les organes analysés, ce sont les feuilles qui semblent contenir le plus de composés lipidiques puis les tiges et les fleurs.

Pour confirmer ces 1^{ères} interprétations visuelles et aller au-delà, des analyses non ciblées de ces familles moléculaires ont été réalisées avec un logiciel en libre-service sur internet, nommé rTLC. Cette application permet de cartographier le regroupement des échantillons directement à partir de la photographie de la ou les plaques d'analyse. Sur la base des 3 systèmes d'analyse TLC présentés (Fig.II-4A, B et C), nous avons ainsi pu nous familiariser avec l'utilisation de ce logiciel et tester sa pertinence en croisant les 1^{ers} résultats obtenus visuellement avec ceux obtenus en appliquant le traitement d'images proposé par le logiciel.

2.3. Analyses non supervisées sur différents organes du colza (feuille, racine, tige, graine)

Les analyses statistiques des échantillons de colza ont été réalisées à partir de l'application web open-source, nommée rTLC qui a été créé par Fichou et al. (2016). Cette application permet de traiter directement les photos des plaques TLC après élution et de réaliser des analyses multivariées comme l'analyse en composante principale (ACP) et la classification ascendante hiérarchique (CAH). Seules les fleurs n'ont pas été incluses dans cette étude par manque de matériel végétal.

Les organes frais et broyés ont été séparés en sept lots de 500 mg. Chaque lot constitue un échantillon qui est extrait dans les mêmes conditions, à l'éthanol par micro-ondes (Annexe 1). Chaque organe est divisé en sept lots indépendants afin d'être le plus représentatif dans les analyses statistiques. Chaque échantillon est déposé une seule fois pour être analysé. Les dépôts ont été réalisés sur deux plaques différentes et répartis de façon à avoir les quatre organes sur chacune des deux plaques (Annexe 3). Les études ont été réalisées séparément par système d'élution et de révélation selon la famille moléculaire à mettre en évidence.

Lors des éluions de plaques, des effets de bord ainsi qu'une élution non linéaire de l'ensemble des dépôts pouvant être observés et entraîner de ce fait une moins bonne reproductibilité d'analyse d'une plaque à une autre, un certain nombre de corrections de lecture sont possibles avec l'application rTLC. Pour chaque chromatogramme associé à chaque dépôt élué, des corrections de ligne de base, de contraste, de bruit de l'image, de lissage et de normalisation des données peuvent être faites ainsi que l'application d'un algorithme de similarité entre différentes plaques suivant la qualité de l'élution. Les données peuvent être centrées, en soustrayant chaque valeur à la moyenne de sa colonne, ainsi le nuage de points est centré sur l'origine. Elles peuvent être aussi réduites, en divisant chaque donnée par l'écart type de sa colonne. Comme Audoin et al. l'ont montré, le pré-traitement des plaques est

important à faire avant de réaliser des analyses statistiques (Audoin 2014). Ainsi dans le système d'élution pour les composés lipidiques et de révélation avec la primuline, un exemple de l'influence des corrections sur l'analyse en composante principale est présenté (Fig. II-5).

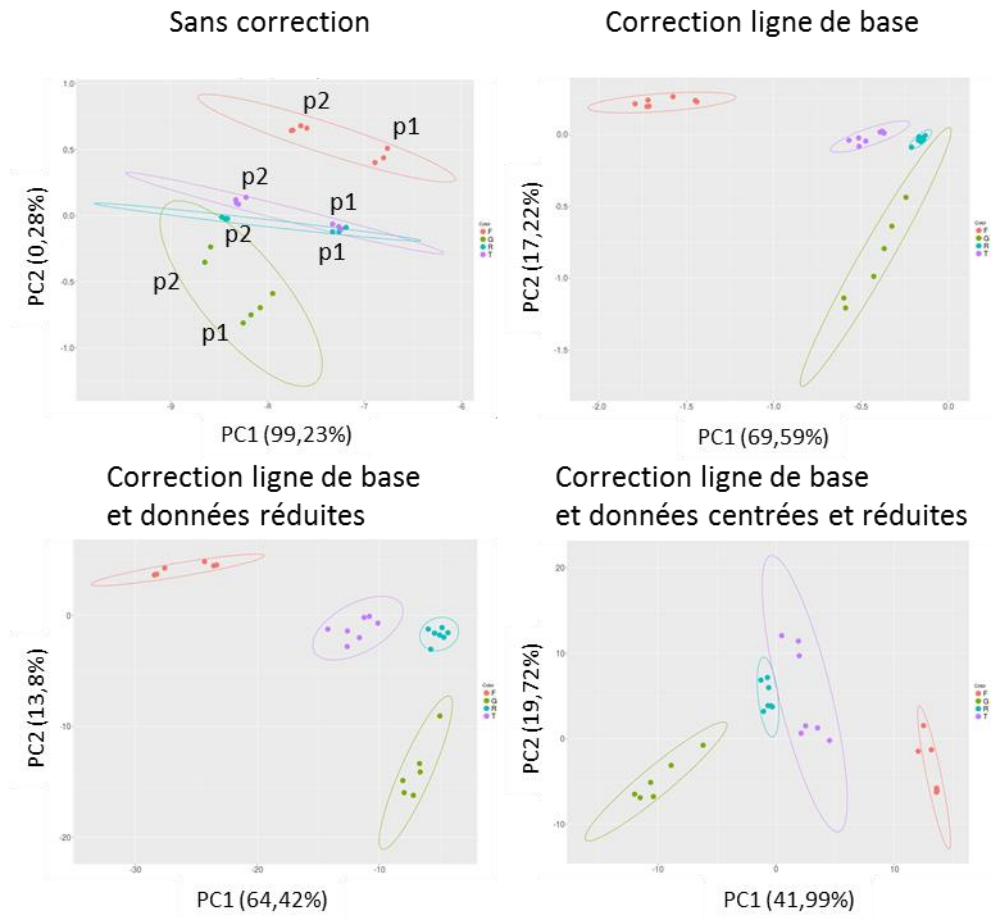


Figure II-5 : Influence des corrections sur l'analyse en composantes principales du système d'élution pour les composés lipidiques et système de révélation avec la primuline. Les échantillons en rouge correspondent aux feuilles, en vert aux graines, en bleu aux racines et en violet aux tiges. Les indications p1 et p2 correspondent respectivement aux plaques 1 et 2.

Ainsi pour ce système d'élution-révélation, les analyses en composantes principales (Fig. II-5) montrent qu'il y a besoin d'appliquer une correction. En effet, sans correction la séparation ne se faisait pas par organes mais plus par groupe d'échantillons et principalement les échantillons d'une plaque (p1) par rapport aux échantillons de l'autre plaque (p2), alors que visuellement sur les deux plaques (Annexe 3) des différences entre organes sont observées. Avec une correction, les groupes sont établis par organe. Le fait d'appliquer les trois corrections (ligne de base et avoir les données centrées et réduites) permet de séparer les quatre organes selon la première composante avec une variance de 42 %, mais la répartition des échantillons par organe semble encore diffuse (ellipse haute). Tandis que les deux autres systèmes de corrections (ligne de base avec ou non les données réduites) permettent d'avoir une meilleure variance sur la première composante (respectivement 69,6 % et 64,4 %). Ainsi les échantillons feuilles (rouge) et graines (vert) sont très bien séparés. Mais c'est avec le système de correction de ligne de base et les données réduites que les échantillons de tiges et de racines semblent

les mieux séparés selon l'axe 1. C'est donc ce système de correction qui sera sélectionné pour ce système de séparation/révélation.

Donc par la suite, les différentes corrections ont été testées sur chaque système d'élution-révélation. Le choix du système de correction s'est fait en fonction de la meilleure séparation et de la meilleure variance par système, comme démontré avec la figure II-5.

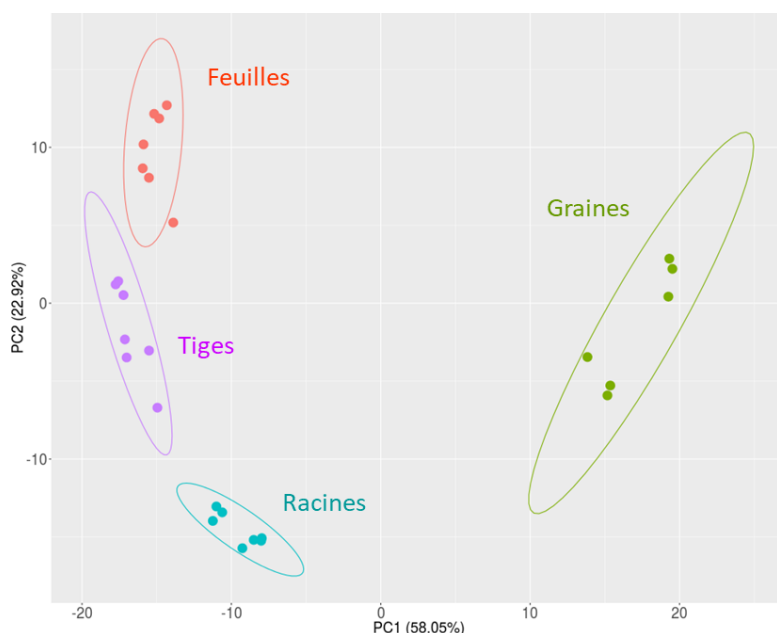


Figure II-6 : Analyse en composantes principales des différentes parties (feuilles, tiges, graines et racines) de la plante de colza. Correction de la ligne de base et données centrées et réduites. Systèmes d'élution et de révélation associés à la famille des polyphénols.

La première analyse multivariée que nous avons réalisée se base sur les données obtenues en appliquant les conditions d'analyse (système d'élution et révélation) des polyphénols (Fig. II-6). Le pourcentage d'explication de la variabilité est de 81 %. Chaque organe est associé à un groupe distinct. Cette première analyse permet de séparer les graines des trois autres organes sur la première composante avec une variance de 58 %. Ce premier constat est en bonne adéquation avec l'observation visuelle des plaques TLC (Annexe 3) qui met en évidence une coloration bleue pour les spots associés à l'analyse des graines et une coloration verte pour les spots associés à l'analyse des trois autres organes (Fig. II-6).

Les trois organes (feuilles, tiges et racines) sont légèrement séparés sur la deuxième composante avec une variance de 23 %. Pour ces trois organes, visuellement (Fig II-4A) les analyses TLC conduisent à des profils très semblables mais avec des intensités de taches différentes d'un organe à un autre. L'analyse visuelle des plaques confirme que ces 3 organes contiendraient les mêmes types de composés polyphénoliques mais avec des abondances relatives variables d'un organe à l'autre. La feuille contient plus de polyphénols que la tige et elle-même en contient plus que la racine (Fig. II-4).

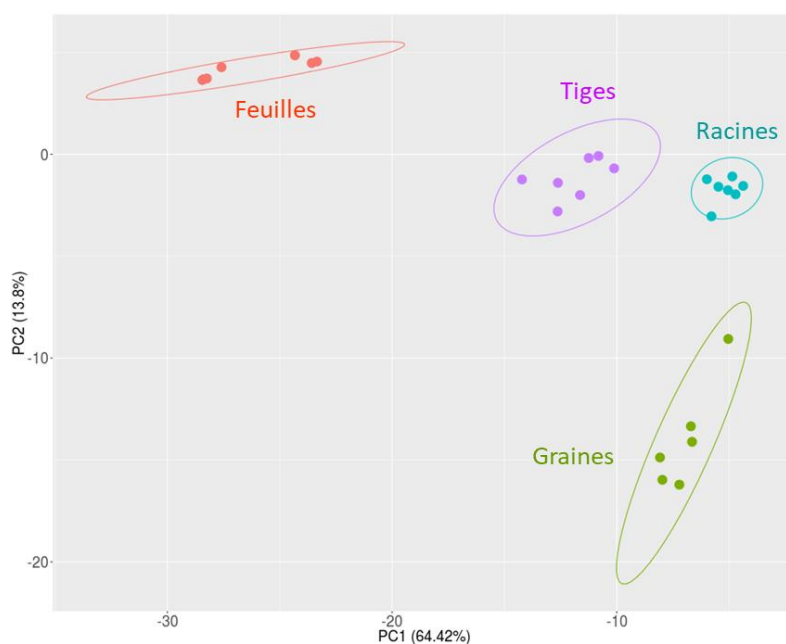


Figure II-7 : Analyse en composantes principales des différentes parties (feuilles, tiges, graines et racines) de la plante de colza. Correction de la ligne de base et données réduites. Système d'élution pour les composés lipidiques et système de révélation avec la primuline.

La deuxième analyse en composantes principales, faite à partir des données sur les composés moins polaires conduit là encore à 4 groupes distincts associés chacun à un organe particulier. Le pourcentage d'explication de la variabilité est dans cette expérience de 78 %. Les groupes sont séparés de façon un peu différente (Fig. II-7) de ce qui est observé lors de la première analyse faite à partir des données sur les composés phénoliques. Ainsi, dans cette deuxième analyse, via la composante 1 et avec une variance de 65 %, ce sont les feuilles qui sont bien séparées des 3 autres organes (tiges, racines et graines). Les graines sont séparées des autres organes via la composante 2 avec une variance de 14 %. Il apparaît que les organes ne sont donc pas discriminés de la même façon selon le type de familles moléculaires d'intérêt. Les tiges et les racines restent cependant dans les deux cas de figure des organes très proches quant à leur composition en ces familles moléculaires.

Lors de l'examen visuel des plaques TLC (Annexe 3), correspondant à l'analyse des 7 lots de 4 organes on observe bien que les échantillons tiges et racines se ressemblent avec à peu près le même nombre de spots observés et que les échantillons de racines et de graines diffèrent davantage.

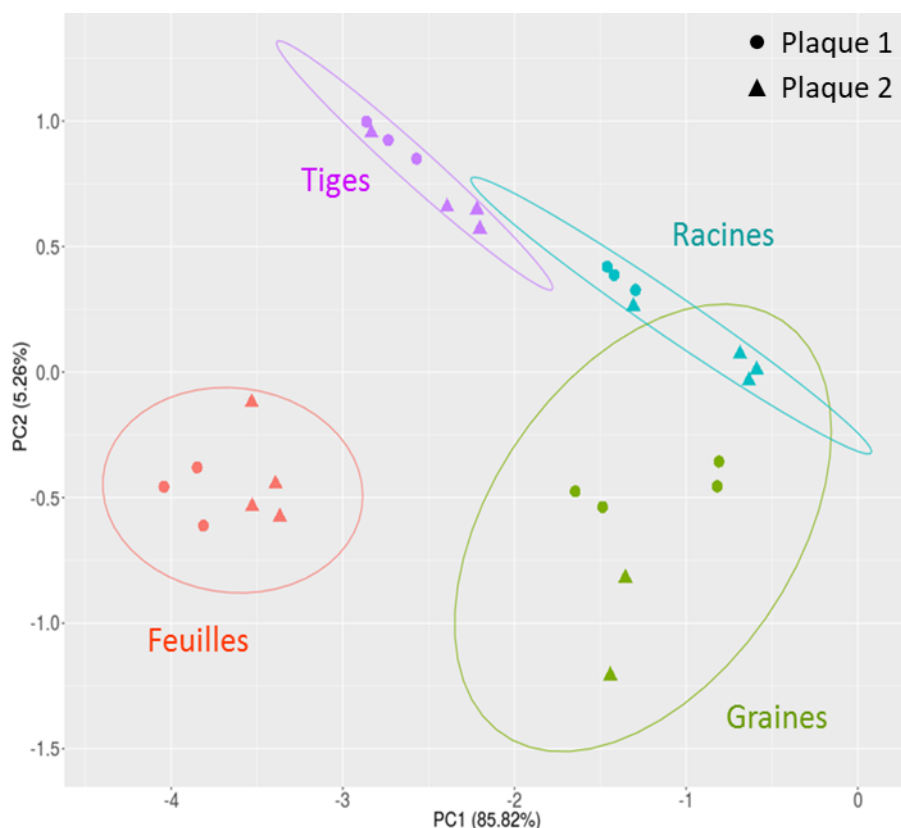


Figure II-8 : Analyse en composantes principales des différentes parties (feuilles, tiges, graines et racines) de la plante de colza avec indication de la plaque d'analyse 1 ou 2 pour les échantillons et une correction de la ligne de base. Système d'élution pour les composés moins polaires et révélation avec l'anisaldéhyde.

La 3^{ème} analyse multivariée que nous avons réalisée se base sur les données obtenues en appliquant les conditions d'analyse (système d'élution) des composés moins polaires et une révélation dans le visible après dérivatisation à l'anisaldéhyde. Par rapport aux 2 précédentes analyses, la première composante porte une variance encore plus importante de 86 % tandis que la variance de la composante 2 est bien plus faible à 5,26 % (figure II-8). Les feuilles et les graines sont toujours bien séparées selon la première composante. Mais étant donné que la variance sur la deuxième composante est faible (5 %), cette deuxième variable ne permet pas de mieux distinguer entre eux les trois groupes racines, tiges et graines. Les tiges et les racines sont légèrement séparées selon la première composante mais restent les 2 organes les plus proches dans leur composition. Visuellement sur les deux plaques TLC (Annexe 3) les pistes des échantillons de feuilles et de graines sont les plus différentes. La légère différence entre les pistes de tiges et de racines est davantage liée à l'intensité des bandes.

Ainsi des traitements de correction (ligne de base, centrage et réduction des données), seuls ou en combinaison, ont pu être appliqués sur les jeux de données des trois systèmes d'élution-révélation que nous avons étudiés. Les corrections appliquées aux traitements des plaques diffèrent d'un système d'élution-révélation à un autre, car les problèmes rencontrés peuvent être différents tels que

contraste de plaque différent suivant si la plaque est prise dans le visible ou à 366 nm, révélation non homogène d'une plaque à une autre, donc des tâches plus ou moins bien révélées, élution avec ou sans effet de bord et linéaire ou non suivant le système d'élution... Toutefois les corrections peuvent ne pas être suffisantes, en effet, sur la figure II-8, l'intervalle de confiance (ellipse) semble plus grand, surtout pour les graines, où la répartition semble plus dispersée, malgré la correction de la ligne de base, donc les profils moléculaires entre les échantillons d'un même organe différent plus. Quand plusieurs plaques doivent être réalisées, elles doivent être éluées de la même façon (même temps de saturation de la cuve, même temps et hauteur d'élution), car le logiciel compare les taches des plaques en relatif en fonction du rapport frontal et non en absolu en fonction de la distance de migration et le logiciel associe un même front de solvant pour l'ensemble des plaques. Les différences apportées peuvent être tellement grandes d'une plaque à une autre qu'un effet plaque peut être mis en évidence, sans que les corrections n'arrivent à diminuer cette différence. Ainsi dans la figure II-8, dans chaque groupe d'organe, deux sous-groupes sont visibles (●, ▲), qui indiquent les échantillons analysés sur la plaque 1 ou 2.

L'analyse en chromatographie sur couche mince permet d'obtenir rapidement des premières empreintes d'extraits complexes et de cibler en première intention les familles moléculaires auxquelles appartiennent les composés par la sélection du système d'élution et de révélation. Une comparaison de la composition des différents extraits peut être réalisée soit visuellement soit via un logiciel de traitements statistiques. L'aide du logiciel permet de réaliser rapidement des traitements statistiques directement à partir de la lecture de plaque et ainsi de comparer en une fois un très grand nombre d'échantillons ce qui n'est pas réalisable manuellement. Comme tout traitement statistique de données issues d'analyses chromatographiques, les corrections peuvent influencer les résultats et donc les conclusions des expériences ce qui nécessite d'exercer un regard critique et circonstancié sur cette approche lors du choix des corrections appliquées.

Le but de ce premier travail présenté sur les différents organes du colza n'étant pas de repérer et d'identifier les molécules permettant de différencier les organes entre eux, mais de cartographier la similitude des échantillons, l'étude statistique que nous avons menée via l'application rTLC a été réalisée avec l'ensemble des quatre canaux de lecture de couleur (rouge, vert, bleu et gris) et sans priorisation de l'un d'entre eux pour avoir la plus grande représentativité de toute l'empreinte. La lecture sur la plaque TLC des pistes associées à chacun des dépôts a ainsi généré différents profils dans les couleurs primaires et le gris (nommé canaux) sans information perdue.

Dans un contexte différent qui serait celui du repérage des spots discriminants de groupes, les analyses multivariées à l'aide de l'application rTLC doivent être réalisées d'une autre façon (que nous n'avons pas testée) en n'utilisant qu'un seul de ces canaux. Les spots permettant de distinguer les groupes seront ainsi indiqués par le logiciel par leurs valeurs de *R_f*. Un couplage TLC/MS ciblé sur ces taches discriminantes, pourra alors ensuite être réalisé afin d'aller jusqu'à l'identification des molécules éluées sous ces taches sur la base des spectres de masse acquis.

Pour aller plus loin dans l'identification, la spectrométrie de masse apparaît donc comme un outil incontournable pour apporter des informations structurales complémentaires. Dans la suite de ce chapitre, après avoir présenté brièvement les instruments, le principe, les avantages et inconvénients

des couplages TLC/MS, nous détaillerons les résultats que nous avons obtenus pour différents échantillons en couplant l'étape de séparation par TLC à une détection par spectrométrie de masse.

3. Etat de l'art de l'analyse par TLC/MS

En TLC, après l'élution, les composés séparés restent adsorbés sur le support solide (gel de silice), contrairement à l'HPLC et la GC où ils sont transportés dans un fluide liquide ou gazeux. Cette particularité de la TLC entraîne des contraintes importantes et spécifiques pour son couplage à la MS : contraintes de désorption des composés de la plaque, de leur transport vers la source et de leur ionisation sous leur forme intacte. Il en résulte de nombreux problèmes (interférence des ions de matrice, ionisation difficile des petites molécules) qui restent à résoudre pour étendre les applications du couplage TLC/MS à l'analyse des échantillons complexes.

3.1. Différentes méthodes de couplage

Il existe plusieurs façons de réaliser le couplage entre la plaque TLC et le spectromètre de masse. Deux catégories de couplage existent : sous vide et à pression atmosphérique (Morlock et Schwack 2010 ; Cheng 2011).

L'analyse sous vide peut être réalisée avec des technologies de type spectrométrie de masse à ionisation par désorption laser (LDI) et à ionisation par désorption laser assistée par matrice (MALDI).

L'analyse avec une ionisation à pression atmosphérique (API) peut concerner les sources comme l'ionisation électrospray par désorption (DESI), l'analyse directe en temps réel (DART), l'électrospray (ESI) et l'ionisation chimique à pression atmosphérique (APCI).

Le travail que nous présenterons dans la suite de ce chapitre a été réalisé en utilisant deux de ces couplages : le *couplage à pression atmosphérique* avec une ionisation par électrospray et par ionisation chimique et le *couplage sous vide* avec une ionisation par désorption laser avec ou sans matrice. Le principe de ces 2 modes de couplage est détaillé ci-après.

3.1.1. Principe du couplage TLC/API-MS

Une première méthode pour ce couplage consiste à utiliser une interface qui désorbe totalement la tache de la plaque de TLC en solubilisant le ou les solutés élués à ce *R_f* dans un solvant approprié (Luftmann 2004) et qui les dirige ensuite vers une source d'ionisation à pression atmosphérique **API-MS** pour une analyse par spectrométrie de masse. Cette approche est très intéressante, car travailler avec une source API-MS permet de pouvoir utiliser différents mécanismes d'ionisation suivant la famille moléculaire à analyser. En effet, l'ionisation peut être effectuée avec une source électrospray (ESI) pour des composés polaires de poids moléculaire faible à élevé, ou en ionisation chimique (APCI) pour des composés moins polaires et de faible poids moléculaire ou encore par photoionisation (APPI)

pour des composés apolaires de même poids moléculaire qu'en APCI. Les inconvénients majeurs de ce type d'introduction dans la source d'ionisation sont que chaque spot séparé ne peut être analysé qu'une seule fois car la buse désorbe toute la tache ; il peut y avoir du bruit de fond dû à la désorption simultanée de la silice et des problèmes de suppressions d'ions (Glavnik 2017). Il faut également trouver le solvant adéquat qui permettra d'une part de désorber les composés de la silice et d'autre part d'aider à leur ionisation quand ils arrivent dans la source. En outre, l'automatisation de l'interface ne semble pas disponible pour le moment pour un criblage automatique de toute la plaque.

Le couplage **TLC/API MS** fonctionne avec une pompe HPLC, l'interface fournie par la société CAMAG et un spectromètre de masse avec une source d'ionisation à pression atmosphérique (ESI ou APCI). La pompe HPLC permet de pousser le solvant vers l'interface à un faible débit (souvent 0,2 mL/min). Une fois le solvant arrivé au niveau de l'interface, il va jusqu'à la plaque et au spot via une buse ronde ou ovale (Fig. II-9). Le solvant permet de désorber le ou les composés du spot. Afin d'éviter que le solvant ne se disperse quand il est en contact avec la plaque, le système entre la buse et la plaque est sous pression pour avoir une bonne étanchéité. Un fritté est situé en sortie de la buse pour éviter que des particules de silice n'aille jusqu'au spectromètre de masse et qu'elles n'obstruent les capillaires. Ainsi, les composés du spot sont acheminés via le solvant jusqu'au spectromètre de masse, afin d'être ionisés et détectés.

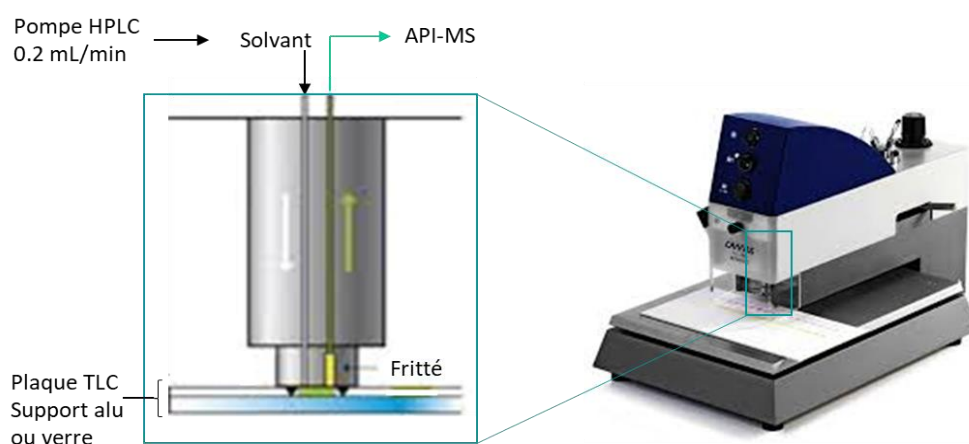


Figure II-9 : Schéma du couplage TLC/API-MS avec l'interface de CAMAG (images extraites du site web de CAMAG).

Pour réaliser ce type d'analyse, des **optimisations** sont donc à réaliser. Le *solvant* qui élue les composés du spot doit être testé pour solubiliser le plus de composés et aider à leur ionisation. Il peut s'agir d'eau, de méthanol, d'acétonitrile ou de mélange, avec ou sans additif, comme l'acide formique. Ainsi par exemple, pour l'estriol, les solvants acétonitrile et méthanol éluent convenablement la molécule. Avec l'acétonitrile l'ion moléculaire intact ainsi que des ions fragments sont obtenus (Fig. II-10) tandis qu'avec le méthanol seuls les ions fragments sont observables. Le méthanol étant un solvant protique, il y a formation de l'ion $[M+H]^+$ par transfert de proton que l'on n'observe pas sur le spectre, car il se fragmente en source. Concernant la réponse avec l'acétonitrile, ce solvant étant aprotique, il y a un échange de charge, ce qui donne l'ion $[M]^+$. L'ajout d'acide formique dans l'acétonitrile permet le transfert de proton, formant l'ion $[M+H]^+$, qui comme pour le méthanol se fragmente en source.

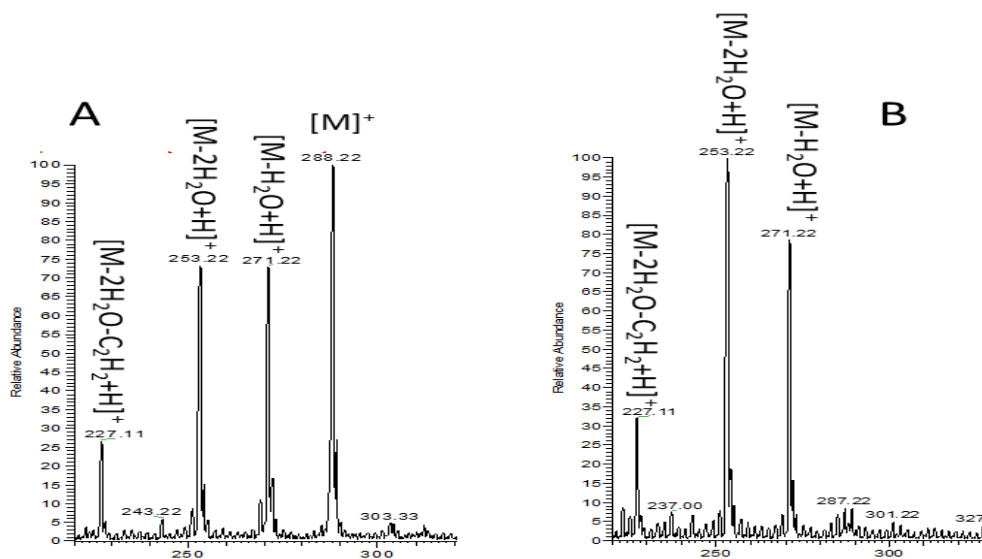


Figure II-10 : Comparaison des spectres de masse de la molécule estriol en TLC/APCI TQ-MS. (A) Solubilisation dans l'acétonitrile acidifié. (B) Solubilisation dans du méthanol acidifié.

Le mode d'ionisation (positif ou négatif), la source d'ionisation (ESI ou APCI), les paramètres de tensions, les températures, les débits de gaz au niveau de la source et la gamme de balayage de masse doivent être sélectionnés judicieusement selon le type de composé à identifier.

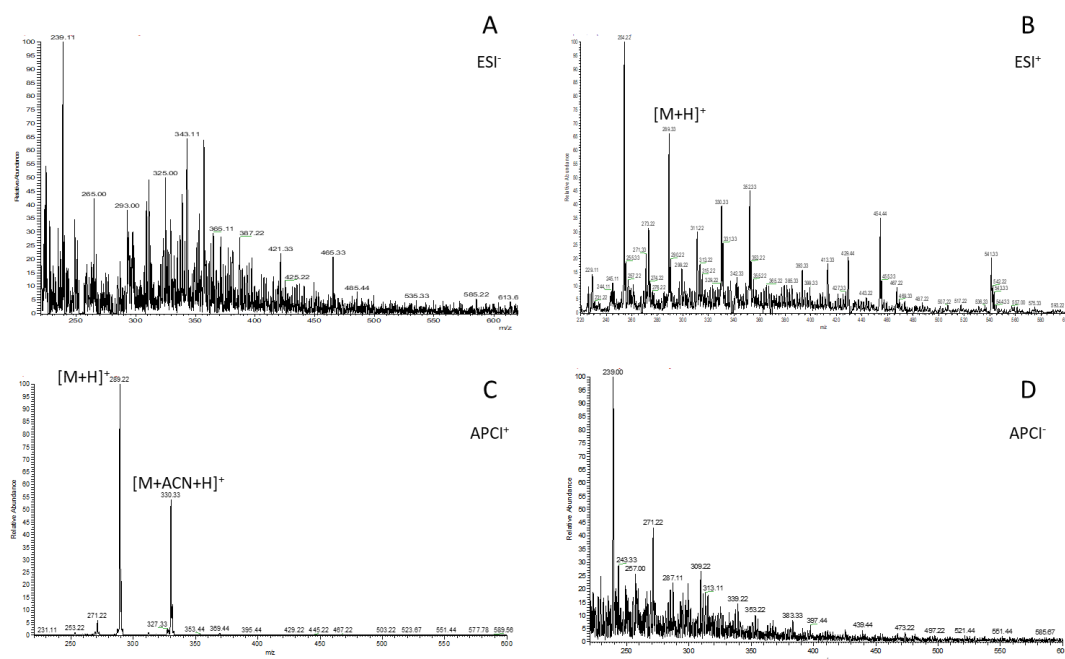


Figure II-11 : Comparaison des spectres de masse de la molécule testostérone en TLC/API-MS avec de l'acétonitrile comme solvant d'élution. (A) En mode d'ionisation négatif avec une source ESI. (B) En mode d'ionisation positif avec une source ESI. (C) En mode d'ionisation positif avec une source APCI. (D) En mode d'ionisation négatif avec une source APCI.

L'exemple de l'analyse de la molécule de testostérone solubilisée dans de l'acétonitrile (Fig. II-11) permet d'illustrer ces propos. Une source électrospray ne permet pas d'ioniser la molécule quel que soit le mode positif ou négatif d'ionisation (Fig. II-11A et 11B) alors qu'une source à ionisation chimique à pression atmosphérique le permet à condition de choisir un mode d'ionisation positif (Fig. II-11C et 11D).

L'avantage de ce type de couplage est de pouvoir analyser une large diversité de composés, avec des mécanismes d'ionisation similaires aux couplages HPLC/MS. En revanche, l'inconvénient majeur est qu'une fois le spot désorbé, il ne peut être analysé qu'une seule fois et dans une seule configuration de conditions expérimentales, car ce type de couplage est destructif. De plus, la résolution spatiale de la buse est faible (hauteur de 4 mm pour une buse ronde ; 2 mm pour une buse ovale), donc si deux composés ne sont pas suffisamment bien séparés, les deux composés peuvent être désorbés et analysés en même temps entraînant ainsi une perte d'informations précieuses dans le cas d'isomères de position par exemple.

3.1.2. Principe du couplage TLC/MALDI-Tof MS

Un autre couplage possible avec la TLC est celui avec la source d'ionisation de spectromètre de masse **MALDI**. Bruker Daltonics TM a développé un adaptateur qui facilite l'introduction d'une plaque TLC dans la source d'ionisation MALDI (Hayen et Volmer 2005) et permet d'analyser directement la plaque TLC par MALDI-Tof MS sans étape supplémentaire. L'adaptateur est pratique à utiliser, cependant le format de la plaque TLC est limité aux dimensions 5 cm par 7.5 cm. Et le support de la plaque doit être en aluminium pour une meilleure conductivité de la plaque (Fuchs 2009). Les plaques non conductrices telles que les plaques en verre causent des problèmes de charge électrostatique. Pour obtenir des spectres de masse des composés présents sur la plaque, ce couplage présente plusieurs avantages : la génération rapide des spectres, la tolérance des impuretés provenant de la silice ou de la phase d'élution et la possibilité de réaliser plusieurs analyses au même endroit (ablation superficielle du spot) dans différents modes d'ionisation (négatif et positif) et différents paramètres (énergie laser ...). Le faisceau laser étant focalisé avec un faible diamètre (environ 20 μm), une bonne résolution spatiale le long d'une bande est possible dans le cas de criblage. Par contre, l'utilisation d'une matrice déposée sur la plaque TLC est souvent nécessaire pour faciliter l'ionisation des composés et obtenir une réponse plus sensible en spectrométrie de masse. Son utilisation demande une optimisation et nécessite de contrôler au préalable l'absence d'interférence entre les ions de la matrice et les ions des analytes.

Le couplage **TLC/MALDI-Tof MS** avec l'interface de Bruker Daltonic TM (Schürenberg 2009) fonctionne sur le même principe que les analyses en MALDI-Tof MS. La source MALDI fait partie des sources d'ionisation douce. Elle a été développée dans les années 1980 pour l'ionisation de protéines de masse supérieure à 10 000 Da (Karas and Hillenkamp 1988, Tanaka 1988). Le support permet d'introduire la plaque TLC directement dans la source MALDI.

Le principe de l'ionisation MALDI a souvent été discuté (Karas 1987). Plusieurs mécanismes de formation des ions primaires et de réactions d'ionisation secondaires ont été détaillés par Zenobi et Knochenmuss (1998) et également discutés par Karas et Krüger (2003). Nous détaillerons plus

spécifiquement le déroulement du processus dans le cas où l'échantillon à analyser est présent sur une plaque TLC.

Une matrice, permettant d'ioniser les composés séparés sur la plaque TLC est déposée sur la totalité de la plaque ou juste sur les spots d'intérêt à analyser. Lors de l'application de la matrice, une cristallisation se forme entre l'analyte et la matrice après évaporation du solvant. Un faisceau laser pulsé irradie ensuite les points d'intérêt à analyser. Le laser pulsé est focalisé avec une certaine énergie d'ionisation sur une petite partie du spot (20 μm). La matrice absorbe l'énergie du faisceau laser pulsé pour aider la désorption puis l'ionisation du dépôt échantillon-matrice, un transfert de charge passe de la matrice aux analytes et provoque ou améliore l'ionisation des analytes (Fig. II-12).

La matrice utilisée doit avoir une forte absorbance à la longueur d'onde d'excitation du laser. Les composés, après être sortis de la plume d'ablation, entrent ensuite via des différences de potentiel appliquées sur des lentilles dans l'analyseur à temps de vol (ToF). Les molécules ionisées sont accélérées puis transférées dans l'analyseur. Cet analyseur va permettre la séparation des molécules ionisées en fonction de leur rapport masse sur charge (m/z). Cette ionisation s'est avérée très utilisée pour l'analyse de grosses molécules, comme les protéines (Liu and Schey 2005), les polymères (Montaudo 2006), les oligosaccharides (Rohmer 2010)... Néanmoins il reste encore des développements à mettre au point pour l'analyse des petites molécules et pour l'utilisation de cette source d'ionisation pour l'ionisation de molécules déposées sur une plaque TLC.

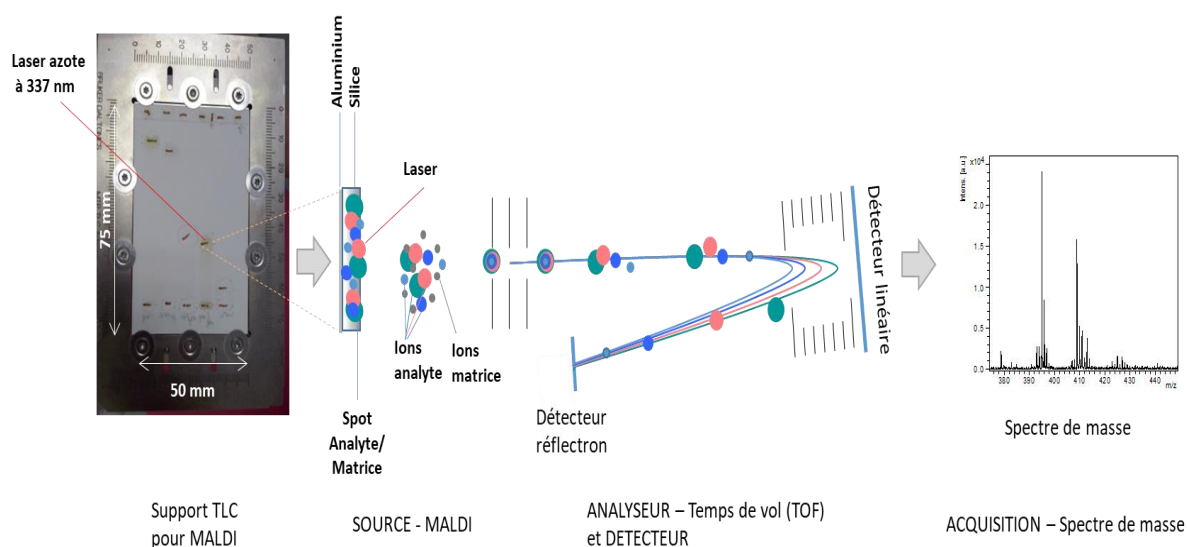


Figure II-12 : Schéma du couplage TLC/MALDI-ToF MS (Support ; Source ; Analyseur ; Détecteur et enregistreur).

En MALDI, le choix du *mode de dépôt* de la matrice est important afin d'améliorer la sensibilité et la reproductibilité des analyses. Les modes en goutte séchée ou sublimée sous vide (Jaskolla 2009), ou en mode sandwich existent. En couplage TLC/MALDI-ToF MS, la matrice ne peut être appliquée, après élution, qu'au-dessus des taches de la plaque. La matrice peut donc être déposée par mode pulvérisation ciblée directement sur le spot d'intérêt, par pulvérisation de la matrice sur la plaque et par immersion de la plaque dans la matrice pour un couplage TLC/MALDI-ToF MS. Tous ces modes de

dépôts ont leurs avantages et inconvénients. Le mode d'immersion utilise une grande quantité de matrice, mais imbibe toute la plaque TLC. Le mode pulvérisation ciblée et la pulvérisation permettent d'utiliser moins de matrice. Le mode pulvérisation ciblée permet de se focaliser si besoin uniquement sur certains spots. La pulvérisation permet d'imprégner toute la plaque de matrice, mais elle ne rentre pas en profondeur, contrairement à l'immersion. De plus, un mode de dépôt pourra convenir à un type d'analyse, mais pas à un autre. Le choix de la méthode de dépôt de la matrice reste encore empirique, dans le sens où il faut toujours faire des optimisations, sans avoir une procédure universelle qui conviendrait à toutes les analyses.

Il est difficile d'évaluer la *quantité de matrice* à déposer pour l'analyse d'un échantillon complexe, car le nombre et la quantité de composés présents diffère d'un échantillon à l'autre et de plus, on peut se trouver confronter à la présence d'une molécule en grande teneur en mélange avec des composés en très faibles teneurs. Lors de l'utilisation de certaines matrices organiques, comme l'acide 2,5-dihydroxybenzoïque (DHB), la co-cristallisation entre la matrice et l'analyte n'est pas homogène, ce qui ne permet pas d'avoir une forte reproductibilité d'une analyse à une autre. Bien que l'ionisation des analytes soit améliorée par la matrice, l'ajout d'une solution de matrice organique à la surface de la plaque peut diminuer la résolution chromatographique et la concentration des composés à analyser dû à une re-solubilisation et une diffusion latérale du spot sur la plaque.

Afin de pouvoir analyser correctement les composés avec ce type de couplage, différentes matrices peuvent être testées, aussi bien des matrices organiques (Zenobi and Knochenmuss 1998) qu'inorganiques (Shi and Deng 2016) et l'ionisation sans matrice, dite LDI. Le type, la quantité de matrice à déposer et la façon de la déposer (par immersion, par pulvérisation ciblée ou non, avant ou après élution ...) sont dans tous les cas à évaluer. Les essais doivent permettre d'obtenir une bonne ionisation des composés avec le moins de pics matrice possible et ces derniers ne doivent pas être dans la zone de rapport m/z des composés à analyser.

Une des matrices organiques les plus couramment utilisées (Annexe 4) est l'acide 2,5-dihydroxybenzoïque (DHB) qui possède une masse moléculaire de 154 Da. Or dans les extraits végétaux, beaucoup de métabolites secondaires sont de faibles masses moléculaires et peuvent répondre dans la même gamme de masse que la matrice DHB (Fig. II-13). Ainsi, dans le spectre de masse de la matrice DHB seule, plusieurs ions sont présents dans une gamme de masse allant de 100 à 600 u, dont l'ion à la masse 307 u (en mode d'ionisation négatif), qui correspond à l'ion moléculaire déprotonné aggloméré à une deuxième molécule de DHB. La présence inévitable de ces ions peut donc gêner l'identification d'autres molécules à caractériser.

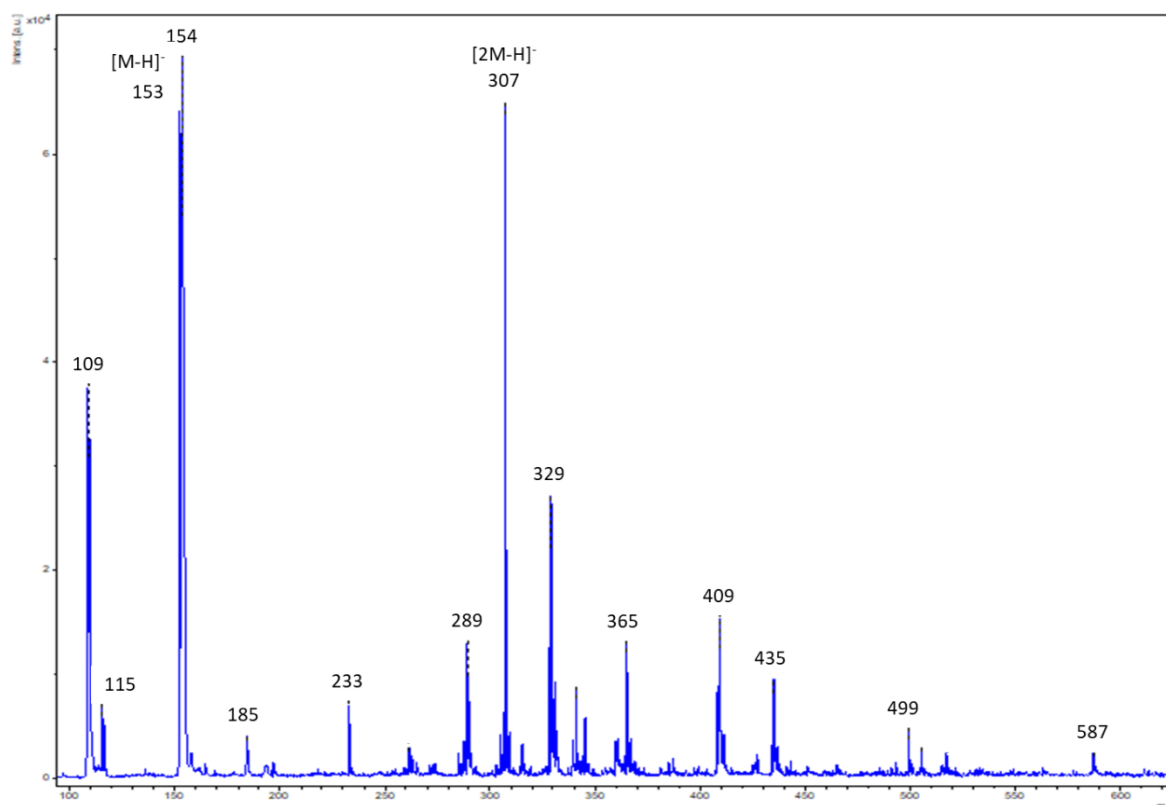


Figure II-13 : Spectre de masse en MALDI-Tof MS de la matrice DHB en mode d'ionisation négatif.

Afin d'éliminer les ions interférents de la zone d'analyse d'intérêt, plusieurs possibilités existent. De nouvelles matrices organiques peuvent être créées pour bien ioniser les composés de faibles masses sans apporter des ions en plus. Ainsi la matrice 9-aminoacridine a été testée dans le cadre d'analyse en imagerie de lipides dans des coupes de tissu cérébral de rat (Cerruti 2012). Une étape de dérivation peut être aussi menée au préalable pour augmenter la masse du composé à ioniser, comme pour des petites molécules contenant un groupement amino (Gao 2012). La matrice organique peut être remplacée par une matrice inorganique avec des nanoparticules (Tanaka 1988). Certains auteurs ont montré l'intérêt d'utiliser des matrices inorganiques (or, argent, nanoparticules à base de carbone oxyde métallique) pour minimiser les pics de matrice et les interférences dans la zone de faible rapport m/z par une étude systématique de comparaison de plusieurs matrices de nanoparticules et de matrices organiques pour l'analyse de différentes classes de petites molécules (Yagnik 2016).

Le couplage avec une matrice inorganique de nanoparticules **NALDI** (Désorption/ionisation laser assistée par nanoparticules), apporte la possibilité d'analyser des composés de faible masse (<700 Da) sans visualiser d'ions interférents de matrice dans la zone de masse où les molécules sont détectées. Beaucoup de nanoparticules et des surfaces nanostructurées ont déjà été développées (Arakawa et Kawasaki 2010) pour le mode NALDI. Selon le métal utilisé, les analyses peuvent être réalisées en mode d'ionisation positif (Chiang 2011) ou négatif (Lin et Cai 2018) et peuvent donner de meilleur résultat d'ionisation qu'avec des matrices organiques sur de petits composés, tels que des phospholipides ou des sucres et des acides aminés. De plus, avec ce type de matrice et en mode d'ionisation négatif, les composés sont seulement déprotonés et ne forment pas d'adduits. L'interprétation des spectres de masse s'en retrouve donc facilitée.

Le couplage avec une matrice inorganique dit **SALDI** (Désorption/ionisation laser assistée en surface) (Sunner 1995) apporte aussi les mêmes possibilités qu'avec une matrice NALDI pour les petites molécules. Dans une analyse en mode SALDI à base de carbone, à base de métal ou de semi-conducteurs, une suspension dans un liquide est appliquée à la surface de la plaque. Crecelius et al. (2002) ont développé le couplage TLC/MALDI-ToF MS en utilisant une matrice à suspension de particules, ce qui a permis d'analyser des antibiotiques de tétracyclines (M aux alentours de 440 Da). Dans tous les cas, la matrice inorganique absorbe l'énergie des rayons UV du laser, et transfère cette énergie aux analytes, qui sont ainsi désorbés puis ionisés (Guinan 2015).

L'énergie du laser et les tensions dans la source peuvent également être ajustées. L'optimisation doit permettre de ne pas saturer le détecteur. Comme le laser n'ablate pas toute la tache, plusieurs analyses peuvent être faites sur un même spot. De plus, le diamètre d'impact du laser étant d'environ 20 µm, l'information obtenue est plus sélective dans le cas de composés non complètement résolus.

Les deux modes de couplages ont des avantages et des inconvénients. Dans les deux cas, avant de réaliser le couplage avec le spectromètre de masse, il faut préalablement pouvoir repérer visuellement sur la plaque TLC les zones à analyser et/ou avoir déterminé les *R_f* des spots d'intérêt. Plusieurs cas peuvent se présenter (Fuchs 2009) :

- Le cas le plus simple et le plus direct est celui pour lequel les analytes sont colorés, tels que les colorants, le couplage peut donc se faire directement.
- Le deuxième cas correspond à celui pour lequel les analytes n'absorbent pas la lumière dans le visible, mais dans l'UV. Ils devront donc être repérés au préalable sous lampe UV avant de réaliser le couplage.
- Le dernier cas correspond à celui pour lequel les analytes n'ont pas de chromophores et devront être dérivés pour être détectables (exemple des sucres).

Pour bien positionner la cible laser juste au-dessus des zones à ioniser, les analyses sont ainsi dupliquées sur chaque plaque TLC. Après la séparation TLC, l'étape de révélation est réalisée sur une seule des deux pistes. Les zones d'intérêt sont repérées par la mesure du *R_f* et reportée sur l'autre partie de la plaque non révélée qui sera seule analysée en spectrométrie de masse.

3.2. Applications des couplages en phytochimie. Focus sur les flavonoïdes

Avec ces deux approches de couplage TLC/MS des polyamidoamines (Leriche 2014), des glycérophospholipides (Krüger 2015), des triterpénoïdes et des phytostérols (Naumoska and Vovk 2015), des phospholipides (Fuchs 2012, Wegener 2013) et des acides glycosphingolipides (Torretta 2014) ont déjà été analysés. Plusieurs articles font état d'applications pour certains groupes de flavonoïdes spécifiques tels que des flavones (Sajewicz 2011), des anthocyanes (Cretu et Morlock 2014), des flavanols (Smrke et Vovk 2013) avec l'utilisation de l'interface TLC/API-MS (CAMAG) ou MALDI-ToF MS sans dimension de séparation au préalable.

En 2016, Krosiakova et al. ont présenté une preuve de concept d'un couplage TLC/MALDI-ToF MS pour une analyse ciblée et automatique de quatre étalons (flavones et flavanols) et d'un extrait dopé de ces

standards pour imiter un extrait complexe de plante. La possibilité de détecter ces quatre étalons par une analyse directe de la plaque TLC par MALDI-Tof MS après son élution a donc été démontrée. Néanmoins, certaines limitations telles que la présence de nombreux pics ou adduits interférents dus à l'utilisation de la matrice organique (acide 2,5-dihydroxybenzoïque (2,5-DHB)) en mode d'ionisation positif limitent l'utilisation de cette méthodologie dans le cadre d'une caractérisation phytochimique de flavonoïdes d'extraits de plante complexes. Afin de corriger ces défauts, les flavonoïdes peuvent être ionisés en mode d'ionisation négatif afin que les composés se retrouvent sous forme déprotonée sans adduits. De plus, Il a été rapporté que les nanoparticules à base de graphène constituaient également une matrice utile pour l'analyse de standards de flavonoïdes et de coumarines en mode d'ionisation négatif avec un MALDI-Tof MS (Liu 2012).

Le couplage TLC/MS reste peu développé dans le cadre des métabolites secondaires polaires. La littérature indique peu d'exemples d'analyse de caractérisation de mélanges complexes provenant de matrices végétales, seulement quelques analyses de molécules standards. En effet, ils restent encore de nombreux problèmes (interférence des ions de matrice, ionisation difficile des petites molécules) à résoudre pour étendre les applications du couplage TLC/MS à l'analyse des échantillons complexes.

La suite de ce chapitre présentera les développements TLC/MS que nous avons réalisés pour la caractérisation de flavonoïdes dans deux échantillons (vin rouge et jus de pomme), dont le contenu moléculaire déjà décrit dans la littérature indique leur richesse en composés phénoliques.

4. Développement du couplage TLC/MALDI-Tof MS pour la caractérisation de flavonoïdes

Les travaux que nous présentons dans ce paragraphe ont été réalisés avec l'objectif de combiner les informations obtenues classiquement par une analyse TLC avec dérivation spécifique des flavonoïdes (polarité, structure) avec les spectres de masse de qualité (peu d'ions interférents, signal sur bruit satisfaisant) obtenus par TLC/MALDI-Tof MS, pour la caractérisation avancée de flavonoïdes (anthocyanides, flavonols (dérivés de quercétine ou de kaempférol), dihydrochalcone et xanthones) dans des échantillons de plantes non dopés. La présente étude vise à examiner l'efficacité d'une matrice inorganique pour l'analyse en MALDI-Tof MS avec ou sans élution au préalable, afin d'identifier les flavonoïdes présents dans les échantillons. Cette matrice, utilisée pour minimiser l'interférence des pics de matrice dans la zone de faible rapport m/z (Fig. II-14), est une matrice magnétique de nanoparticules d'oxyde de fer revêtues de silice à noyau-enveloppe, développée au laboratoire par Aleksander Salwiński (2014).

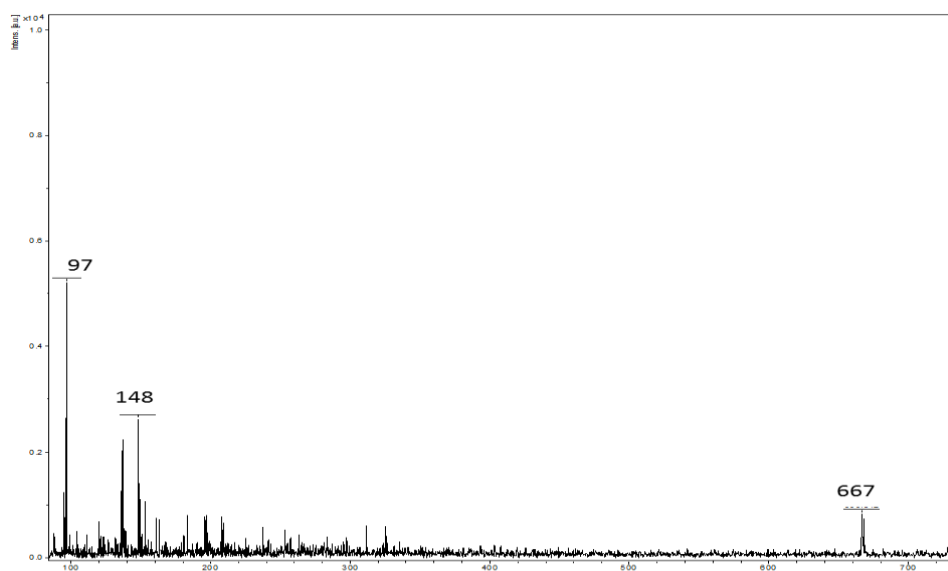


Figure II-14 : Spectre de masse en mode d'ionisation négatif de la matrice inorganique par MALDI-Tof MS (Annexe 6).

Afin d'évaluer l'efficacité du couplage TLC/MALDI-Tof MS pour l'ionisation des flavonoïdes, notre étude a tout d'abord commencé par évaluer la nécessité (ou non) d'utiliser une matrice pour ioniser ces composés avant de s'intéresser au choix de la matrice et à son dépôt sur la plaque TLC, d'abord non éluee, puis après élution des composés.

4.1. Optimisation des paramètres MALDI-Tof-MS

4.1.1. Besoin ou non d'une matrice pour l'ionisation ?

Les flavonoïdes étant constitués de cycles benzéniques, ces molécules absorbent, avec un maximum d'absorbance pour des longueurs d'onde comprises entre 280 et 370 nm. Le laser à N₂ de notre appareillage émettant à une longueur d'onde de 337 nm, comprise dans la gamme d'absorbance maximale des flavonoïdes, nous avons voulu tout d'abord évaluer si les flavonoïdes pouvaient être ionisés directement sans avoir recours à l'ajout d'une matrice. La figure II-15 montre l'ionisation de deux composés standards de flavonoïdes (acide ellagique et isorhamnétine 3-glucoside) en présence ou non de matrice.

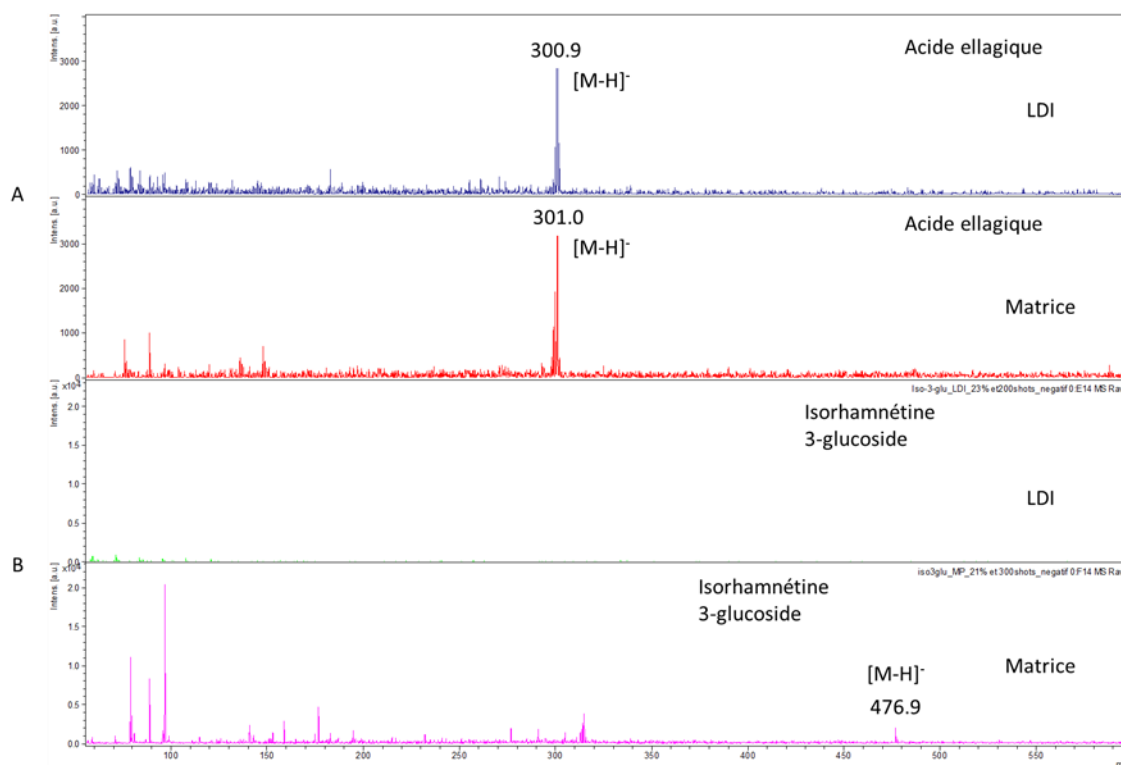


Figure II-15 : Spectre de masse par MALDI-Tof-MS de 2 standards : l'acide ellagique (A) et l'isorhamnétine 3-glucoside (B) en mode d'ionisation négatif en l'absence de matrice (LDI) ou en présence d'une matrice inorganique. Dépôt de 1 μ L de solution standard à 1000 ppm effectué sur une plaque aluminium TLC silice 60 F₂₅₄ (5 x 7,5 cm), application d'1 μ L de matrice inorganique à 10 mg.mL⁻¹ en mode goutte séchée.

Sans dépôt de matrice, l'acide ellagique (Fig. II-15A) s'ionise (m/z 301) mais pas la molécule d'isorhamnétine 3-glucoside (m/z 477) (fig. II-15B). En revanche, ces deux standards s'ionisent après avoir appliqué la matrice inorganique toutefois avec une faible intensité pour l'ion déprotonné de l'isorhamnétine 3-glucoside. Bien que le maximum d'absorption de ces molécules ($\lambda = 366$ nm) soit proche de la longueur d'onde du laser ($\lambda = 337$ nm), le transfert d'énergie entre le laser et les molécules ne s'avère pas suffisant pour ioniser correctement les molécules.

Pour la suite des analyses, une matrice sera donc appliquée sur les spots à analyser. Le choix de la matrice et le mode d'ionisation s'est donc posé.

4.1.2. Choix de la matrice

Pour analyser les flavonoïdes en mode d'ionisation positive, Krosiakova et al. (2016) ont déposé la matrice acide 2,5-dihydroxybenzoïque (2,5-DHB ou DHB). Dans ces conditions, les ions observés sont sous forme d'adduit souvent sodium et certains ions de la DHB apparaissent dans la même gamme de masse que certains flavonoïdes. Comme évoqué précédemment, l'utilisation d'une matrice

inorganique peut permettre de minimiser les ions interférents. C'est pourquoi l'ionisation de standards de flavonoïdes a été évaluée en comparant les réponses obtenues après avoir appliqué sur les standards soit une matrice organique (DHB) soit la matrice inorganique d'oxyde de fer réalisée au laboratoire dans les deux modes d'ionisation négatif et positif.

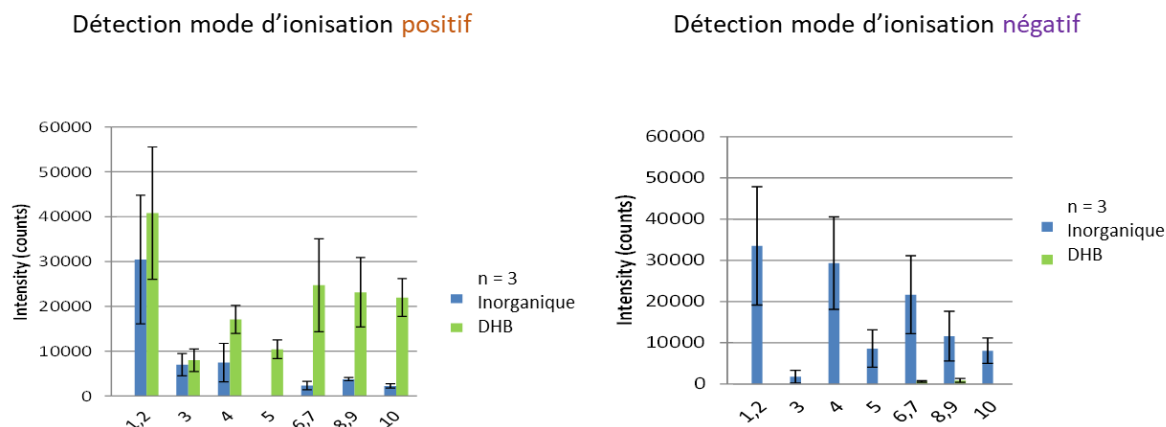


Figure II-16 : Analyse MALDI-ToF MS de 10 standards de flavonoïdes en mode d'ionisation positif et négatif en présence de 2 μL d'une matrice DHB (vert) ou inorganique (bleu) à 10 mg.mL^{-1} appliquée en goutte séchée. Dépôt sur la plaque TLC silice 60 F₂₅₄ de 1 μL du mélange des 10 composés à 100 ppm chacun. (1) Kaempférol ; (2) Cyanidine ; (3) Catéchine ; (4) Quercétine ; (5) Isorhamnétine ; (6) Quercitrine ; (7) Cyanidine 3-O-glucoside ; (8) Kaempférol 7-O-néohespéroside ; (9) Cyanidine 3-O-rutinoside ; (10) Rutine. Le dépôt est analysé sans séparation préalable du mélange.

En mode d'ionisation positif, les ions des molécules standards sont observés sous forme d'adduits (sodium). La matrice DHB permet globalement d'obtenir une meilleure intensité de signal que la matrice inorganique (Fig. II-16), bien que l'identification des composés reste plus complexe avec ce type de matrice organique due aux ions interférents de la matrice.

En mode d'ionisation négatif, les molécules sont détectées sous forme d'ions déprotonés. Et dans ces conditions, la matrice inorganique donne de meilleurs résultats que la matrice DHB (Fig. II-16) avec une meilleure intensité de signal et une meilleure qualité de spectre (moins de pics interférents). La majorité des composés ne sont pas ionisée en mode d'ionisation négatif avec la matrice organique, sauf les composés de 6 à 9, qui le sont légèrement.

A ce stade, ces premières expériences ne permettent pas toutefois de choisir entre la matrice organique ou inorganique pour l'ionisation de l'ensemble des molécules de cette famille. L'étude du dépôt de la matrice se fera alors avec les deux types de matrices.

Parmi les différents types de dépôts réalisables pour une analyse en mode MALDI, seuls les dépôts en mode immersion (Leriche 2014), ou en mode pulvérisation avec la matrice déposée au-dessus (Crecelius 2004) de l'échantillon (mode pulvérisation ciblée) ou sur toute la plaque, sont utilisables

pour le couplage TLC/MALDI-Tof MS. Seuls les modes d'immersion et de pulvérisation ciblés ont été évalués par la suite.

4.1.3. Mode de dépôt de la matrice

Les deux modes de dépôt de la matrice par immersion de la plaque ou par pulvérisation ciblée sur le spot ont été testés sur la molécule rutine et sur un mélange de standards de flavonoïdes avec les deux matrices (DHB et inorganique) et en utilisant le mode d'ionisation le plus favorable pour chacune de ces matrices (respectivement positif et négatif). La rutine a été déposée seule à une concentration de 1 mg.mL^{-1} pour être certain de la détecter et permettre de confirmer que notre couplage fonctionne, puis elle a été déposée à $0,1 \text{ mg/mL}$ en mélange avec les autres standards pour être comparés.

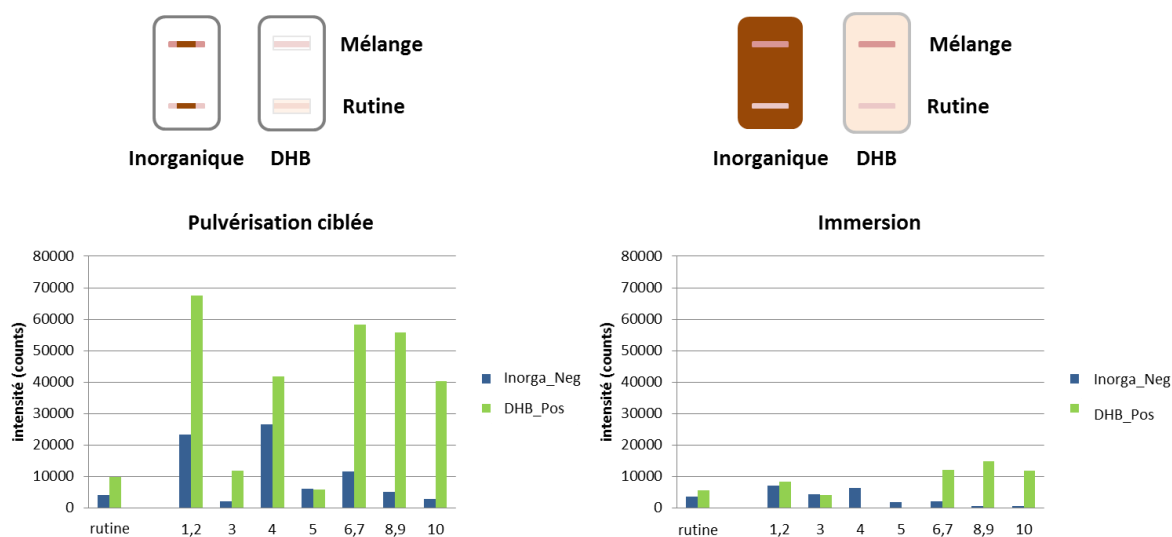


Figure II-17 : Comparaison du type de dépôt de la matrice (pulvérisation ciblée ou immersion) sur une plaque TLC silice 60 F₂₅₄ (5 x 7,5 cm). En mode spray les deux matrices ont été déposées avec un volume de $10 \mu\text{L}$ à la concentration de 10 mg.mL^{-1} . En mode immersion les concentrations étaient de 10 mg.mL^{-1} pour la matrice inorganique et de 200 mg.mL^{-1} pour la DHB (Kondakova 2017). Dépôt de $10 \mu\text{L}$ de Rutine à 1 mg.mL^{-1} et dépôt de $10 \mu\text{L}$ du mélange des 10 molécules chacune à la teneur de $0,1 \text{ mg.mL}^{-1}$. 1-Kaempférol ; 2-Cyanidine ; 3-Catéchine ; 4-Quercétine ; 5- Isorhamnétine ; 6-Quercitrine ; 7- Cyanidine 3-O-glucoside ; 8-Kaempférol 7-O-néohespéroside ; 9-Cyanidine 3-O-rutinoside ; 10-Rutine.

Quelle que soit la matrice utilisée, les molécules ont mieux répondu avec le dépôt directement sur les spots (pulvérisation ciblée) plutôt que par immersion (Fig. II-17). Globalement la quantité de matrice déposée est plus importante sur le spot quand le dépôt se fait en mode pulvérisation ciblée ce qui favorise l'ionisation des composés testés.

Compte tenu de l'ensemble des résultats obtenus quant au choix de la matrice et de son mode de dépôt, les conditions retenues pour la suite des analyses seront l'utilisation en mode d'ionisation négatif d'une matrice inorganique déposée en mode pulvérisation ciblée sur les spots. L'emploi de la matrice inorganique a été privilégié pour obtenir des spectres de masse avec moins d'ions interférents et ainsi simplifier l'identification des composés présents dans les spots.

Le couplage TLC/MALDI-Tof MS a ensuite été évalué avec l'élution d'un mélange de standards. Le type de dépôt par pulvérisation ciblée choisi entraîne *de facto* une analyse ciblée des composés puisque la matrice devra être déposée à la main sur chaque zone d'intérêt.

4.2. Mise au point du couplage TLC/MALDI-Tof MS à partir de l'analyse de standards

4.2.1. Le système TLC pour la séparation des flavonoïdes

Afin de séparer les flavonoïdes sur la plaque TLC, le système d'élution optimisé au laboratoire pour cette famille de composés a été choisi. Cette élution est réalisée avec le mélange de solvants suivant : acétate d'éthyle (AcOEt), méthanol (MeOH) et acide formique (HCOOH). Dans ces conditions chromatographiques sur support polaire, plus une molécule est polaire moins elle migrera durant la phase d'élution. Ainsi pour les différents dérivés d'une même génine, la valeur de *R_f* peut être corrélée directement au nombre de motifs glycosylés (Fig. II-18) : la rutine (9), qui est une quercétine diglycosylée, est peu éluee (*R_f* 0.12) tandis que la quercitrine (13) qui est une quercétine monoglycosylée, migre en milieu de plaque (*R_f* 0.50) et la quercétine (7), qui est l'aglycone (ou génine), migre vers le haut de la plaque (*R_f* 0.87), car moins polaire. Pour des flavonoïdes de structure différente, la valeur du *R_f* peut dépendre du nombre de fonctions polaires (hydroxyle, acide) de leur génine ainsi par exemple le kaempférol et l'isorhamnétine de structures proches de celle de la quercétine mais avec respectivement un OH de moins (kaempférol) ou un Me en plus (isorhamnétine) migrent avec des *R_f* supérieurs à celui de la quercétine (Annexe 5).

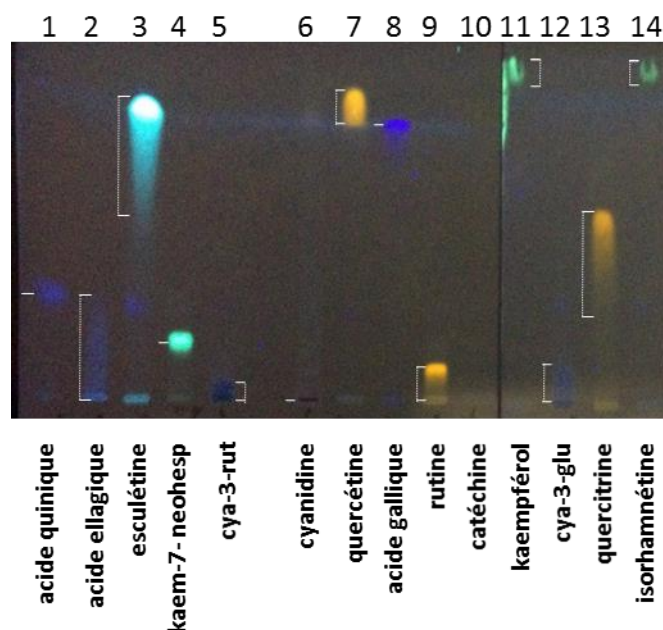


Figure II-18 : Elution en TLC de standards de différentes sous-familles de flavonoïdes. Phase stationnaire : plaque TLC aluminium silice 60 F₂₅₄ (20 x 7,5 cm). 10 µL de standards déposés à une concentration de 100 ppm. Solvant de migration : AcOEt-MeOH-HCOOH (90 :10 :1, v/v/v). Révélation NEU-PEG avec détection à 366 nm.

Une fois l'élution réalisée, une pulvérisation du réactif d' amino-2-éthyle diphénylborate dit NEU, puis du PEG est effectuée sur la plaque afin de visualiser les composés. Cette révélation permet de mettre en évidence sous UV à 366 nm les différentes couleurs de fluorescence des sous-familles de flavonoïdes (Fig. II-18). Ainsi, les dérivés de quercétine (quercétine (7), quercitrine (13) et rutine (9)) prennent une coloration jaune-orangé tandis que les dérivés de kaempférol et d'isorhamnétine (kaempférol (11), kaempférol 7-néohespéridine (4), isorhamnétine (14)) ont une couleur verte et les acides phénoliques (acide quinique (1), acide ellagique (2), acide gallique (8)) et les coumarines (esculétine (3)) sont bleus. En revanche, la molécule de catéchine (10) n'a pas pu être observée dans ces conditions de dérivation.

L'analyse globale de la plaque TLC met en lumière plusieurs limites à la détection visuelle par dérivation : certaines molécules diffusent le long de la plaque (comme la quercitrine et l'esculétine) et les couleurs des composés ayant des valeurs de R_f proches (tels que l'esculétine (bleu cyan), l'acide gallique (bleu foncé) et la quercétine (orange) au R_f voisin de 0.85), peuvent se retrouver masquées les unes par les autres lors de l'analyse d'un mélange ou d'un échantillon complexe. Le couplage TLC/MALDI-Tof MS peut alors permettre de mettre en évidence ces co-élutions.

4.2.2. Couplage du système TLC au MALDI-Tof-MS

Un mélange de dix standards (Cyanidine, cyanidine-3-glucoside, cyanidine-3-rutinoside, quercétine, quercitrine, rutine, kaempférol, kaempférol-7-néohespéridine, catéchine, isorhamnétine) a été

constitué afin d'avoir une représentativité en réponse colorée (bleu, orange, vert), en polarité (différentes valeurs de R_f), en valeurs de rapport m/z et en sous-familles moléculaires de flavonoïdes. Dans un premier temps le mélange standard a été déposé deux fois sur une même plaque TLC. Après élution de la plaque, une seule des deux pistes est révélée par les révélateurs NEU puis PEG pour mettre en évidence la position des flavonoïdes (Fig. II-19A) après migration. L'autre piste éluee simultanément mais non révélée, est utilisée pour réaliser les analyses par spectrométrie de masse (Fig. II-19B). La matrice inorganique a alors été déposée manuellement sur la piste 2 et sur chaque zone (R_f) où un flavonoïde avait été visualisé sur la piste 1 puis la plaque a été analysée par MALDI-ToF MS en mode négatif.

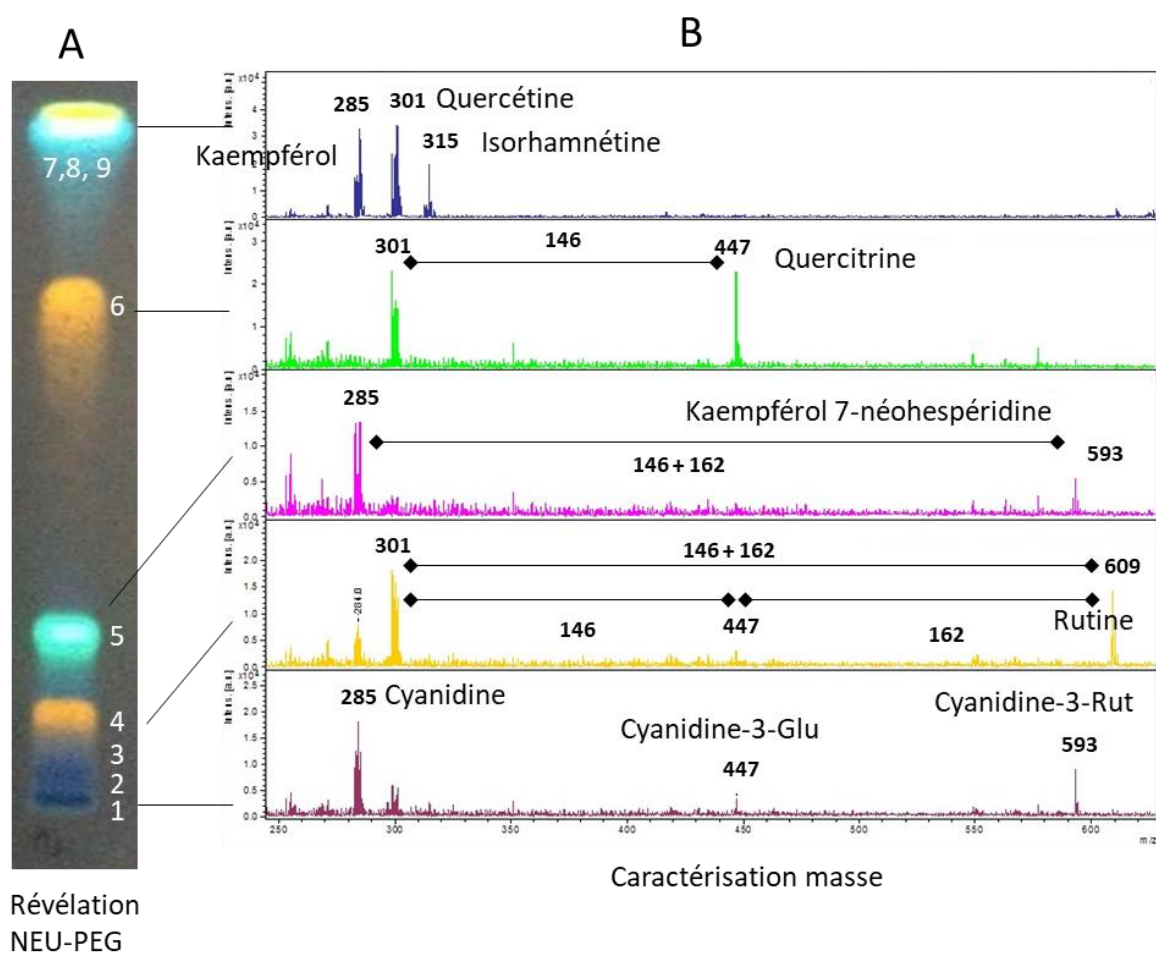


Figure II-19 : Analyse en TLC/MALDI-ToF MS d'un mélange de 10 standards de flavonoïdes. Dépôt de 10 μL d'un mélange des 10 standards à 50 ppm chacun. (Cyanidine (1) ; cyanidine-3-glucoside (2) ; cyanidine-3-rutinoside (3) ; rutine (4) ; kaempférol-7-néohespéridine (5) ; quercitrine (6) ; catéchine (7) ; quercétine (8) ; kaempférol (9) ; isorhamnétine (10)). Phase stationnaire : plaque TLC aluminium silice 60 F₂₅₄ (20 x 7,5 cm). Solvant de migration : AcOEt-MeOH-HCOOH (90 :10 :1, v/v/v). (A) Révélation NEU-PEG avec détection à 366 nm. (B) Spectres de masse obtenus en mode d'ionisation négatif des différents spots après élution et dépôt de la matrice inorganique.

Le couplage TLC/MALDI-Tof-MS permet ainsi d'associer à chaque spot révélé après dérivation, un spectre de masse (Fig. II-19B) dont l'interprétation est une aide pour identifier les composés élués. La fragmentation dans la source MALDI est semblable à celle observée dans une source ESI avec dissociation induite par collision (CID) (Silva and Lopes 2015). Pour les flavonoïdes *O*-glycosylés, la fragmentation consiste en la coupure de la liaison CO, qui donne des pertes de neutres caractéristiques du type de sucre. Ainsi, pour l'identification, la littérature existante peut aider à l'interprétation de la fragmentation (Cuyckens 2000). Par exemple, le spectre de masse en mode d'ionisation négatif du composé élué au *R_f* égal à 0,50 (Spot 6) est caractérisé par un ion dont la valeur du rapport *m/z* est égale à 447, ce qui permet l'identification de la molécule de quercitrine sous la forme de son ion moléculaire déprotoné. Dans ce spectre de masse, l'ion de rapport *m/z* 301, correspond à un fragment de la molécule formé en source et correspondant au motif de base de la génine, quercétine.

L'analyse du spectre de masse du spot élué en haut de plaque met en évidence trois ions de masses proches (*m/z* 285 ; 301 et 315) qui correspondent aux ions moléculaires déprotonés de trois molécules co-éluées non glycosylées, qui sont les molécules de kaempférol, de quercétine et d'isorhamnétine. Donc avec l'information des spectres de masse, des coélutions peuvent être mises en évidence, ce que la révélation chimique n'avait pas pu résoudre. Par ailleurs, dans tous les spectres de masse obtenus, sont observés des ions dont les rapports *m/z* ont des valeurs inférieures à une valeur de 150, ces ions correspondant à des ions générés par la matrice inorganique ne gênent pas l'analyse des flavonoïdes qui ont des masses supérieures à 210 Da.

Après avoir démontré que cette méthode de couplage TLC/MALDI-Tof MS avec une matrice inorganique en mode négatif est adaptée à l'analyse des flavonoïdes en mélange standard, la suite de notre travail a consisté à évaluer la transposition de ces conditions à l'analyse de divers produits issus d'une transformation agro-alimentaire, tels que le vin rouge et un jus de pomme enrichi en dérivés de quercétine ou d'extraits de fruits tels que les baies d'argousier et le mangoustan. Cette étude d'échantillons complexes permettra de mettre en évidence tout l'intérêt de ce type de couplage et ce qui peut être envisagé en cas de problème.

4.3. Application à l'analyse d'anthocyanes dans le vin rouge

Le vin rouge est connu pour être riche en molécules d'anthocyanes. Ces molécules sont naturellement colorées dans le visible et sont donc facilement observables sur une plaque TLC sans étape de dérivation. De plus, en milieu acide, les anthocyanes se présentent sous la forme d'ions oxonium. Ces molécules chargées positivement sont donc supposées être plus facilement détectables en spectrométrie de masse en mode d'ionisation positif. Le mode LDI (désorption/ionisation laser) a donc d'abord été évalué pour désorber et ioniser les anthocyanes directement à partir de la plaque TLC. Sans élution, le mode LDI devrait être d'autant plus efficace pour la détection des anthocyanes, que les molécules chargées pourraient elles-mêmes se comporter comme une matrice favorisant l'ionisation du mélange. Malheureusement, aucun signal MS n'a été détecté dans ces conditions. Aucune désorption de molécule n'a pu être obtenue directement et par conséquent aucune ionisation. Le spectromètre de masse MALDI-Tof MS est équipé d'un laser à azote émettant une longueur d'onde à 337 nm, or le spectre d'absorbance des anthocyanes montre une bande intense d'absorption vers

520 nm et une très faible absorption à 337 nm. On peut donc supposer que la désorption et l'ionisation des anthocyanes ne sont pas optimales dans les conditions d'utilisation de ce laser à azote. Par ailleurs, comme le vin a été déposé sur la plaque sans préparation préalable (ni filtration, ni pré-concentration), la quantité d'anthocyanes présents dans l'échantillon peut être trop faible pour être détectée directement dans ces conditions. En effet, si la détection de la molécule standard de cyanidine déposée à la concentration de 1000 ppm a pu être obtenue dans ces conditions de LDI, elle ne l'a pas été pour une concentration de 100 ppm. En teneur faible, la détection de la molécule de cyanidine nécessite l'utilisation d'une matrice facilitant son ionisation. Cumuler à la fois une faible teneur en composé et une faible possibilité d'absorption de l'énergie du laser, n'est donc pas favorable à une ionisation directe.

Pour la suite de nos travaux, l'utilisation d'une matrice appropriée pour faciliter l'ionisation des anthocyanes en mode positif s'est imposée. La matrice inorganique ayant montré d'une part, de bonnes performances sur différents standards de flavonoïdes en mode d'ionisation négatif et d'autre part, ayant été utilisée avec succès dans l'équipe pour ioniser les sucres en mode positif lors de leur analyse par TLC/MALDI-Tof MS (Ferey 2017), nous avons choisi d'évaluer ses performances en vue d'ioniser en mode positif les anthocyanes contenus dans un échantillon de vin rouge. L'analyse a été réalisée tout d'abord sans élution de la plaque TLC pour évaluer la compatibilité de la matrice et détecter d'éventuelles interférences (dus à la matrice ou à la plaque de silice) tout en s'affranchissant d'une diffusion possible de l'échantillon en lien avec son élution.

Dans la plage de rapports de masse m/z 300 à 700, plage adaptée pour l'analyse des anthocyanes, ni la matrice ni la plaque de silice ne produisent d'ions parasites. Dans ces conditions d'analyse (analyse directe du dépôt de vin sur la plaque TLC non élue), les molécules d'anthocyanes ont été facilement ionisées en mode positif sous forme d'ions $[M]^+$ avec une intensité de signal appropriée. L'empreinte de masses obtenue après l'ablation directe de la tache colorée du vin montre quelques pics d'intérêt (Fig. II-20). Les valeurs des rapports m/z des ions détectés sont en corrélation avec la présence de malvidine (m/z 331) ou de malvidine hexoside (m/z 493), communément décrite dans le vin rouge (Wang et Sporns 1999). Sept candidats dont quatre dérivés monoglycosylés pouvant appartenir à la famille des anthocyanes, ont été identifiés grâce à leur ion moléculaire $[M]^+$. Les ions de rapport m/z 301, 317 et 331 pourraient correspondre respectivement (i) à des composés de péonidine, de pétunidine et de malvidine (forme aglycone) ou (ii) à des fragments obtenus à partir de ces molécules glycosylées 5, 6, 7 et 8 par perte de 132 u (pentose), 146 u (désoxyhexose), 162 u (hexose) ou 176 u (acide glucuronique), ou (iii) à la présence des deux formes (molécules d'aglycone et fragments d'ions de molécules glycosylées). Cependant, sans séparation, il n'est pas possible d'attribuer une identification plus précise. De plus, en fonction de leur teneur plus ou moins faible dans l'échantillon, d'autres composés peuvent être présents dans l'échantillon mais non détectés.

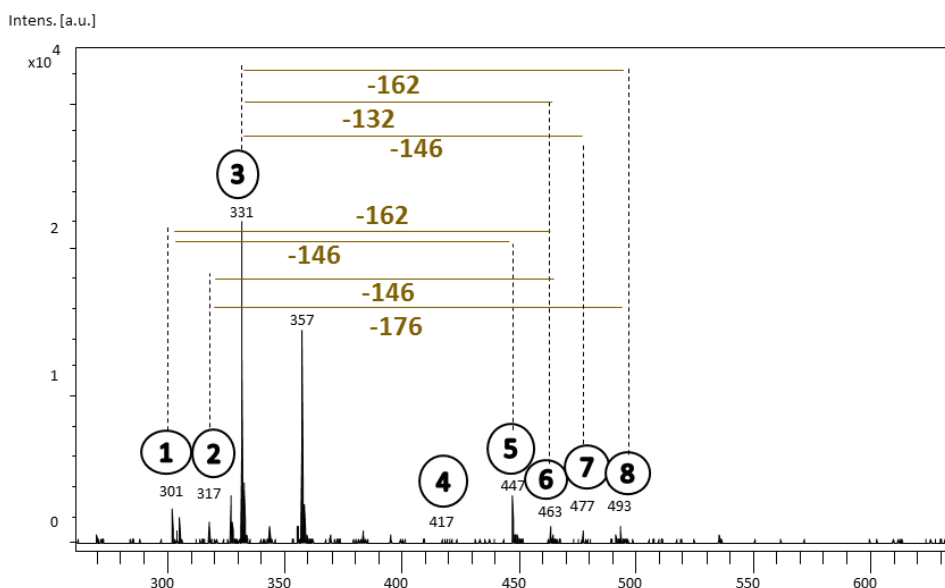


Figure II-20 : Analyse directe d'un échantillon de vin rouge par MALDI-TOF/MS en mode d'ionisation positif. Dépôt de 10 μL de vin rouge sur plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm) puis dépôt au-dessus de cette tache de 2 μL de matrice inorganique. Plaque TLC non éluee.

Cette première approche met en évidence le potentiel de l'analyse directe par MALDI-Tof MS de l'extrait brut sans élution. La présence d'anthocyanes est bien détectée dans le vin, mais sans séparation, la forme précise sous laquelle ces composés sont présents dans l'échantillon (forme glycosylée ou forme aglycone) ne peut pas être caractérisée formellement. L'étape suivante consiste à améliorer la méthode en ajoutant la contribution d'une séparation par TLC avant la détection par spectrométrie de masse afin d'obtenir une connaissance plus approfondie de la composition en anthocyanes de l'échantillon de vin.

Après la séparation sur TLC du dépôt de vin rouge (Fig. II-21A), plusieurs taches de couleur rose, caractéristiques des anthocyanes, sont visibles à la lumière blanche. Aux vues des valeurs de R_f de ces différentes taches et des spectres de masse qui leur sont associés, le couplage TLC/MALDI-Tof MS (Fig. II-21B) confirme la présence simultanée d'anthocyanes monoglycosylés (Spot 1 à 4) et de composés aglycones (Spot 5 et 6) dans l'échantillon de vin rouge et facilite l'identification de ces composés.

Dans les spectres de masse des spots 1 à 4, l'ion de rapport m/z le plus élevé (respectivement 493, 463, 477 et 447) est compatible avec la présence d'un dérivé glycosylé correspondant à une ou deux combinaisons entre une génine et un sucre (Tableau II-1).

Un ion de rapport m/z égal à 331 est présent dans tous les spectres de masse et correspond à l'ion de rapport m/z le plus élevé dans le spectre de masse du spot 6. Cet ion peut ainsi correspondre à l'ion $[M]^+$ associé à l'aglycone malvidine et dans ce cas, son observation dans les spectres de masse des spots 1,2 et 3 confirmerait la présence dans l'échantillon de vin de différents dérivés glycosylés de malvidine qui seraient élués selon leur polarité avec différents rapports frontaux. Néanmoins, pour les spots 4 et 5 cette attribution d'ion de rapport m/z égal à 331 n'est pas en cohérence avec la nature du dérivé glycosylé correspondant à l'ion parent du spectre. Pour le spot 4 par exemple, l'ion $[M]^+$ au

rapport m/z 447 n'est pas associé à une malvidine glycosylée, mais plus à une péonidine déoxyhexoside (Tableau II-1).

L'analyse du spectre de masse du spot 1 indique une co-migration de deux anthocyanes mono-glycosylés de masses différentes (493 et 463 u) avec la présence dans le spectre de deux ions fragments (m/z 331 et 317) associés respectivement aux deux génines correspondantes (Malvidine et Pétunidine) pour confirmer cette attribution.

Dans le cas du spectre de masse du spot 2, un seul ion (celui de rapport m/z égal à 463) peut être attribué à un composé mono-glycosylé mais deux ions $[M]^+$ de rapport m/z 317 et 331 sont présents et peuvent correspondre à deux génines d'anthocyanes (Malvidine et Petunidine). Il est donc impossible de dire dans ce cas, si le spot 2 contient une ou deux molécules glycosylées co-migrées de même masse (463 u).

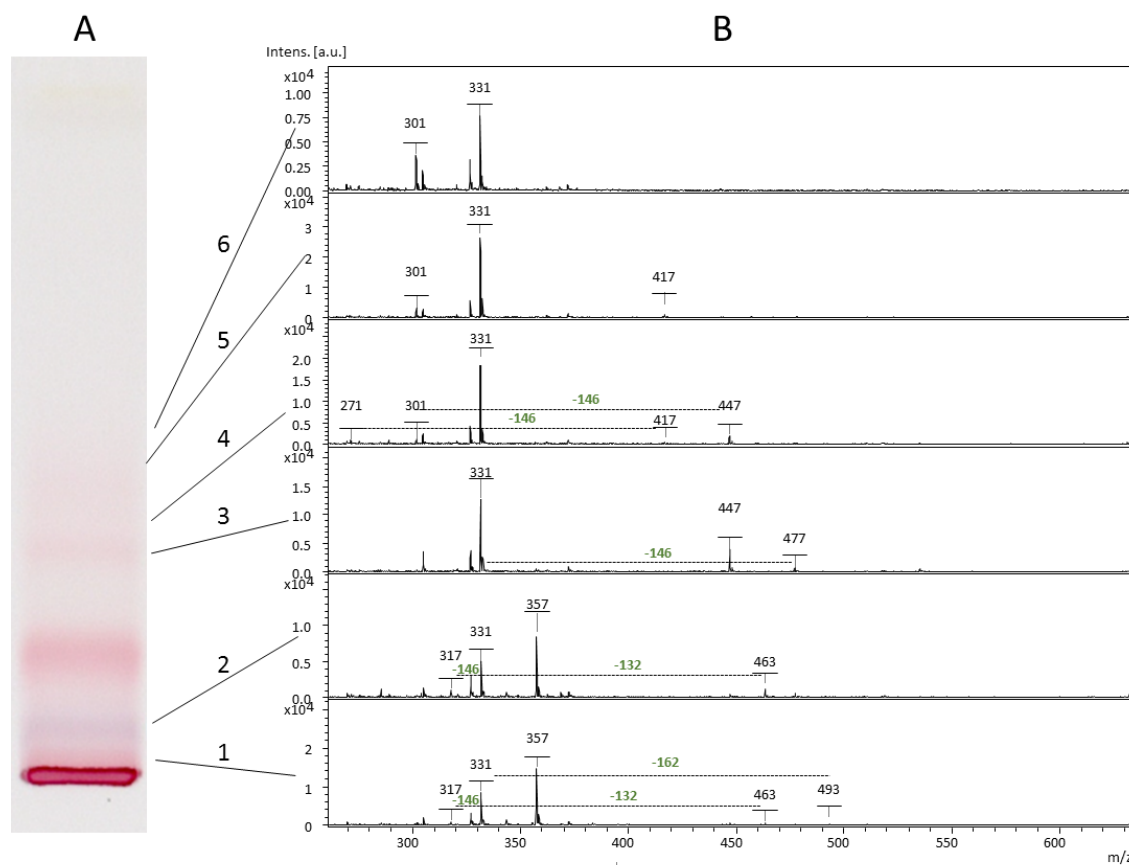


Figure II-21 : Analyse par TLC/MALDI-ToF MS d'un échantillon de vin rouge avec un volume de 2 μ L de matrice inorganique déposée sur chacun des spots. Phase stationnaire : plaque TLC en support aluminium silice 60 F₂₅₄ (5 x 7,5 cm). Solvant de migration : AcOEt/MeOH/HCOOH (90 : 10 : 1, v/v/v). A) Plaque TLC dans le visible. B) Spectres de masse de chaque spot élué en mode d'ionisation positif.

Tableau II-1 : Caractérisation par TLC/MALDI-Tof MS de la composition en anthocyanes du vin rouge déposé sur la plaque TLC après élution.

N° du spot	Valeur m/z [M] ⁺	Décomposition de masse	Proposition d'identification
6	331		Malvidine
	301		Péonidine
5	417		/
4	447	301 + 146	Péonidine déoxyhexose
	417	271 + 146	Pélagonidine déoxyhexose
3	477	331 + 146	Malvidine déoxyhexose
2	463	331 + 132	Malvidine pentose
		317 + 146	Pétunidine déoxyhexose
1	493	331 + 162	Malvidine hexose
		317 + 176	Pétunidine acide glucuronique
	463	331 + 132	Malvidine pentose
		317 + 146	Pétunidine déoxyhexose

Ces premiers résultats montrent le potentiel du couplage TLC/MALDI-Tof MS en mode d'ionisation positif pour l'identification de molécules appartenant à une même famille moléculaire ionisée positivement dans l'échantillon. La dimension séparation chromatographique permet de s'assurer de la présence simultanée dans l'échantillon des formes glycosylées et aglycones correspondantes. Cette analyse a mis en évidence néanmoins un premier cas d'indétermination de structure lié à l'observation d'un ion de même rapport m/z (331 u) dans l'ensemble des spectres de masse associés à l'ensemble des spots séparés. De même, l'ion de rapport m/z 357 (spots 1 et 2) n'a pas pu être attribué à une structure d'anthocyane.

Le couplage TLC/MALDI-Tof MS a permis d'ioniser en mode d'ionisation positif la famille moléculaire des anthocyanes. Par la suite, l'intérêt du couplage va être mis en évidence avec l'analyse d'un extrait de baies d'argousier en mode d'ionisation négatif.

4.4. Intérêt du couplage TLC/MALDI-Tof-MS pour caractériser les flavonoïdes présents dans un extrait de baies d'argousier

Un extrait éthanolique de baies d'argousier (*Hippophae rhamnoides*) a été analysé par MALDI-Tof MS avant et après séparation par TLC. Ces deux analyses permettent de comparer les spectres de masse obtenus et d'illustrer l'intérêt du couplage TLC/MALDI-Tof-MS en vue d'une identification du contenu en flavonoïdes de ce type d'extrait. Le choix d'étudier cet extrait s'est basé sur les connaissances que nous avons acquises au laboratoire, sur sa composition en composés phénoliques, lors de travaux antérieurs effectués par HPLC/MS (Michel 2011).

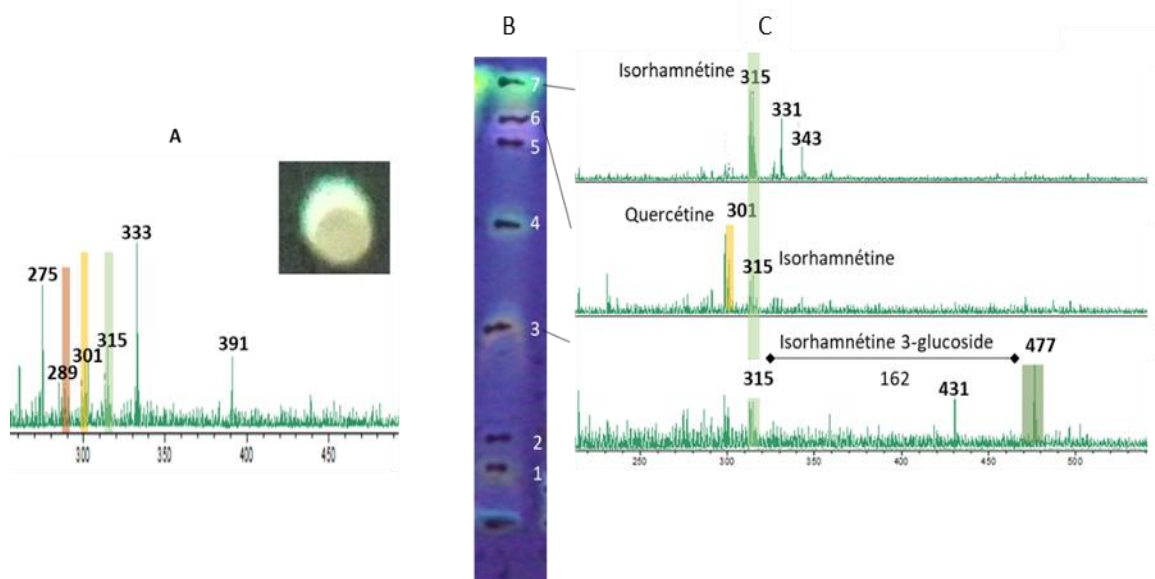


Figure II-22 : Analyse d'un extrait éthanolique de baies d'argousier (*Hippophae rhamnoides*) par A) MALDI-ToF MS en mode d'ionisation négatif et par B) TLC/MALDI-ToF MS sur plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm). Volume déposé 10 μ L. Solvant de migration : AcOEt/MeOH/AF (90/10/1 ; v/v/v). Révélation NEU-PEG à 366 nm. 2 μ L de matrice inorganique déposée sur chacun des spots. C) Spectres de masse associés à quelques spots de la plaque TLC.

Le spectre de masse obtenu par analyse directe en MALDI-ToF MS (Fig. II-22A) de l'extrait d'argousier indique la présence d'ions de rapport m/z n'allant pas au-dessus de 400 Da. Trois de ces ions (m/z 289 ; 301 et 315) peuvent correspondre aux ions $[M-H]^-$ de 3 génines de la famille des flavonoïdes (respectivement catéchine, quercétine et isorhamnétine), composés dont la présence a été précédemment décrite dans la littérature (Michel 2011). L'information obtenue à partir de cette analyse directe d'un mélange complexe reste néanmoins très approximative et très peu riche.

L'analyse de l'extrait après séparation TLC est un peu plus informative. Après la séparation sur une plaque de silice (Fig. II-22B), sept spots de différentes couleurs (vert, jaune orangé et jaune bleu) ont été révélés après l'application des réactifs NEU, puis PEG indiquant la présence de différents flavonoïdes qui au regard des valeurs de R_f observées peuvent être pour certains glycosylés et pour d'autres sous leur forme aglycone.

Ainsi les spots 5 et 7 situés en haut de la plaque ont des masses de faible rapport m/z (respectivement 301 et 315 pour le 5 et 315 ; 331 et 343 pour le spot 7). Deux de ces ions pourraient correspondre à des flavonoïdes aglycones de type quercétine (m/z 301) et isorhamnétine (m/z 315). Ces molécules étant sans fonction sucre, correspondent aux formes les moins polaires donc celles qui ont le moins d'affinité avec la phase stationnaire, elles migrent rapidement. Il est alors logique de les identifier dans les spots en haut de plaque.

Quant au spot 3, qui se situe en milieu de plaque, la masse de rapport m/z 315 est retrouvé en plus de l'ion de rapport massique m/z 477. Cette molécule pourrait correspondre à la molécule isorhamnétine 3-O-glucoside. Or cette masse n'était pas visible dans le spectre de masse de l'extrait brut initial (avant élution).

L'analyse directe par MALDI-Tof-MS de l'extrait brut d'argousier a apporté peu d'information. On peut évoquer à ce stade la possibilité d'un effet de compétition d'ionisation entre les nombreux composés présents ensemble dans un même spot, se traduisant par un défaut de détection de certains composés et donc à une perte d'information sur la composition de l'échantillon. La séparation de l'extrait par TLC avant son analyse en spectrométrie de masse peut permettre d'éviter ces phénomènes de compétition en diminuant le nombre de composés co-élus dans un spot donné et ainsi conduire à une meilleure représentativité de l'échantillon. Par contre cette analyse est moins riche en information que celle obtenue par Michel (2011) après fractionnement, mais ce couplage met en évidence rapidement les composés majoritaires.

Pour continuer dans l'évaluation du couplage TLC/MALDI-Tof-MS, un deuxième type d'échantillon complexe a fait l'objet d'étude. Ainsi, un jus de pomme riche en dérivés de molécule quercétine a été analysé en mode d'ionisation négative après séparation par TLC.

4.5. Application à l'analyse d'un jus de pomme enrichi en dérivés de la quercétine

La pomme est riche en quatre sous-groupes de flavonoïdes : dihydrochalcone (phloridzine), flavonol (quercétine), flavanol (catéchine) et des acides phénoliques (acide chlorogénique) (Francini et Sebastiani 2013). Un échantillon de jus de pomme enrichi en dérivés de quercétine à la fabrication a été sélectionné pour déterminer si le couplage TLC/MALDI-Tof MS était approprié pour identifier différents dérivés glycosylés basés sur une même génine (quercétine) et simultanément pour identifier certains composés en plus faible teneur et représentatifs d'autres sous-groupes de flavonoïdes.

Par ailleurs, étant donné que le jus de pomme a juste été resolubilisé, aucune préparation d'échantillon hormis une filtration n'a été effectuée avant son dépôt sur la plaque.

4.5.1. Caractérisation de dérivés de quercétine

Après dépôt de l'échantillon de jus de pomme sur une plaque de silice de type Silica gel F₂₅₄, la séparation des composés a été réalisée par élution de la plaque avec un mélange AcOEt :MeOH :HCOOH (90 :10 :1, v/v/v). La pulvérisation des réactifs NEU-PEG a permis la révélation de nombreuses bandes jaune orangé caractéristiques de la présence de dérivés de quercétine dans l'échantillon de jus de pomme (Fig. II-23A). L'étude s'est donc dans un premier temps concentrée sur les dérivés de quercétine.

Comme observé lors de l'analyse des molécules standard, les dérivés de quercétine sont séparés dans ces conditions chromatographiques en fonction de la nature de leur partie glycosylée (quercétine Spot 8, R_f 0.87), quercitrine (quercétine monoglycosylée Spot 6, R_f 0.50) et rutine (quercétine di-glycosylée Spot 2, R_f 0.12). La révélation chimique montre la présence de dérivés de quercétine tout au long de la plaque situés entre la rutine (Spot 2) et la quercétine (Spot 8) suggérant la présence de différents composés monoglycosylés (Fig. II-23A) dans l'échantillon de jus.

Un couplage TLC/MALDI-ToF-MS a ensuite été mis en œuvre afin d'associer à chaque spot révélé après dérivation, un spectre de masse. Le mode d'ionisation négatif a été sélectionné pour l'analyse car ce mode permet une meilleure ionisation des dérivés de quercétine que le mode positif et surtout permet une détection sans adduits. L'analyse MS de chaque spot confirme la présence de divers dérivés de quercétine dans le jus de pomme. L'ion de rapport m/z 301, obtenu par fragmentation en source, correspondant à l'ion de la molécule déprotonée $[M-H]^-$ de la quercétine (forme aglycone) est en effet présent dans tous les spectres de masse associés aux taches jaunes (Fig. II-23B).

Les dérivés mono-glycosylés de quercétine sont séparés avec ce système d'élution de TLC selon la nature et le nombre de sucres liés à leur partie aglycone (Fig. II-23C). Comme un fragment hexose est plus hydroxylé qu'un fragment désoxyhexose ou pentose, les dérivés hexose sont davantage retenus sur la plaque de silice. La structure du substituant glycosylé est déterminée à partir de la polarité de la molécule et également à partir des ions de rapport m/z observés sur le spectre de masse. Comme vu dans le chapitre 1, une perte de 132 u correspond à la perte d'un motif pentose (taches 5 et 7), celle de 146 u indique une perte de désoxyhexose (spot 6) alors que la différence de 162 u désigne une perte d'hexose (spot 4) pour les quercétines monoglycosylées.

En spot 2, la rutine étant diglycosylée, sa masse moléculaire est supérieure à celle des composés monoglycosylés et dans son spectre de masse la présence de l'ion fragment (m/z 301) correspondant à la génine quercétine est bien observée. La différence en masse entre l'ion fragment et l'ion moléculaire correspond à un motif rutinose.

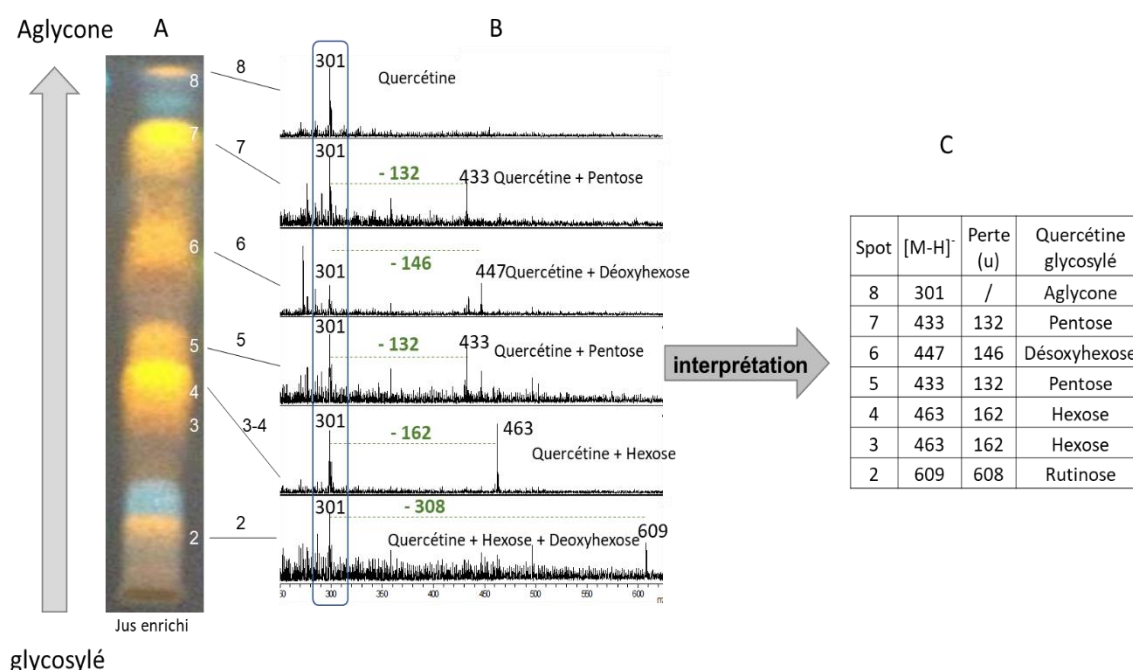


Figure II-23 : Analyse en TLC/MALDI-ToF MS en mode d'ionisation négatif d'un jus de pomme enrichi en dérivés de quercétine avec 2 μ L de matrice inorganique déposé sur chaque spot. (A) Plaquette TLC après dérivation. Phase stationnaire : plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm). Volume déposé 10 μ L. Solvant de migration : AcOEt/MeOH/HCOOH (90 :10 :1, v/v/v). Système de révélation : NEU-PEG à 366 nm. (B) Spectres de masse en MALDI-ToF MS en mode d'ionisation négatif associé à chaque spot, obtenus à partir d'une plaque TLC non dérivée. (C) Identification des dérivés de quercétine.

4.5.2. Détermination de la position du sucre sur la quercétine et cas d'isoméries

Une molécule quercétine mono-*O*-glycosylée est majoritairement glycosylée sur les positions 3 ou 7 ou parfois 4' de la quercétine. Donc pour une génine donnée, une même masse peut caractériser différents mono-*O*-glycosylés isomères de position.

En TLC, la séparation des isomères est possible (spots 5 et 7). Mais si aucun standard n'a été positionné en parallèle de l'extrait, il est difficile de caractériser formellement l'isomère associé à chacun des spots. Comme indiqué dans le chapitre 1 (6.3), l'interprétation des spectres de masse peut permettre néanmoins de donner des indications sur la position du sucre sur la génine. Ainsi lors de la coupure de la liaison entre le sucre et la génine, il peut y avoir un clivage hétérolytique conduisant à un fragment de type $[Y_0-H]^-$ ou un clivage homolytique correspondant à un fragment de type Y_0^- . Le rapport d'intensité de ces deux ions permet alors de distinguer pour une génine et un sucre donnés, la position 3 de la position 7-*O*-glycosylée.

Ainsi, pour le spot 6 identifié comme correspondant à une quercétine déoxyhexose, l'ion $[Y_0-H]^-$ est bien de plus petite intensité que l'ion Y_0^- , comme dans l'exemple donné dans la figure I-17 dans le chapitre 1, ce qui correspondrait à une quercétine 3-*O*-rhamnose. Pour le spot 7 associé à une quercétine *O*-pentose, le rapport d'intensité des ions $[Y_0-H]^-$ et Y_0^- est inversé par rapport au spot 6, mais n'ayant pas d'illustration donnée dans la littérature avec des pentoses, il est difficile d'apporter une conclusion sur la position du sucre pour ce composé.

L'analyse TLC a mis en évidence la présence de deux isomères de position, correspondant à deux quercétine *O*-pentose (taches 5 et 7, Fig. II-23), ce qui n'aurait pas pu être détecté sans la dimension séparative de la TLC. N'ayant pas déposé de standard de quercétine *O*-pentoside, dans ce cas, l'information sur le rapport des intensités ne peut pas lever l'indétermination sur les deux isomères. En l'absence de standard, le couplage TLC/MALDI-Tof MS ne peut donner d'informations plus précises quant à la structure du pentose et sa position sur la génine.

De plus, en zoomant sur les masses de la génine 301, les ions $[Y_0-H]^-$ et Y_0^- des spots 4 et 5 sont sensiblement de même intensité, donc il est difficile de conclure sur la position respective des sucres (respectivement hexose et pentose).

Cette première analyse de résultats confirme que le couplage TLC/MALDI-Tof MS permet de distinguer les différents dérivés de quercétine, composés majoritaires dans l'échantillon de jus de pomme enrichi. Un examen plus approfondi des spectres de masse s'avère aussi intéressant pour évaluer si la présence d'autres composés en teneur plus faible et correspondant à une génine différente de la quercétine pourrait être décelée.

4.5.3. Détection de composés minoritaires

Pour aller plus loin dans la connaissance de la composition chimique de l'échantillon de jus de pomme enrichi nous avons cherché à repérer dans les spectres de masse associés à chacun des spots obtenus

lors de la séparation TLC (Fig. II-23) des indices de présence d'autres flavonoïdes que les dérivés de quercétine. La figure II-24 illustre l'abondance relative de tous les ions détectés dans chaque spectre de masse. Ce schéma a été établi en repérant sur un axe vertical chaque spot selon la valeur de son rapport frontal et en reportant pour chaque spot sur un axe horizontal, les valeurs du rapport m/z de tous les ions présents. Cette figure met en évidence les composés mineurs présents dans le spot 6 mais non visibles par la couleur du spot révélé (jaune orangé), car présents en trop faible quantité.

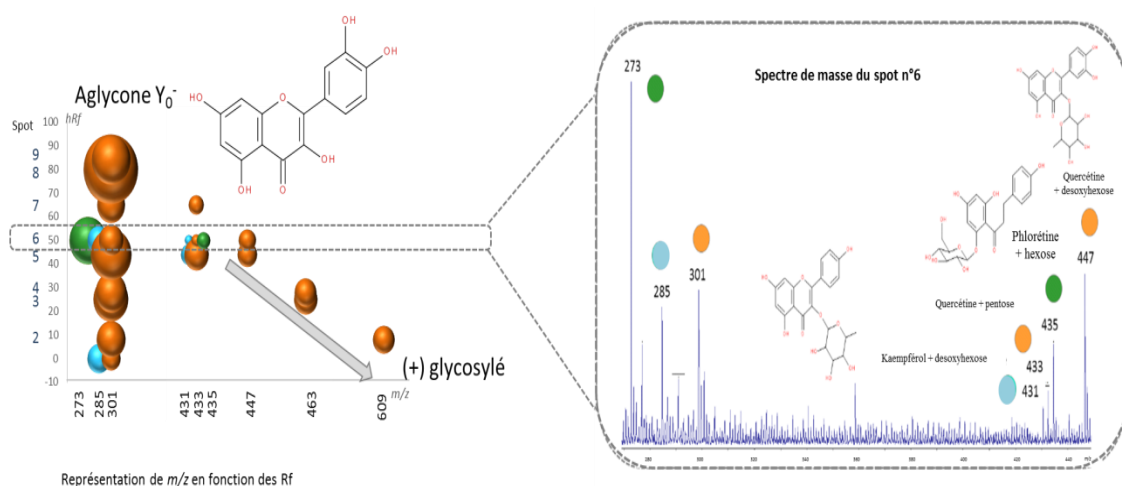


Figure II-24 : Analyse TLC/MALDI-ToF-MS du jus de pomme enrichi en dérivés de quercétine. Représentation des valeurs de rapport m/z en fonction des zones ciblées et des rapports frontaux R_f . La taille des bulles correspond à l'abondance des ions. La couleur des bulles est associée à la structure aglycone des composés détectés et associée à la couleur des taches observées après dérivation au NEU-PEG (orange pour les dérivés de quercétine, vert pour les dérivés de phlorétine et bleu pour les dérivés de kaempférol).

Bien que la tache 6 de couleur jaune soit principalement associée à un dérivé de quercétine (ici la quercitrine), la lecture approfondie du spectre de masse met en évidence la présence en plus de l'ion $[M-H]^-$ spécifique à la quercitrine au rapport m/z 447, des ions de masses proches mais d'intensité plus faible pouvant être associés à d'autres génines glycosylées. Le spectre de masse de cette tache 6 indique la présence potentielle de trois génines différentes : phlorétine (m/z 273), kaempférol (m/z 285) et quercétine (m/z 301). A partir des ions spécifiques de ces 3 génines, il est alors possible de mettre en correspondance 3 autres ions observés de rapport m/z 431, 433, 435 avec une structure possible de flavonoïde *O*-glycosylé. Ainsi l'ion m/z 431 est compatible avec une structure de quercétine *O*-pentoside, l'ion m/z 433 avec la structure kaempférol *O*-désoxyhexoside et l'ion m/z 435 avec la molécule de phlorétine *O*-hexoside. La présence de ces types de composés et de génines concorde avec les résultats disponibles dans la littérature (Reis 2012, Grigoros 2013). L'ion au rapport m/z 435 peut correspondre de plus à la molécule phloridzine, molécule très présente dans les échantillons de pomme (Wojdyło 2008).

Ce couplage TLC/MALDI-Tof MS a permis ainsi de mettre en évidence les différents dérivés de quercétine contenus dans le jus de pomme et la présence d'autres composés d'autres familles de flavonoïdes, ainsi que les composés minoritaires et les molécules isomères de position.

Avec ce modèle de jus de pomme enrichi, il a été possible d'évaluer la puissance du couplage TLC/MALDI-Tof MS dans le cadre d'analyse ciblée en phytochimie. La dimension apportée par la séparation chromatographique a confirmé la richesse de l'échantillon en dérivés de quercétine grâce à la lecture en fluorescence. Une empreinte spectrale informative a pu être associée à chaque spot grâce au couplage avec le MALDI-Tof MS. Ainsi, pour chaque zone d'intérêt, il a été possible de proposer une structure pour les différents dérivés de la quercétine. De plus, l'interprétation fine des spectres de masse a mis en évidence d'autres composés bien connus, présents en plus faibles quantités que les molécules de dérivés de quercétine.

La suite des travaux présentés a permis d'élargir le champ d'applications de ce type de couplage à celui de l'analyse d'une autre famille moléculaire proche de celle des flavonoïdes mais moins polaires à savoir les xanthones prénylés.

4.6. Transposition de la méthode TLC/MALDI-Tof MS à l'analyse d'une famille moléculaire de structure proche de celle des flavonoïdes : les xanthones prénylés

Compte tenu des résultats prometteurs obtenus par TLC/MALDI-TOF MS pour la caractérisation des flavonoïdes, un extrait éthanolique de mangoustan, composé de nombreuses xanthones prénylées (Annexe 5) avec des masses très proches (2 Da de différence) a été analysé dans ces conditions. Cet échantillon a été choisi car sa composition moléculaire nous est connue grâce à des travaux antérieurs de notre laboratoire en LC/MS (Destandau 2015). Les xanthones prénylés sont structurellement proches des flavonoïdes cependant, la substitution prényle diminue la polarité des molécules et peut modifier le comportement de l'ionisation et de la fragmentation en source des molécules.

4.6.1 Analyse directe de l'extrait par MALDI-Tof MS avant l'élution

Afin d'évaluer la complexité du travail sur cette famille de composés, la faisabilité d'identifier ces molécules par TLC/MALDI-Tof MS a été évaluée d'abord sans étape d'élution, puis avec élution afin de réaliser une identification plus spécifique des différentes molécules. L'analyse directe par MALDI-Tof MS en mode d'ionisation négatif a été réalisée sur l'extrait déposé sur la plaque de TLC et non élué. Il a été observé que la matrice inorganique était moins efficace pour ioniser les constituants de cet extrait que pour ioniser les flavonoïdes des échantillons précédents. C'est pourquoi la matrice inorganique a été remplacée par une matrice organique, la DHB. Le dépôt de cette matrice a été réalisé par immersion, car la réponse MS était plus sensible.

Cette première analyse directe (Fig. II-25) a clairement mis en évidence la présence dans le spectre de masse de dix ions $[M-H]^-$ pouvant correspondre à des molécules appartenant à la famille de xanthones prénylés telles que le garcinone C (m/z 413), le garcinone D (m/z 427), ainsi que les α , β et γ mangostines (respectivement m/z 409 ; 423 et 395), déjà identifiés dans l'extrait de *Garcinia*

mangostana (tableau II-2). Mais sans élution, une masse de 395 u peut correspondre aussi bien à l' α -mangostine, qu'au BR-xanthone A ou au smeathxanthone A. De plus une molécule possédant une différence de 2 u en masse peut être confondue avec le profil isotopique de la molécule de masse M-2 u.

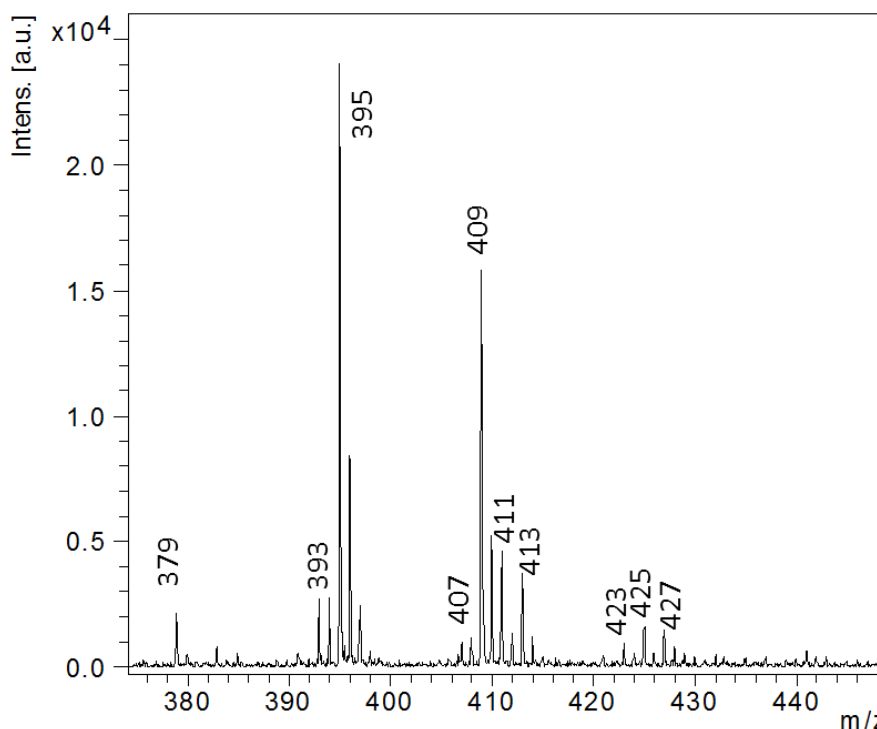


Figure II-25 : Spectre MALDI-Tof MS sans élution de l'extrait de *Garcinia mangostana* en mode d'ionisation négatif. 10 μ L de l'extrait déposé sur une plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm) non élué. Dépôt de la matrice DHB à 200 mg.mL⁻¹ par immersion.

4.6.2 Analyse de l'extrait par couplage TLC/MALDI-Tof-MS

La séparation TLC de l'extrait de mangoustan a été réalisée dans des conditions chromatographiques que nous avons ajustées par rapport à celles du système précédemment utilisé pour analyser les flavonoïdes du fait de la polarité moindre des xanthones prénylés. Le système d'élution choisi est de composition CHCl₃/n-heptane/MeOH avec des proportions 70/20/10 (v/v/v). L'élution et la révélation après dérivation avec les réactifs NEU-PEG ont montré lors d'une lecture à 366 nm qu'il y avait beaucoup plus de bandes d'élution (15) que de nombre d'ions (10) détectés lors de l'analyse directe par MALDI-Tof MS (Fig. II-26A) du même échantillon.

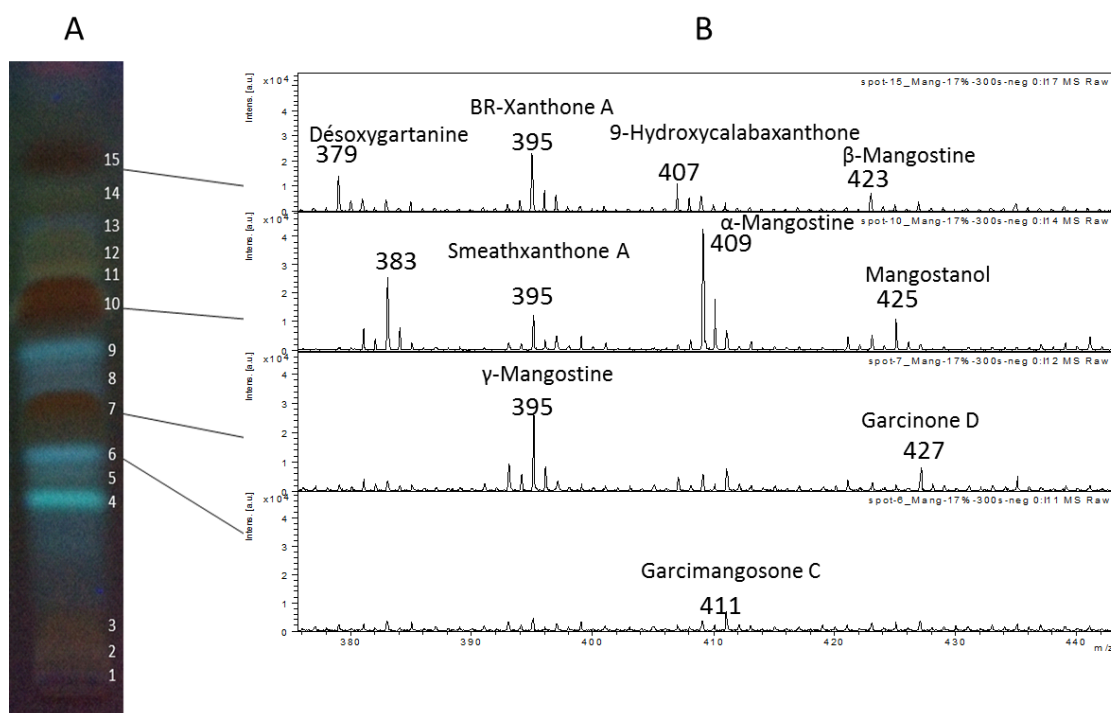


Figure II-26 : Analyse par TLC/MALDI-ToF MS en mode d'ionisation négatif de l'extrait de *Garcinia mangostana* (Annexe 1) avec la matrice DHB déposée par immersion à 200 mg.mL⁻¹. Phase stationnaire : plaque TLC aluminium silice 60 F₂₅₄ (5 x 7,5 cm). Volume déposé de 10 µL. Système de migration : CHCl₃/n-heptane/MeOH (70 :20 :10, v/v/v). Révélation : NEU-PEG à 366 nm (A) Plaque TLC révélée après élution et dérivation. (B) Spectres de masse associés à plusieurs spots.

Tableau II-2 : Analyse TLC/MALDI-ToF MS d'un extrait de *Garcinia mangostana*. Attribution des masses détectées en se basant sur la publication de Destandau (2015).

Nom	[M-H] ⁻	Formule brute
Désoxygartanine	379	C ₂₃ H ₂₄ O ₅
Garcinone B	393	C ₂₃ H ₂₂ O ₆
γ-Mangostine	395	C ₂₃ H ₂₄ O ₆
Smeathxanthone A	395	C ₂₃ H ₂₄ O ₆
BR-xanthone A	395	C ₂₃ H ₂₄ O ₆
9-Hydroxycalabaxanthone	407	C ₂₄ H ₂₄ O ₆
α-Mangostine	409	C ₂₄ H ₂₆ O ₆
Garcimangosone C	411	C ₂₃ H ₂₄ O ₇
Garcinone C	413	C ₂₃ H ₂₆ O ₇
β-Mangostine	423	C ₂₅ H ₂₈ O ₆
Mangostanol	425	C ₂₄ H ₂₆ O ₇
Garcinone D	427	C ₂₄ H ₂₈ O ₇

La dimension séparative permet donc de révéler la présence dans l'extrait de plus de composés que ne le laissait présager l'analyse directe par spectrométrie de masse. Cette différence peut être due à la présence de différents isomères dans l'extrait de mangoustan et/ou à une suppression éventuelle du signal de certains composés minoritaires entraînant leur absence de détection lors de l'analyse directe par MALDI-Tof MS. L'élution par TLC a conduit à la séparation de trois composés caractérisés par la même valeur de rapport m/z (395 u) pour leur ion moléculaire déprotoné (Fig. II-26B, spots 7, 10 et 15) et pouvant correspondre aux molécules smeathxanthone A, BR-xanthone A ou γ -mangostine. Ces structures moléculaires ont été proposées pour les différents spots en fonction de l'association d'informations différentes, telles que l'empreinte spectrale, la polarité par rapport à la valeur de R_f et des études bibliographiques. L'ion de rapport m/z 383 observé dans le spectre du spot 10 (Fig. II-26B) n'a pas été détecté dans cet extrait lors de l'analyse sans élution et n'avait pas été observé lors des études précédentes en LC/MS. Certains composés restent toujours co-migrés comme le révèle le spectre de masse du spot 15. Ce spectre présente en effet trois ions de masses correspondant aux molécules smeathxanthone A (m/z 395), 9-Hydroxycalabaxanthone (m/z 407) et β -mangostine (m/z 423). Les trois masses proches (423, 425 et 427 Da) obtenues en MALDI-Tof MS sont distinctes et bien attribuées à une bande chacune (bandes 15, 10 et 7 respectivement) avec une intensité élevée.

Bien que l'élution ait minimisé les risques d'extinction du signal, les composés associés à certaines bandes très visibles à 366 nm comme la bande 4 n'ont pas pu être ionisés et identifiés. Plusieurs possibilités peuvent être envisagées à ce problème d'ionisation et être illustrées. En effet, les spots 6 et 9 semblent avoir une couleur identique et de même intensité, et le spot 6 donne une information en masse, contrairement au 9. Ce composé non ionisé peut être dû à un ratio matrice/analyte qui ne convient pas ou à un mauvais mode d'ionisation. De même pour les spots 3, 7, 10 et 15, qui sont eux aussi de même couleur, sans forcément être de même intensité, ils n'ont pas la même polarité car positionnés à des R_f différents les uns des autres, et le spot 3 ne donne aucune information massique, contrairement aux autres. Ce manque d'ionisation peut dans ce cas-là être dû à une quantité de molécules présente trop faible pour être détectable en masse, mais suffisante pour être visualisée par fluorescence, spectrométrie connue pour sa grande sensibilité de détection ; à un ratio matrice/analyte qui ne convient pas, ou au mauvais mode d'ionisation. Certains ions observés avant l'élution (m/z 413) ne sont plus retrouvés après l'élution. Ce problème peut être dû à la matrice qui ne convient pas pour ioniser cette molécule, qui n'est pas forcément de la même famille moléculaire que les autres composés. Et que c'est la présence des autres composés qui a permis de l'ioniser avant élution.

Pour plus d'informations sur la caractérisation phytochimique de cet extrait de mangoustan, une combinaison de plusieurs approches analytiques complémentaires est requise. L'analyse MALDI-Tof MS sans élution donne un aperçu des masses des composés présents dans l'extrait et de la matrice à utiliser, l'élution avec la révélation NEU-PEG montre la répartition des molécules en fonction de leur polarité, et le couplage TLC/MALDI-Tof MS fournit des informations sur l'identification des composés, tout en permettant de distinguer plusieurs molécules différentes ayant la même masse. Des études complémentaires sur l'utilisation de la matrice inorganique en mode d'immersion ou en la greffant directement sur la silice de la plaque, pourraient permettre d'améliorer l'ionisation des composés et seraient à réaliser pour élargir le panel de familles moléculaires potentiellement caractérisables par

couplage TLC/MALDI-Tof MS. Il serait alors possible de proposer une nouvelle méthode générique à la disposition des phytochimistes.

Le couplage TLC/MALDI-Tof MS est un couplage efficace pour l'analyse des flavonoïdes et qui permet de croiser différentes informations (polarité, famille moléculaire, masse moléculaire). La taille du faisceau laser permet d'avoir une bonne résolution spatiale et il pourrait être envisagé de réaliser des analyses non ciblées en automatique.

Ces résultats montrent que la méthode d'analyse ciblée avec la matrice inorganique est appropriée pour l'analyse rapide par TLC/MALDI-Tof MS en mode d'ionisation négatif des familles moléculaires de flavonoïdes. Ces analyses mettent en évidence qu'il reste encore à résoudre certains problèmes d'ionisation des composés dans les cas suivants : un rapport entre la quantité de matrice déposée et la quantité d'analyte non appropriée pour ioniser le composé, une teneur trop faible de l'analyte ou le choix non approprié du mode d'ionisation.

Ainsi dans le cas de problème d'ionisation, il peut être envisagé une approche complémentaire en passant par un autre système de couplage, mettant en jeu d'autres modes d'ionisation, comme le couplage à pression atmosphérique TLC/API MS avec une source électrospray ou à ionisation chimique, suivant la polarité des composés.

5. Etude exploratoire du couplage TLC/API MS

Le couplage TLC/API MS peut apporter une complémentarité au couplage TLC/MALDI-Tof MS. En effet, avec le couplage TLC/API MS, l'utilisation d'une matrice pour assister l'ionisation n'a plus lieu d'être ce qui supprime alors les problèmes éventuels d'interférences avec les ions de la matrice dans le cas d'analyse de molécules de faible masse moléculaire (< 200 Da). De plus, selon les molécules à étudier, deux sources principales d'ionisation à pression atmosphérique peuvent être envisagées, une source électrospray qui permet d'ioniser une majorité de composés et une source APCI qui permet d'ioniser des composés moins polaires.

Le couplage TLC/API MS, comme le précédent, demande à réaliser un certain nombre d'optimisations. Après avoir sélectionné la source d'ionisation la mieux adaptée aux composés d'intérêt, il est en effet nécessaire d'optimiser les paramètres de tensions et de températures de la source ainsi que tester les modes d'ionisation (positif ou négatif). De plus, les solutés restant après la séparation TLC adsorbés sur la plaque de silice, le choix d'un solvant approprié pour solubiliser et désorber la tache de la plaque ainsi que pour transporter les composés solubilisés jusqu'à la source d'ionisation tout en favorisant leur ionisation s'impose et doit être optimisé.

5.1. Réponse en ESI-MS de molécules de masse inférieure à 200 Da déposées sur gel de silice

Le système TLC/API MS que nous avons testé implique l'utilisation d'une interface placée entre la plaque TLC et la source du spectromètre de masse. Cette interface de la société CAMAG permet de désorber les analytes de la plaque de silice, via un solvant adéquat apporté à la surface de la plaque par une pompe HPLC et de les introduire ensuite dans la source d'ionisation à pression atmosphérique (ESI ou APCI). L'analyseur du spectromètre de masse est un triple quadripôle fonctionnant en mode balayage sur le premier quadripôle.

Les premières molécules que nous avons analysées en couplage TLC/ESI MS correspondent à quatre petits acides organiques (acide citrique, acide tartrique, acide succinique, acide benzoïque). Il est en effet bien établi que ces molécules de par leur caractère acide sont facilement ionisables en mode négatif dans une source de type ESI (Flores 2012). Ces molécules nous ont donc servi de molécules modèles pour nous familiariser avec l'appareillage et évaluer la faisabilité du couplage (Fig. II-27).

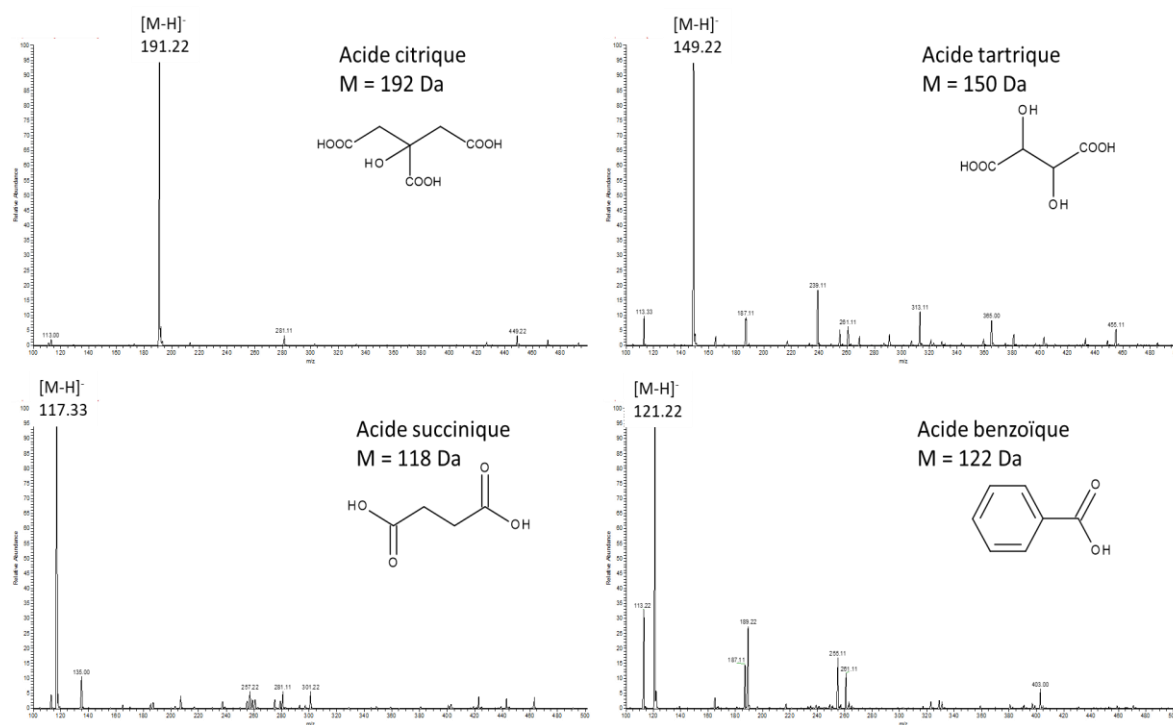


Figure II-27 : Spectres de masse de quatre acides organiques analysés en TLC/ESI TQ-MS en mode d'ionisation négatif après élution avec une phase éluante $\text{NH}_4\text{OH}/\text{EtOH}/\text{H}_2\text{O}$ 12,5 : 75,5 : 12. Dépôt de 5 μL des 4 solutions standards à 10 000 $\text{mg}\cdot\text{L}^{-1}$. Plaque TLC silice 60 F_{254} (20 x 10 cm), Solvant de désorption : $\text{ACN}/\text{H}_2\text{O}$ +0.1% AF (50/50) à 0.2mL/min.

Les premiers essais d'analyse en TLC/ESI TQ-MS ont été effectués directement après le dépôt de ces 4 standards sur la plaque TLC en silice à différentes concentrations (10 000 et 1 000 $\text{mg}\cdot\text{L}^{-1}$) (sans élution)

afin d'optimiser les conditions expérimentales de désorption et de transport des molécules vers l'entrée de la source ainsi que les réglages intrinsèques du spectromètre de masse.

La figure II-27 permet de mettre en évidence la détection des 4 acides organiques sous forme déprotonée $[M-H]^-$, comme attendu pour ce type de molécules portant une fonction acide. Bien qu'il y ait d'autres ions interférents, comme ceux dus à la plaque de silice, les ions moléculaires déprotonés des composés restent majoritaires dans les spectres de masse avec très peu de fragmentation observée pour ces ions.

Pour obtenir une sensibilité de détection suffisante il a fallu cependant déposer sur la plaque une forte quantité de ces acides organiques (5 μ L d'une solution à 10 000 mg.L^{-1}). Cette limite rencontrée peut être liée au choix du solvant de désorption et de transport des composés insuffisamment optimisé pour favoriser la désorption et/ou l'ionisation. L'ajout dans le solvant de désorption d'un additif tel qu'une base pourrait être testé pour aider à la déprotonation des molécules dans la source.

Cette première étape de prise en main de l'appareillage étant concluante, nous avons testé ensuite d'autres types de molécules moins polaires.

5.2. Analyse de molécules peu polaires : les oxystérols et les stéroïdes

Les oxystérols et les stéroïdes sont des composés moyennement à peu polaires très importants pour la détection de différentes maladies et sont utilisés pour réaliser des diagnostics thérapeutiques. Habituellement, ces composés sont analysés par chromatographie en phase liquide (HPLC), ou par chromatographie en phase gazeuse couplée à la spectrométrie de masse. Une analyse par chromatographie sur couche mince (Makin 2010) est également possible mais cette dernière technique n'a pas encore été couplée à la spectrométrie de masse. Par ailleurs, la faisabilité de l'analyse par MALDI-Tof MS de différents stéroïdes a déjà été décrite dans la littérature (Son and Cha 2014).

Par conséquent, il nous a paru intéressant d'évaluer les conditions d'analyse de ces composés peu polaires, par couplage TLC/MS avec les trois types de source d'ionisation à notre disposition: MALDI, ESI et APCI et selon les 2 modes d'ionisation possibles (positif et négatif). Les conditions expérimentales ont tout d'abord été optimisées sur la base d'une analyse directe du dépôt de ces composés sur la plaque TLC sans élution chromatographique préalable, puis la séparation HPTLC d'un mélange de 23 standards a été réalisée conformément aux conditions couramment utilisées dans le laboratoire de Tours Nutrition, Croissance et Cancer (N2C) pour le criblage de stérols et de stéroïdes dans des tissus humains enfin, l'analyse TLC/MS a été effectuée.

La structure de l'ensemble des composés standards de stérols et de stéroïdes que nous avons étudiés est donnée en Annexe 5 (tableau II-3).

5.2.1. Evaluation des trois sources d'ionisation

Les premiers essais ont porté sur les conditions d'ionisation des oxystérols. La figure II-28 illustre les résultats obtenus à partir de l'un de ces composés, le 24-OH-Cholestérol. Tous ces composés ont été

déposés sur plaque de silice, non élués, et analysés avec les deux types de couplage, les sources MALDI, ESI et APCI en modes d'ionisation négatif et positif.

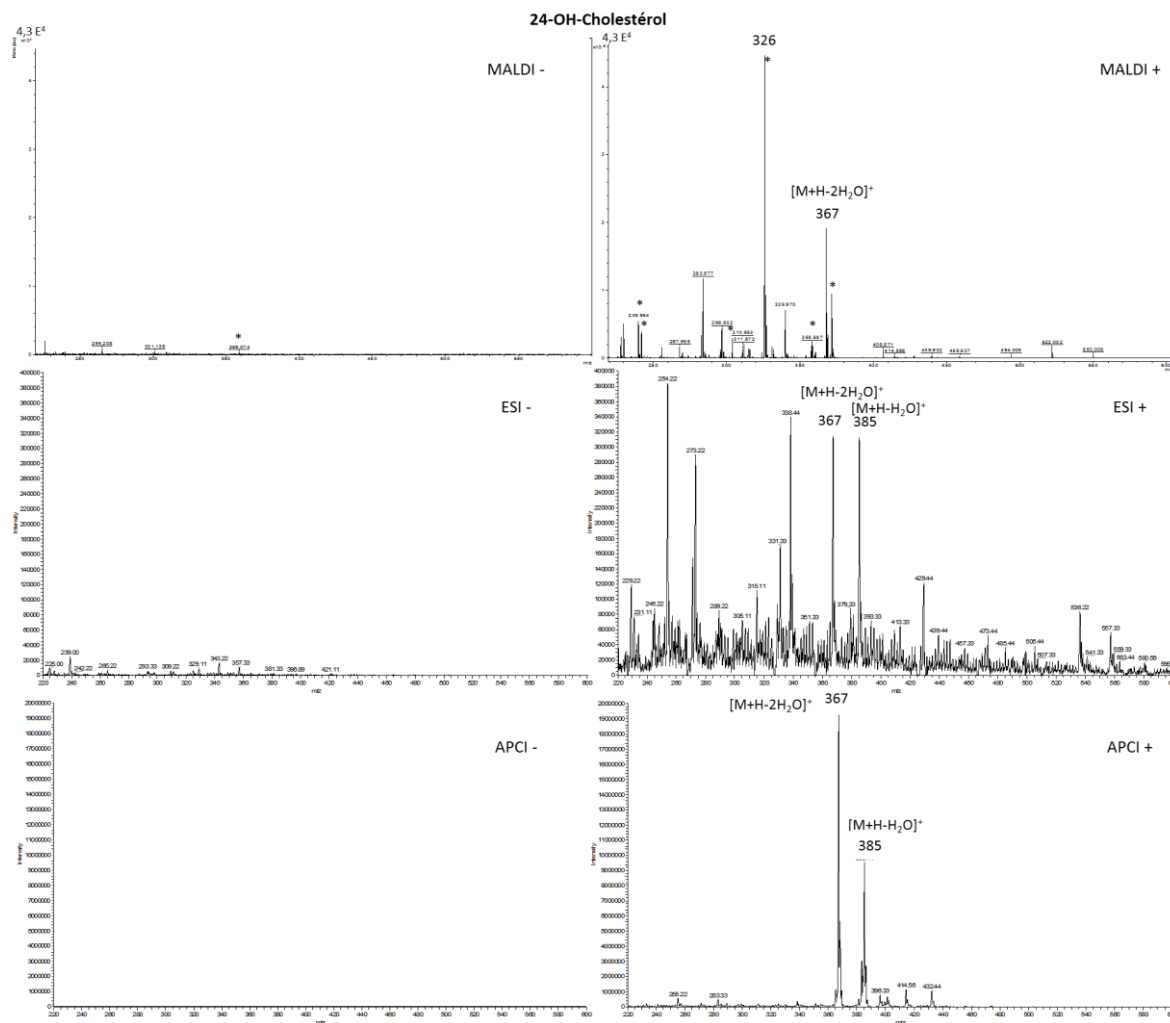


Figure II-28 : Spectres de masse du standard de 24-OH-Cholestérol obtenus avec les sources APCI, MALDI et ESI en mode d'ionisation négatif et positif. Dépôt de 2 μL d'une solution de standard à 1mg.mL^{-1} sur une plaque TLC silice 60 F₂₅₄ (5 x 7,5 cm ou 10 x 10 cm suivant le couplage). Résultats obtenus en couplage avec les trois sources sans élution chromatographique préalable. Le symbole * correspond à des ions de la matrice inorganique.

De toutes les conditions testées, seule la source d'ionisation APCI (Fig. II-28) en mode positif permet d'identifier la molécule 24 OH-cholestérol, avec la plus grande intensité (2×10^7). Lors de son ionisation, la molécule perd une ou deux molécules d'eau, les ions détectés correspondent aux ions $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$ et $[\text{M}+\text{H}-2\text{H}_2\text{O}]^+$. Dans les autres conditions de source ou de polarité, plus ou moins d'ions sont présents dans les spectres de masse, correspondant aux ions de la matrice ou de la silice ou du système et aucun ion ne permet de caractériser la molécule. Pour les autres standards d'oxystérols et les phytostérols, le même constat est fait. Ils ne s'ionisent qu'avec la source APCI en mode d'ionisation positif et sont détectés grâce à leur ion $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$.

Différents stéroïdes ont ensuite été testés dont la testostérone pour laquelle les résultats obtenus sont présentés figure II-29.

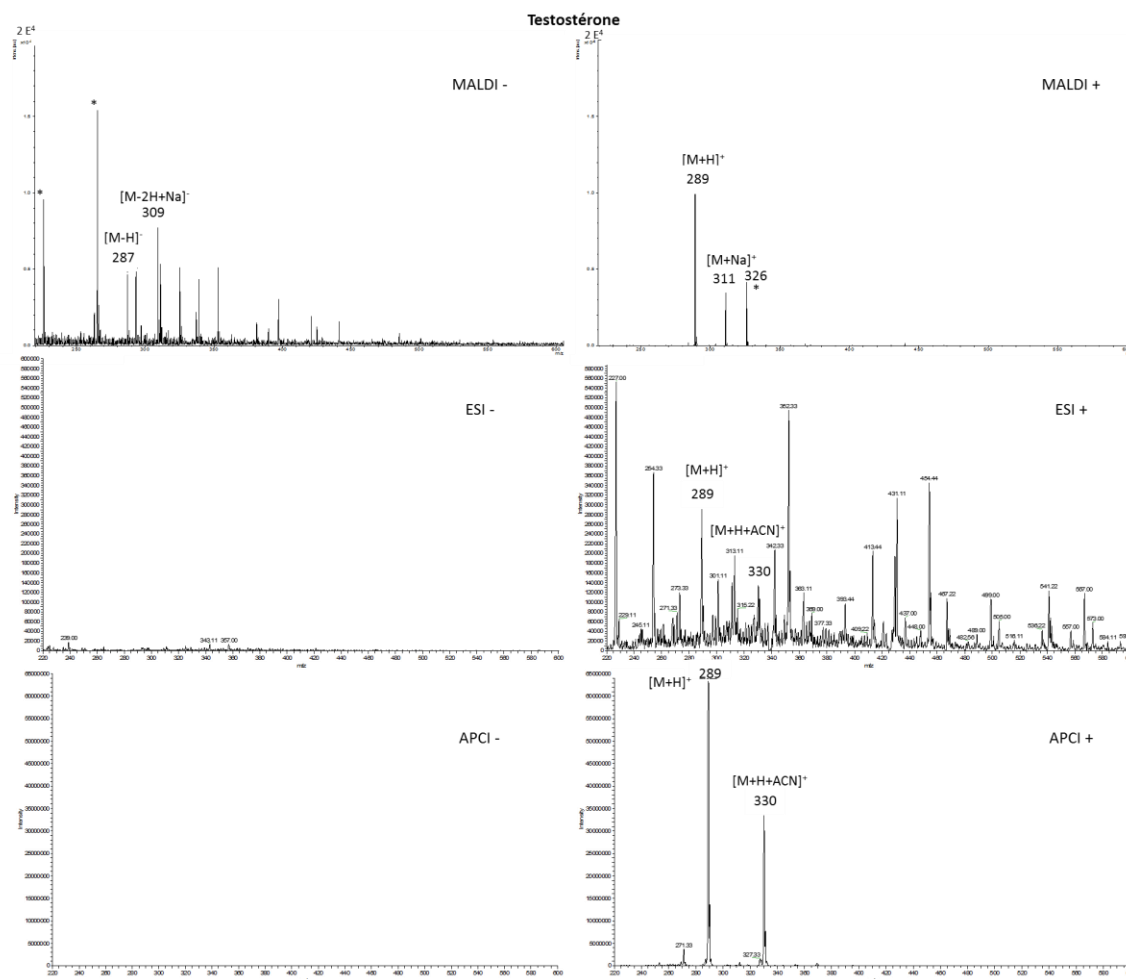


Figure II-29 : Spectres de masse du standard de testostérone obtenus avec les sources APCI, MALDI et ESI en mode d'ionisation négatif et positif. Dépôt de 2 μL d'une solution standard à 1mg.mL^{-1} sur une plaque TLC de silice 60 F₂₅₄ (10 x 10 cm ou 5 x 7,5 cm suivant le couplage). Résultats obtenus en couplage avec les trois sources sans élution chromatographique préalable. Le symbole * correspond à des ions de matrice inorganique.

La molécule de testostérone s'ionise dans toutes les conditions testées mais le rendement d'ionisation est meilleur en mode d'ionisation positif (Fig. II-29). Quelle que soit la source d'ionisation, l'ion observé en mode d'ionisation positif est la molécule avec un adduit protoné $[\text{M}+\text{H}]^+$ ou un adduit sodium $[\text{M}+\text{Na}]^+$. Tous les stéroïdes standards que nous avons testés s'ionisent avec les trois sources dans les deux modes d'ionisation (positif ou négatif). Malgré l'observation des ions caractéristiques des molécules avec les trois sources d'ionisation en mode positif, la source APCI permet d'avoir la meilleure réponse (APCI : $6,3 \text{ E}^7$; ESI : $2,66 \text{ E}^5$; MALDI : 1 E^4). Selon la structure des molécules, l'un des deux modes d'ionisation peut être privilégié, la testostérone (Fig. II-29) et la progestérone (autres stéroïdes) ont été ionisées de manière préférentielle en mode d'ionisation positif tandis que l'estriol, l'estrone et l'estradiol (œstrogènes) ont été détectés plus efficacement en mode d'ionisation négatif.

Etant donné que la source d'ionisation MALDI n'a donné que peu de résultats satisfaisants pour l'ensemble des standards, la suite de l'étude a été réalisée en utilisant l'ionisation à pression atmosphérique. Comme il a été décrit précédemment, un solvant est nécessaire pour désorber les composés de la plaque et les entraîner jusqu'à la source à pression atmosphérique du spectromètre de masse. Il est bien connu en HPLC/MS que la nature de la phase mobile a une grande influence sur l'ionisation des composés, c'est pourquoi une étude sur la nature du solvant de désorption a été menée avec les deux sources d'ionisation ESI et APCI dans les deux modes d'ionisation avec un à deux standards par famille moléculaire (24-OH, 7 β -OH, testostérone, estriol et stigmastérol).

5.2.2. Etude du solvant de désorption

Les composés à analyser étant moyennement polaires, l'eau n'a pas été sélectionnée comme solvant de désorption des molécules de la plaque TLC en silice. Le méthanol et l'acétonitrile ont été évalués en tant que solvant de désorption, car ce sont des solvants souvent utilisés en UHPLC/MS pour l'analyse de ces composés. De plus, il a été ajouté un additif à ce solvant suivant le mode d'ionisation. L'acide formique (AF) a été ajouté pour l'étude en mode d'ionisation positif pour avoir un apport de proton, et l'acétate d'ammonium (AcAm) a été ajouté pour l'ionisation en mode négatif pour faciliter la perte de proton lors de l'ionisation. L'étude des solvants a été réalisée sur des dépôts de standards non élués.

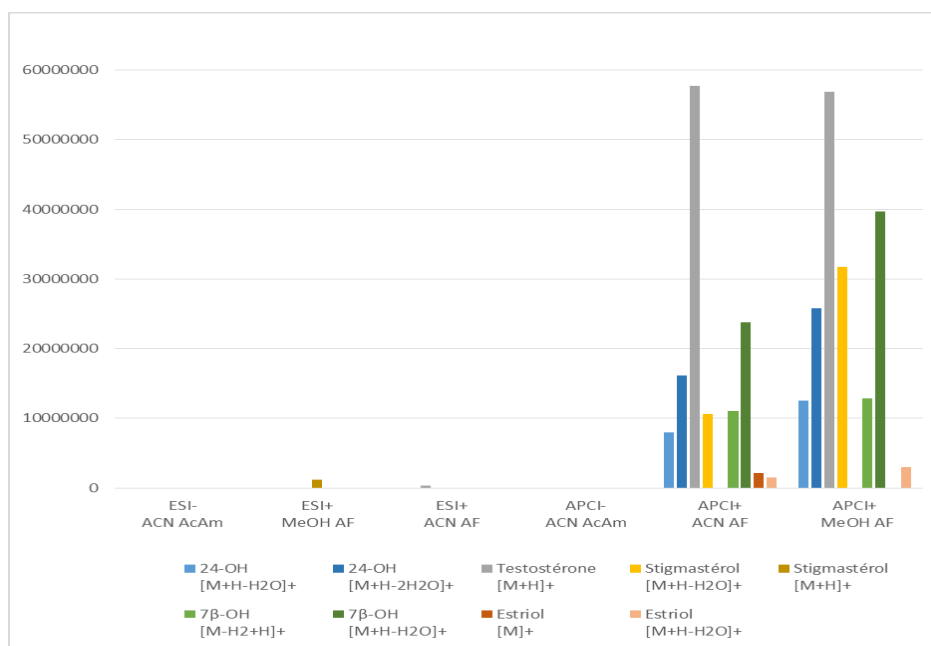


Figure II-30 : Comparaison de la réponse de cinq standards en couplage TLC/API MS, avec deux sources (ESI et APCI) en mode d'ionisation négatif et positif. Dépôt de 5 μ L de chaque solution standard à 1 mg.mL⁻¹ sur plaque TLC silice 60 F₂₅₄ (10 x 10 cm). Solvant de désorption : acétonitrile avec 5 mM d'acétate d'ammonium (ACN AcAm) ou méthanol avec 0.1% d'acide formique (MeOH AF) ou acétonitrile avec 0.1% d'acide formique (ACN AF). Résultats obtenus en couplage avec les deux sources sans élution chromatographique préalable.

Les résultats présentés Figure II-30 confirment que quel que soit le solvant de désorption utilisé ces molécules s'ionisent mieux en APCI qu'en ESI et s'ionisent mieux en mode d'ionisation positive. La réponse des composés est sensiblement plus importante avec le solvant méthanol cependant pour la suite de l'étude le solvant acétonitrile a été préféré, car il permet pour certains composés, comme l'estriol, d'obtenir l'ion moléculaire de la molécule en plus de l'ion fragment correspondant à une perte d'eau $[M+H-H_2O]^+$. Lors de l'analyse de mélanges inconnus, cette information est précieuse car elle permet plus facilement de confirmer la masse des molécules présentes.

5.2.3. Analyse HPTLC/APCI-TQ MS d'un mélange de 23 standards

En s'appuyant sur ces résultats, l'analyse par couplage HPTLC/APCI-TQ MS d'un mélange de 23 standards a été menée après avoir réalisé la séparation chromatographique du mélange sur la plaque. Parmi les 23 standards étudiés, quatre cas différents de molécules isobares sont observés pour les masses moléculaires respectivement de 402 ; 386 ; 360 ou 288 Da (valeurs repérées par une couleur spécifique dans le tableau II-3). Sans séparation chromatographique, l'identification de ces composés semble difficile. Après la séparation par HPTLC (Fig. II-31A et tableau II-3), neuf bandes d'élution sont distinguables.

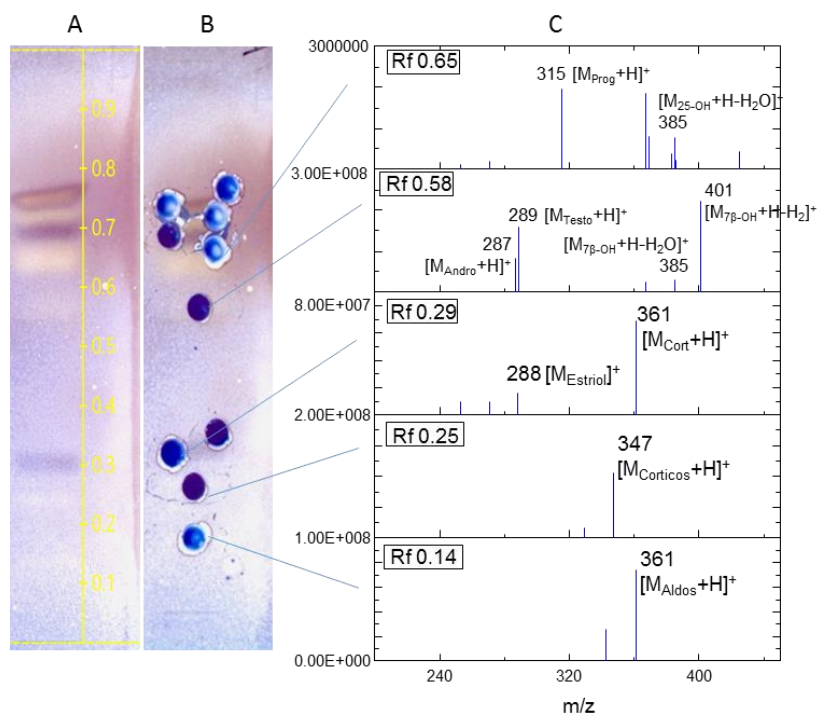
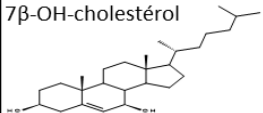
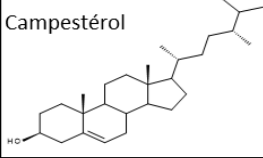
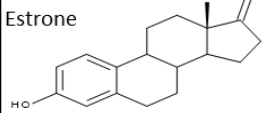
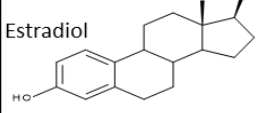
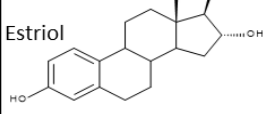


Figure II-31 : Analyse HPTLC/APCI-MS en mode d'ionisation positif pour l'identification d'un mélange de 23 standards d'oxystérols et de stéroïdes. Dépôt de 5 μ L de solutions standard à 1 mg.mL⁻¹ sur une plaque HPTLC verre silice 60 F₂₅₄ (20 x 10 cm). Phase mobile pour l'élution : n-hexane/Iso-octane/Isopropanol 69 :24 :7 v/v/v, avec 1% AcOH. La désorption des composés a été fait à 0.2 mL/min avec de l'acétonitrile acidifié à 0.1% avec de l'acide formique. (A) Plaque du mélange élué à 366 nm. (B) Plaque après analyse en HPTLC/APCI-MS. (C) Spectres de masses en mode d'ionisation positif avec le couplage HPTLC/APCI MS.

Tableau II-3 : Liste des oxystérols, phytostérols et stéroïdes (structures en annexe 5) étudiés avec masse et exemple de structure par famille. Elution sur plaque HPTLC silice F₂₅₄ avec une composition d'éluant de n-hexane/Iso-octane/Isopropanol (69 :24 :7 v/v/v +1%AcOH). Indication des masses isobares par une même couleur.

Composés non dérivés (1mg/mL)		Masse théorique (M en Da)	R _f expérimental	Exemples de structures
Oxystérols	Cholestérol	386	0,75	7β-OH-cholestérol 
	7β-OH-Cholestérol	402	0,57	
	22,24,25 et 27-OH-Cholestérol	402	0,72 et 0,67	
Phytostérols	Campestérol	400	0,74	Campestérol 
	Desmostérol	384	0,74	
	Stigmastérol	412	0,74	
	β-sitostérol	414	0,75	
	Bétuline	442	0,73	
Stéroïdes	Cortisone	360	0,28	Estrone 
	Aldostérone	360	0,14	
	Corticostérone	346	0,25	
	Lathostérol	386	0,75	Estradiol 
	DeHydroEpiAndrostérone (DHEA)	288	0,66	
	Androstènedione	286	0,6	
	Estrone	270	0,74	Estriol 
	Estradiol	272	0,7	
	Estriol	288	0,31	
	Testostérone	288	0,57	
	Progestérone	314	0,66	
	DiHydroTestostérone (DHT)	290	0,68	

Les principaux composés du mélange ont été identifiés pour chaque rapport frontal (Fig. II-31C et 32, tableau II-3) à partir de leurs ions $[M+H]^+$. Des co-migrations sont confirmées après la détection en masse (Fig. II-31C). Par exemple, au rapport frontal R_f 0,58, les 3 composés : testostérone, androstènedione et 7β-OH-cholestérol sont co-élus et au rapport frontal R_f 0,65, les 2 composés, progestérone et 25-OH-cholestérol, sont co-élus. Certains spectres de masse présentant plusieurs ions différents (Fig. II-31C, spot R_f 0,65) peuvent laisser supposer qu'il y a co-élution de plusieurs composés mais cela peut être dû aussi à un artefact engendré par la tête de migration de l'interface lors de la désorption. En effet cette tête de migration étant sphérique avec un diamètre de 4 mm, elle ne permet pas d'obtenir une bonne résolution spatiale ainsi pour deux bandes séparées sur la plaque HPTLC mais avec des R_f très proches, cette buse peut désorber des composés provenant des deux bandes. Pour vérifier la co-élution, une analyse en zigzag a été effectuée sur la plaque (Fig. II-31B). Pour obtenir une analyse plus fine il est maintenant aussi possible d'utiliser un autre type de buse telle qu'une buse ovale de 2 mm de hauteur.

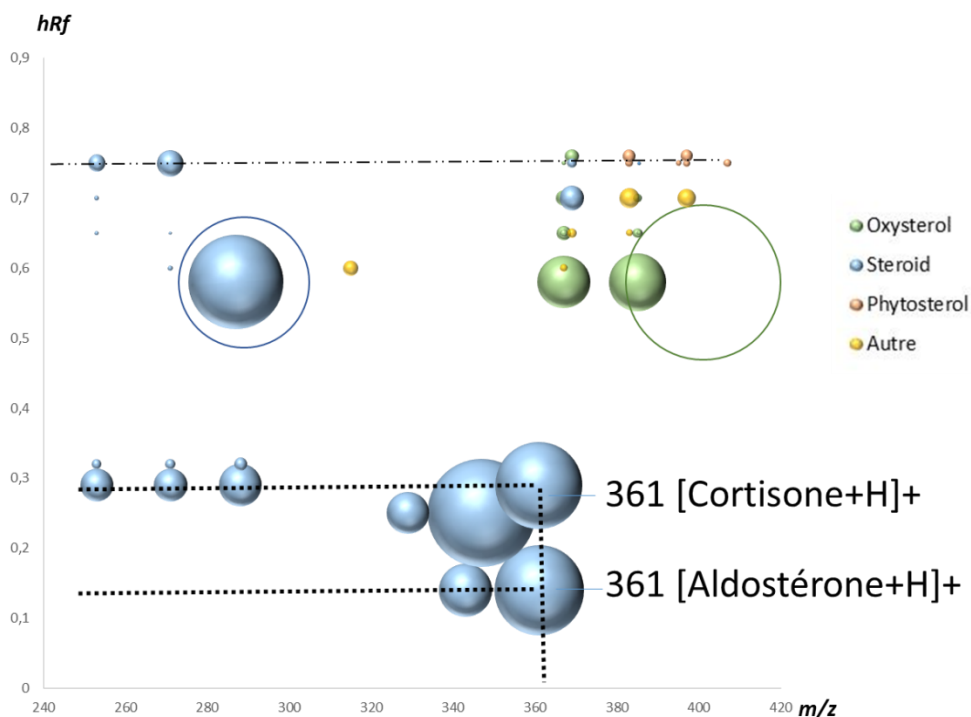


Figure II-32 : Analyse HPTLC/APCI MS d'un mélange de standards d'oxystérols et de stéroïdes. Graphique à bulles corrélant l'empreinte spectrale massique et la migration (R_f) des standards. La taille des bulles indique l'intensité des ions.

Les molécules ayant une même masse, comme les molécules de cortisone et d'aldostérone ($M = 360$ Da) sont distinguables par leur différence de rapport frontal et par l'intensité de la réponse de leur ion $[M+H]^+$ (Fig. II-31C et 32). Les molécules, comme le cholestérol et le lathostérol, sont élués avec le même rapport frontal (R_f 0,75) et la même masse moléculaire ($M = 386$ Da) mais peuvent cependant être distinguées. En effet, la molécule de lathostérol étant un stéroïde, son ion moléculaire sous forme $[M]^+$ (386 u) est présent en plus de son ion fragment (369 u) correspondant à la perte d'eau, alors que le cholestérol, qui est un oxystérol, n'est observable que sous la forme $[M+H-H_2O]^+$ (369 u) avec la perte d'eau.

L'analyse du spot au rapport frontal R_f de 0,75 permet de mettre en évidence la présence de molécules des trois familles moléculaires (stéroïdes, oxystérols et phytostérols). Cette analyse permet d'illustrer que pour un rapport frontal donné, même si les composés ont été déposés initialement sur la plaque à la même concentration (1 mg.mL^{-1}), ils n'ont pas tous le même rendement d'ionisation. En effet, les stéroïdes (lathostérol et estrone) s'ionisent mieux que les phytostérols (campestérol et β -sitostérol) et les oxystérols (cholestérol).

Bien que l'ionisation avec la source APCI ait produit une fragmentation (les alcools peuvent perdre une molécule d'eau), la capacité à ioniser les composés stéroïdiens a permis d'identifier confortablement les molécules par rapport aux sources ESI et MALDI. Ces résultats obtenus à partir d'un mélange standard prouvent l'intérêt du couplage HPTLC/APCI MS pour la caractérisation rapide des stérols et des stéroïdes dans des matrices complexes.

Le couplage TLC/API MS est un couplage rapide, de première intention, qui peut permettre de réaliser des analyses en routine de caractérisation d'échantillons dès lors que les molécules recherchées sont connues ainsi que leurs propriétés (solubilité, polarité, conditions les plus favorables d'ionisation (source et mode d'ionisation positif ou négatif)). Par contre, le couplage reste limité dans le cas d'analyses non ciblées dû à l'utilisation semi-manuelle de l'interface et au manque de résolution de l'emporte-pièce pour désorber les taches.

6. Conclusions et perspectives

Ces différentes approches démontrent que la chromatographie sur couche mince permet de faire des comparaisons entre des variétés ou des organes de plante que cela soit visuellement, mais aussi via des traitements de données, tels que l'analyse en composante principale (ACP). Les études avec les différents couplages TLC/MS dans des domaines comme la phytochimie et le médical mettent en évidence la complémentarité des sources MALDI, ESI et APCI et des couplages sous vide et à pression atmosphérique.

Le traitement statistique des profils phytochimiques obtenus sur plaques TLC en analyse en composante principale permet d'avoir rapidement une comparaison d'un grand nombre d'échantillons par famille moléculaire basée sur la couleur et l'intensité des différentes bandes et donc de proposer un regroupement des extraits par similitude de profil en peu de temps. Il peut permettre de s'affranchir de passer en chromatographie en phase liquide pour obtenir le même type d'information. Le couplage avec la spectrométrie de masse peut permettre l'identification des molécules qui influencent la séparation des groupes.

L'élution, la zone de prélèvement de la tache, ou de la désorption et l'ionisation sont les principaux facteurs qui influencent la réponse en spectrométrie de masse.

Par exemple, l'utilisation d'une source à vide avec un laser pulsé MALDI permet d'avoir une petite zone de désorption et donc une bonne résolution spatiale. Comme le laser n'ablate qu'en superficie les composés de la tache et uniquement sur une petite zone, il est parfaitement envisageable de réaliser plusieurs analyses successives d'une même tache. La matrice doit par contre être bien sélectionnée pour ioniser au mieux l'ensemble des composés. La quantité de matrice déposée joue un rôle important sur l'intensité de la réponse de l'analyte. Un bon équilibre entre cette quantité déposée et la teneur du soluté est à respecter. Les plaques TLC peuvent être scannées tout au long des pistes pour obtenir des informations chromatographiques et moléculaires, afin d'obtenir une image des extraits déposés selon le système d'élution. Il est possible de programmer le logiciel pour balayer le laser sur la largeur d'une bande et tout au long de la plaque avec un pas défini et régulier et faire des acquisitions successives. Cette acquisition automatisée permettrait de réaliser des analyses ciblées des extraits, mais avec une usure du laser plus importante. Ce type d'acquisition n'est pas réalisable avec tout type de matrice, car la matrice doit pouvoir répondre au déplacement du laser programmé, ce qui n'est pas le cas actuellement pour la matrice inorganique développée au laboratoire. En effet, le laser doit taper la matrice inorganique régulièrement et très rapidement, avec une certaine fréquence, mais surtout avec des impacts proches les uns des autres, pour que la matrice puisse émaner de l'énergie photonique et permettre une bonne ionisation des composés, ce que le logiciel ne permet pas de faire

dans les déplacements pré-enregistrés. Le manque de matrice générique en MALDI peut limiter ce type de couplage, et son installation en technique de routine.

L'échantillonnage direct à pression atmosphérique sur la plaque TLC avec solubilisation du composé ciblé fournit une plus grande surface de prélèvement de la tache et peut donc avoir une réponse plus importante, néanmoins dans l'état actuel, ce procédé est limité par une mauvaise résolution spatiale et donc le risque de désorber deux taches en même temps. Lorsque la plaque est positionnée sous l'emporte-pièce, il peut arriver que l'analyse se fasse à côté de la tache et non précisément sur la tache. Ce cas de figure peut arriver quand on se sert comme référentiel d'une plaque révélée pour positionner la buse sur la plaque non révélée et dans ce cas-là, il faut tout recommencer, y compris la plaque, car l'analyse est destructive. Le couplage TLC/API MS est un couplage facile à mettre en œuvre, si on connaît le solvant adéquat de solubilisation et de transport des composés, le bon mode d'ionisation et la bonne source d'ionisation. L'inconvénient dans l'utilisation de l'interface Camag est que l'utilisateur doit réaliser les différentes étapes lui-même sans automatisation. Ainsi il doit positionner la plaque pour le prélèvement de chaque spot, descendre le positionnement du piston, changer la position d'une vanne pour effectuer la désorption et lancer l'acquisition. Ce temps passé à cette manutention peut devenir long quand il y a plus d'une dizaine de taches à analyser, ce qui peut être souvent le cas lors d'analyses d'échantillons complexes.

Dans les deux types de couplage, la sensibilité de l'analyse va dépendre de la focalisation du composé sur la plaque TLC, de son rendement d'ionisation et de sa capacité à se fragmenter en source. De plus, dans un mélange complexe biologique ou provenant d'extrait de plante, les molécules constituant le mélange peuvent être présentes dans des quantités très variables. Bien que séparées, cela peut gêner la qualité de l'analyse que cela soit en couplage TLC/MALDI-Tof MS ou en TLC/API MS. En effet, pour le premier couplage, il faudra ajuster la quantité de la matrice à déposer. Et dans le cas du deuxième couplage, l'analyse de composés en trop forte concentration peut polluer les analyses suivantes. Les deux types de couplage sont complémentaires et permettent d'avancer dans la caractérisation d'échantillon complexe.

Lors d'une analyse d'une plaque TLC en couplage à la spectrométrie de masse, quelques questions doivent être étudiées avant de sélectionner une méthode, si on ne peut pas faire les deux. Ainsi il doit être connu les familles moléculaires attendues, si elles ont déjà été ionisées avec une des trois sources disponibles, si elles ont un mode d'ionisation préférentiel, leurs solubilités et comment on peut les repérer sur la plaque. Ces réponses permettront de sélectionner au plus vite l'un des deux couplages. Il est après préférable d'analyser des standards élués dans les mêmes conditions, s'ils sont disponibles, pour connaître leur réponse et leur comportement avec la méthode, et ensuite de passer à l'analyse des échantillons.

Quel que soit le domaine d'activité, le couplage TLC/MS permet d'avoir une information moléculaire rapide. En chimie de synthèse, par exemple, il permet d'identifier la position sur la plaque du composé d'intérêt synthétisé, de vérifier la présence éventuelle de co-produits et de les caractériser. Cette opération permet de réaliser par la suite une transposition rapide de la plaque TLC à la colonne ouverte pour purifier les composés en utilisant peu de solvant. Dans le domaine de la pharmaceutique, le couplage permet d'avoir une méthode complémentaire au suivi des impuretés et de donner une identification de ces dernières. Dans le cadre de l'utilisation de plante pour des applications thérapeutique, cosmétique ou agroalimentaire, l'analyse statistique des données TLC permettrait de faire le suivi plus rapidement et de façon plus fiable, de la qualité des plantes et de leur préparation.

Afin d'aller plus loin dans ces couplages TLC/MS, il faudrait aussi les évaluer d'un point de vue quantitatif. Ces études permettraient de connaître les limites de détection selon les molécules à détecter et donc de connaître la gamme dynamique qui peut être étudiée dans un mélange complexe. En TLC, des courbes d'étalonnage et des dosages de molécule peuvent déjà être réalisées avec un densitomètre. Mais lors de co-élution, les réponses peuvent être faussées. Dans ce cas, la dimension masse permettrait de ne doser que la molécule ciblée. Pour les limites de détection insuffisantes, il peut être envisagé de développer une méthode au préalable du couplage, qui permettrait d'attirer et de focaliser les composés à la surface de la plaque pour que l'ionisation soit plus importante, ou de développer une méthode en mode MRM avec le couplage TLC/API-TQ MS qui améliorerait la sensibilité de l'analyse.

Comme pour les xanthones, l'analyse TLC/MS pourrait être élargie à d'autres familles moléculaires appartenant à la voie de biosynthèse des polyphénols, mais aussi aux métabolites secondaires de types terpénoïdes et alcaloïdes pour avoir un large panel d'analyses et couvrir les composés obtenus dans différents extraits de plantes. Il serait également intéressant de regarder si en changeant la phase stationnaire TLC en utilisant par exemple de la silice greffée C18, il interviendrait un changement de réponse ou des interférents.

Afin de généraliser les analyses TLC/MALDI-Tof MS, la matrice inorganique pourrait être greffée à toute la phase stationnaire. En effet, contrairement à la matrice DHB, la matrice inorganique peut imprégner la plaque avant dépôt des échantillons et élution sans aucune modification des analyses. Ainsi les utilisateurs pourraient l'utiliser sans réaliser de préparation supplémentaire. Cette matrice inorganique pourrait également être combinée à d'autres matrices pour améliorer son efficacité.

Références bibliographiques

- R. Arakawa and H. Kawasaki**, Functionalized nanoparticles and nanostructured surfaces for surface-assisted laser desorption/ionization mass spectrometry, *Analytical sciences*, 2010, 26, 1229-1240
- C. Audoin, S. Holderith, K. Romari, O. Thomas, G. Genta-Jouve**, Development of a work-flow for high-performance thin-layer chromatography data processing for untargeted metabolomics, *Journal of planar chromatography*, 2014, 27(5), 328-332
- C. D. Cerruti, F. Benabdellah, O. Laprévote, D. Touboul, A. Brunelle**, MALDI imaging and structural analysis of rat brain lipid negative ions with 9-aminoacridine matrix, *Analytical chemistry*, 2012, 84, 2164-2171
- SC. Cheng, MZ. Huang, J. Shiea**, Thin layer chromatography/mass spectrometry, *Journal of chromatography A*, 2011, 1218, 2700-2711
- CK. Chiang, WT. Chen, HT. Chang**, Nanoparticle-based mass spectrometry for the analysis of biomolecules, *Chemical society reviews*, 2011, 40, 1269-1281
- A. Crecelius, MR. Clench, DS. Richards, V. Parr**, Thin-layer chromatography-matrix-assisted laser desorption ionization-time-of-flight mass spectrometry using particle suspension matrices, *Journal of chromatography A*, 2002, 958, 249-260
- A. Crecelius, M. R. Clench, D. S. Richards**, TLC-MALDI in pharmaceutical analysis, *Current trends in mass spectrometry*, 2004, 19 (5S), 28-34
- GC. Cretu and GE. Morlock**, Analysis of anthocyanins in powered berry extracts by planar chromatography linked with bioassay and mass spectrometry, *Food chemistry*, 2014, 146, 104-112
- F. Cuyckens, Y. L. Ma, G. Pocsfalvi, M. Claeys**, Tandem mass spectral strategies for the structural characterization of flavonoid glycosides, *Analisis*, 2000, 28, 888-895
- E. Destandau, T. Michel, A. Toribio, C. Elfakir**, Different ways to on-line hyphenate centrifugal partition chromatography and mass spectrometry: Application to prenylated xanthenes from *Garcinia mangostana*, *Planta medica*, 2015, 81(17), 1597-1603
- H. E.-D. S. El Beltagi and A. Amin Mohamed**, Variations in fatty acid composition, glucosinolate profile and some phytochemical contents in selected oil seed rape (*Brassica napus* L.) cultivars, *Grasas y aceites*, 2010, 61(2), 143-150
- D. Fichou, P. Ristivojević, G. E. Morlock**, Proof-of-principle of rTLC, an open-source software developed for image evaluation and multivariate analysis of planar chromatograms, *Analytical chemistry*, 2016, 88, 12494-12501
- J. Ferey, D. Da Silva, P. Lafite, R. Daniellou, B. Maunit**, TLC-UV hyphenated with MALDI-TOFMS for the screening of invertase substrates in plant extracts, *Talanta*, 2017, 170, 419-424
- P. Flores, P. Hellín, J. Fenoll**, Determination of organic acids in fruits and vegetables by liquid chromatography with tandem-mass spectrometry, *Food chemistry*, 2012, 132, 1049-1054
- A. Francini and L. Sebastiani**, Phenolic compounds in apple (*Malus x domestica* Borkh.): Compounds characterization and stability during postharvest and after processing, *Antioxidants (Basel)*, 2013, 2(3), 181-193
- B. Fuchs, R. Süß, A. Nimptsch, J. Schiller**, MALDI-TOF-MS directly combined with TLC: a review of the current state, *Chromatographia*, 2009, 69, S95-S105
- B. Fuchs**, Analysis of phospholipids and glycolipids by thin layer chromatography matrix assisted laser desorption and ionization mass spectrometry, *Journal of chromatography A*, 2012, 1259, 62-73

- X. Gao, Z. Tang, L. Lu, H. Liu, Y. Jiang, Y. Zhao, Z. Cai**, Suppression of matrix ions by N-phosphorylation labeling using matrix-assisted laser desorption-ionization time-of-flight mass spectrometry, *Chemical communications*, 2012, 48, 10198-10200
- V. Glavnik, I. Vovk, A. Albrecht**, High performance thin layer chromatography mass spectrometry of Japanese knotweed flavan-3-ols and proanthocyanidins on silica gel plates, *Journal of chromatography A*, 2017, 1482, 97-108
- C. G. Grigoras, E. Destandau, L. Fougère, C. Elfakir**, Evaluation of apple pomace extracts as a source of bioactive compounds, *Industrial crops and products*, 2013, 49, 794-804
- T. Guinan, P. Kirkbride, P. E. Pigou, M. Ronci, H. Kobus, N. H. Voelcker**, Surface-assisted laser desorption ionization mass spectrometry techniques for application in forensics, *Mass spectrometry reviews*, 2015, 34, 627-640
- H. Hayen and D. A. Volmer**, Rapid identification of siderophores by combined thin-layer chromatography/matrix-assisted laser desorption/ionization mass spectrometry, *Rapid communications in mass spectrometry*, 2005, 19, 711-720
- TW. Jaskolla, M. Karas, U. Roth, K. Steinert, C. Menzel, K. Reihs**, Comparison between vacuum sublimed matrices and conventional dried droplet preparation in MALDI-TOF mass spectrometry, *Journal American society for mass spectrometry*, 2009, 20, 1104-1114
- W. G. Jing, J. Fu, Y. Guo, A. Liu**, Phytochemical screening of flavonoids with their antioxidant activities from rapeseed (*Brassica napus* L.), *Phytochemistry letters*, 2015, 13, 239-245
- M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp**, Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *International journal of mass spectrometry and ion processes*, 1987, 78, 56-68
- M. Karas and F. Hillenkamp**, Laser desorption ionization of proteins with molecular masses exceeding 10 000 daltons, *Analytical chemistry*, 1988, 60 (20), 2299-2301
- M. Karas and R. Krüger**, Ion formation in MALDI: the cluster ionization mechanism, *Chemical reviews*, 2003, 103, 427-439
- T. Kartnig and I. Göbel**, Effect of fluorescence intensifiers on the fluordensitometric determination of flavones and flavonols after detection with diphenylboric acid 2-aminoethyl ester, *Journal of chromatography A*, 1996, 740, 99-107
- T. Kondakova, N. M. Machour, C. D. Poc**, HPTLC-MALDI TOF imaging analysis of phospholipids, *Lipidomics*, chapter 12, 2017, 163-173
- T. Kowalska, K. Kaczmarek, W. Prus**, Theory and mechanism of thin-layer chromatography, *Handbook of thin-layer chromatography*, chapter 2, 2003, 1-53
- I. Krosiakova, S. Pedrussio, E. Wolfram**, Direct coupling of TLC with MALDI-TOF MS for qualitative detection of flavonoids on phytochemical fingerprints, *Phytochemical analysis*, 2016, 27, 222-228
- S. Krüger, L. Bürmann, GE. Morlock**, Comparison and characterization of soybean and sunflower lecithins used for chocolate production by TLC-FLD-ESIMS, *Journal of agricultural and food chemistry*, 2015, 63, 2893-2901
- D. Kumar and U. Sharma**, High-performance thin-layer chromatography: an economical alternative for the quality control of medicinal plants and derived products, *Separation science plus*, 2018, 1, 100-134
- ED. Leriche, M. Hubert-Roux, MC. Grossel, CM. Lange, C. Afonso, C. Loutelier-Bourhis**, Direct TLC/MALDI-MS coupling for modified polyamidoamine dendrimers analyses, *Analytica chimica acta*, 2014, 808, 144-150
- Z. Lin and Z. Cai**, Negative ion laser desorption/ionization time-of-flight mass spectrometric analysis of small molecules by using nanostructured substrate as matrices, *Mass spectrometry reviews*, 2018, 37, 681-696

- Z. Liu and KL. Schey**, Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis, *American society for mass spectrometry*, 2005, 16, 482-490
- C.-W. Liu, M.-W. Chen, C.-Y. Su, H. Y. Chen, L.-J. Li, C.-C. Lai**, Analysis of flavonoids by graphene-based surface-assisted laser desorption/ionization time-of-flight mass spectrometry, *Analyst*, 2012, 137, 5809-5816
- H. Luftmann**, A simple device for the extraction of TLC spots: direct coupling with an electrospray mass spectrometer, *Analytical and bioanalytical chemistry*, 2004, 378 (4), 964-968
- H. L. J. Makin, J. W. Honour, C. H. L. Shackleton, W. J. Griffiths**, Chapter 3 – General methods for the extractions, purification, and measurement of steroids by chromatography and mass spectrometry, *Steroids analysis*, 2010, 163-282
- T. Michel, E. Destandau, C. Elfakir**, On-line hyphenation of centrifugal partition chromatography and high pressure liquid chromatography for the fractionation of flavonoids from *Hippophaë rhamnoides* L. berries, *Journal of chromatography A*, 2011, 1218 (36), 6173-6178
- T. Michel, E. Destandau, G. Le Floch, M. E. Lucchesi, C. Elfakir**, Antimicrobial, antioxidant and phytochemical investigations of sea buckthorn (*Hippophaë rhamnoides* L.) leaf, stem, root and seed, *Food chemistry*, 2012, 131 (3), 754-760
- G. Montaudo, F. Samperi, M. S. Montaudo**, Characterization of synthetic polymers by MALDI-MS, *Progress in polymer science*, 2006, 31, 277-357
- G. Morlock and W. Schwack**, Coupling of planar chromatography to mass spectrometry, *Trends in analytical chemistry*, 2010, 29, 1157-1171
- K. Naumoska and I. Vovk**, Analysis of triterpenoids and phytosterols in vegetables by thin-layer chromatography coupled to tandem mass spectrometry, *Journal of chromatography A*, 2015, 1381, 229-238
- SF. Reis, DK. Rai, N. Abu-Ghannam**, Water at room temperature as a solvent for the extraction of apple pomace phenolic compounds, *Food chemistry*, 2012, 82 (9), 3719-3726
- L. Riffault, E. Destandau, L. Pasquier, P. André, C. Elfakir**, Phytochemical analysis of *Rosa hybrida* cv. 'Jardin de Granville' by HPTLC, HPLC-DAD and HPLC-ESI-HRMS: Polyphenolic fingerprints of six plant organs, *Phytochemistry*, 2014, 99, 127-134
- M. Rohmer, B. Meyer, M. Mank, B. Stahl, U. Bahr, M. Karas**, 3-aminoquinoline acting as matrix and derivatizing agent for MALDI MS analysis of oligosaccharides, *Analytical chemistry*, 2010, 135, 1991-1993
- M. Sajewicz, D. Staszek, M. Natić, Ł. Wojtal, M. Waksmundzka-Hajnos, T. Kowalska**, TLC-MS versus TLC-LC-MS fingerprints of herbal extracts. Part II. Phenolic acids and flavonoids, *Journal of liquid chromatography & related technologies*, 2011, 54, 864-887
- A. Salwiński, D. Da Silva, R. Delépée, B. Maunit**, Enzyme-coupled nanoparticles-assisted laser desorption ionization mass spectrometry for searching for low-mass inhibitors of enzymes in complex mixtures, *Journal American society for mass spectrometry*, 2014, 25, 538-547
- J. P. Sang, I. R. Minchinton, P. K. Johnstone, R. J. W. Truscott**, Glucosinolate profiles in the seed, root and leaf tissue of cabbage, mustard, rapeseed, radish and swede, *Canadian journal of plant science*, 1984, 64(1), 77-93
- M. Schürenberg, D. Suckau, B. Fuchs, J. Schiller**, Direct read-out of thin layer chromatography (TLC) using MALDI-TOF, *Bruker Daltonics application note*, 2009, MT-94
- Y. Shao, J. Jiang, L. Ran, C. Lu, C. Wei, Y. Wang**, Analysis of flavonoids and hydroxycinnamic acid derivatives in rapeseeds (*Brassica napus* L. var. *napus*) by HPLC-PDA-ESI(-)-MSⁿ/HRMS, *Journal of agricultural and food chemistry*, 2014, 62, 2935-2945

- C. Y. Shi and C. H. Deng**, Recent advances in inorganic materials for LDI-MS analysis of small molecules, *Analyst*, 2016, 141, 2816-2826
- DB. Silva and NP. Lopes**, MALDI-MS of flavonoids: a systematic investigation of ionization and in-source dissociation mechanisms, *Journal of mass spectrometry*, 2015, 50, 182-190
- S. Smrke and I. Vovk**, Comprehensive thin-layer chromatography mass spectrometry of flavanols from *Juniperus communis* L. and *Punica granatum* L., *Journal of chromatography A*, 2013, 1289, 119-126
- J. Son and S. Cha**, MALDI mass spectrometry analysis of nonderivatized steroids using cyclodextrin-supported 2,5-dihydrobenzoic acid as matrix, *Bulletin of the Korean chemical society*, 2014, 35 (5), 1409-1412
- J. Sunner, E. Dratz, Y.-C. Chen**, Graphite surface-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides and proteins from liquid solutions, *Analytical chemistry*, 1995, 67, 4335-4342
- K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo**, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid communications in mass spectrometry*, 1988, 2 (8), 151-153
- E. Torretta, M. Vasso, C. Fania, D. Capitanio, S. Bergante, M. Piccoli, G. Tettamanti, L. Anastasia, C. Gelfi**, Application of direct TLC-MALDI for the qualitative and quantitative profiling of neutral and acidic glycosphingolipids: The case of NEU3 overexpressing C2C12 murine myoblasts, *Electrophoresis*, 2014, 35, 1319-1328
- J. Wang and P. Sporns**, Analysis of anthocyanins in red wine and fruit juice using MALDI-MS, *Journal of agricultural and food chemistry*, 1999, 47, 2009-2015
- J. Wegener, K. Zschörnig, K. Onischke, B. Fuchs, J. Schiller, K. Müller**, Conservation of honey bee (*Apis mellifera*) sperm phospholipids during storage in the bee queen- A TLC/MALDI-TOF MS study, *Experimental gerontology*, 2013, 48, 213-222
- A. Wojdyło, J. Oszmiański, P. Laskowski**, Polyphenolic compounds and antioxidant activity of new and old apple varieties, *journal of agricultural and food chemistry*, 2008, 56, 6520-6530
- G. B. Yagnik, R. L. Hansen, A. R. Korte, M. D. Reichert, J. Vela, Y. J. Lee**, Large scale nanoparticle screening for small molecule analysis in laser desorption ionization mass spectrometry, *Analytical chemistry*, 2016, 88, 8926-8930
- R. Zenobi and R. Knochenmuss**, Ion formation in MALDI mass spectrometry, *Mass spectrometry reviews*, 1998, 17, 337-366

**Approche métabolomique :
Comparaison ciblée
de la composition
en flavonoïdes glycosylés
de différentes variétés
de soies de maïs par UHPLC/MS**

Les métabolites secondaires sont des molécules importantes dans la résistance des plantes face aux agresseurs et leurs concentrations varient considérablement selon la variété, l'âge et les stress biotique et abiotique subis, tels que l'exposition à la lumière, le climat, les agressions... Les interactions plante-insectes sont souvent impactées par l'émission de mélanges complexes de substances volatiles, contenant un grand nombre de mono et sesquiterpènes en proportion variable selon les variétés et les plants (Gouinguéné 2001, Degen 2004). Les métabolites secondaires peuvent également avoir des effets directs sur l'interaction plante-insecte. Glauser et al. (2011) ont ainsi montré par une approche métabolomique que les acides hydroxamiques, dont HDMBOA-Glc (2-β-D-glucopyranosyloxy-4,7-diméthoxy-1,4-benzooxazin-3-one) sont produits en grande quantité lors d'attaques d'insectes, la molécule HDMBOA libérée agissant alors comme fort dissuasif contre certains herbivores. Mais ces molécules peuvent avoir aussi un effet indirect sur les insectes puisque des composés organiques volatils peuvent être libérés de la plante après des dommages d'herbivores pour attirer les ennemis naturels (Gouinguéné 2001, Degen 2004). Stockel et al. (1987) ont émis des hypothèses utilisant ces deux types d'effets pour expliquer le comportement de la ponte des femelles *O. nubilalis* sur les soies de maïs. En effet, les mites seraient attirées par des composés volatils depuis une distance éloignée et l'oviposition serait induite au contact de plantes dépourvues de certains composés tels que les alcanes. Marti et al. (2013) ont développé par ailleurs une approche métabolomique non ciblée sur différents organes de maïs (feuilles, sève, racines, exsudats de racines) afin d'évaluer les changements locaux et systémiques induits par les herbivores. L'utilisation d'une approche métabolomique pour mieux comprendre les liens existants entre les phénomènes biologiques observés suite à un stress et la composition moléculaire d'une plante est en plein essor et semble très prometteuse quant aux réponses apportées.

Le travail de thèse de Béatrice RHINO (2015, université de Montpellier) intitulé « Gestion agroécologique de la noctuelle de la tomate *Helicoverpa zea* par l'utilisation du maïs comme plante piège », nous a donné l'opportunité d'initier une première collaboration entre notre équipe de recherche et le CIRAD de Martinique et de mettre en place au sein de notre laboratoire une première démarche métabolomique ciblée pour mieux comprendre les interactions plante/insecte. Le travail de thèse de B. Rhino a consisté principalement à étudier les volatils produits par les soies de maïs et l'efficacité de différentes variétés de maïs en tant qu'attracteur de l'*Helicoverpa zea* (*H. zea*) tandis que de notre côté, nous nous sommes intéressées à caractériser le profil métabolomique en flavonoïdes de cinq variétés de soies de maïs afin de déterminer si certains flavonoïdes pouvaient être discriminants d'une de ces variétés et expliquer la meilleure inhibition de croissance de larves observée pour certaines variétés. En effet, l'activité inhibitrice de la croissance des larves *H. zea* a déjà été démontrée pour trois flavones C-glycosylés (Elliger 1980a).

Un profilage des métabolites secondaires des soies de maïs a donc été mené par UHPLC/HRMS avec une première caractérisation des flavonoïdes dans les extraits de soies de maïs et une cartographie de l'empreinte moléculaire. Puis une méthode de quantification relative a été réalisée par UHPLC/MRM MS sur les flavonoïdes issus de cinq variétés de soies de maïs pour réaliser l'étude métabolomique ciblée. Différentes approches statistiques pour le traitement des données recueillies ont ensuite été mises en œuvre afin d'étudier si une corrélation entre nos profils phytochimiques et les observations faites sur le terrain en Martinique pouvait être établie. Cette approche métabolomique ciblée par UHPLC/MRM MS va ainsi nous permettre d'apporter une connaissance plus complète de la composition en flavonoïdes des soies de maïs et devrait permettre de mettre en évidence les différences de composition des différentes variétés testées.

1. Etat de l'art

1.1. Les plantes pièges

Dans le contexte actuel où des solutions alternatives à l'emploi massif de pesticides sont recherchées, la mise en place d'agroécosystèmes antiparasitaires reste une approche intéressante pour contrôler un ravageur par un ennemi naturel. Il est connu que certaines plantes repoussent les insectes tandis que d'autres au contraire les attirent. Une association raisonnée de différentes cultures végétales peut donc participer à réguler l'attaque de ravageurs.

Shelton et al. (2006) ont décrit plusieurs procédés pour travailler avec des plantes pièges plus attrayantes pour les insectes que la culture à protéger. Elles peuvent être attrayantes et limiter la survie des insectes ou de leur progéniture; elles peuvent être basées sur des plantes génétiquement modifiées, ou être assistées biologiquement ou semi-chimiquement. Les plantes pièges sont cultivées à des endroits ou à des moments spécifiques du développement de la culture à protéger et être composées d'un seul type de plante ou d'une combinaison.

Une étude sur l'utilisation du Sorgho comme plante piège de l'*Helicoverpa zea* (*H. zea*) aux abords des champs de coton a déjà été menée (Tillman and Mullinix 2004) et a permis de diminuer l'utilisation d'insecticide.

Sur ce principe, le CIRAD en Martinique a entrepris d'évaluer le maïs en tant que plante piège pour protéger les cultures de tomates contre ce même ravageur. Le choix de tester plutôt cette céréale est basé sur le fait qu'elle est très peu cultivée en Martinique et ainsi, dans le cas où une attractivité avérée pour *H. zea* et une action inhibitrice sur la croissance de ses larves seraient confirmées, son utilisation en tant que plante piège ne sera pas perturbée avec d'autres champs de maïs plantés en tant que culture céréalière.

1.2. Le maïs : Plante piège de l'*Helicoverpa zea*

Helicoverpa zea (*H. zea*) est une noctuelle, insecte connu pour ravager différents types de plantes comme le maïs (Fig. III-1A), le soja, le sorgho ou les tomates ... Il sévit principalement en Amérique du Nord. En Martinique, il est la cause de dévastation importante des cultures de tomates, car ses larves se nourrissent des fruits.

Dans le cas de la protection des plants de tomate, le maïs en tant que plante piège doit non seulement attirer les ravageurs *H. zea* (Fig. III-1A) mais aussi diminuer l'infestation. Pour que les ravageurs ne prolifèrent pas sur la plante à protéger, la plante piège sélectionnée doit donc pouvoir aussi attirer les prédateurs naturels (Fig. III-1B) de ses insectes ravageurs. Elle peut aussi produire une ou plusieurs molécules toxiques pour le ravageur limitant de ce fait sa croissance voire le supprimant.

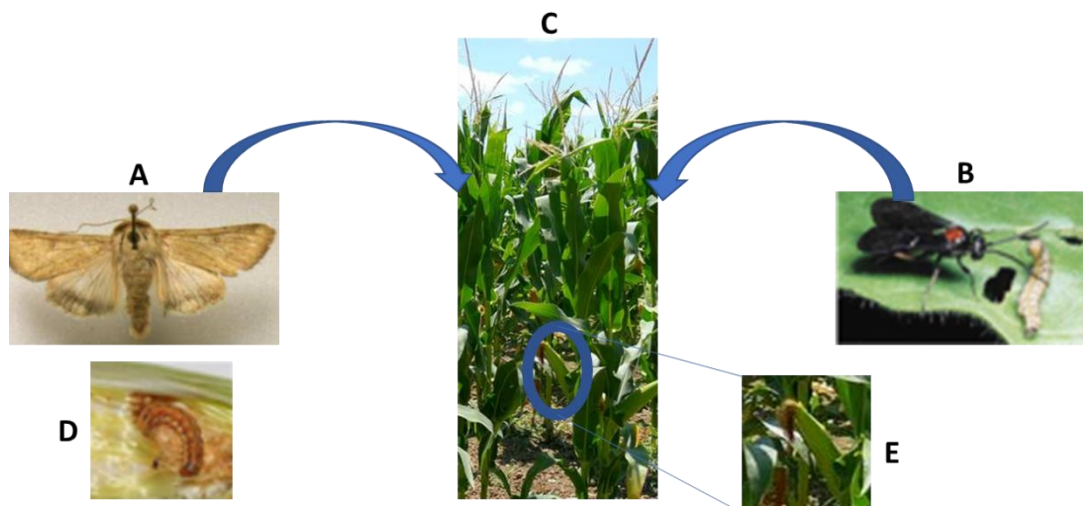


Figure III-1 : Schéma d'attraction de *H. zea* (A et D) et des prédateurs (B) sur le maïs (C). Larve de *H. zea* (D) se nourrissant de soie de maïs (E).

1.3. La soie de maïs

La soie (Fig. III-1E) de maïs (*Zea mays* L.) correspond aux stigmates femelles de la fleur de maïs. Fraîche, elle ressemble à des fils de soie pouvant être de différentes couleurs suivant la variété et la maturité de la plante.

Les soies de maïs sont utilisées en médecine pour diverses applications. Traditionnellement, la soie de maïs a été utilisée pour le traitement des troubles de l'urine, les gonflements, l'asthme, l'hypertension (Wynn and Fougère 2007). Les soies de maïs ont montré différentes activités d'intérêt. Ainsi elles peuvent être antidiabétiques (Rau 2006), anticoagulantes (Kim 2004), inhibitrices des anticorps IgE par les glycoprotéines dans le cadre de maladies allergiques (Namba 1993), antifongiques (Zeringue 2000) et anti-oxydantes (El-Ghorab 2007 ; Maksimović and Kovačević 2003).

Parmi les flavonoïdes présents dans la soie de maïs, 2 molécules (maysin et 3'-methoxymaysin) ont été décrites comme actives sur 11 bactéries telles que *Bacillus cereus* (Nessa 2012). Les soies de maïs ont également prouvé un effet antibiose sur *H. zea*. Elliger et al. (1980a) ont rapporté que les flavones-C-glycosylés des soies de maïs, tels que la maysin, l'apimaysin et la 3'-methoxymaysin, réduisaient la croissance des larves et que la puissance de l'inhibition de la croissance larvaire était directement liée au nombre de groupes hydroxyle attachés au cycle phényle (B ; Fig I-5) des flavones (Elliger 1980b). L'isoorientine serait une autre molécule potentiellement active sur la croissance des larves (Widstrom and Snook 1998). Ces flavones présentent donc un intérêt dans l'étude de l'inhibition de la croissance larvaire.

Dans la littérature des études métabolomiques le maïs, se basant sur la composition totale en composés phénoliques, en flavonoïdes, en anthocyanes et en activité oxydante, ont déjà été réalisées pour mettre en évidence l'effet de la maturité des soies de maïs et de leur variété (Sarepoua 2015) sur la composition phytochimique ou pour caractériser les différences de composition en fonction des organes de la plante (Dong 2014). La couleur de la soie et le nombre de jours après son émergence ont été comparés à ceux d'autres herbes médicinales (Žilić 2016). Différents hybrides de Serbie

(Maksimović 2005) ont été évalués pour leur activité antioxydante, certaines variétés plantées en Chine (Tian 2013) comme anti-cancéreux et d'autres comme antidiabétiques (Zhang 2016). Il est donc possible de mettre en évidence des différences d'activité entre variétés en analysant juste une partie de la composition moléculaire.

Les soies de maïs qui seront l'organe d'intérêt pour notre étude constituent une excellente source de composés bioactifs tels que les flavonoïdes, les saponines, les alcaloïdes, les tannins, les phytostérols, les terpénoïdes, les stéroïdes, les saccharides, les cérébrosides, l'allantoïne et les vitamines E et K (Bhaigiyabati 2011). Les soies de maïs contiennent également des huiles volatiles, des résines, des sucres, du mucilage (Milind and Isha 2013).

Dans la suite de ce chapitre, la composition moléculaire globale du maïs sera présentée puis l'étude sera focalisée sur la famille des flavonoïdes mise en évidence comme famille d'intérêt dans la littérature.

1.4. Composition moléculaire des composés non volatils du maïs et de la soie

Les métabolites secondaires proviennent de différentes voies de biosynthèse suivant leur famille moléculaire. La plupart des composés azotés sont synthétisés à partir d'acides aminés. Les terpénoïdes sont des dérivés du composé IPP (isopentényl-pyrophosphate) formé dans la voie de l'acide mévalonique ou par la voie du méthylérythritol phosphate (MEP). Quant aux composés phénoliques, ils sont produits par la voie de l'acide shikimique ou par celle du malonyl CoA. Suivant les voies de biosynthèse privilégiées, les plantes produisent différents types de métabolites secondaires. Différentes études ont déjà décrit des familles moléculaires ou des molécules présentes dans le maïs et également dans ses soies. Un bref récapitulatif de ces études est présenté.

1.4.1. Composés azotés

Les *Poaceae* comme le maïs contiennent de grandes quantités d'**acides hydroxamiques** suivant les variétés (Copaja 1991). Les molécules retrouvées dans le maïs sont des dérivés de 1,4-benzooxazine-3-one telles que 2-β-D-glucopyranosyloxy-4-hydroxy-7-methoxy-1,4-benzooxazin-3-one (DIMBOA-Glc), 2-β-D-glucopyranosyloxy-4-hydroxy-1,4-benzooxazin-3-one (DIBOA-Glc), 2-β-D-glucopyranosyloxy-4-hydroxy-7,8-dimethoxy-1,4-benzooxazin-3-one (DIM₂BOA-Glc). Les dérivés 1,4-benzooxazine-3-one sont principalement présents sous forme glycoside et ont été retrouvés dans les racines et les parties aériennes à des concentrations différentes suivant l'âge du plant et sa variété (Cambier 2000).

Glauser et al. (2011) ont montré que ces molécules avaient un rôle de défense lors d'une attaque d'insectes sur les plants de maïs. Tout d'abord, une grande quantité de 2-β-D-glucopyranosyloxy-4,7-diméthoxy-1,4-benzooxazin-3-one (HDMBOA-Glc) est produite quand la plante est attaquée, puis après la rupture des tissus par les herbivores, l'aglycone est libéré et a un effet dissuasif sur les insectes.

Des **auxines** (ou acide indole 3-acétique) sont présents dans les grains de maïs (Ueda and Bandurski 1969). L'acide indole 3-acétique est une hormone de croissance naturelle des plantes.

1.4.2. Composés terpéniques

La présence de quelques terpénoïdes a été décrite dans le maïs. Ainsi cinq **stéroïdes** ont été trouvés dans les soies de maïs, comme les molécules de stigmast-4-en-3-one, β -sitostérol, stigmastérol, stigmastone, 7- α -hydroxysitostérol (Suzuki 2007).

Les **sesquiterpènes** se retrouvent dans tous les organes du maïs hormis les inflorescences males et les soies (Köllner 2004). Leur abondance diffère suivant l'organe et l'étape de développement. Ils peuvent être répartis en cinq groupes suivant leur abondance. Le groupe A n'apparaît que lorsque les herbivores endommagent les plantules et se trouve principalement dans les feuilles. Le (*E*)- β -farnésène et le α -bergamotène sont les deux composés dominants dans ce groupe. Les principaux composés du groupe B sont le α -copaène, le germacrène D et le δ -cadinène. Ils sont principalement présents dans les semis et dans les enveloppes de plantes mûres. Dans les racines sont retrouvés les composés du groupe C constitué du β -bisabolène. Les feuilles et les cosses des plantes matures contiennent des sesquiterpènes des groupes D et E constitués des composés comme le β -bisabolène et des sesquiterpènes de type bergamotane pour le groupe D et des molécules comme le γ -cadinène, le δ -cadinène et le α -cadinène.

Les maïs, suivant leur variété, peuvent être constitués de caroténoïdes, dont le β -carotène, la zéaxanthine et la lutéine, qui modifient leurs couleurs (Kurilich and Juvik 1999).

1.4.3. Composés phénylpropanoïdes

1.4.3.1. Acides phénoliques et lignanes

Les **acides phénoliques** sont des dérivés de l'acide benzoïque et cinnamique. Le maïs contient différents acides phénoliques et dans les différentes parties de la plante. Ainsi, Liu et al. en ont retrouvé dans les feuilles (2012). Des dérivés d'acide hydroxycinnamique et des saccharides ferruliques ont été retrouvés dans les panicules (Allerdings 2006, Kim 2009). Trois dérivés de vanilline (vanilline, acide vanillique, acétovanillone) ont été décrits par Suzuki dans les soies (2007). Les soies de maïs sont également riches en composés phénoliques tels que l'acide p-coumarique et l'acide férulique, l'acide protocatéchique et des dérivés d'acides hydroxycinnamiques. Elles contiennent aussi l'acide maizenique, l'acide chlorogénique et des dérivés d'acide cinnamique (Milind and Isha 2013). Les acides phénoliques contenus dans les grains de maïs ont été reconnus comme des molécules impliquées dans le processus de résistance contre le charançon et le perce-grain (Classen 1990).

Les tiges renferment un **lignane** de type acide néo-olivique, deux flavonolignanes et les salcolines A et B (Jung 2015).

1.4.3.2. Flavonoïdes

La structure générale des flavonoïdes a été décrite dans le chapitre 1. Ce sont une classe de phénols largement distribués dans le règne végétal. Ils jouent un rôle structurel dans la plante et un rôle crucial dans son processus de reproduction. Ils permettent non seulement de protéger les plantes contre les rayons UV mais aussi de les défendre contre les insectes. C'est une des classes de molécules impliquées dans les réactions de stress de la plante et dans la coloration de la plante par l'intermédiaire des anthocyanes.

- Anthocyanes

Dans le maïs, certains **anthocyanes** se retrouvent dans tous les tissus de la plante sous forme glycosylée en position 3 (Halbwirth 2003), comme la cyanidine 3-glucoside. Alors que d'autres seront spécifiques à des organes, comme la pélargonidine 3-glucoside, qui n'est présente qu'au niveau du péricarpe et du tissu cicatriciel de la fleur pistillée, ou soie (Halbwirth 2003). Beaucoup d'anthocyanes sont présents dans les fleurs et les feuilles, dont la cyanidine 3-glucoside, mais aussi d'autres dérivés de cyanidine avec des motifs 6''-malonylglucoside et 3'',6''-dimalonylglucoside et des dérivés de péonidine avec les mêmes motifs de sucre (Fossen 2001).

Suivant la couleur des grains de maïs, en plus des caroténoïdes (terpènes), les anthocyanes peuvent participer à la couleur des grains (Žilić 2012). Ainsi les grains de maïs jaune ne contiennent pas d'anthocyanes tandis que les grains rouge foncé, bleu ou multicolore en contiennent.

- Flavonols

Dans le pollen de maïs, les principaux flavonoïdes retrouvés sont sous forme de quercétine et d'isorhamnétine mono et di-*O*-glycosylés. Il n'y pas de flavones, qui sont pourtant communs dans d'autres tissus du maïs (Ceska and Styles 1984).

- Flavones

Quelques aglycones de flavones ont déjà été isolés. La molécule de tricine a été retrouvée dans les tiges (Jung 2015) et les molécules de chrysoériol et de génistéine (isoflavone) ont été identifiées dans les soies (Suzuki 2007). Mais la plupart des flavones isolés sont des flavones glycosylés.

Les flavonoïdes isolés et identifiés dans les soies de maïs (Elliger 1980, Kim 2014, Hasanudin 2012, Liu 2011, Ren 2009 et 2013, Snook 1995, Suzuki 2003), dont la maysin, l'apimaysin, le 3-methoxymaysin sont récapitulés dans le tableau III-1 et figure III-2.

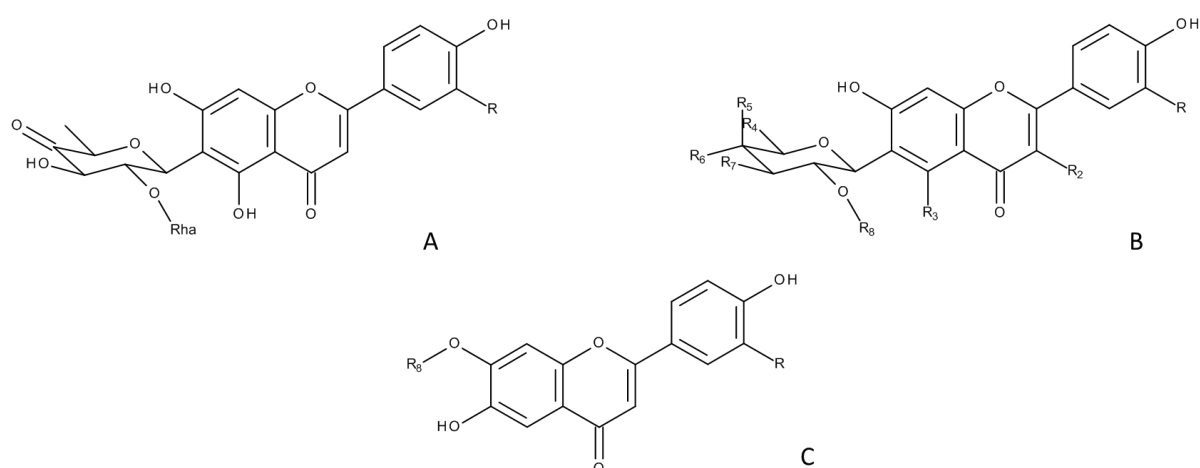


Figure III-2 : Squelettes des différents types de flavonoïdes identifiés dans les soies de maïs. (A) Dérivés de la molécule apimaysin. (B) Dérivés de la molécule apigénine 6-C-glycosylés. (C) Dérivés des molécules lutéoline ou chrysoériol 7-O-glycosylés.

Tableau III-1 : Flavonoïdes identifiés dans les soies de maïs.

Auteur (année)	Molécules	FigII-1	R	R2	R3	R4	R5	R6	R7	R8	M (g/mol)	FB
Snook (1995) Kim (2014) Elliger (1980) Hasanudin (2012)	Maysin 2''-O- α -L-rhamnosyl-6-C-(6-deoxy-xylo-hexose-4-ulosyl)luteolin 2''-O- α -L-rhamnosyl-6-C-quinovosylluteolin	A	OH							Rha	576	C ₂₇ H ₂₈ O ₁₄
Snook (1995) Elliger (1980)	Apimaysin	A	H							Rha	560	C ₂₇ H ₂₈ O ₁₃
Snook (1995) Hasanudin (2012) Liu (2011) Elliger (1980) Ren (2009 et 2013)	3'-Methoxymaysin ax-5-methane-3-methoxymaysin	A	OCH ₃							Rha	590	C ₂₈ H ₃₀ O ₁₄
Elliger (1980a)	6-C-glycosylflavone	A	OH							H	430	C ₂₁ H ₁₈ O ₁₀
Snook (1995)	ax-4''-OH-Maysin	B	OH	H	OH	CH ₃	OH	H	OH	Rha	578	C ₂₇ H ₃₀ O ₁₄
Snook (1995)	eq-4''-OH-Maysin	B	OH	H	OH	CH ₃	H	OH	OH	Rha	578	C ₂₇ H ₃₀ O ₁₄
Hasanudin (2012)	2''-O- α -L-rhamnosyl-6-C-fucosylluteolin	B	OH	H	OH	CH ₃	OH	OH	OH	Rha	594	C ₂₇ H ₃₀ O ₁₅
Liu (2011) Hasanudin (2012)	isoorientin-2''-O- α -rhamnoside	B	OH	H	OH	CH ₂ OH	H	OH	OH	Rha	594	C ₂₇ H ₃₀ O ₁₅
Snook (1995) Hasanudin (2012) Suzuki (2003) Ren (2009 et 2013)	ax-4''-OH-3'-Methoxymaysin	B	OCH ₃	H	OH	CH ₃	OH	H	OH	Rha	592	C ₂₈ H ₃₂ O ₁₄
Hasanudin (2012)	2''-O- α -L-rhamnosyl-6-C-fucosyl-3'-methoxyluteolin	B	OCH ₃	H	OH	CH ₃	OH	OH	OH	Rha	604	C ₂₈ H ₃₂ O ₁₅
Ren (2009 et 2013)	2-O- α -L-rhamnosyl-6-C-3-deoxyglucosyl-3-methoxyluteolin	B	OCH ₃	H	OH	CH ₂ OH	OH	H	H	Rha	592	C ₂₈ H ₃₂ O ₁₄
Ren (2009 et 2013) Hasanudin (2012)	7,4-dihydroxy-3-methoxyflavone-2-O- α -L-rhamnosyl-6-C-fucoside	B	OCH ₃	OH	H	CH ₃	OH	H	OH	Rha	592	C ₂₈ H ₃₂ O ₁₄
Suzuki (2003)	chrysoeriol-6-C- β -fucopyranoside	B	OCH ₃								446	C ₂₂ H ₂₃ O ₉
Ren (2009 et 2013) Hasanudin (2012)	6,4-dihydroxy-3-methoxyflavone-7-O-glucoside	C	OCH ₃							Glu	462	C ₂₂ H ₂₂ O ₁₁
Kim (2014)	luteolin 3'-methyl ether 7-glucuronosyl-(1->2)-glucuronide	C								di-Glucu	652	C ₂₈ H ₂₈ O ₁₈
Kim (2014)	luteolin 7-O-neohesperidoside	C								Glu-Rha	594	C ₂₇ H ₃₀ O ₁₅

Les flavonoïdes décrits dans les soies de maïs ont des motifs de flavones comme l'apigénine, la lutéoline, le chrysoériol et l'isoorientine. Ils sont C- ou O-glycosylés ou di-glycosylés, avec différents motifs de sucre, comme glucose, pyranose, rhamnose, fucose, acide glucuronique et quinovose avec une cétone.

Tous ces flavonoïdes peuvent donc intervenir dans l'inhibition de la croissance des larves, lors de l'ingestion des soies de maïs et ont un potentiel d'activités.

1.5. Activités des flavonoïdes

Les activités pharmacologiques des C-glycosylflavonoïdes sont presque similaires à celles des O-glycosylflavonoïdes et sont principalement axées sur les activités antioxydante, hépatoprotective, anti-inflammatoire, antivirale, anti-glycation et anti-hyperglycémique. Les C-glycosylflavonoïdes ont une meilleure stabilité à l'hydrolyse que les O-glycosylflavonoïdes (Xiao 2016).

Au vu de ce rapide état de l'art, il apparaît donc bien que les soies de maïs contiennent certaines molécules (principalement des flavonoïdes) permettant de diminuer la croissance des larves de *H. zea*. Néanmoins, afin d'obtenir un effet optimum lors de l'utilisation du maïs en tant que plante piège pour la culture des tomates, une connaissance plus approfondie du contenu moléculaire en flavonoïdes de différentes variétés de soies de maïs s'avère nécessaire et pourrait apporter une aide précieuse pour sélectionner la variété de maïs la plus efficace.

C'est dans ce contexte qu'une étude métabolomique ciblée des soies de maïs a été menée par UHPLC/MRM MS, en se focalisant sur les flavonoïdes, afin de mettre en évidence les molécules différenciantes entre variétés et d'établir une potentielle corrélation entre les différences de composition des différentes variétés étudiées et les résultats observés sur l'inhibition de la croissance des larves de *H. zea* lors de l'étude terrain réalisée en parallèle par le CIRAD sur ces mêmes variétés.

2. Etude métabolomique ciblée des soies de maïs par UHPLC/MS

L'étude métabolomique que nous avons mise en place sur les flavonoïdes des soies de maïs s'est déroulée selon les différentes étapes schématisées dans la figure III-3. Après avoir reçu les soies de maïs collectées et conservées suivant le protocole établi par le CIRAD, ces soies ont été extraites, puis analysées par UHPLC/HRMS et UHPLC/MRM MS, selon les méthodes que nous avons au préalable développées. Les résultats ont permis soit de cartographier la composition moléculaire de l'extrait et caractériser les flavonoïdes contenus dans les extraits, soit de réaliser des traitements statistiques. Ces derniers ont permis de mettre en évidence les groupes d'échantillons et les métabolites significatifs.

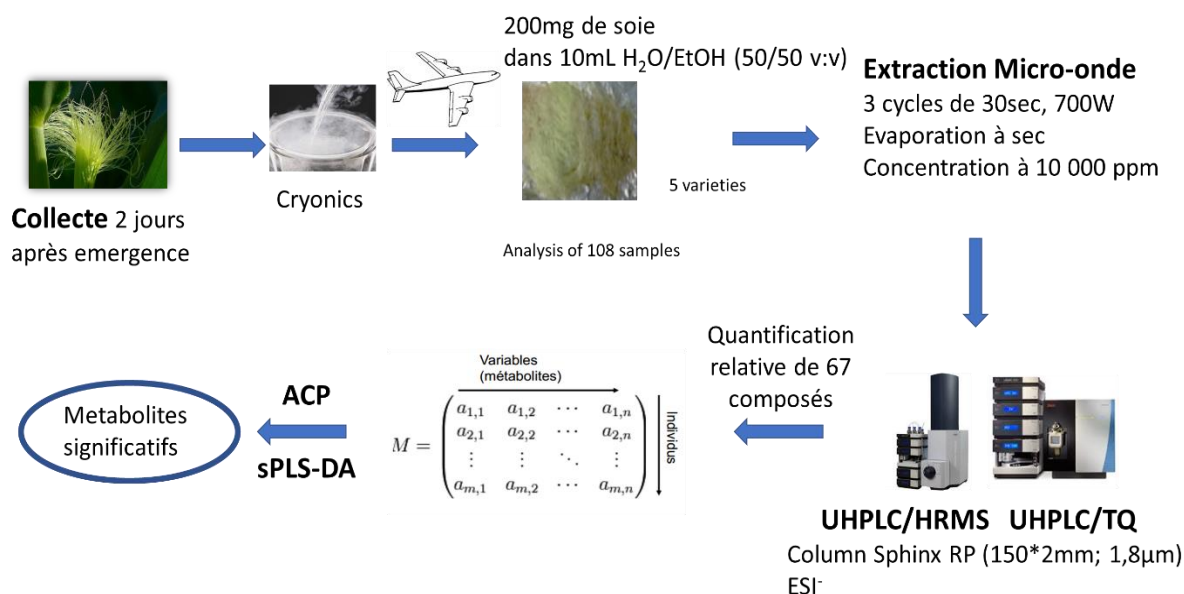


Figure III-3 : Schéma des différentes étapes de l'étude métabolomique appliquée aux soies de maïs.

2.1. Sélection et récolte du matériel végétal

Le CIRAD a cultivé en Martinique 5 variétés commerciales de maïs doux dont le nom du fournisseur est indiqué entre parenthèses : Golden bantam (Germinance), Sugar Jean (Agrinova), Java (Technisem), Nova (Agri-semences) et Garrison (Syngenta). Ces variétés ont été sélectionnées d'après des études bibliographiques et se sont révélées toutes attrayantes pour le ravageur *H. zea* (Rhino 2016).

Ainsi, la variété Golden Bantam est sensible à *H. zea*. La variété Java favorise une ponte abondante des œufs d'*H. Zea* contrairement à la variété Sugar Jean. Alors que le poids des larves de *H. zea* est faible

pour la variété Java et moyenne pour toutes les autres variétés. Les soies de ces 5 variétés ont des textures et des couleurs différentes (Fig. III-4) et donc potentiellement des molécules ou des quantités de molécules différentes.



Figure III-4 : Photographie des 5 variétés de soies de maïs après réception à l'ICOA.

Les soies de maïs ont été récoltées 2 jours après leur émergence de l'épi et tout de suite plongées dans de l'azote liquide pour stopper une possible évolution de leur profil métabolomique. Elles ont ensuite été envoyées à l'ICOA à Orléans par avion dans des bacs de carboglace et stockées dès leur arrivée à -20°C. Ainsi 5 variétés de soies de maïs codées alpha-numériquement (Gb : Golden Bantam ; Sj : Sugar jean ; Jav : Java ; Nov : Nova ; Gar : Garrison) ont été reçues, avec 6 à 8 lots par variété, soit un total de 36 soies de maïs (Tableau III-2). Ces échantillons ont été analysés trois fois.

Tableau III-2 : Echantillonnage des soies de maïs.

Variété	Golden Bantam	Sugar jean	Java	Nova	Garrison
Acronyme	Gb	Sj	Jav	Nov	Gar
nbre de lots	7	6	8	8	7
nbre de répliquats d'analyse	3	3	3	3	3

2.2. Préparation des échantillons

2.2.1. Méthode d'extraction

Sur la base des travaux antérieurs du laboratoire (Michel 2011), l'extraction assistée par micro-ondes a été choisie pour extraire les faibles quantités d'échantillons de soies de maïs dont nous disposions (moins de 500 mg pour chaque lot) et ce, pour plusieurs raisons : extraction efficace des flavonoïdes, facilité d'utilisation de l'appareil, possibilité de travailler sur de petites quantités de végétal et de solvant, temps réduit d'extraction et possibilité d'utiliser un carrousel pouvant contenir 12 réacteurs permettant de ce fait de réaliser simultanément un plus grand nombre d'extraits. Afin d'obtenir une empreinte moléculaire la plus riche possible en flavonoïdes glycosylés, un mélange de solvants polaires (Eau/Ethanol) adapté à la polarité des molécules a également été sélectionné.

Ainsi 200 mg de chaque soie ont été pesés et extraits avec 10 mL d'un mélange Eau/EtOH (50/50) (Riffault 2014) pour permettre d'obtenir un maximum de flavonoïdes glycosylés (Annexe 1). Le

rendement d'extraction est en moyenne de $7,0 \pm 1,5$ %. Cette extraction a été réalisée à une puissance de 700W en 3 cycles de 30 sec.

2.2.2. Etude de répétabilité de la méthode d'extraction

Une première étude de répétabilité de l'étape d'extraction (Tableau III-3 et Fig. III-5) a été menée sur trois individus des deux variétés de maïs Garrison et Java (Gar 1-03 ; Gar 1-04 et Gar 1-05 pour la variété Garrison et Jav 1-04 ; Jav 1-05 et Jav 5-05 pour la variété Java). Pour évaluer la variabilité de la méthode, le même protocole d'extraction a été reproduit 3 fois de façon distincte (MO1 ; MO2 et MO3) pour trois individus (Jav1-04 ; Jav5-05 et Gar 1-05). Par ailleurs, tous les extraits obtenus pour les 6 individus ont été analysés avec trois répétitions d'injection (A, B et C).

Tableau III-3 : Nomenclature des différents essais de répétabilité.

Variété	Gar			Jav		
Lot	1-03	1-04	1-05	1-04	1-05	5-05
répétition extraction			MO1 ; MO2 ; MO3	MO1 ; MO2 ; MO3		MO1 ; MO2 ; MO3
répétition injection	A ; B ; C	A ; B ; C	A ; B ; C	A ; B ; C	A ; B ; C	A ; B ; C

Pour évaluer l'influence de l'extraction micro-ondes, mais également la variabilité des individus au sein d'une même variété, une analyse en composante principale (Fig. III-14) a été réalisée sur l'ensemble de ces données, après avoir extrait et analysé chacun de ces échantillons en spectrométrie de masse (Annexe 6).

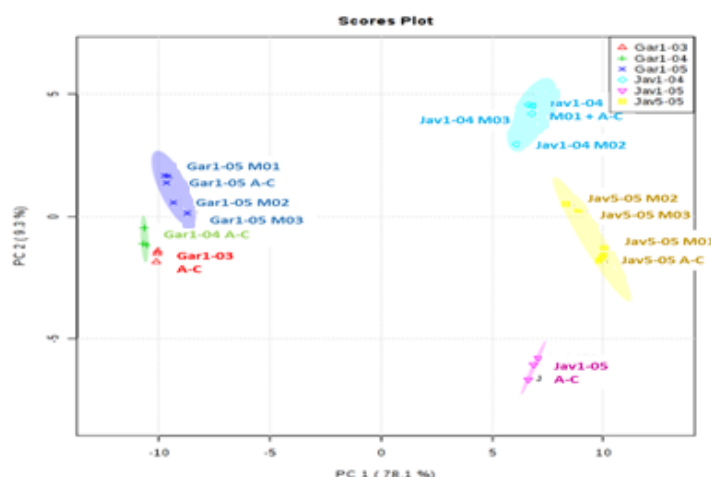


Figure III-5 : Répartition des individus (1-05 ; 1-04 ; 1-03 ; 5-05) en fonction de la répétabilité d'extraction (MO1, MO2, MO3) et de la variété (Garrison (Gar) et Java (Jav)).

Ainsi cette représentation permet de mettre en évidence la répartition des individus suivant les différentes analyses (injection, extraction, variété). La variabilité d'injection (A à C dans chaque lot) est très peu distinguée dans cette analyse en composante principale. La dispersion engendrée par des extractions différentes d'un même lot est faible (MO1 à 3 en bleu et jaune) tandis que la variabilité des

lots/individus (variétés Garrison (Gar 1-05 en violet, Gar 1-04 en vert et Gar 1-03 en rouge) à gauche de l'ACP et Java (Jav 1-04 en bleu, Jav 5-05 en jaune et Jav 1-05 en rose) à droite de l'ACP) est plus importante. Comme l'indique Summer (2007), il est préférable d'avoir des répliques biologiques à des répliques analytiques, car la variance biologique dépasse presque toujours la variance analytique. Cet écart entre individu se manifeste essentiellement sur l'axe 2 de l'ACP qui ne représente que 9,3% de variance. Cette variabilité individu est sûrement due à des répartitions de molécules différentes d'un pied de maïs à un autre, suivant l'exposition et l'emplacement sur le terrain. La variation observée entre les individus d'une même variété reste néanmoins faible et ne masque pas la variabilité entre variétés qui est la plus importante. Chacune des deux variétés peut être clairement distinguée de l'autre selon l'axe 1 avec 78,1% de variance.

Au vu de l'ensemble de ces résultats, il apparaît que la répétabilité de la méthode de préparation d'échantillons peut être considérée comme acceptable et la méthode est conservée dans ces conditions pour la suite des analyses. Néanmoins, pour mieux prendre en compte la variabilité observée entre individus d'une même variété, et ainsi minimiser les erreurs d'interprétation des résultats, 6 à 8 lots par variété seront extraits et analysés pour les études plus approfondies.

L'étude métabolomique sur les flavonoïdes de soies de maïs a donc été menée à partir des résultats obtenus lors de l'analyse UHPLC/MS de 6 à 8 lots collectés (Fig. III-6) pour chacune des 5 variétés de maïs. Chaque lot correspond à la soie d'un épi de maïs. Ainsi l'étude comprenant 5 variétés, il a été collecté et traité 36 soies de maïs.

Etant donné par ailleurs que les répétitions d'extraction sur les 6 premiers individus ont montré des écarts acceptables et que la quantité de matériel végétal à notre disposition était faible, chaque lot n'a été extrait qu'une seule fois. Toutefois, chaque extrait a été injecté 3 fois (Fig. III-6) à des moments différents de la séquence, afin de prévenir d'un problème éventuel de l'appareillage. Ainsi 108 analyses UHPLC/MS ont été effectuées.

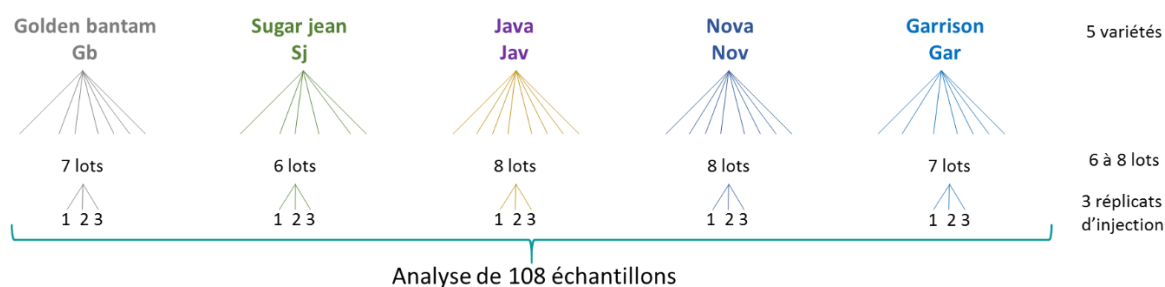


Figure III-6 : Schéma de la composition de l'échantillonnage.

Une fois la méthode d'extraction validée, la méthode d'analyse en UHPLC/MS a été développée, afin d'avoir une séparation convenable dans un temps raisonnable, sans perdre d'information. En effet, dans les plantes, des molécules isomères de position sont souvent retrouvées. Il faut donc pouvoir les séparer pour identifier la présence de ces molécules différentes. Etant donné que la méthode chromatographique est couplée à un appareil de spectrométrie de masse, si des molécules de masses différentes sont coéluees, elles peuvent le rester tant que leur ionisation en source ne s'en trouve pas perturbée, car elles pourront être distinguées via leur masse.

2.3. Optimisation des conditions chromatographiques en vue de l'analyse UHPLC/MS/MS des flavonoïdes dans les extraits de soies de maïs

Pour réaliser l'étude métabolomique ciblée, la chromatographie liquide couplée à un spectromètre de masse triple quadripôle en mode MRM a été choisie pour analyser les molécules. Les flavonoïdes ont été séparés sur un système de chromatographie de partage à polarité de phases inversée. Le suivi des molécules a été réalisé en suivant une transition spécifique (ion parent -> ion fragment) pour chaque flavonoïde. Puis une quantification relative par rapport à un standard de flavonoïde en étalonnage externe a été réalisée afin de permettre la comparaison des concentrations relatives des molécules d'une variété à une autre.

2.3.1. Choix de la phase mobile

D'après la publication de Cajka et Fiehn (2016) et l'expérience du laboratoire sur l'analyse des flavonoïdes glycosylés (Michel 2011), la phase mobile sélectionnée est constituée d'un mélange de deux solvants : eau et acétonitrile en proportions variables au cours de l'analyse. Une optimisation du gradient a été réalisée pour séparer les molécules de l'extrait de soie de maïs au mieux, afin que le détecteur du spectromètre de masse n'analyse pas en même temps trop de composés. Une attention particulière a été portée à la séparation des molécules de même masse, au cas où elles ne pourraient pas être distinguées par des transitions MRM différentes.

Deux conditions d'additifs compatibles avec la spectrométrie de masse ont été testées pour la phase mobile (Annexe 2) : acide formique 0,1%, acétate d'ammonium à 5 mM et comparées aux conditions sans additif. L'ajout d'acide formique a permis pour certaines molécules (telles que rutine, peltatoside ou orientine), une réponse plus intense et de meilleures efficacités et résolutions de pics qu'en l'absence d'additif. L'ajout d'acétate d'ammonium conduit globalement à une élution plus rapide des molécules et à de moins bonnes conditions d'ionisation pour certaines d'entre elles.

2.3.2. Choix de la colonne

Deux phases stationnaires en silice bifonctionnalisée C18 et avec des greffages secondaires ont été évaluées pour la séparation des composés moyennement polaires que sont les flavonoïdes glycosylés des soies de maïs. Il s'agit de deux colonnes Nucleodur C18 de chez Macherey-Nagel, une nommée Sphinx RP et l'autre Pyramid (Fig. III-7).

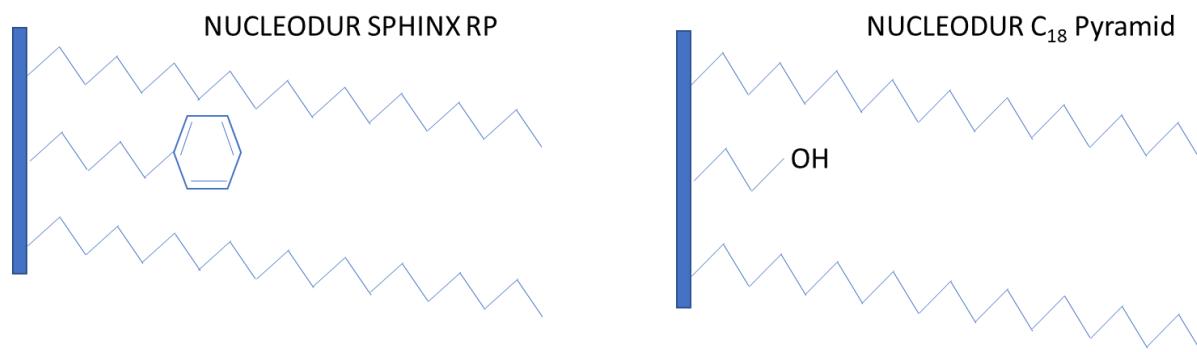


Figure III-7 : Structures des greffons des deux types de colonnes Nucleodur testés.

Ainsi ces deux colonnes étant des colonnes C18, elles permettent d'avoir des interactions majoritaires hydrophobes. La colonne Nucleodur Sphinx RP (Macherey-Nagel) possède en plus de chaînes C18 des groupements phényle greffés et intercalés sur la phase stationnaire (Fig. III-7), ce qui permet d'établir également des interactions π - π avec les molécules. Cette colonne est donc intéressante pour renforcer les interactions avec les molécules ayant des noyaux aromatiques tels que les flavonoïdes. Macherey-Nagel (Fig. III-8) a montré qu'il est alors possible de séparer sur ce type de support des flavonoïdes isomères de position, comme la fisétine (3) et le kaempférol (5) (Annexe 7) qui ne diffèrent que par la position d'un hydroxyle. En ce qui concerne la colonne Nucleodur C18 Pyramid (Macherey-Nagel), elle possède un end-capping polaire (Fig. III-7) offrant des possibilités d'interactions polaires supplémentaires avec les solutés tout en conservant le caractère hydrophobe majoritaire de la phase. Ce type de support renforce la sélectivité vis-à-vis des molécules polaires, comme les flavonoïdes avec beaucoup de groupements hydroxyle et/ou sucre. Ces deux colonnes ont donc un potentiel pour séparer les flavonoïdes et ont été testées pour effectuer la séparation des composés des extraits de soies de maïs.

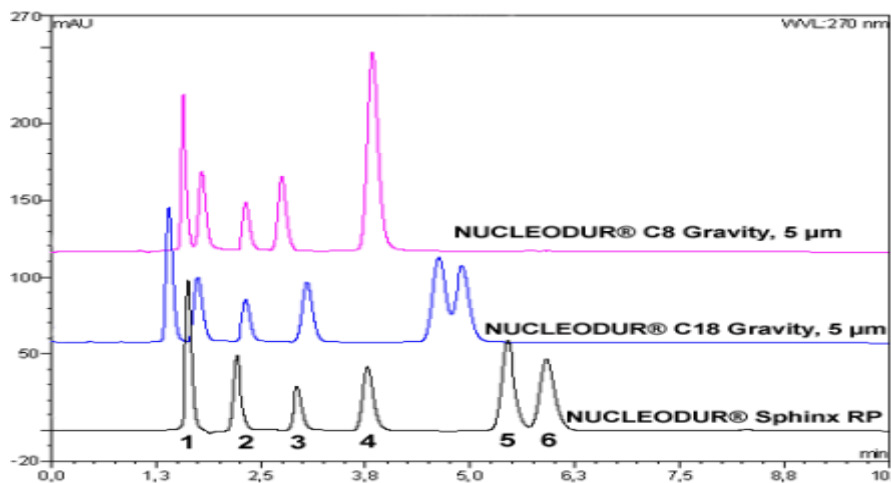


Figure III-8 : Séparation de 6 flavonoïdes standards (1 : Catéchine ; 2 : Rutine ; 3 : Fisétine ; 4 : Quercétine ; 5 : Kaempférol ; 6 : Isorhamnétine) sur trois colonnes Nucléodur® différentes. Phase mobile Eau/MeOH (40/60 v/v) ; débit 1 mL/min ; Température colonne 30°C ; Volume d'injection de 3 μ L ; Détection UV à 270 nm. D'après la note d'application Macherey-Nagel 119830.

Ainsi la figure III-9 montre la comparaison de la séparation de quelques molécules de l'extrait obtenue avec ces deux colonnes chromatographiques. On observe que ces colonnes ont des sélectivités

différentes avec des ordres de rétention différents. Le pic 5 (575->411) est plus retenu sur la colonne Nucleodur Pyramid, ainsi que le pic 4 (487 ->341) mais dans une moindre mesure. Pour les autres composés, les temps de rétention sont similaires et les symétries de pic convenables. Pour le pic 6, une forme non gaussienne plus prononcée sur le support Nucleodur Sphinx que sur le support Pyramid laisse supposer une coélution de plusieurs solutés de même transition. D'après le spectre de masse associé, l'attribution de ce pic correspondrait à la methoxymaysin (Fig. III-2A), un des composés d'intérêt dans cette étude. Limiter la déformation de ce pic s'avère nécessaire pour ne pas diminuer l'exactitude de la quantification de ce composé.

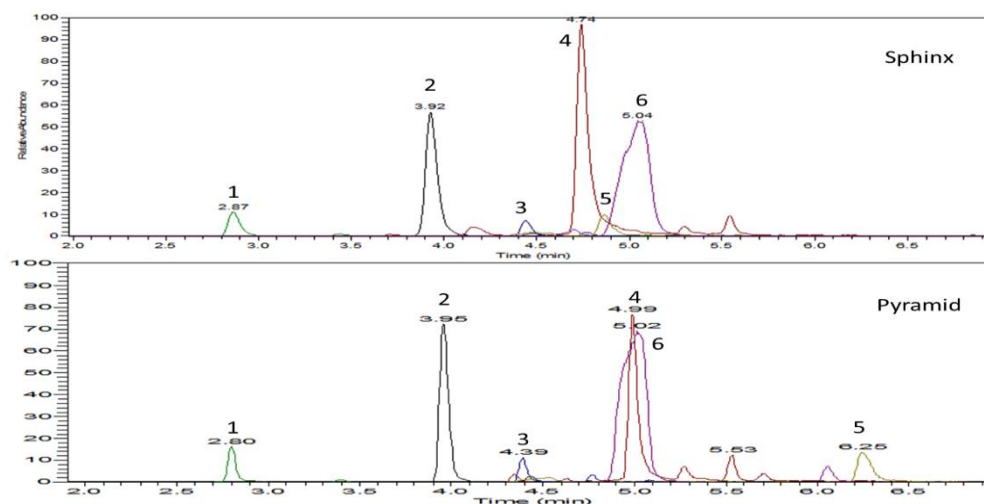


Figure III-9 : Analyse UHPLC/MRM MS de flavonoïdes ((1) 563 -> 473 ; (2) 461 -> 299 ; (3) 591 -> 487 ; (4) 487 -> 341 ; (5) 575 -> 411 ; (6) 589 -> 351) contenus dans un extrait de soies de maïs sur une colonne Nucleodur Sphinx RP et une colonne Nucleodur Pyramid. Colonnes 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% d'acide formique en gradient ; débit de 0,6 mL/min ; T° four à 60°C ; volume d'injection de 3 μ L.

Au vu de ces résultats et pour la suite de notre étude, c'est la colonne Nucleodur Pyramid que nous avons retenue comme phase stationnaire pour la séparation chromatographique afin d'être dans des conditions plus favorables pour une quantification par UHPLC/MS. Globalement de meilleurs résultats et une meilleure séparation ont été obtenus pour l'ensemble des molécules contenues dans l'extrait de soie en utilisant cette colonne. La coélution obtenue avec la colonne Pyramid pour les pics 4 et 6 n'est pas gênante dans ces conditions, car l'étude sera faite en mode MRM, donc les pics seront séparés par leur différence de masse et de fragment.

2.3.3. Choix du diamètre de particules du support chromatographique

Il a été évalué l'apport de la diminution de la taille des particules en passant d'une granulométrie moyenne de 5 μ m à 1,8 μ m (Fig. III-10) sur la résolution d'un mélange de six standards de flavonoïdes (Shaftoside (1) : Apigénine 6-C-glucoside-8-C-arabinoside ; Orientine (2) : Lutéoline 8-C-glucoside ; Peltatoside (3) : Quercétine 3-O-arabinoglucoside ; Vitexine 2''-O-rhamnoside (4) : Apigénine 8-C-glucoside-2''-O-rhamnoside ; Rutine (5) : Quercétine 3-O-rhamnoside-7-O-glucoside ; Phloridzine (6) :

Phlorétine-2'-*O*-glucoside). Ces standards ont été sélectionnés suivant des critères de structures moléculaires. En effet, il a été choisi aussi bien des standards de flavonoïdes *C*- que *O*-glycosylés en mono ou di-glycosylés.

Ainsi comme attendu (Walter and Andrews 2014), la diminution de la taille de particules a permis d'augmenter l'efficacité et la résolution des pics (facteur 2,5), et de diminuer le temps d'analyse en moyenne d'un facteur 2 (Figure III-10). C'est la granulométrie de 1,8 µm qui sera retenue pour la suite de notre étude.

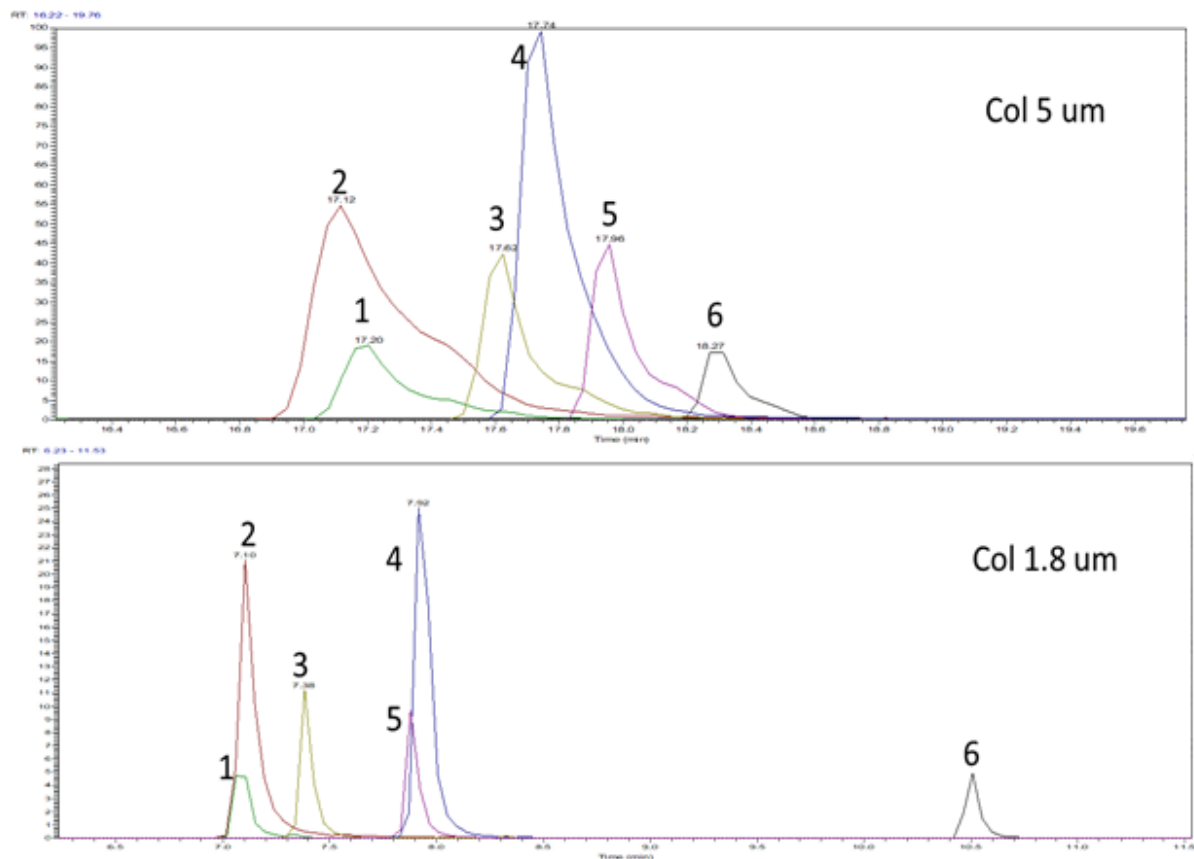


Figure III-10 : Effet de la réduction du diamètre des particules du support chromatographique sur la séparation d'un mélange de six standards de flavonoïdes (Shaftoside (1) : Apigénine 6-*C*-glucoside-8-*C*-arabinoside ; Orientin (2) : Lutéoline 8-*C*-glucoside ; Peltatoside (3) : Quercétine 3-*O*-arabinoglucoside ; Vitexine 2''-*O*-rhamnoside (4) : Apigénine 8-*C*-glucoside-2''-*O*-rhamnoside ; Rutine (5) : Quercétine 3-*O*-rhamnoside-7-*O*-glucoside ; Phloridzine (6) : Phlorétine-2'-*O*-glucoside). Colonne Nucleodur Sphinx RP 150 x 2 mm (1,8 et 5 µm). Phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% d'acide formique en gradient ; débit de 0,6 mL/min ; T° four à 60°C ; volume d'injection de 3 µL. Détection MS en mode d'acquisition MRM.

2.3.4. Choix de la température de la colonne

Comme Grata et al. l'ont montré en 2009, un profilage de métabolites d'extrait de plante à température plus élevée que la température ambiante peut conduire à de meilleures conditions d'analyse, mais cela dépend du type d'échantillon en présence. Etant donné que les colonnes Nucleodur supportent une élévation de la température, deux températures 30 et 60°C (Fig. III-11) ont été testées pour observer le gain en temps d'analyse et en efficacité de pics.

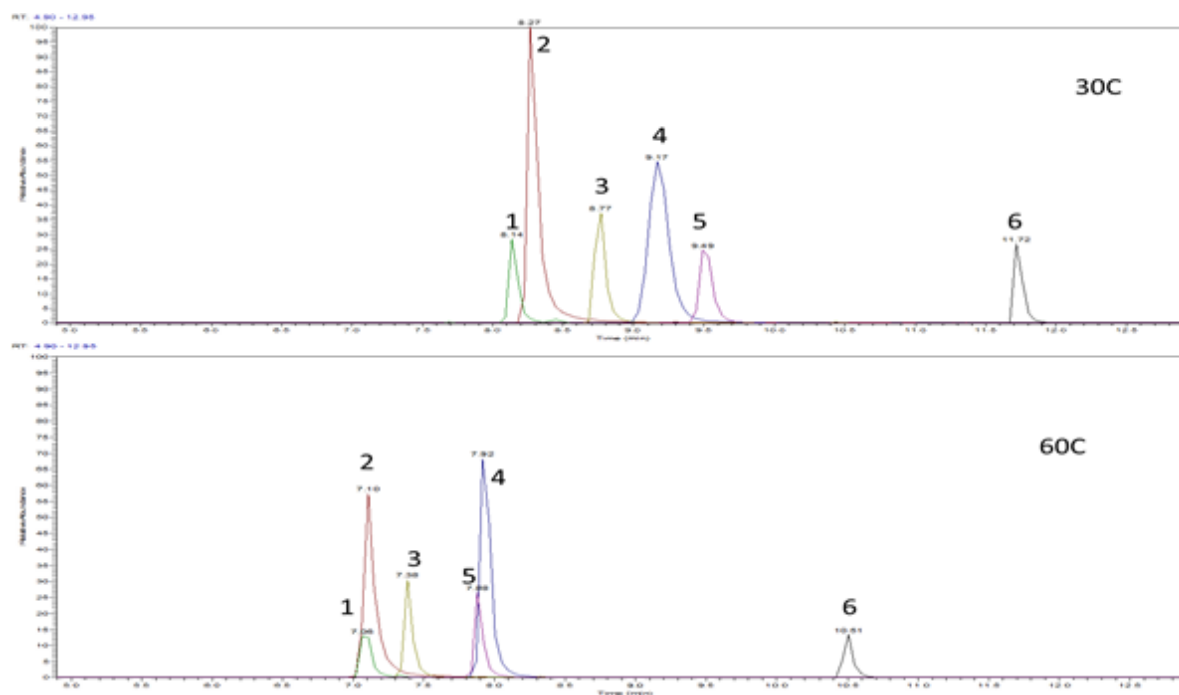


Figure III-11 : Effet de la température de la colonne sur la séparation d'un mélange de six standards de flavonoïdes (Shaftoside (1) : Apigénine 6-C-glucoside-8-C-arabinoside ; Orientin (2) : Lutéoline 8-C-glucoside ; Peltatoside (3) : Quercétine 3-O-arabinoglucoside ; Vitexine 2''-O-rhamnoside (4) : Apigénine 8-C-glucoside-2''-O-rhamnoside ; Rutine (5) : Quercétine 3-O-rhamnoside-7-O-glucoside ; Phloridzine (6) : Phlorétine-2'-O-glucoside). Colonne Nucleodur Sphinx RP 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% d'acide formique en gradient ; débit de 0,6 mL/min ; volume d'injection de 3 μ L. Détection MS en mode d'acquisition MRM.

Comme attendu, en augmentant la température à 60°C, le temps de rétention des standards diminue et les pics chromatographiques sont plus fins. Malgré la perte de résolution observée à 60°C pour la séparation du mélange, le suivi des transitions d'ions par MRM permet toujours de distinguer les solutés. C'est donc cette température qui a été choisie.

2.3.5. Optimisation du débit de la phase mobile

Afin de diminuer le temps total d'analyse (Fig. III-12), deux débits de phase mobile : 0,3 et 0,6 mL/min ont été testés pour analyser le même mélange de standards.

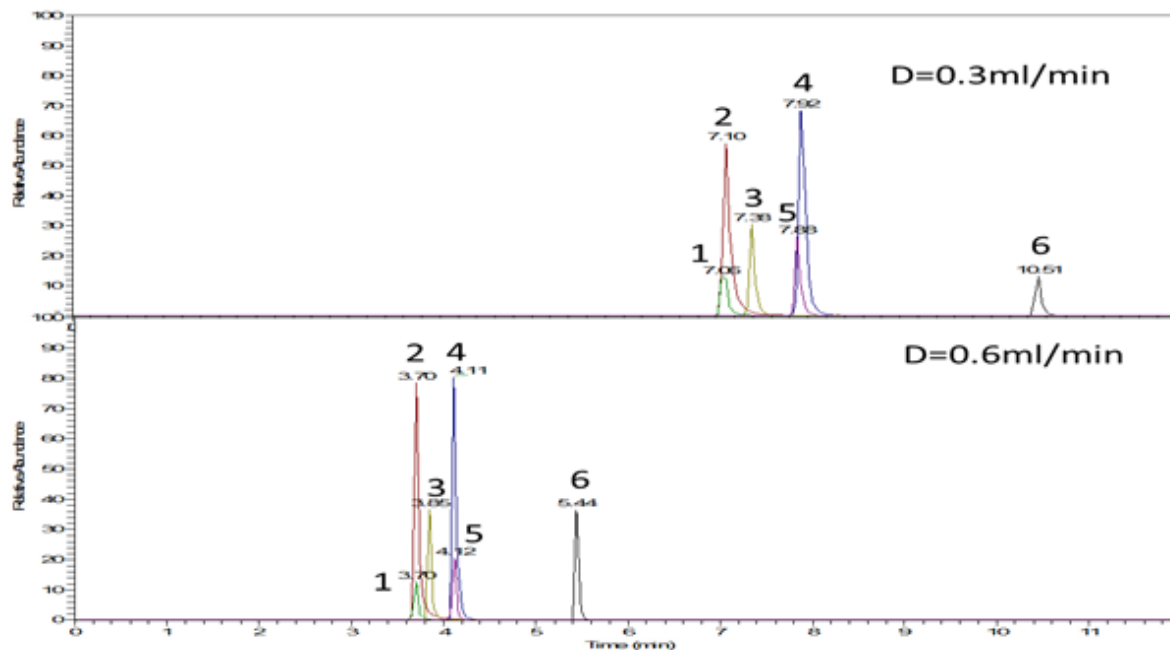


Figure III-12 : Effet du débit de la phase mobile (0,3 et 0,6 mL/min) sur la séparation d'un mélange de six standards de flavonoïdes (Shaftoside (**1**) : Apigénine 6-C-glucoside-8-C-arabinoside ; Orientin (**2**) : Lutéoline 8-C-glucoside ; Peltatoside (**3**) : Quercétine 3-O-arabinoglucoside ; Vitexine 2''-O-rhamnoside (**4**) : Apigénine 8-C-glucoside-2''-O-rhamnoside ; Rutine (**5**) : Quercétine 3-O-rhamnoside-7-O-glucoside ; Phloridzine (**6**) : Phlorétine-2'-O-glucoside). Colonne Nucleodur Sphinx RP 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% d'acide formique en gradient ; T° four à 60°C ; volume d'injection de 3 μ L. Détection par MS en mode d'acquisition MRM.

Comme attendu le fait d'augmenter le débit d'un facteur deux, permet de diviser par deux le temps d'analyse tout en augmentant un peu l'intensité des pics ce qui permet d'être plus sensible sans détériorer la séparation (Figure III-12).

En conclusion de ce paragraphe, **les conditions chromatographiques optimisées retenues** pour l'analyse des flavonoïdes contenues dans les extraits de soies de maïs sont donc les suivantes : Colonne : Nucleodur Pyramid (150 x 2 mm, 1,8 μ m) ; Phase mobile : mélange d'eau et d'acétonitrile contenant 0,1% d'acide formique en gradient d'élution ; Débit de la phase mobile : 0,6 mL/min ; Température de la colonne 60°C. (Annexe2).

Une fois toutes ces optimisations effectuées, une analyse en UHPLC/HRMS a été effectuée sur un mélange d'extraits des cinq variétés de maïs étudiées pour avoir une empreinte globale de l'ensemble des composés que peuvent contenir ces extraits de soie de maïs quelle que soit la variété. Cette étude permettra également de proposer une identification des flavonoïdes contenus dans les extraits.

2.4. Sélection des composés d'intérêt à suivre pour l'approche métabolomique ciblée des soies de maïs

Afin de pouvoir sélectionner les flavonoïdes, la composition de l'extrait de soie de maïs doit être connue. Pour cela, une première analyse UHPLC/HRMS/MS d'un mélange des extraits des différentes variétés est menée afin de détecter un grand nombre de composés sans a priori et d'obtenir une information de masse exacte, de formule brute et de schéma de fragmentation des composés.

2.4.1. Analyse totale en UHPLC/HRMS en vue de la détection d'un maximum de composés

L'analyse en UHPLC/HRMS d'un mélange d'extraits de soies de maïs fournit l'empreinte chromatographique suivante (Fig. III-13), obtenue avec la source d'ionisation électrospray en mode d'ionisation négatif (Annexe 2) :

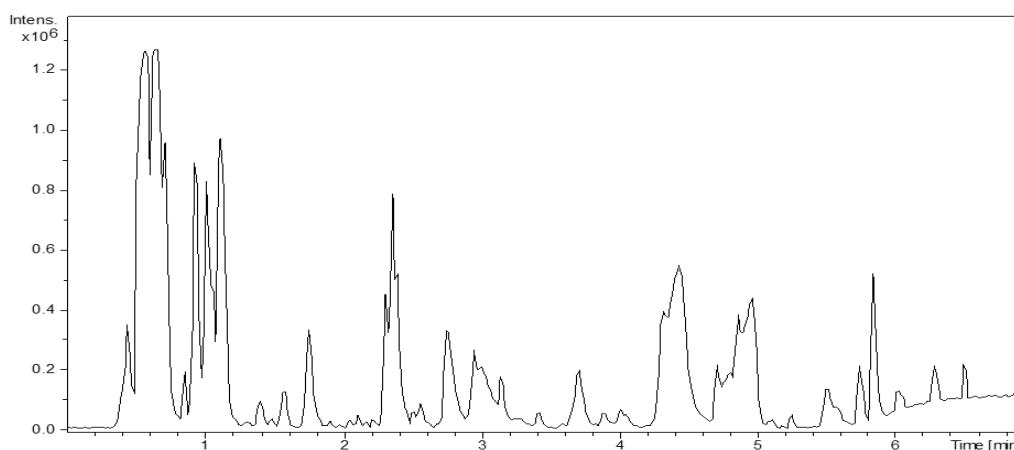


Figure III-13 : Base Peak Chromatogram (BPC) d'un mélange de soies de maïs obtenu en UHPLC/HRMS. Colonne Nucleodur Pyramid 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% et 0,08% d'acide formique respectivement en gradient ; T° four à 60°C ; débit de 0,6 mL/min ; volume d'injection de 3 μ L.

Ce mode d'analyse par spectrométrie de masse à haute résolution permet d'obtenir des propositions de formule brute pour chaque pic moléculaire détecté avec une précision en masse inférieure à 2 ppm pour les composés majoritaires et 5 ppm pour ceux qui ont une intensité faible. Une acquisition dépendante des données (DDA), dite autoMS² a également été effectuée. Cette analyse a été réalisée en fixant un seuil d'intensité minimum de 1 000 qui déclenche une acquisition des ions précurseurs pendant 400 ms, puis deux balayages MS² de 500 ms sont réalisés. Pendant chaque balayage, un ion précurseur est isolé, fragmenté et ses ions produits sont détectés. Une fenêtre d'exclusion dynamique de 7 sec permet de s'assurer que les composés récemment analysés par MS² ne sont pas de nouveau fragmentés (Annexe 2). Cette analyse permet d'obtenir les spectres de fragmentation de chaque

molécule, d'avoir une proposition de formule brute pour l'ion précurseur, mais aussi pour les ions fragments produits et ainsi de pouvoir proposer une identification des flavonoïdes de l'extrait.

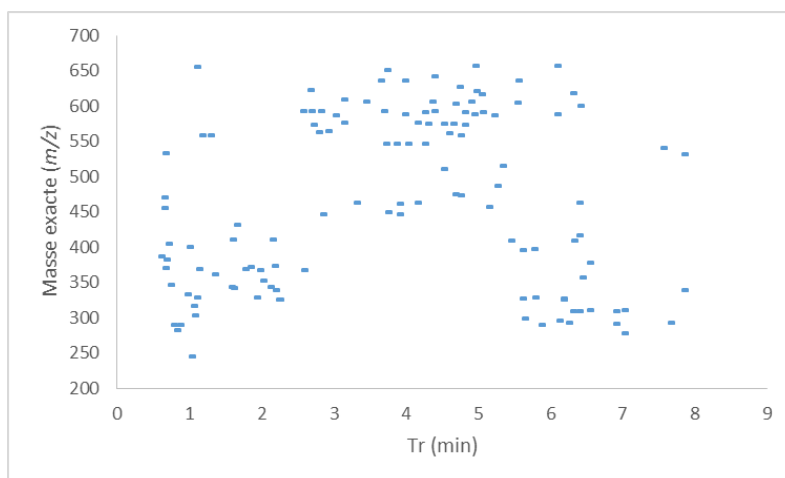


Figure III-14 : Représentation des masses exactes de l'ensemble des composés détectés dans les extraits de soies de maïs en fonction de leur temps de rétention. Colonne Nucleodur Pyramid 150 x 2 mm, 1,8 μm ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% et 0,08% d'acide formique respectivement ; T° four à 60°C ; débit de 0,6 mL/min ; volume d'injection de 3 μL .

La figure III-14 illustre la répartition des masses des composés au cours du temps. Beaucoup de composés de masses différentes sortent en tout début d'analyse vers les 1 min. L'étude de ces composés sera difficilement exploitable car leur rétention est trop faible et ils sont trop différents entre eux. Un deuxième groupe de composés se distingue dans la zone de rétention autour des 2 min et dans une plage de rapports de masse m/z entre 300 et 400 u, ces composés pourraient correspondre à des acides phénoliques et des alcaloïdes (comme DIMBOA). De nombreux composés sont élués ensuite entre 3 et 5 min avec des rapports de masse m/z autour de 600 u, ces composés pourraient correspondre à des flavonoïdes di-glycosylés. Et enfin les composés les plus retenus pourraient correspondre à des mono-glycosylés de rapport de masse m/z vers 450 u et des génines avec des rapports de masse m/z entre 250 et 300 u.

Dans les mêmes conditions d'analyse, un mélange de standards a également été fragmenté pour confirmer le schéma de fragmentation des flavonoïdes glycosylés.

2.4.2. Etude de fragmentation de quelques flavonoïdes glycosylés standards

2.4.2.1. Comparaison de la fragmentation avec la littérature

Etant donné que la fragmentation de composés connus peut être différente suivant l'appareil sur lequel elle est effectuée en raison de différence de géométrie des sources et des analyseurs d'un constructeur à l'autre, il peut y avoir des chemins de fragmentation préférentiels et des ratios d'ions

différents (Wolfender 2000). Par conséquent, une étude de fragmentation (Annexe 3) de cinq standards (Peltatoside : Quercétine 3-*O*-arabinoglucoside ; Orientine : Lutéoline 8-*C*-glucoside ; Homoorientine : Lutéoline 6-*C*-glucoside ; Schaftoside : Apigénine 6-*C*-glucoside-8-*C*-arabinoside ; Vitexine 2''-*O*-rhamnoside : Apigénine 8-*C*-glucoside-2''-*O*-rhamnoside) de flavonoïdes glycosylés est réalisée sur l'appareil de l'ICOA utilisé pour faire cette étude métabolomique et confrontée à la littérature. Ces standards représentent aussi bien des flavonoïdes *O*-glycosylés que des *C*-glycosylés, des mono et des di-glycosylés et avec un exemple d'isomérisation de position du sucre. Dans la section du chapitre 1, une partie de l'état de l'art de la fragmentation des flavonoïdes glycosylés a été présentée à partir de la bibliographie (Abad-Garcia 2008 et 2009, Ferreres 2007 et 2008, Figueirinha 2008, Vukics and Guttman 2010). Cette étude sur les standards connus permet ainsi de vérifier les types de rupture obtenus sur le système de spectrométrie de masse à notre disposition (Maxis qTOF (Bruker) avec une source électrospray).

Concernant la fragmentation du *flavonoïde mono-O-glycosylé-O-glycosylé* (peltatoside), on retrouve bien en mode d'ionisation négatif directement la perte des deux sucres, alors qu'en mode d'ionisation positif, il y a en plus des ions intermédiaires correspondant à la perte d'un seul des sucres, comme décrit par Cuyckens (2000). La fragmentation des flavonoïdes-*O*-glycosylés apporte beaucoup d'informations, le type de sucre sur la génine et le type d'aglycone, quand l'énergie de collision est bien adaptée.

La fragmentation des *flavonoïdes mono-C-glycosylés* (orientine et isoorientine) en mode d'ionisation négatif semble identique à ce qui est décrit dans la littérature, alors qu'en mode d'ionisation positif, le spectre de fragmentation semble plus riche en ions que celui de la littérature (Vukics and Guttman 2010). Deux isomères de position ont été fragmentés (orientine et homoorientine). Cette étude de fragmentation a permis de confirmer qu'en mode d'ionisation négatif seul le fragment correspondant à une perte d'eau permettait de distinguer l'isomère mono-*C*-glycosylé en position 6 de celui mono-*C*-glycosylé en position 8. Les deux spectres de fragmentation des deux isomères de position montrent également des ions d'intensités différentes. En effet, pour l'orientine, dont l'hexose est en position C-8, l'ion $[Ag+41]^-$ est plus abondant que l'ion $[Ag+71]^-$, alors que pour l'isoorientine les abondances pour ces deux ions sont inversées. En mode d'ionisation positif, l'ion E_1^+ est présent pour l'isomère en C-8. Pour l'isomère en C-6, l'ion $^{0,1}X^+$ est plus intense que l'ion $^{0,2}X^+$ et inversement pour C-8. Ainsi, que ce soit en mode positif ou négatif, l'isomérisation en C-6 ou C-8 pourra être élucidée bien que cela soit plus évident dans le cas du mode négatif.

La fragmentation du *flavonoïde mono-C-glycosylé-O-glycosylé* (vitexine-2''-*O*-rhamnoside) est un mélange des fragmentations *O* et *C*-glycosylés vu précédemment. Les spectres de masse ressemblent à ceux de la littérature, avec en mode d'ionisation négatif, la présence des ions $^{0,2}X^-$, Z_1^- et les ions $[Ag+41]^-$ et $[Ag+71]^-$. En mode d'ionisation positif, la fragmentation commence bien par la perte du motif *O*-glycosylé, puis il y a la fragmentation dans le sucre en *C*-glycosylé.

Quant à la fragmentation du *flavonoïde di-C-glycosylé* (schaftoside), elle correspond bien à la fragmentation des mono-*C*-glycosylés, avec comme ions pour aider à l'identification les ions $[Ag+83]^-$ et $[Ag+113]^-$ en mode d'ionisation négatif. Cependant le spectre de fragmentation en mode d'ionisation positif semble plus riche que celui de la littérature, car il y a tous les fragments correspondant aux combinaisons possibles des différentes pertes des deux sucres *C*-glycosylés.

Dans les précédentes études concernant les soies de maïs, il a été publié qu'il y avait des flavonoïdes de types flavones-C-glycosylés, un tableau (III-4) récapitulatif des ions diagnostiques pour identifier la génine en mode d'ionisation négatif est donné ci-dessous.

Tableau III-4 : Flavones aglycones et ions diagnostiques pour des flavones di-C-glycosylés et des flavones mono-C-glycosylés en mode d'ionisation négatif.

Aglycone	M (Da)	di-C-glycosylés		mono-C-glycosylés	
		$^{0,2}X_1^{0,2}X_2$	$^{0,2}X_1^{0,3}X_2$	$^{0,2}X$	$^{0,3}X$
		$[Ag+83]^-$	$[Ag+113]^-$	$[Ag+41]^-$	$[Ag+71]^-$
Apigénine	270	353	383	311	341
Lutéoline	286	369	399	327	357
Chrysoériol	300	383	413	341	371
Diosmétine	300	383	413	341	371

2.4.2.2. Cas de sucres particuliers

Dans le cadre de notre étude, la bibliographie indique que les soies de maïs possèdent des molécules dont la maysin et ses dérivés ayant d'autres types de sucres, tels que le sucre 6-deoxy-xylo-hexos-4-ulopyranosyl, qui est un désoxyhexose avec une cétone (voir Fig. III-2A) qui est lié en C-glycosylé avec un désoxyhexose sur son hydroxyle. La maysin faisant partie des composés d'intérêt, l'identification de ce type de structure doit être connue. Ne disposant pas de standard de cette famille de molécules il est montré un exemple de spectre en mode négatif acquis lors de l'injection d'un des extraits de soies de maïs (Fig. III-15).

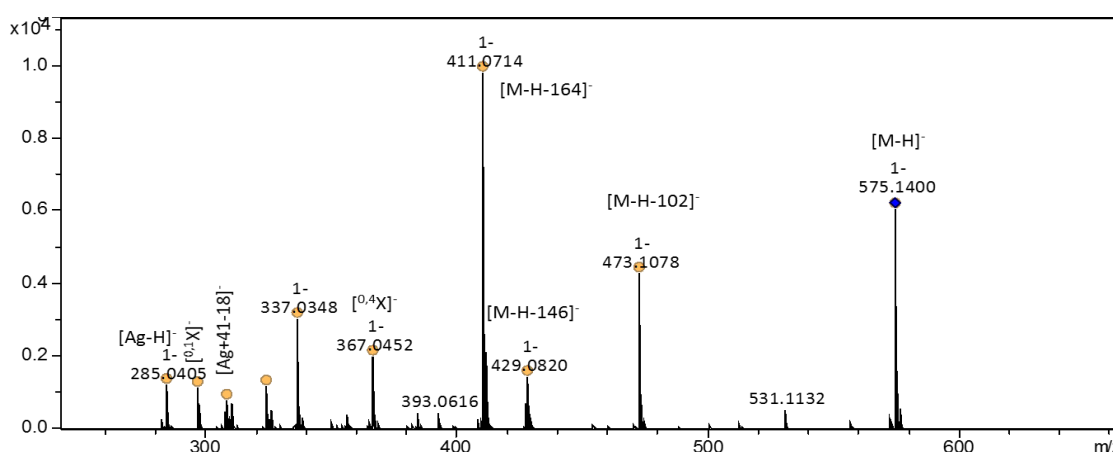


Figure III-15 : Spectre de fragmentation de la molécule maysin (possédant un 6-deoxy-xylo-hexos-4-ulopyranosyl) en mode d'ionisation négatif, obtenu sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV. Spectre obtenu à partir de la maysin présente dans un extrait de soie de maïs.

Ainsi à chaque fois que ce motif de sucre est présent en liaison C-C sur l'aglycone, une perte caractéristique de 102 u en mode négatif, au lieu de 120, 104 ou 60 u est observable (Figueirinha 2008).

Une fois l'empreinte non ciblée d'un extrait de mélange de soies de maïs obtenue, et la fragmentation des flavonoïdes confirmée sur l'appareil qTOF (Maxis de Bruker), différentes approches peuvent être utilisées pour identifier les molécules d'intérêt. Selon si la molécule a déjà été décrite ou est complètement inconnue, son niveau d'identification différera.

2.4.3. Démarche appliquée à l'identification d'une molécule dans l'extrait de soies de maïs

2.4.3.1. Cas d'une molécule identifiée à un niveau 1

Les standards utilisés pour vérifier la réponse de l'appareil qTOF (Maxis de Bruker) pour la fragmentation des flavonoïdes ont été injectés dans les mêmes conditions que l'extrait. Ainsi on peut vérifier si ces composés sont présents ou non dans l'extrait.

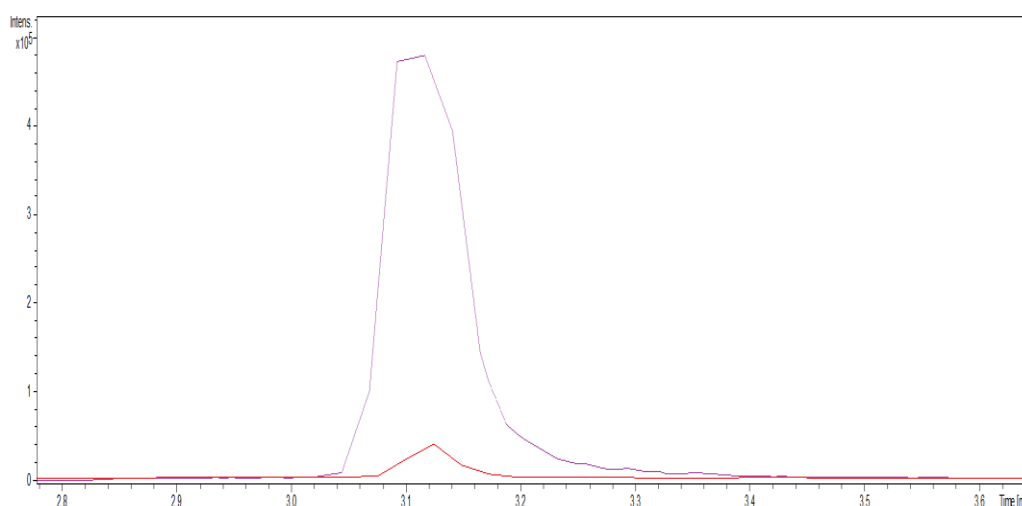


Figure III-16 : Extractions d'ion de rapport de masse m/z 609,1461 u dans les analyses de mélange de standards (violet) et d'un extrait de soie (rouge).

Ainsi la molécule rutine (m/z 609,1461) se retrouve en petite quantité (Fig. III-16) dans l'extrait de soies analysé. En effet, on retrouve un pic chromatographique au même temps de rétention que ce standard et avec la même masse moléculaire déprotonée. De plus, le spectre de fragmentation du composé (Fig. III-17) ressemble au spectre de fragmentation du standard, mais le composé étant en faible teneur dans l'extrait, le spectre correspondant est moins intense, donc un peu plus bruiteux.

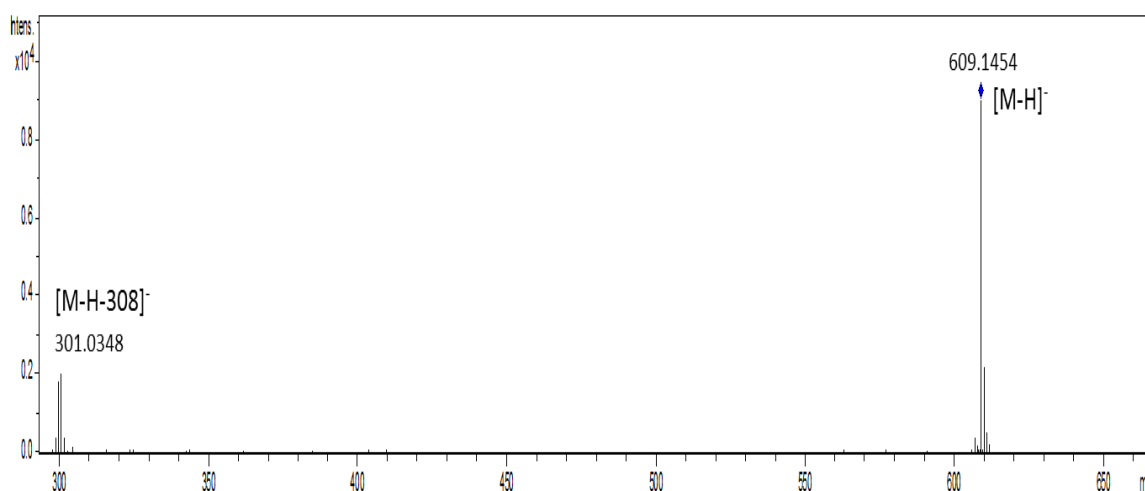


Figure III-17 : Spectre de fragmentation du composé élué au temps de rétention 3,12 min.

Les formules brutes proposées par SmartFormulaManually, outil du logiciel DataAnalysis (v. 4.4) avec une tolérance en masse de 2 ppm et un rapport d'atome d'hydrogène sur atome de carbone (H/C) maximum de 3, sont aux nombres de 5, dont une seule formule sans atome d'azote, avec une précision en masse de 1 ppm. La formule proposée pour l'ion de la molécule déprotonée est $C_{27}H_{29}O_{16}$, ce qui correspond bien à la formule brute de la molécule rutine moins un atome d'hydrogène.

Par contre, lorsque la formule brute $C_{27}H_{30}O_{16}$ est soumise dans les bases de données pour avoir une proposition de structure moléculaire, une multitude de propositions est faite. En effet, dans CompoundCrawler (v 3.0), logiciel de Bruker, 260 correspondances sont données dans les bases de données interrogées (KEGG, ChemSpider, PubChem et ChEBI). La base de données Metlin propose 120 composés, dont deux molécules avec des spectres acquis expérimentalement (cyanine et rutine) et 107 molécules avec des spectres de fragmentation in silico. Certaines de ces propositions sont identiques, seul l'identifiant (ID) de Metlin ou le nom diffère. L'inconvénient des spectres de fragmentation in silico est qu'ils sont générés souvent pour un mode d'ionisation positif, donc il est difficile de pouvoir les comparer avec les spectres expérimentaux acquis en mode d'ionisation négatif. En effet, comme il a été montré dans les parties d'étude de fragmentation de flavonoïdes, la réponse n'est pas identique dans les deux modes d'ionisation. Donc il est utile à ce niveau de pouvoir interpréter le spectre de fragmentation (Fig. III-17) de cette molécule, afin d'émettre des hypothèses de structure.

Ce spectre de fragmentation ressemble à celui du standard rutine. Les formules brutes proposées pour le fragment et pour l'ion moléculaire (Tableau III-5 et Fig. III-17) correspondent aux fragments décrits pour ce standard. Le même ratio $[Y_0-H]^-/Y_0^-$ est également retrouvé.

Tableau III-5 : Proposition de formules brutes pour le fragment et l'ion moléculaire du composé élué au temps de rétention de 3,12 min.

m/z mesuré	Formule brute	m/z théorique	err [ppm]	mSigma	DBE
609,145504	C ₂₇ H ₂₉ O ₁₆	609,146108	1,0	22,1	13,0
301,034839	C ₁₅ H ₉ O ₇	301,035376	1,8	18,9	11,0
300,027235	C ₁₅ H ₈ O ₇	300,027551	1,1	494,2	11,5

Dans cet exemple, le standard et le composé dans l'extrait ont le même temps de rétention et le même spectre de masse avec les mêmes ions. Ce composé est donc confirmé comme étant la molécule de rutine.

L'analyse par UHPLC/HRMS-MS permet d'obtenir une liste de masses avec les spectres de fragmentation associés et d'obtenir une proposition de formules brutes pour chaque ion détecté. En l'absence de standard, la démarche de tentative d'identification de molécule présentée ci-dessous est plus délicate à mener à bien.

2.4.3.2. Cas d'indétermination

Malheureusement tous les composés ne correspondent pas à une molécule standard que l'on possède ou qui peut facilement être obtenue. Par ailleurs, si la séparation du mélange analysé n'est pas optimale, la réponse des composés peut différer un peu de celle qu'elle serait s'ils étaient purs.

Ainsi le composé élué au temps de rétention de 4,80 min avec un rapport m/z de 573,1241 en mode d'ionisation négatif est associé à une formule brute C₂₇H₂₆O₁₄. L'interrogation de CompoundCrawler ne conduit qu'à neuf propositions (Annexe 4), mais aucune structure correspondant à un flavone.

Le spectre de fragmentation de ce composé (Fig. III-18) permet de mettre en évidence certains fragments qui peuvent être caractéristiques de la molécule, bien que les ions soient peu intenses.

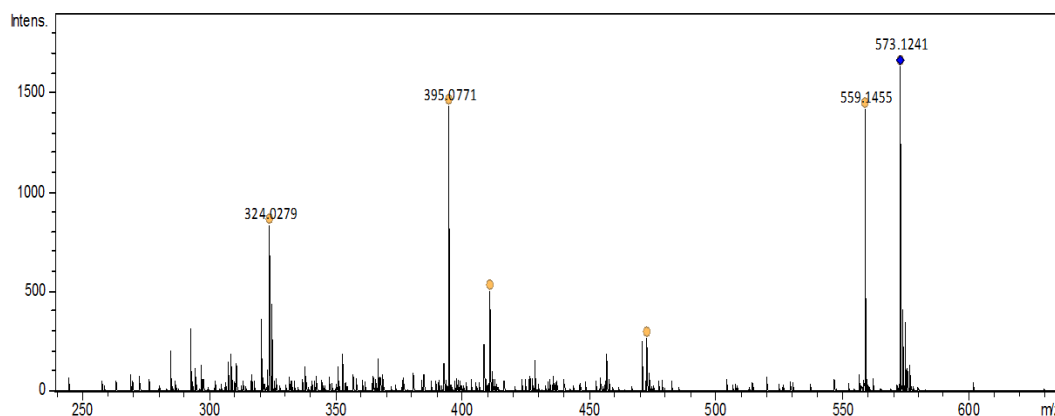


Figure III-18 : Spectre de fragmentation du composé élué au temps de rétention 4,80 min en mode d'ionisation négatif.

Les formules brutes de chacun de ces fragments peuvent être proposées via le logiciel SmartFormulaManually (Tableau III-6).

Tableau III-6 : Proposition de formules brutes des ions observés sur le spectre de fragmentation de l'ion m/z 573,1241.

m/z mesuré	Formule brute	m/z théorique	err [ppm]	DBE
573,124106	$C_{27}H_{25}O_{14}$	573,124979	1,5	15,0
559,145546	$C_{27}H_{27}O_{13}$	559,145715	0,3	14,0
473,107900	$C_{23}H_{21}O_{11}$	473,108935	2,2	13,0
411,073400	$C_{21}H_{15}O_9$	411,072156	-3,0	14,0
395,077059	$C_{21}H_{15}O_8$	395,077241	0,5	14,0
324,027898	$C_{17}H_8O_7$	324,027551	-1,1	13,5

Le fragment de rapport m/z 559,1455 pourrait correspondre à une perte de méthyle par rapport au premier ion détecté (m/z 573,1241). Ce fragment est associé à une proposition de formule brute qui ne correspond pas à cette perte, mais à une perte d'un atome d'oxygène avec un gain de deux protons. De plus quand on zoome sur cet ion, il ne possède pas d'isotope, contrairement aux autres ions. Donc cet ion pourrait provenir d'un autre ion parent qui aurait été isolé sans ses ions isotopes. Pour confirmer cette observation, l'analyse de ce même échantillon dans les mêmes conditions a été effectuée en mode d'ionisation positif. Dans le spectre de fragmentation obtenu dans ces conditions, on ne retrouve pas cette même perte, mais on observe la présence d'autres fragments qui ne possèdent également pas de profil isotopique. Un zoom a donc été effectué au temps de rétention (4,79 min) de l'ion de rapport m/z 573,1241 (Fig. III-19) en mode de balayage full scan.

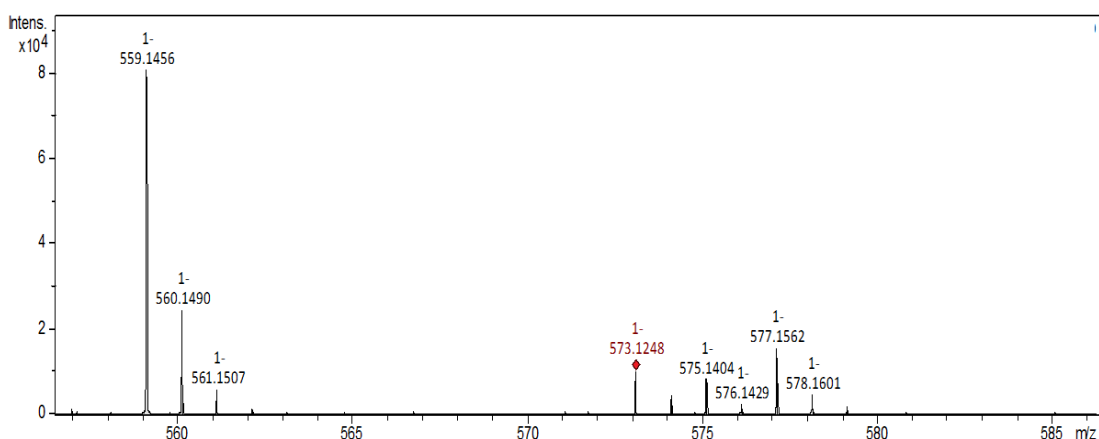


Figure III-19 : Spectre de masse en full scan associé au temps de rétention 4,79 min.

Dans ce spectre de masse (Fig. III-19), on observe l'ion $[M-H]^-$ de rapport m/z 573, ainsi que l'ion de rapport m/z 559 déjà observé auparavant en tant que fragment de l'ion 573. A côté de l'ion de rapport m/z 573, les ions de rapport m/z 575 et 577 sont également présents dans la même gamme d'intensité.

Etant donné qu'à ce rapport m/z , la fenêtre d'isolement est d'environ 5,2 u, il est donc possible que lors de l'isolement de l'ion de rapport m/z 573, les ions 575 et 577 aient été englobés dans la fenêtre d'isolement et qu'ainsi dans le spectre de fragmentation (Fig. III-18), un mélange de fragments de ces différents ions moléculaires apparaisse. Comme certains fragments ne possèdent pas de profil isotopique, une des masses moléculaires a été prise sans son massif isotopique, sûrement l'ion de rapport m/z 577.

Afin d'attribuer les ions spécifiques à la fragmentation de l'ion de rapport m/z 573, une comparaison des différents spectres de fragmentation de ces 3 ions moléculaires (m/z 573, 575 et 577) ainsi que de l'ion fragment (m/z 559) peut être effectuée (Fig. III-20).

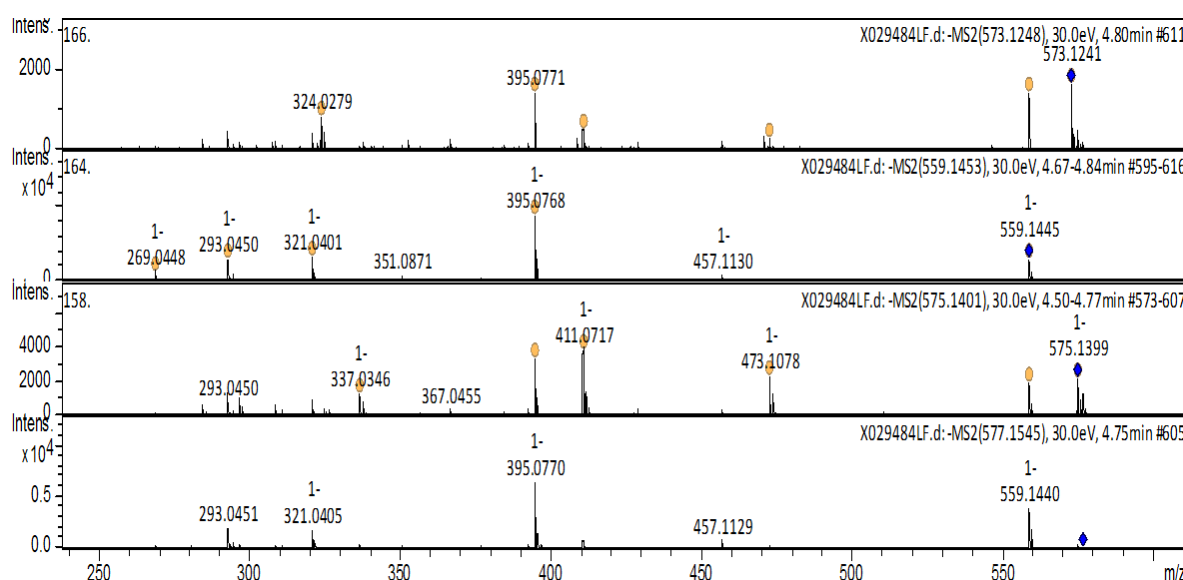


Figure III-20 : Spectres de fragmentation en mode d'ionisation négatif des quatre ions de rapport m/z 573,1241 ; 559,1445 ; 575,1399 ; 577,1562.

Ainsi les composés de rapport m/z 575 et 577 donnent comme fragments les ions de rapport m/z 559 et 395, qui n'avaient pas de profil isotopique lors de la fragmentation de l'ion de rapport m/z 573. Ces ions fragments ne sont donc pas forcément affiliés à l'ion de rapport m/z 573,1241.

La même comparaison a été réalisée en mode d'ionisation positif (Fig. III-21) pour confirmer ces observations.

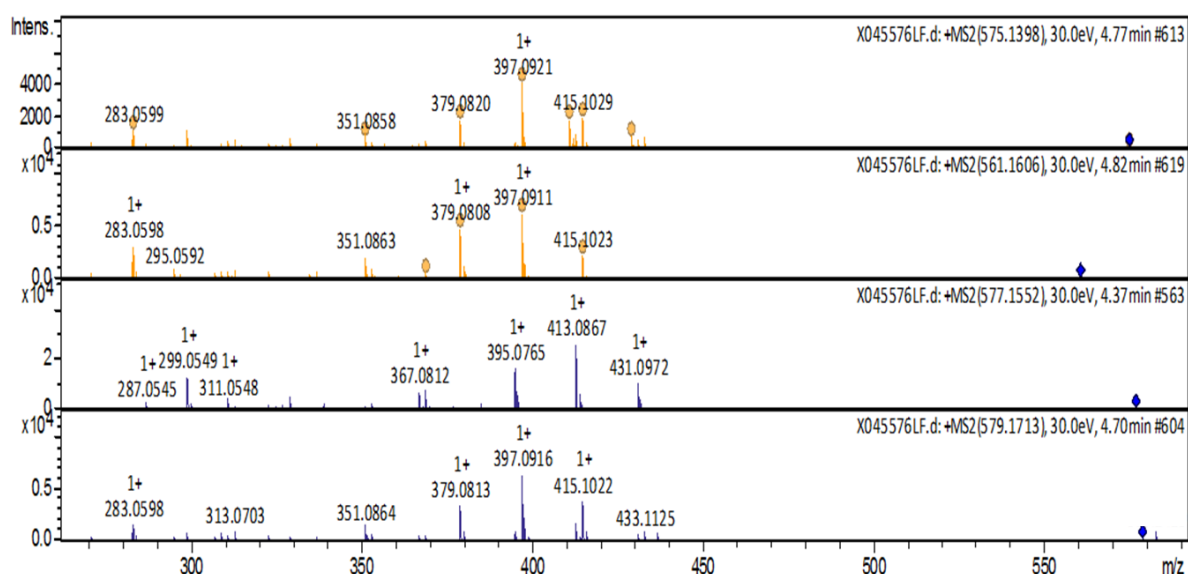


Figure III-21 : Spectres de fragmentation en mode d'ionisation positif des quatre ions de rapport m/z 575,1398 ; 561,1606 ; 577,1552 ; 579,1713.

Le fragment de rapport m/z 561 n'est pas obtenu dans ce mode d'ionisation (Fig. III-21), mais les ions de rapport m/z 283 et 313 n'ont pas de profil isotopique, ils peuvent donc provenir de l'ion de rapport m/z 579. En revanche, dans ce mode d'ionisation, l'ion de rapport m/z 575 donne, en plus, des fragments spécifiques que l'on ne retrouve pas dans les spectres de fragmentation des autres composés (Tableau III-7). Ces ions sont de rapport m/z 411 et 429, ils correspondraient respectivement à des formules brutes en $C_{21}H_{15}O_9$ et $C_{21}H_{17}O_{10}$, soit une différence de perte d'eau entre les deux. La différence de masse entre l'ion moléculaire m/z 575 et l'ion fragment m/z 429 est de 146 u. Ces fragments pourraient donc correspondre à une perte de sucre *O*-déoxyhexose lié à un C-glycosylé. Pour connaître le type de sucre en C-glycosylé, il faudrait d'autres ions spécifiques différents des trois ions parents, mais qui n'ont pas pu être déterminés sur les spectres obtenus. Pour cette molécule, il n'est donc pas possible actuellement d'identifier précisément le sucre C-glycosylé et par conséquent, également la génine. Toutefois, il a pu être déterminé que cette molécule (m/z 575,1398 en mode d'ionisation positif) possède un sucre *O*-déoxyshexose lié à un C-glycosylé.

Tableau III-7 : Ions spécifiques à la fragmentation des ions moléculaires en mode d'ionisation positif.

$[M+H]^+$	-146	-146-18	-146-2x18	-146-3x18	-146-64	-146-132
575,1398	429	411				
577,1552	431	413	395		367	299
579,1713	433	415	397	379		
561,1606	415	397	379		351	283

Les structures proposées par CompoundCrawler (Annexe 4) ne sont donc pas cohérentes avec les interprétations de ces ions fragments et les formules brutes qui leur sont associées. Donc on est peut-être en présence d'une structure inconnue. Pour confirmer les fragments de cette molécule (m/z

575,1398 en mode d'ionisation positif), l'analyse devrait être réalisée avec une fenêtre d'isolement réduite, mais avec l'analyseur qTOF, plus la fenêtre d'isolement est faible plus l'abondance d'ions isolés diminue ce qui engendre des spectres avec beaucoup de bruit de fond. Pour aller plus loin dans la structure de la molécule, il faudrait utiliser d'autres techniques, comme la RMN, mais il faudrait pouvoir obtenir le composé pur. Comme cette molécule est éluée avec d'autres composés de masses très proches, il est difficile d'isoler ce composé avec les outils actuels à disposition à l'institut.

Avec ces deux exemples de spectre de masse, il a pu être montré deux cas de figure extrêmes. D'une part, le cas où le spectre de fragmentation est clair et le schéma de fragmentation logique permettant une proposition de nom à partir des bases de données cohérente avec celle d'un standard de flavonoïde. Et d'autre part, le cas où le spectre de masse est constitué d'ions en faible intensité et/ou présente des ions interférents dus à une mauvaise séparation conduisant à une absence de correspondance avec les bases de données. Il est donc important de comprendre les schémas de fragmentation des molécules d'intérêt pour comprendre le problème et avoir un début d'identification. Bien que les spectres de fragmentation soient parfois complexes et difficiles à interpréter, la spectrométrie de masse haute résolution donne accès aux masses exactes des molécules et donc à leurs formules brutes. Les outils de représentation graphique peuvent alors être utilisés pour classer les formules obtenues avec la totalité du chromatogramme et aider à déterminer la famille d'une molécule ou d'un groupe de molécules.

2.4.4. Construction du diagramme de Van Krevelen

Dans les conditions d'extraction et d'analyse UHPLC/HRMS que nous avons appliquées pour analyser l'extrait représentatif de soies de maïs, il a été détecté environ 124 composés différents. Pour une représentation plus globale des molécules dans l'extrait, les formules brutes des composés ont été reportées sur un diagramme de Van Krevelen (Fig. III-30). L'intérêt de ce diagramme est de décrire visuellement la distribution des composés d'un système complexe en familles chimiques.

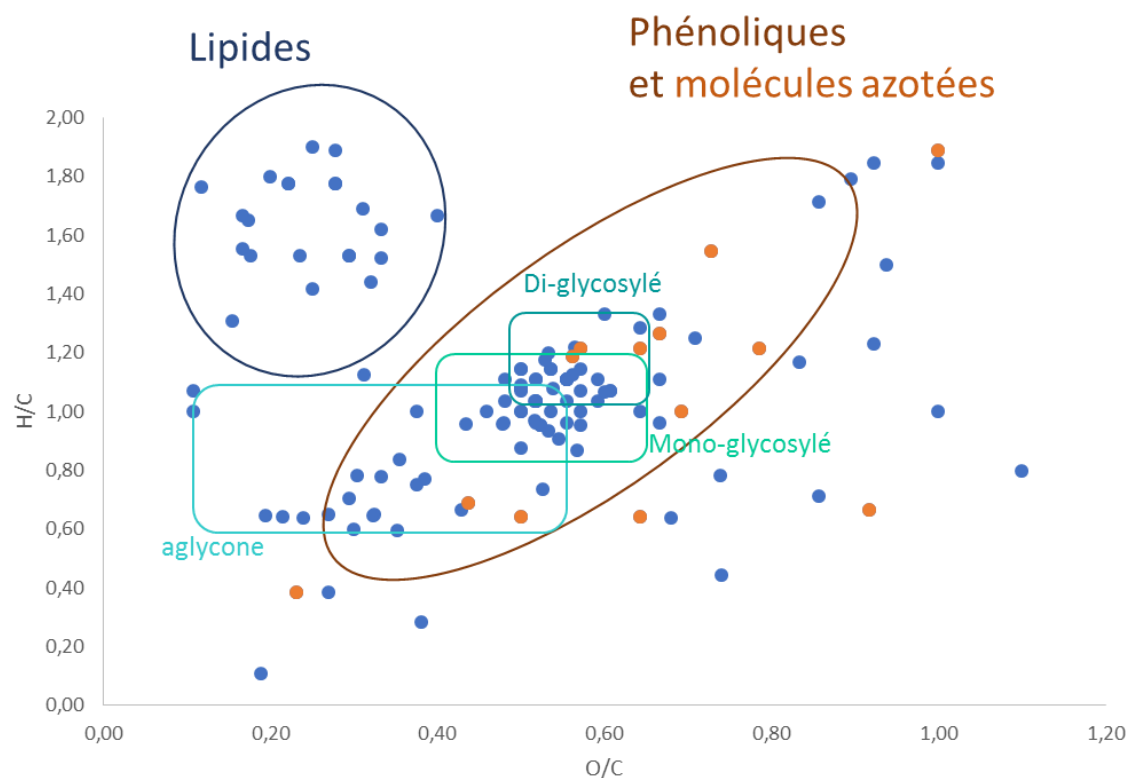


Figure III-22 : Diagramme de Van Krevelen répartissant les 124 molécules extraites des soies de maïs analysées en UHPLC/HRMS en mode d'ionisation négatif. Sur le diagramme, les cercles positionnent les familles moléculaires possibles et les rectangles visualisent les zones correspondant aux flavonoïdes aglycone, mono et di-glycosylés. Ces attributions ont été effectuées sur la base du diagramme de Van Krevelen tracé à partir de 260 formules de flavonoïdes extraites du catalogue Extrasynthèse.

D'après le diagramme (Fig. III-22), il y a deux zones principales de répartition des molécules. Une zone correspond aux molécules lipidiques avec des ratios H/C élevés et une autre correspond aux composés phénoliques et aux molécules azotées pour lesquels les ratios O/C sont supérieurs. Ce diagramme permet de bien mettre en évidence la répartition de l'ensemble des composés obtenus après extraction et analyse à partir des formules brutes proposées par le logiciel du spectromètre de masse. La zone principale (entourée en brun) contient les flavonoïdes. Après avoir positionné sur ce diagramme, les formules des 260 flavonoïdes du catalogue Extrasynthèse, il a pu être localisé sur la figure III-22 approximativement trois sous-zones (rectangle) de flavonoïdes correspondant aux flavonoïdes aglycones, mono ou di-glycosylés. On peut donc supposer que dans la zone repérée comme celle des composés phénoliques présents dans l'extrait, il y a une majorité de composés mono et di-glycosylés. Dans la suite de notre étude, l'identification et la quantification des composés de l'extrait ont été restreintes à la partie phénolique contenant le plus de molécules (O/C de 0,4 - 0,7 et H/C de 0,7 - 1,4).

Ce diagramme a permis d'accéder rapidement à une première estimation des familles moléculaires présentes dans l'extrait de soie de maïs. L'extraction des soies par un mélange eau / éthanol a permis

d'isoler un contenu majoritairement de composés phénoliques. Ainsi les composés d'intérêt à quantifier ont été vite repérés par ce diagramme avant même l'identification de tous les composés.

2.4.5. Identification des flavonoïdes dans les extraits de soies de maïs

Dans ces travaux, une caractérisation complète des flavonoïdes a été menée sans isolement préalable, afin d'élargir les connaissances sur la composition phytochimique des soies de maïs et de compléter la liste des molécules majoritaires déjà décrites dans la littérature après un long travail de purification, isolement et caractérisation par RMN.

L'analyse obtenue en UHPLC/HRMS-MS permet de proposer une formule brute pour les ions moléculaires et leurs fragments dans l'extrait de soie de maïs représentatif des 5 variétés. Afin de pouvoir réaliser l'identification des flavonoïdes détectés, il est nécessaire de prendre en compte la nomenclature des flavonoïdes tel que décrite dans le chapitre 1, ainsi que les schémas de fragmentation des flavonoïdes *O* et *C*-glycosylés donnés dans la littérature et ceux plus spécifiquement confirmés avec l'analyse de molécules standards avec l'appareil UHPLC/HRMS Maxis (de Bruker) (cf paragraphe 3.4.2.).

Ainsi, l'identification des différentes molécules détectées est présentée par type de flavonoïdes que l'on peut rencontrer : *O*-glycosylés, mono-*C*-glycosylés, mono-*C*-glycosylés-*O*-glycosylés et di-*C*-glycosylés. Les tableaux (III-8 à III-11) rassemblent pour chacun de ces sous-groupes, les ions caractéristiques observés ainsi que leur abondance. Seules les molécules suffisamment abondantes pour avoir une bonne précision sur la formule brute proposée et les abondances de leurs fragments ont été caractérisées.

Les propositions de formules brutes ont été réalisées via SmartFormulaManually, logiciel de Bruker, qui prend en compte la masse exacte à 2 ppm près, le rapport isotopique (Kind and Fiehn 2006), la règle de Lewis et le nombre d'insaturations, avec les éléments C, H, O.

2.4.5.1. Flavonoïdes-*O*-glycosylés

Parmi les flavonoïdes détectés en HRMS dans l'extrait représentatif de soies, 11 composés ont été identifiés comme des *O*-glycosylés mono-*O*-glycosylés ou di-*O*-glycosylés. Le tableau III-8 présente pour chacun de ces composés, son temps de rétention chromatographique, les informations obtenues à partir des spectres de masse full scan, des spectres de fragmentation MS/MS et le niveau d'identification selon Sumner (2007). Il n'y a pas de famille d'aglycone qui prédomine, il y a aussi bien des flavones que des flavonols.

On peut retrouver dans le tableau III-8, deux molécules ayant la même masse nominale (m/z 463) mais qui diffèrent dans leur formule brute. Ainsi grâce à la précision en masse de l'analyse haute résolution de la spectrométrie de masse, la quercétine (C₁₅H₁₀O₇) peut être distinguée de l'hespérétine (C₁₆H₁₄O₆).

La séparation chromatographique a permis de différencier deux isomères de rapport m/z 637,1414. Ces deux molécules ont soit deux génines différentes (chrysoériol ou diosmétine), soit des formes d'hexose ou glucuronique différentes ou des positions d'hexose différentes sur la génine ou une liaison entre les deux sucres différente. Dans ce cas particulier, ni la précision en masse, ni la fragmentation ne permettent d'informer davantage sur la structure fine de ces deux molécules. Dans le tableau III-8, il a été privilégié une génine de type chrysoériol, car c'est ce type de structure qui a déjà été identifié dans la soie de maïs. En effet, parmi les 16 flavonoïdes décrits dans la littérature (Tableau III-1), il y a trois flavonoïdes *O*-glycosylés. Ces trois composés ont bien été retrouvés parmi les 11 flavonoïdes *O*-glycosylés que nous avons identifiés. Les trois masses correspondantes sont respectivement associées aux temps de rétention 3,67, 3,71 et 3,88 min. Donc ces trois composés pourraient correspondre respectivement aux molécules lutéoline 7-*O*-néohesperidoside (Kim 2014), lutéoline 3'-méthyl éther 7-glucuronosyl-(1->2)-glucuronide (Kim 2014) et 6,4-dihydro-3-methoxyflavone-7-*O*-glucoside (Ren 2009 et 2013, Hasanudin 2012). Cette attribution est confirmée par le fait que les fragmentations observées dans nos conditions d'expérience pour ces trois composés sont en bon accord avec la fragmentation théorique de ces molécules.

Tableau III-8 : Caractérisation des flavonoïdes-*O*-glycosylés d'un extrait de soies de maïs. Le symbole * correspond à une masse décrite dans la littérature.

Tr (min)	[M-H] ⁻ (<i>m/z</i>)	Formule moléculaire	Erreur Masse (ppm)	Y ₀ ⁻ ou [Y ₀ -H] ⁻					Proposition de molécule	Niveau d'identification et Nom provenant de la littérature
				Perte -162	Perte -176	Perte -308	Perte -338	Perte -352		
3,12	609,1455	C ₂₇ H ₂₉ O ₁₆	1			301,0348 (C ₁₅ H ₉ O ₇)			Quercétine <i>O</i> -rutinoside	1 (Rutine)
3,29	463,0862	C ₂₁ H ₁₉ O ₁₂	4,4	301,0342 (C ₁₅ H ₉ O ₇)					Quercétine <i>O</i> -hexose	3
3,62	637,1414	C ₂₈ H ₂₉ O ₁₇	-0,7				299,0557 (C ₁₆ H ₁₁ O ₆)		Chrysoérior mono- <i>O</i> -hexose- <i>O</i> -glucuronide	3
3,67	593,1508*	C ₂₇ H ₂₉ O ₁₅	0,6			285,0403 (C ₁₅ H ₉ O ₆)			Kaempférol ou Lutéoline <i>O</i> -rutinoside	3 Lutéoline 7- <i>O</i> -néohesperidoside
3,71	651,1199*	C ₂₈ H ₂₇ O ₁₈	0,6					299,0559 (C ₁₆ H ₁₁ O ₆)	Chrysoérior mono- <i>O</i> -glucuronide- <i>O</i> -glucuronide	3 Lutéoline 3'-methyl éther 7- glucuronosyl-(1->2)-glucuronide
3,72	449,1029	C ₂₈ H ₁₇ O ₆	0,4	287,0563 (C ₁₅ H ₁₁ O ₆)					Eriodictyol ou Fustin <i>O</i> -hexose	3
3,88	461,107*	C ₂₂ H ₂₁ O ₁₁	4,2	299,0562 (C ₁₆ H ₁₁ O ₆)					Chrysoérior <i>O</i> -hexose	3 6,4-dihydro-3-methoxyflavone- 7- <i>O</i> -glucoside
3,89	447,0927	C ₂₁ H ₁₉ O ₁₁	1,2	284,0326 (C ₁₅ H ₈ O ₆)					Kaempférol ou Lutéoline <i>O</i> -hexose	3
3,96	637,1413	C ₂₈ H ₂₉ O ₁₇	-0,4				299,0552 (C ₁₆ H ₁₁ O ₆)		Chrysoérior mono- <i>O</i> -hexose- <i>O</i> -glucuronide	3
4,13	463,1231	C ₂₂ H ₂₃ O ₁₁	3,2	301,0712 (C ₁₆ H ₁₃ O ₆)					Hespérétine <i>O</i> -hexose	3
4,65	475,0868	C ₂₂ H ₁₉ O ₁₂	3		299,0560 (C ₁₆ H ₁₁ O ₆)				Chrysoérior <i>O</i> -désoxyhexose	3

2.4.5.2. Flavonoïdes mono-C-glycosylés

Les mono-C-glycosylés détectés dans l'extrait sont au nombre de deux (Tableau III-9). Ce sont des dérivés de flavones avec comme aglycone une lutéoline. Dans la littérature (Tableau III-1), deux flavonoïdes mono-C-glycosylés ont été décrits dans les soies de maïs. Une des deux masses correspondantes (m/z 429) est retrouvée parmi les composés décrits avec la bonne formule brute et le même type de fragmentation. Cette molécule ($Tr = 5,24$ min) pourrait correspondre à la molécule décrite par Elliger (1980a) à savoir le 6-C-glycosylflavone. Par contre, la deuxième molécule décrite dans la littérature n'a pas été retrouvée, et c'est une autre structure qui a été identifiée, comme étant une lutéoline 6-C-hexose.

Tableau III-9 : Caractérisation des flavonoïdes mono-C-glycosylés d'un extrait de soies de maïs. Le symbole * correspond à une masse décrite dans la littérature.

Tr (min)	[M-H] ⁻ (m/z)	Formule moléculaire	Erreur masse (ppm)	Ag+71 (FB) Proportion	Ag+41 (FB) Proportion	Aglycone dédruit	Proposition de molécule	Niveau d'identification
2,82	447,0942	C ₂₁ H ₁₉ O ₁₁	-2	357,0617 (C ₁₈ H ₁₃ O ₈) 100%	327,0507 (C ₁₇ H ₁₁ O ₇) 78%	286	Lutéoline 6-C-hexose	1 (Isoorientine)
5,24	429,0880*	C ₂₁ H ₁₇ O ₁₀	-12,3		327,0502 (C ₁₈ H ₁₃ O ₇) 100%	286	Lutéoline C-oxodéoxyhexose	3

2.4.5.3. Flavonoïdes mono-C-glycosylés-O-glycosylés

Enfin, on retrouve 11 flavonoïdes C-glycosylés-O-glycosylés (Tableau III-10), ce qui constitue le sous-groupe le plus important de flavonoïdes dans les soies. C'est dans ce sous-groupe, que l'on retrouve les molécules maysin, apimaysin et méthoxymaysin, qui ont déjà été décrites en tant qu'inhibitrices de la croissance des larves. La majorité des composés décrits dans les soies de maïs (Tableau III-1) sont aussi des flavonoïdes mono-C-glycosylés-O-glycosylés. Parmi eux, sept ont été identifiés dans l'extrait que nous avons analysé, mais certains ions (m/z 575 et 593), présentent des isomères qui ne sont pas décrits dans la littérature et que la fragmentation ne permet pas de distinguer. Au contraire les masses 578 et 592 Da sont citées deux fois dans la littérature mais n'ont été identifiées qu'une seule fois dans cet extrait. Ainsi, la fragmentation ne permettant pas de distinguer ces isomères, soit il n'y a qu'une seule forme présente soit les deux formes sont présentes dans l'extrait mais co-éluées. De nouveaux flavonoïdes C-glycosylés-O-glycosylés non mentionnés dans la littérature, ont donc été identifiés dans l'extrait analysé.

Tableau III-10 : Caractérisation des flavonoïdes mono-C-glycosylés-O-glycosylés d'un extrait de soies de maïs. Le symbole * correspond à une masse décrite dans la littérature.

Tr (min)	[M-H] ⁻ (m/z)	Formule moléculaire	Erreur masse (ppm)	Ag+41 (FB) proportion	Ag+71 (FB) proportion	Ag+71-18 (FB) proportion	Ag+41-18 (FB) proportion	^{0,2} X ⁻ (FB) proportion	Z ₁ ⁻ (FB) proportion	Ag-H (FB) proportion	Aglycone dédié	Proposition de molécule	Niv. d'id. et Nom de la littérature
2,66	593,1511*	C ₂₇ H ₂₉ O ₁₅	0,2	327,0503 (C ₁₇ H ₁₁ O ₇) 4%	357,0616 (C ₁₈ H ₁₃ O ₈) 5%		309,0398 (C ₁₇ H ₉ O ₆) 4%	473,1082 (C ₂₃ H ₂₁ O ₁₁) 39%	429,0818 (C ₂₁ H ₁₇ O ₁₀) 12%		286	Lutéoline C-hexose 2"-O-désoxyhexose	3
2,79	593,1491*	C ₂₇ H ₂₉ O ₁₅	3,6	327,0503 (C ₁₇ H ₁₁ O ₇) 7%	357,0616 (C ₁₈ H ₁₃ O ₈) 10%	339,0493 (C ₁₈ H ₁₁ O ₇) 5%	309,0397 (C ₁₇ H ₉ O ₆) 8%	473,108 (C ₂₃ H ₂₁ O ₁₁) 52%	429,0818 (C ₂₁ H ₁₇ O ₁₀) 23%		286	Lutéoline C-hexose 2"-O-désoxyhexose	3
3,12	577,1556	C ₂₇ H ₂₉ O ₁₄	1,2	311,054 (C ₁₇ H ₁₁ O ₆) 5%	341,0646 (C ₁₈ H ₁₃ O ₇) 15%	323,053 (C ₁₈ H ₁₁ O ₆) 8%	293,0454 (C ₁₇ H ₉ O ₅) 57%	457,1115 (C ₂₃ H ₂₁ O ₁₀) 6%	413,088 (C ₂₁ H ₁₇ O ₉) 80%		270	Apigénine C-hexose 2''-O-désoxyhexose	3
3,42	607,1658	C ₂₈ H ₃₁ O ₁₅	1,8	341,0660 (C ₁₈ H ₁₃ O ₇) 7%	371,0777 (C ₁₉ H ₁₅ O ₈) 7%	353,0670 (C ₁₉ H ₁₃ O ₇) 7%	323,0556 (C ₁₈ H ₁₁ O ₆) 39%	487,1242 (C ₂₄ H ₂₃ O ₁₁) 6%	443,0975 (C ₂₂ H ₁₉ O ₁₀) 63%		300	Chrysoériol C-Hexose 2"-O-désoxyhexose	3
4,27	575,1399*	C ₂₇ H ₂₇ O ₁₄	1,2	327,051 (C ₁₇ H ₁₁ O ₇) 6%			309,0399 (C ₁₇ H ₉ O ₆) 8%	473,1078 (C ₂₃ H ₂₁ O ₁₁) 44%	411,0714 (C ₂₁ H ₁₅ O ₉) 100%	285,0404 (C ₁₅ H ₉ O ₆) 12%	286	Lutéoline C-désoxyhexose C=O 2"-O-désoxyhexose	3
4,37	593,1499*	C ₂₇ H ₂₉ O ₁₅	2,2				309,0396 (C ₁₇ H ₉ O ₆) 9%	473,1079 (C ₂₃ H ₂₁ O ₁₁) 31%	429,0828 (C ₂₁ H ₁₇ O ₁₀) 6%	285,0404 (C ₁₅ H ₉ O ₆) 7%	286		
4,5	575,1402*	C ₂₇ H ₂₇ O ₁₄	0,8	327,0508 (C ₁₇ H ₁₁ O ₇) 6%			309,0393 (C ₁₇ H ₉ O ₆) 8%	473,1079 (C ₂₃ H ₂₁ O ₁₁) 50%	411,0715 (C ₂₁ H ₁₅ O ₉) 100%	285,0399 (C ₁₅ H ₉ O ₆) 13%	286	Lutéoline C-désoxyhexose C=O 2''-O-deoxyhexose	3
4,63	575,1399*	C ₂₇ H ₂₇ O ₁₄	1,3				309,0400 (C ₁₇ H ₉ O ₆) 14%	473,1078 (C ₂₃ H ₂₁ O ₁₁) 62%	411,0717 (C ₂₁ H ₁₅ O ₉) 100%	285,0391 (C ₁₅ H ₉ O ₆) 12%	286	Lutéoline C-désoxyhexose C=O O-désoxyhexose	3
4,73	559,1445*	C ₂₇ H ₂₇ O ₁₃	2,2				293,0450 (C ₁₇ H ₉ O ₅) 35%	457,113 (C ₂₃ H ₂₁ O ₁₀) 8%	395,0768 (C ₂₁ H ₁₅ O ₈) 100%	269,0448 (C ₁₅ H ₉ O ₅) 14%	270	Apigénine C-désoxyhexose C=O 2''-O-désoxyhexose	3
4,91	589,155*	C ₂₈ H ₂₉ O ₁₄	1,3				323,0557 (C ₁₈ H ₁₁ O ₆) 28%	487,1246 (C ₂₄ H ₂₃ O ₁₁) 7%	425,0873 (C ₂₂ H ₁₇ O ₉) 100%	299,0563 (C ₁₆ H ₁₁ O ₆) 10%	300	Chrysoériol C-désoxyhexose C=O 2''-O-désoxyhexose	3 3'-methoxymaysin
5,03	591,1706*	C ₂₈ H ₃₁ O ₁₄	2,2				323,0556 (C ₁₈ H ₁₁ O ₆) 31%	487,1246 (C ₂₄ H ₂₃ O ₁₁) 10%	427,0873 (C ₂₂ H ₁₉ O ₉) 57%		300	Chrysoériol C-désoxyhexose -O-désoxyhexose	3 ax-4''-OH-3'- methoxymaysin

2.4.5.4. Flavonoïdes di-C-glycosylés

On retrouve 6 flavonoïdes di-C-glycosylés (Tableau III-11) avec principalement des molécules ayant comme aglycone une apigénine, dont quatre ayant des sucres positionnés en 6-C-pentose et en 8-C-désoxyhexose. La position des sucres est déterminée suivant l'abondance des deux fragments $^{0,2}X^-$ ($^{0,2}X_6^- > ^{0,2}X_8^-$). Ces molécules se différencient sûrement dans la structure du sucre, ce que l'on pourrait vérifier après isolement des molécules et identification par RMN. Toutes ces molécules n'ont pas été évoquées dans la littérature.

Tableau III-11 : Caractérisation des flavonoïdes di-C-glycosylés d'un extrait de soies de maïs.

Tr (min)	[M-H] ⁻ (m/z)	Formule moléculaire	Erreur masse (ppm)	-60 (FB) proportion	-90 (FB) proportion	-120 (FB) proportion	-104 (FB) proportion	Ag+113 (FB) proportion	Ag+83 (FB) proportion	Aglycone dédit	Proposition de molécule	Niveau d'identification
2,54	593,1511	C ₂₇ H ₂₉ O ₁₅	-0,7			473,1085 (C ₂₃ H ₂₁ O ₁₁) 25%		383,0772 (C ₂₀ H ₁₅ O ₈) 7%	353,0667 (C ₁₉ H ₁₃ O ₇) 9%	270	Apigénine di-C-hexose	3
2,76	563,1397	C ₂₆ H ₂₇ O ₁₄	1,6		473,1084 (C ₂₃ H ₂₁ O ₁₁) 19%	443,0979 (C ₂₂ H ₁₉ O ₁₀) 22%		383,0779 (C ₂₀ H ₁₅ O ₈) 9%	353,0658 (C ₁₉ H ₁₃ O ₇) 15%	270	Apigénine 6-C-hexose 8-C-pentose	1 (Schaftoside)
3,05	577,1559	C ₂₇ H ₂₉ O ₁₄	-9,5			457,1119 (C ₂₃ H ₂₁ O ₁₀) 18%	473,1123 (C ₂₃ H ₂₁ O ₁₁) 11%	383,0761 (C ₂₀ H ₁₅ O ₈) 14%	353,0653 (C ₁₉ H ₁₃ O ₇) 20%	270	Apigénine 6-C-hexose 8-C-désoxyhexose	3
3,69	547,1454	C ₂₆ H ₂₇ O ₁₃	0,5		457,1073 (C ₂₃ H ₂₁ O ₁₀) 5%		443,0985 (C ₂₂ H ₁₉ O ₁₀) 23%	383,077 (C ₂₀ H ₁₅ O ₈) 6%	353,0653 (C ₁₉ H ₁₃ O ₇) 20%	270	Apigénine 6-C-désoxyhexose 8-C-pentose	3
4	547,1453	C ₂₆ H ₂₇ O ₁₃	0,7		457,1139 (C ₂₃ H ₂₁ O ₁₀) 7%		443,0983 (C ₂₂ H ₁₉ O ₁₀) 18%	383,0768 (C ₂₀ H ₁₅ O ₈) 9%	353,0651 (C ₁₉ H ₁₃ O ₇) 12%	270	Apigénine 6-C-désoxyhexose 8-C-pentose	3
4,24	547,1433	C ₂₆ H ₂₇ O ₁₃	4,4	487,1243 (C ₂₄ H ₂₃ O ₁₁) 18%	457,1188 (C ₂₃ H ₂₁ O ₁₀) 21%		443,0957 (C ₂₂ H ₁₉ O ₁₀) 22%	383,0768 (C ₂₀ H ₁₅ O ₈) 23%	353,0651 (C ₁₉ H ₁₃ O ₇) 14%	270	Apigénine 6-C-désoxyhexose 8-C-pentose	3

Ainsi à partir de l'interprétation des spectres de fragmentation, 30 flavonoïdes ont pu être identifiés comme présents dans les extraits de soies de maïs. La plupart d'entre eux sont identifiés pour la première fois dans les soies de maïs et ne sont pas décrits dans la littérature. D'autres composés restent encore à identifier, car leur schéma de fragmentation diffère un peu de ceux décrits dans la littérature.

Cette identification réalisée à partir d'un extrait représentatif des 5 variétés de soies de maïs permet maintenant de sélectionner des ions moléculaires et leurs fragments et de choisir les transitions à suivre afin de mettre en place une méthode d'analyse quantitative par spectrométrie de masse basse résolution en mode MRM des molécules présentes dans chacune des 5 variétés de soies de maïs.

2.5. Mise en place de la méthode de quantification

Les chromatogrammes présentés figure III-23 révèlent globalement la présence des mêmes composés dans les 5 variétés de soies de maïs mais avec des abondances propres à chaque variété. On peut observer notamment que la variété Java présente des pics à Tr 4,5 et 5 min bien plus intenses que les autres variétés. En revanche les variétés Sugar jean et Nova semblent contenir davantage de molécules éluées à 3,2 min. A ce stade de l'observation visuelle, il reste difficile de repérer des molécules spécifiques permettant de distinguer les variétés.

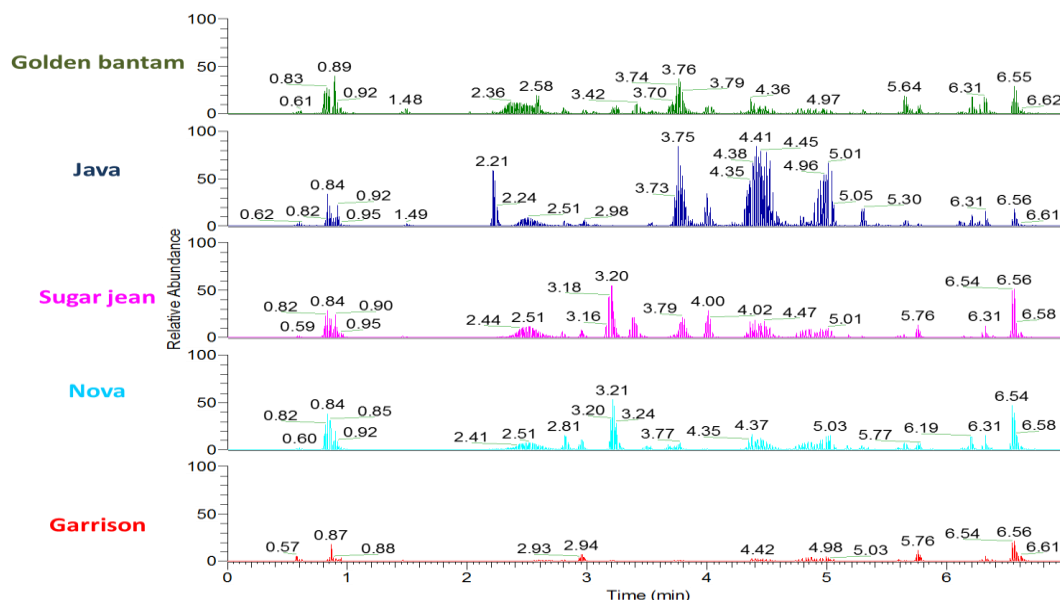


Figure III-23 : Comparaison des chromatogrammes acquis en mode MRM pour les 5 variétés de soies de maïs. Colonne Nucleodur Pyramid 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% d'acide formique en gradient ; débit de 0,6 mL/min ; T° four à 60°C ; volume d'injection de 3 μ L. Détection par spectrométrie de masse basse résolution en mode d'acquisition MRM des 67 transitions.

Une analyse multivariée ciblée a donc été mise en œuvre pour rechercher les molécules différenciantes de chaque variété. Pour cela une sélection plus globale de 67 composés, pas tous identifiés mais potentiellement phénoliques d'après leur position dans le diagramme de Van Krevelen (Fig. III-22), a été faite afin de les quantifier et d'évaluer les différences de concentration de ces composés d'intérêt en fonction des 5 variétés. Ces composés ont des rapports de masse m/z compris entre 400 et 700 Da et se retrouvent dans le cercle brun du diagramme de Van Krevelen (Fig. III-22).

Cette étude de quantification est réalisée par un couplage UHPLC avec un triple quadripôle. En 1979, Yost et Enke montraient déjà la versatilité d'un tel instrument, qui permet aussi bien de faire de l'élucidation structurale par fragmentation d'ions spécifiques, que de mettre en évidence des fonctions communes par perte de neutre ou alors comme ici de faire du dosage quantitatif en suivant des couples d'ions parents/ ions fils spécifiques, dit mode d'acquisition MRM. Pour mener à bien ce type d'analyse, il faut également réaliser des optimisations, telles que le bon choix des transitions, l'énergie de collision, le temps passé par transition et la tension de focalisation de l'ion afin d'être le plus sensible possible pour chaque molécule et de ne pas faire d'erreur d'attribution de valeur dans le cas d'isomères insuffisamment séparés par le système chromatographique.

2.5.1. Optimisation de la méthode MRM

Les conditions MS ont été d'abord optimisées sur le premier quadripôle (Q1) afin que le premier quadripôle ne transmette qu'un seul ion de rapport m/z spécifique par composé. Cet ion, dit ion parent, correspond à la molécule déprotonée puisque l'ionisation se fait en mode négatif. Ensuite une étude de fragmentation de l'ion parent a été menée (Fig. III-24), afin de sélectionner la meilleure énergie de collision pour obtenir via le troisième quadripôle, la réponse optimale pour le fragment choisi, dit ion fils. Pour les flavonoïdes glycosylés, cet ion correspond généralement à la perte du sucre ou d'une partie du sucre. Dans le cas des flavonoïdes C-glycosylés, leur absence de liaison labile et leur schéma particulier de fragmentation principale intra-sucre, feront qu'une plus grande énergie de collision sera nécessaire pour les fragmenter (Abad-Garcia 2009).

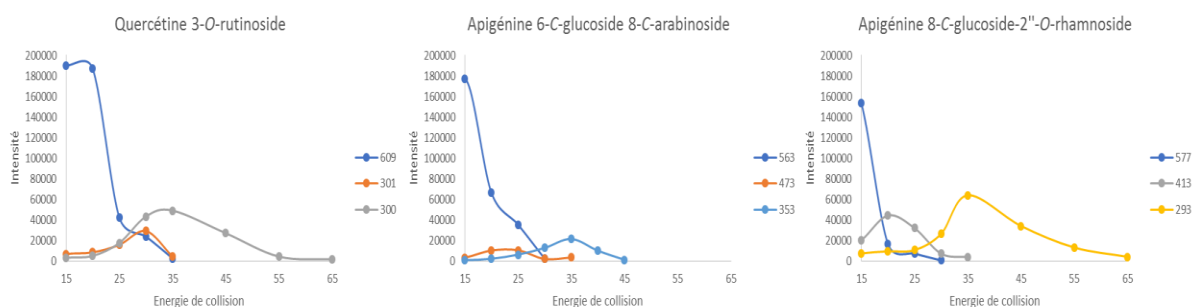


Figure III-24 : Abondance relative des ions (Parents et fragments) en fonction de l'énergie de collision pour trois standards représentatifs des trois sous-groupes de flavonoïdes présents dans les soies de maïs. Détection par spectrométrie de masse basse résolution en mode d'ionisation négatif.

Une étude de fragmentation en fonction des énergies de collision peut être illustrée comme présentée dans la figure II-24. L'intensité des différents ions (moléculaire et fragments) est reportée en fonction de la valeur de l'énergie de collision appliquée. Ainsi pour une énergie de collision donnée, l'abondance des ions est tout de suite observée. Lorsque l'on augmente l'énergie de collision l'abondance de l'ion parent diminue au profit de certains ions fragments du fait de sa fragmentation. La rutine (ou quercétine 3-*O*-rutinoside) se fragmente rapidement pour donner deux ions (m/z 301 et 300) associés à la génine de quercétine. La partie glycosylée (rutinoside) se sépare directement de la molécule dès que l'énergie de collision excède 20 eV. Concernant les molécules en mono-*C*-glycosylées (Apigénine 8-*C*-glucoside-2''-*O*-rhamnoside) ou di-*C*-glycosylées (Apigénine 6-*C*-glucoside 8-*C*-arabinoside), elles perdent un premier sucre (ions de masses m/z 473 et 413 respectivement) à des énergies de collision autour de 20 eV, puis le deuxième sucre (ions de masses m/z 353 et 293 respectivement) à des énergies de collision autour de 35 eV. Il faut donc fournir plus d'énergie de collision pour obtenir la fragmentation des deux sucres d'une molécule *C*-glycosylée par rapport à une molécule *O*-glycosylée.

Ainsi pour chacune des molécules à quantifier, la transition la plus adaptée peut être sélectionnée, différentes énergies de collision ont été testées et celle conduisant à la meilleure abondance d'ions a été choisie. Pour que la quantification soit optimale, d'autres paramètres ont été optimisés, comme la tension de focalisation de l'ion, la fenêtre de résolution pour le premier quadripôle et le temps d'acquisition par transition. Ce dernier paramètre est évalué dans la figure III-25 avec un exemple pour deux temps d'acquisition (200 et 1 ms/transition). Ce paramètre permet à l'analyseur de rester un temps d'acquisition suffisant par transition pour avoir des pics chromatographiques suffisamment résolus sans perdre en résolution.

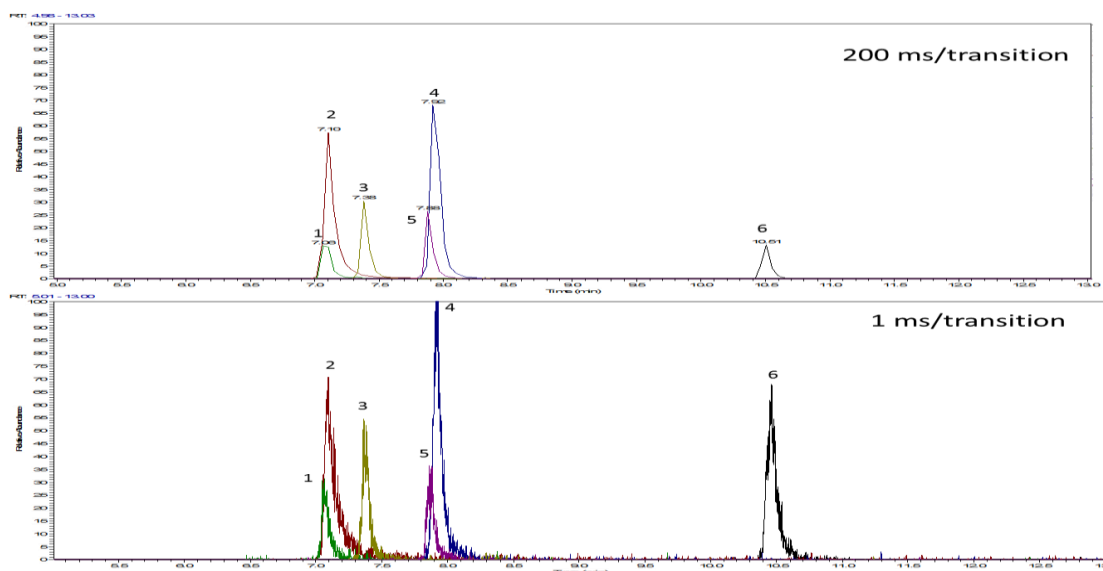


Figure III-25 : Effet du temps d'acquisition des transitions lors de l'analyse UHPLC-MS/MS d'un mélange de six standards de flavonoïdes (Shaftoside (1) : Apigénine 6-*C*-glucoside-8-*C*-arabinoside ; Orientin (2) : Lutéoline 8-*C*-glucoside ; Peltatoside (3) : Quercétine 3-*O*-arabinoglucoside ; Vitexine 2''-*O*-rhamnoside (4) : Apigénine 8-*C*-glucoside-2''-*O*-rhamnoside ; Rutine (5) : Quercétine 3-*O*-rhamnoside-7-*O*-glucoside ; Phloridzine (6) : Phlorétine-2'-*O*-glucoside) en mode MRM. Colonne Nucleodur Sphinx RP 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% d'acide formique en gradient ; débit de 0,3 mL/min ; T° four à 60°C ; volume d'injection de 3 μ L.

Plus le temps d'acquisition (dwell time) par transition est élevé, soit ici quand on passe à 200 ms, meilleur est le rapport signal/bruit des pics (Fig. III-25). Toutefois, si plusieurs transitions sont suivies sur une même période de temps, le nombre de points d'acquisition par pic diminue et la définition du pic s'avère moins bonne. Il faut donc en fonction du nombre de molécules suivies à un temps de rétention donné, trouver le juste milieu entre le nombre de points acquis pour définir correctement le pic et le rapport signal/bruit sans perdre d'information (Annexe 5 Dwell time, valeur allant de 1 à 250 ms).

De même la tension de focalisation et la résolution du premier quadripôle sont ajustées en fonction de la meilleure réponse des molécules (Annexe 5 Rf Lens, valeur allant de 140 à 250 V et Q1 Res, valeur de 0,7 ou 2).

Le tableau d'acquisition résumant les transitions des molécules suivies à des temps de rétention donnés, avec leur énergie de collision, leur temps d'acquisition, la tension de focalisation et la résolution du premier quadripôle est récapitulé en Annexe 5.

2.5.2. Quantification relative par rapport à un standard

Toutes les molécules de l'extrait n'étant pas disponibles en standards commerciaux, une molécule représentative de la famille (Vitexine 2''-O-Rhamnoside) et éluée dans la même zone de temps de rétention que les autres flavonoïdes ($T_r = 3,05$ min) a été sélectionnée comme étalon externe pour réaliser la quantification relative des composés. Par ailleurs cette molécule est à la fois *C* et *O*-glycosylée. Bien que le rendement d'ionisation de cette molécule par rapport aux autres composés puisse être différent, cette inégalité de réponse d'ionisation sera la même pour tous les échantillons ce qui permettra tout de même de comparer les concentrations relatives calculées pour un composé donné dans les différents échantillons. Les molécules dosées seront donc exprimées en concentration équivalente par rapport à la vitexine 2''-O-rhamnoside.

Lors de la séquence d'analyse des extraits de soies de maïs, une gamme d'étalonnage de cette molécule standard a été injectée à différents moments de la séquence pour encadrer les échantillons. Cette gamme est composée de 10 points de concentration allant de $0,01 \text{ mg.L}^{-1}$ à 20 mg.L^{-1} pour couvrir l'amplitude d'intensité de toutes les molécules à doser (Fig. III-26). Dans cette gamme de concentration, la linéarité de la gamme d'étalonnage a un coefficient de régression de 0,9968, ce qui permet d'avoir une bonne confiance dans nos résultats. Les points à différents moments de la séquence et pour une concentration donnée restent bien regroupés, avec un coefficient de variance inférieure à 6% pour une concentration donnée, donc le calcul des concentrations des molécules dans les échantillons injectés à différents moments peut se faire avec l'une des gammes ou la moyenne de toutes les gammes de la séquence. L'équation de l'intensité du signal en fonction de la concentration sera donc celle de la figure III-26 ($y = 67408x + 2553,4$).

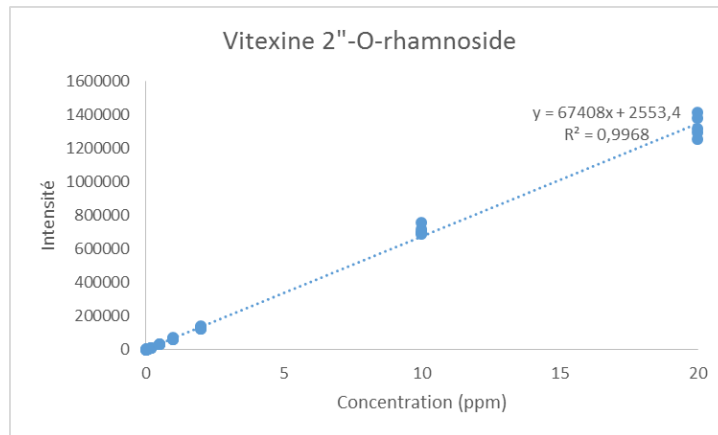


Figure III-26 : Gamme d'étalonnage (n = 5) du standard Vitexine 2''-O-Rhamnoside.

Une fois la méthode de quantification mise en place, deux séries d'analyses ont été réalisées (Annexe 8). Les 36 échantillons de soies de maïs ont été analysés trois fois chacun, des échantillons de contrôle (QC) ainsi que des points de gamme d'étalonnage externe ont également été injectés régulièrement. Pour contrer tout biais instrumental qui pourrait survenir au cours de l'ensemble des analyses, les échantillons issus d'une même variété n'ont pas été injectés successivement, mais aléatoirement. Une série d'analyses (une séquence) représente entre 80 et 90 injections. L'ensemble des analyses a représenté 470 injections d'une durée de 15 minutes chacune ; soit une durée totale d'analyse de 118 heures répartie sur les deux séquences d'analyse. Un suivi de la stabilité du système analytique (UHPLC et MS), a été effectué pendant 75 injections et n'a pas montré de déviation de la réponse pour les molécules de shaftoside et peltatoside en mode d'ionisation négatif. De plus, pour vérifier la stabilité du système tout au long des analyses, un échantillon contrôle (QC) est injecté régulièrement. Ce QC est composé d'un mélange équivolumique de 15 échantillons d'extrait de soies de maïs prélevés aléatoirement. Des blancs de solvant de préparation des échantillons sont également injectés régulièrement pour vérifier l'absence de contamination du système.

2.6. Traitement de données

Après avoir analysé et quantifié dans tous les échantillons, les molécules de rapport m/z compris dans la gamme de masse 400 à 700 Da et correspondant aux molécules glycosylées, un traitement de données est réalisé via metaboanalyst 3.0, qui est une plateforme en accès libre sur le web (Xia and Wishart 2016). Cette plateforme permet d'effectuer les traitements statistiques uni et multivariés.

Une fois l'analyse de tous les échantillons effectuée avec les deux séquences, les résultats sur les différents composés et dans toutes les variétés sont visualisés dans leur globalité pour voir d'abord l'influence des deux séquences. Un prétraitement des données est également effectué pour homogénéiser les réponses (Liland 2011). Et des analyses non supervisées et supervisées sont également réalisées pour mettre en évidence les différences entre les variétés et les molécules qui influencent cette distinction.

2.6.1. Prétraitement des données

2.6.1.1. Influence des 2 séquences

Un suivi des analyses au cours du temps en fonction de la somme de la concentration de l'ensemble des composés dosés dans chaque variété est présenté dans la figure III-27 ci-dessous. Ce graphique consiste à représenter tous les échantillons analysés dans leur ordre d'injection au cours des deux séquences. Les échantillons contrôle (QC en orange) permettent de mettre en évidence un éventuel problème d'analyse de notre système UHPLC/MRM MS dans une même séquence et d'une séquence à une autre. Ainsi dans une même série, les échantillons contrôle varient très peu (Cv 4,12%). Mais la variation entre les deux séries est un peu plus importante et correspondant à une augmentation d'un facteur 2. Donc pour les deux séquences d'analyse une correction des concentrations est effectuée en divisant la concentration de chaque composé de chaque échantillon d'extrait de soie par la valeur moyenne de sa propre concentration dans l'échantillon QC injecté plusieurs fois au cours de la même série d'analyse. Avec cette correction, la variation des QC n'est plus que de 3,5% pour l'ensemble des deux séquences (Fig. III-27B).

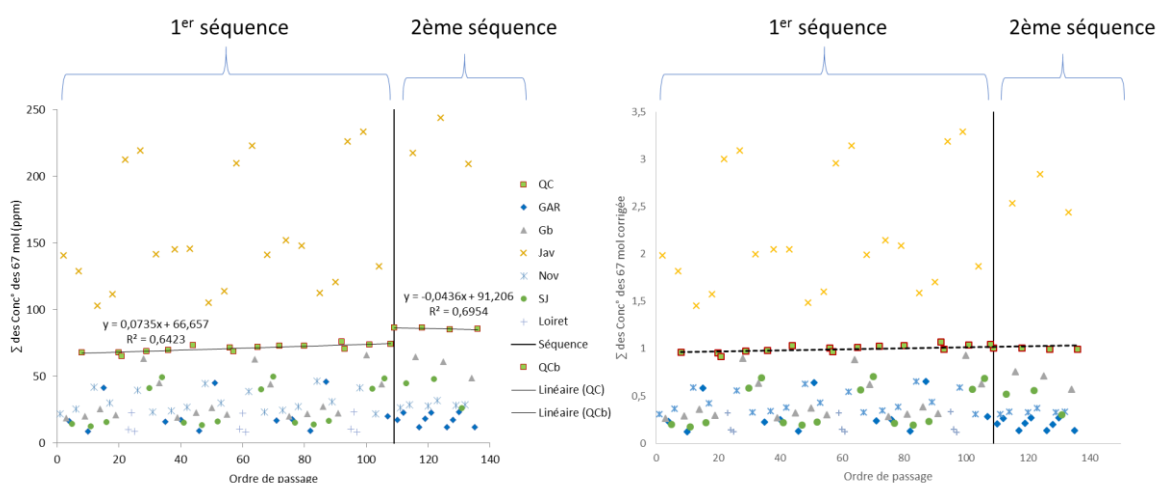


Figure III-27 : Suivi des analyses au cours du temps ; somme des concentrations des 67 composés quantifiés sans correction (à gauche) et après correction des concentrations (à droite).

On voit que grâce à la correction les concentrations totales pour toutes les injections de QC ont un même ordre de grandeur sur l'ensemble des deux séquences d'analyse. L'ensemble des analyses est donc maintenant comparable. D'après le graphique, la variété Java (croix jaunes) possède les concentrations les plus élevées et les plus dispersées en molécules, alors que les autres variétés semblent plus proches en termes de concentration pour la globalité des composés.

2.6.1.2. Normalisation

Avant de réaliser des études métabolomiques, les données peuvent être transformées, afin que les concentrations aient le même poids dans chaque variété et dans chaque lot. En effet, il existe différents types d'opération. Les données peuvent être centrées, en soustrayant chaque valeur à la moyenne de sa colonne, ainsi le nuage de points est centré sur l'origine. Elles peuvent être aussi réduites, en divisant chaque donnée par l'écart type de sa colonne. Chaque colonne a ainsi la même variation globale, qui a le même poids. On peut également faire des combinaisons de traitement. Les variations des variables peuvent donc être comparées plus aisément. Metaboanalyst propose d'autres opérations telles qu'une transformation logarithmique, par pareto (centrée et divisée par la racine carrée), par exemple. Ainsi on peut adapter les transformations en fonction du jeu de données pour avoir des données qui suivent au plus près une gaussienne. Dans notre cas, les données ont été transformées suivant la combinaison d'une normalisation quantile pour les échantillons (lignes) et également une transformation en racine cubique avec une mise à l'échelle des données (range scaling ; moyenne des données centrées et divisée par l'intervalle de chaque variable) pour les variables (colonnes) en enlevant les valeurs non significatives ce qui permet d'avoir une répartition proche d'une gaussienne des données pour les individus et les variables (Fig. III-28). La mise à l'échelle des données permet de comparer les variables bien que les réponses des molécules par variétés soient très différentes. En effet, les concentrations des composés constituant la variété Java sont environ quatre fois plus abondantes. Ainsi même s'il y a de grandes différences de concentration entre des métabolites (comme ici pour Java), ils ont tous le même poids. La normalisation pour les échantillons permet de réduire les biais dus aux individus qui sont différents lors de la collecte des échantillons.

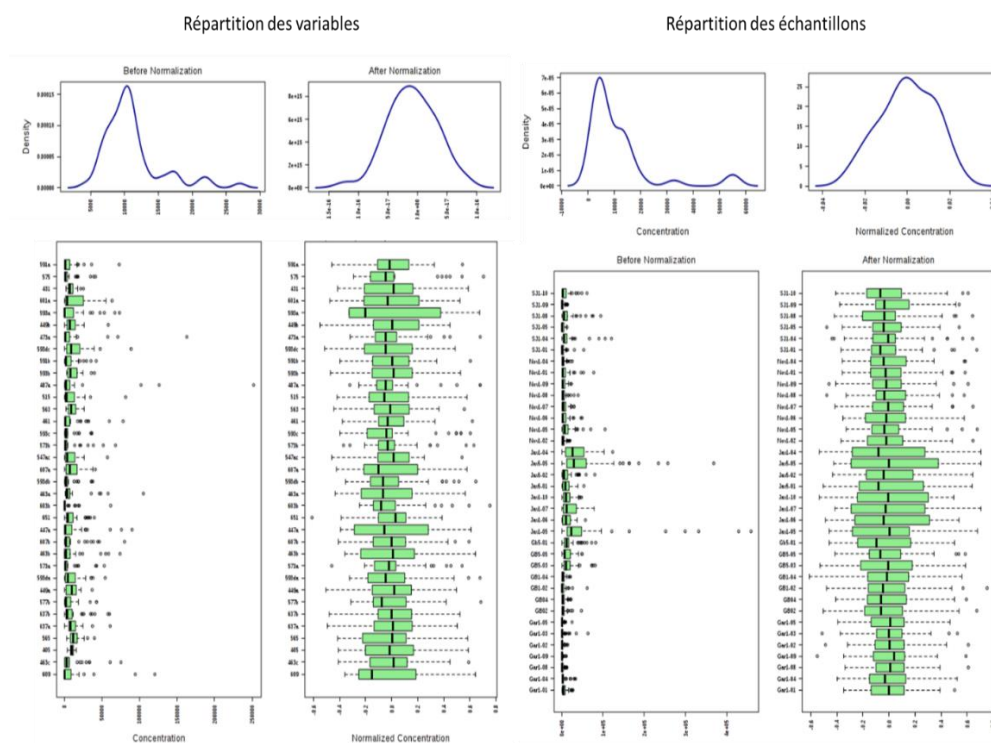


Figure III-28 : Répartition des 35 premières variables (molécules) et des échantillons (variétés) avant et après transformation (normalisation quantile ; range scaling).

Ainsi les données sont prétraitées tout d'abord en effectuant la division par la moyenne des QC, puis par un traitement de normalisation, de transformation et de mise à l'échelle des données pour avoir des données qui se rapprochent le plus de la loi normale pour les variables et les individus. Une fois le prétraitement des données réalisé, des analyses non supervisées et supervisées sont effectuées. Les analyses multivariées utilisées permettent d'avoir une représentation graphique plus visuelle pour représenter un grand nombre d'échantillons et de variables en seulement 2 à 3 dimensions.

2.6.2. Classification Ascendante Hiérarchique : CAH

La Classification Ascendante Hiérarchique est une méthode non supervisée qui permet de classer les individus par mesure de la similarité.

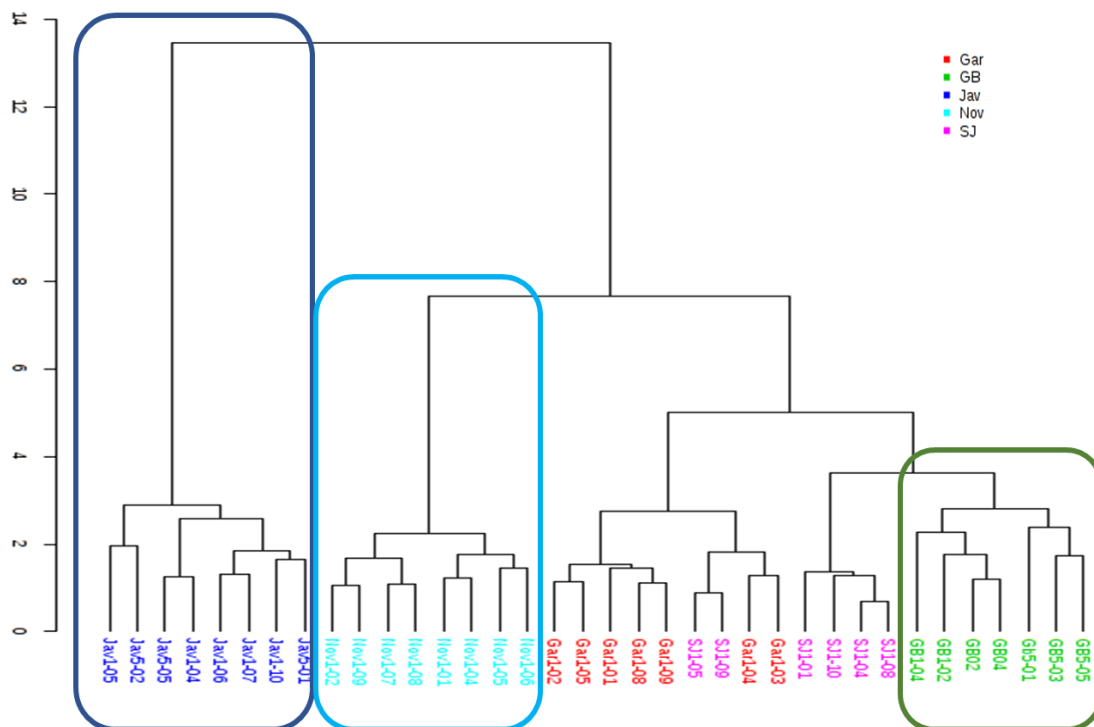


Figure III-29 : Classification ascendante hiérarchique des individus (variétés) établie à partir des résultats de dosage quantitatif de 67 composés dans les 36 soies de maïs provenant de 5 variétés.

Cette classification ascendante hiérarchique a été effectuée en réalisant la mesure des distances euclidiennes et la réalisation des groupes via l'algorithme de Ward. Sans indiquer quel échantillon appartenait à quel groupe, on peut observer que le dendrogramme répartit les échantillons suivant leur appartenance aux variétés pour la plupart (Fig. III-29). Ainsi ce type d'analyse permet de montrer que la variété Java (bleu foncé) se différencie le plus des autres variétés, puis la variété Nova (bleu clair), alors que les variétés Garrison (rouge) et Sugar Jean (violet) sont mélangées, la variété Golden Bantam (vert) est bien rassemblée en un groupe, mais très similaire à une partie des échantillons de la variété Sugar Jean.

Une autre méthode non supervisée a ensuite été testée pour confirmer ce résultat, observer les corrélations avec les variables (molécules) et visualiser s'il n'y a pas de données aberrantes, que ce soit pour un lot d'une variété ou pour une molécule dosée.

2.6.3. Analyse en Composantes Principales (ACP)

L'ACP permet d'illustrer les données par un graphique des variables (molécules dosées) et par un graphique des individus (échantillon de soies) (Fig. III-30). Elle permet de mettre en valeur des tendances ou des comportements inhabituels des individus et des variables, et de mettre en évidence des variables corrélées aux individus. Ces deux graphiques sont liés et sont utilisés ensemble pour leurs interprétations.

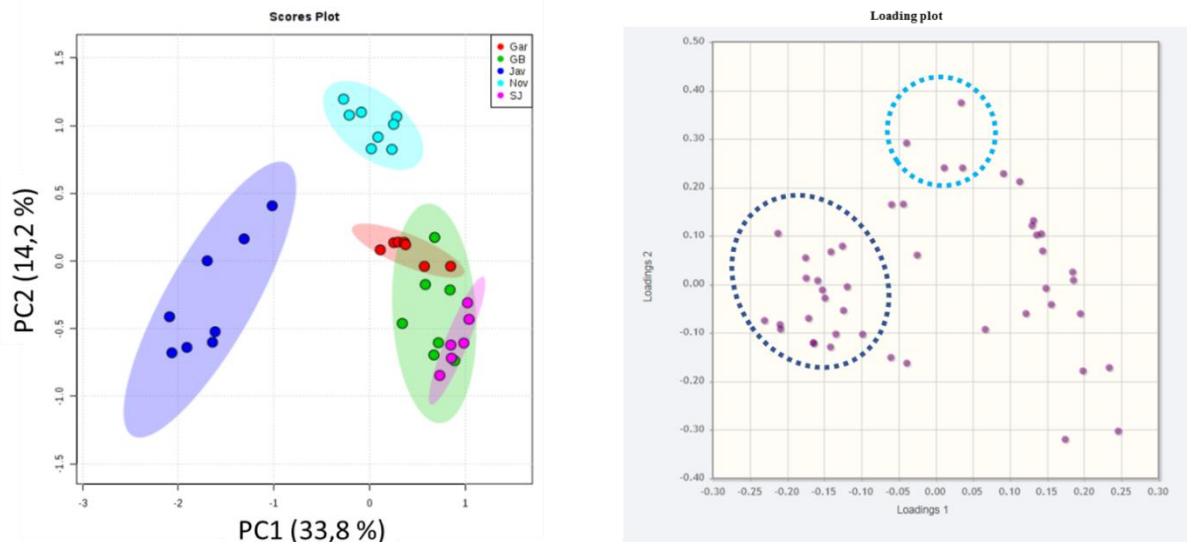


Figure III-30 : Score plot (variétés) et loading plot (molécules) de l'ACP.

Les deux premières composantes principales représentent 48% de la variance initiale (Fig. III-30). Elles permettent de mettre en évidence que la variété Java (bleu foncé) se distingue des autres variétés comme pour la CAH. Cette distinction se fait principalement suivant la composante principale 1 avec une variance de 33,8%. La variété Nova (bleu clair) se différencie via la composante principale 2, mais avec une variance seulement de 14,2%. Le graphique de la représentation factorielle des variables met en évidence les molécules qui sont corrélées et anti-corrélées pour ces deux mêmes composantes principales et celles qui n'ont pas d'impact sur ces deux dimensions. Un groupe de molécules (entourées par le cercle bleu foncé) va dans le même sens que la variété Java, ces molécules commencent également à se distinguer ainsi que quatre molécules (entourées dans le cercle bleu clair) qui vont dans le sens de la variété Nova.

Les analyses non supervisées permettent de valider que la teneur en flavonoïdes glycosylés choisis comme molécules d'intérêt dépend de la variété de maïs. Ainsi les variétés de maïs peuvent donc être différenciées selon leur composition en flavonoïdes glycosylés.

Pour aller plus loin dans l'étude et identifier des composés significatifs qui permettent de différencier la variété Java, une étude supervisée a été menée.

2.6.4. sPLS-DA

L'algorithme sparse PLS-DA (sPLS-DA) permet de réduire le nombre de variables dans les données métabolomiques pour produire un modèle plus robuste et plus facile à interpréter. Les groupes sont forcés via la méthode de sPLS-DA et la méthode utilise un nombre restreint de variables qui permettent de les séparer. Suite aux résultats de l'ACP, les composés permettant de séparer la variété Java était d'environ une vingtaine, donc il a été fixé de prendre les 20 premières variables de chaque composante.

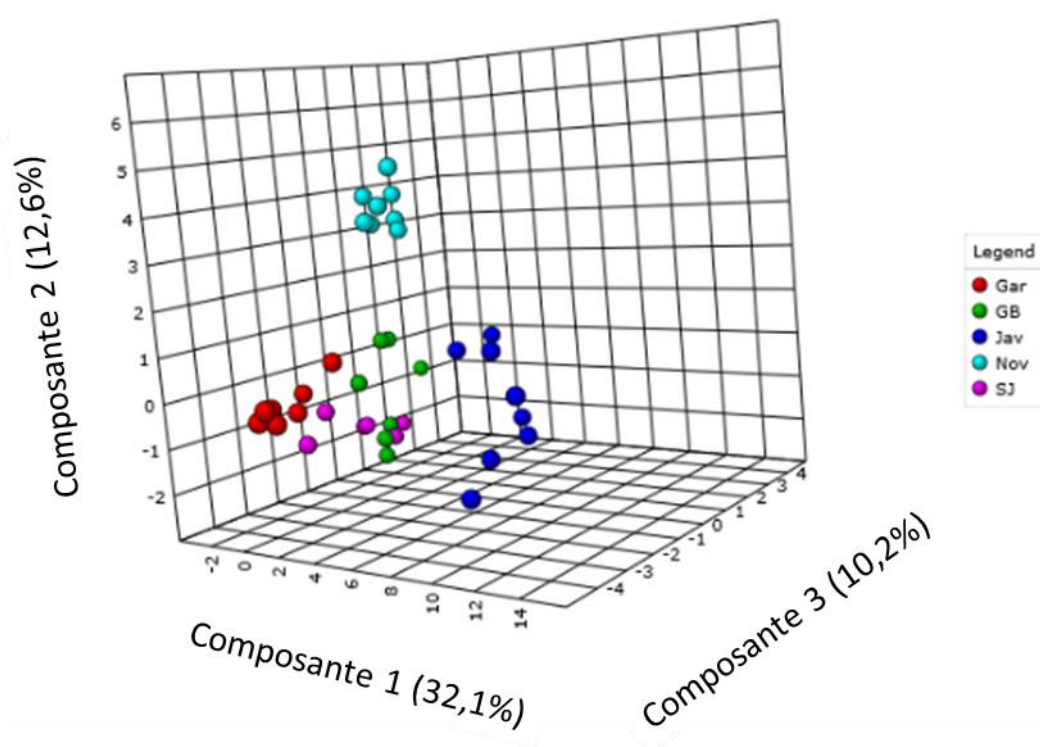


Figure III-31 : Score plot (variétés) du sPLS-DA.

Ce modèle de régression (Fig. III-31) permet de décrire 68,4 % de la variance en utilisant les 5 premières composantes et 44,9% de la variance avec trois composantes. Les performances du modèle sPLS-DA ont été évaluées en utilisant une validation croisée des données qui montre un taux d'erreur de 11%. Le faible taux d'erreur de ce modèle permet d'avoir une certaine confiance dans la détermination des molécules qui permettent de distinguer les variétés par composante. La variété Java est séparée des autres variétés via la première composante avec une variance de 32,1%, alors que la variété Nova est quant à elle séparée par la deuxième composante avec une variance de 12,6%.

Ces différents traitements de données multivariées en mode supervisé et non supervisé ont permis de distinguer des variétés plus que d'autres. En effet, la variété Java se différencie très bien des autres variétés quelle que soit la méthode statistique utilisée. Quant aux variétés Golden bantam, Sugar Jean et Garrison, elles ont des compositions en flavonoïdes glycosylés très similaires et sont peu distinguables. Enfin la variété Nova se différencie légèrement, mais sur des axes moins représentatifs de la variance globale.

2.6.5. Interprétation des résultats et comparaison avec les résultats terrain

2.6.5.1. Etudes terrain du CIRAD

En parallèle, le CIRAD a réalisé des études terrain sur ces mêmes variétés, et la variété Java s'est montrée la variété la plus efficace dans l'inhibition de la croissance des larves de *H. zea* (Fig. III-32). La variété Nova et les trois autres variétés étudiées (Garrison, Golden Bantam et Sugar Jean) présentent un comportement similaire et moins efficace sur l'inhibition de la croissance des larves de *H. zea*.

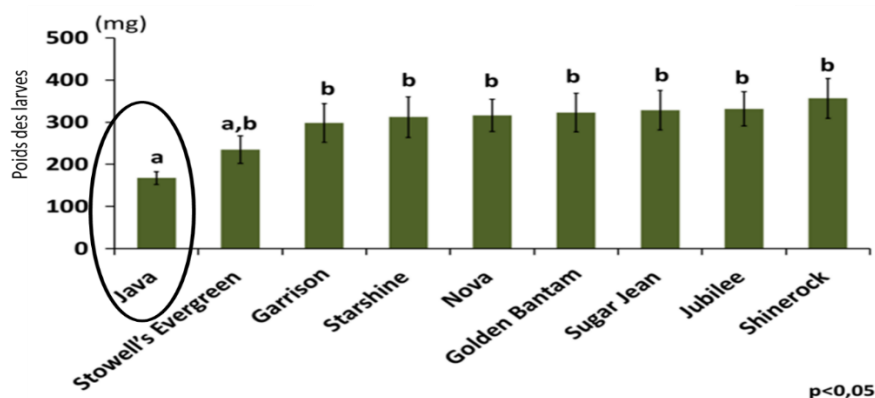


Figure III-32 : Poids moyen des stades larvaires de *H. zea*, tiré de la publication Rhino 2016.

Que ce soit dans l'étude de terrain (variété la plus efficace) ou dans l'étude métabolomique ciblée des flavonoïdes glycosylés, la variété Java se distingue (selon la composante 1) le mieux des autres variétés.

L'information du modèle de régression sPLS-DA peut permettre d'avoir une interprétation moléculaire des différences entre les variétés par la projection des variables importantes (VIP) données via une valeur de score. Ainsi les résultats de la figure III-33 permettent de mettre en évidence les molécules qui influencent les composantes 1 et 2, et donc qui différencient la variété Java et la variété Nova respectivement.

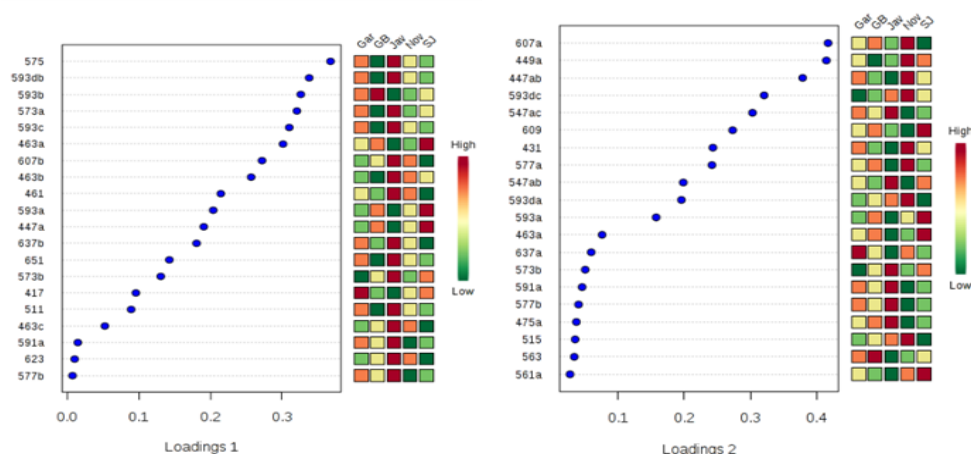


Figure III-33 : Loading plot des composantes 1 et 2 avec leur valeur de score, ainsi que leur abondance (rouge = forte ; vert = faible) dans la variété.

2.6.5.2. Molécules VIP de la composante 1

Les premières molécules ayant le meilleur score dans une composante correspondent à une variété où elles sont bien représentées. Ainsi pour la composante 1 où la variété Java était bien séparée des autres variétés, pour les quatre premières variables, on retrouve que les molécules de masses m/z 575 ($Tr = 4,27$ min), 593 ($Tr = 4,41$ min) et 573 ($Tr = 4,47$ min) sont surexprimées dans la variété Java et la molécule de masse m/z 593 ($Tr = 2,54$ min) est sous exprimée dans cette même variété. Ceci est illustré par les boîtes à moustache obtenues pour ces variables (Fig. III-34).

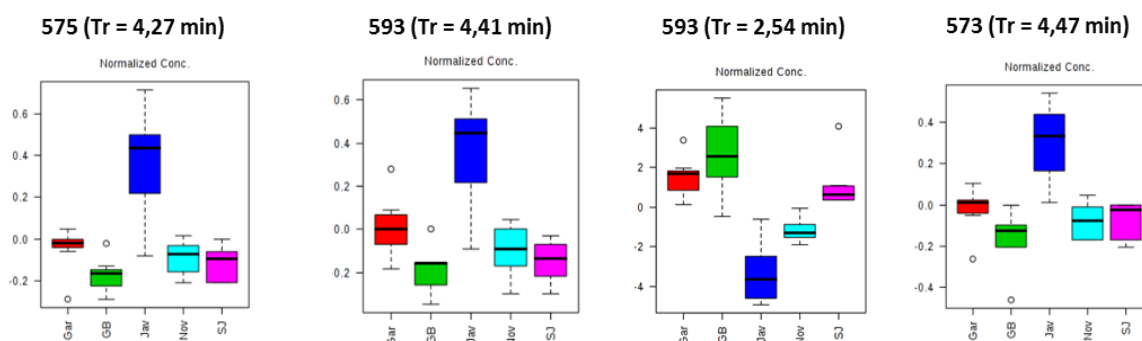


Figure III-34 : Boîte à moustache des 4 principaux VIP de la composante 1 (p-value < 0,05).

Cette représentation confirme que ces quatre molécules sont corrélées à la variété Java et que la molécule de masse m/z 593 est anti-corrélée à cette même variété.

Comme le montre les extraits d'ions des molécules VIP présentés sur la figure III-35, les trois molécules surexprimées sortent dans la même gamme de temps de rétention (Tr 4-4,5 min), contrairement à celle qui est sous exprimée. Ce sont des molécules qui ont des abondances dans l'extrait différentes : les molécules de m/z 575 et 593 (sur exprimées) font partie de molécules majoritaires tandis que les

molécules de m/z 593 (sous exprimée) et 573 (sur exprimée) sont présentent en plus faible quantité (Fig. III-35).

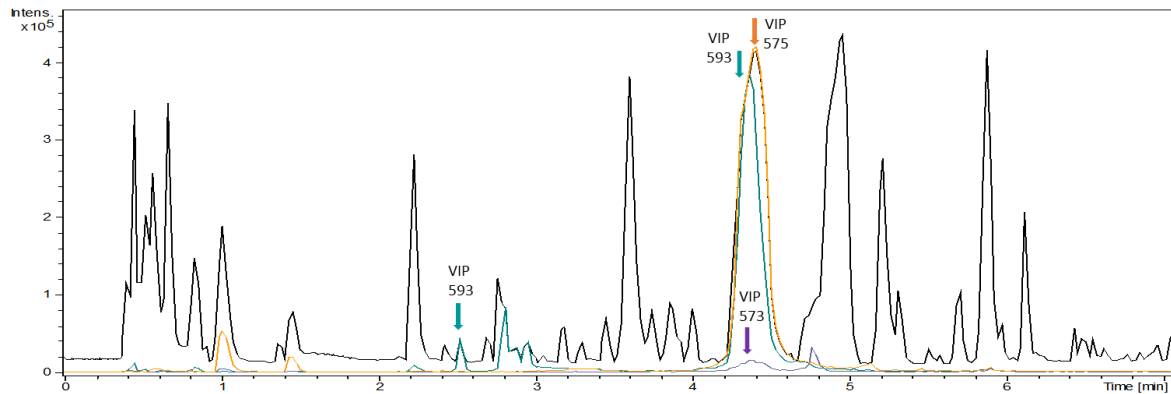


Figure III-35 : Base Peak Chromatogram (BPC) d'un mélange de soies de maïs de la variété Java obtenu en UHPLC/HRMS, avec les extraits d'ions des molécules VIP de la composante 1. Colonne Nucleodur Pyramid 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% et 0,08% d'acide formique respectivement en gradient ; T° four à 60°C ; débit de 0,6 mL/min ; volume d'injection de 3 μ L.

Ainsi ces quatre molécules peuvent correspondre respectivement à la lutéoline C-oxodeoxyhexose-O-deoxyhexose, soit la Maysin (m/z 575 - Tr = 4,27 min), à la lutéoline C-hexose-O-désoxyhexose (m/z 593 - Tr = 4,41 min), et à une molécule (m/z 573 - Tr = 4,47 min) partiellement identifiée, comme ayant un sucre O-deoxyhexose lié à un sucre C-glycosylé (cf 3.4.3.2.). L'autre molécule de masse m/z 593 (Tr = 2,54 min), qui est anti corrélée correspondrait à une apigénine C-hexose-C-hexose. Les molécules surexprimées sont des flavonoïdes mono-C-glycosylés-O-glycosylés dont la molécule Maysin (lutéoline C-oxodeoxyhexose-O-deoxyhexose) a déjà été décrite comme inhibitrice de la croissance des larves *H. zea* (Elliger 1980a). Ainsi notre approche de métabolomique ciblée en flavonoïdes glycosylés est en cohérence avec la littérature.

2.6.5.3. Molécules VIP de la composante 2

Concernant les VIP influençant la composante 2, les trois premières molécules (m/z 607 - Tr = 3,42 min ; m/z 449 - Tr = 3,72 min et m/z 447 - Tr = 4,39 min) sont surexprimées dans la variété Nova, variété qui est différenciée par cette composante. La variété Nova n'est pas liée par les observations de terrain à une activité inhibitrice particulière, mais elle se différencie des autres variétés dans les analyses statistiques et nous avons souhaité déterminer les molécules qui induisent cette différence.

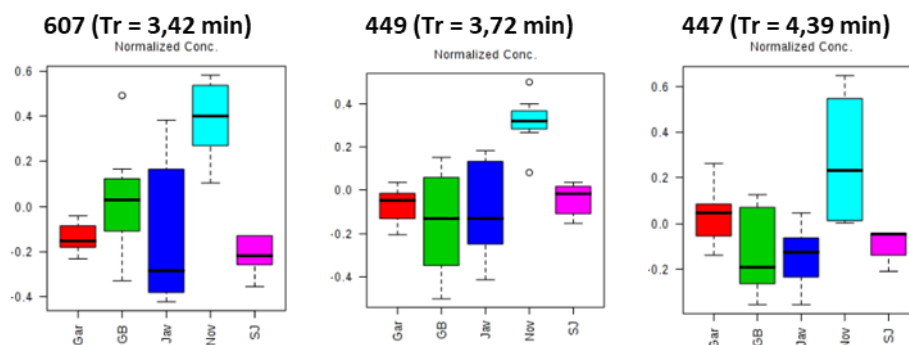


Figure III-36 : Boîte à moustache des 3 principaux VIP de la composante 2 (p -value < 0,05).

Les boîtes à moustaches concernant les molécules de la deuxième composante (Fig. III-36) montrent que la différenciation est moins nette qu'avec la première composante. En effet, la variance de la composante 2 n'est que de 12,6%, au lieu de 32,1% pour la composante 1. Néanmoins les molécules qui permettent de différencier la variété Nova des autres pourraient correspondre à des molécules de chrysoeriol *C*-hexose-2''-*O*-deoxyhexose (m/z 607 - Tr = 3,42 min), ériodictyol *O*-hexose (m/z 449 - Tr = 3,72 min) et lutéoline *O*-hexose (m/z 447 - Tr = 4,39 min). Ces flavonoïdes sont mono-glycosylés avec un ou deux sucres, aussi bien en *C* qu'en *O*-glycosylation. Ces molécules ont des temps de rétention différents (Fig. III-37) des VIP identifiés pour la composante 1. En effet, elles ont des structures différentes. Ce sont des molécules dont la teneur est faible dans la variété. Les molécules *O*-glycosylés ne sont pas décrites comme ayant des activités inhibitrices. La molécule chrysoeriol *C*-hexose-2''-*O*-deoxyhexose (m/z 607 - Tr = 3,42 min) pourrait avoir un potentiel d'activité d'inhibition, car elle a la même génine que la molécule 3'-méthoxymaysin, décrite comme une des molécules inhibitrice, mais le composé est à une faible teneur (0,82 mg.L⁻¹ équivalent de vitexine 2''-*O*-rhamnoside) dans la variété Nova. Cette valeur est très faible par rapport à la teneur de la molécule 3'-méthoxymaysin (Nova : 7,75 mg.L⁻¹ équivalent de vitexine 2''-*O*-rhamnoside ; Java : 41,15 mg.L⁻¹ équivalent de vitexine 2''-*O*-rhamnoside).

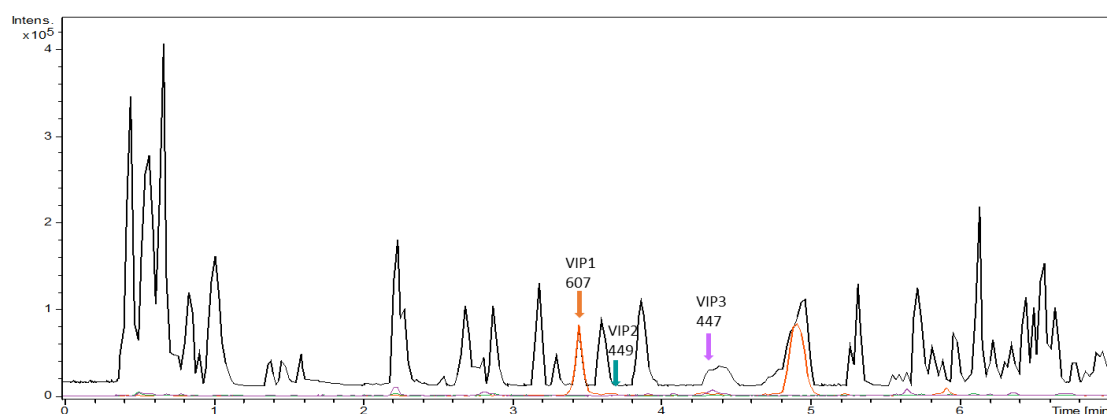


Figure III-37 : Base Peak Chromatogram (BPC) d'un mélange de soies de maïs de la variété Nova obtenu en UHPLC/HRMS, avec les extraits d'ions des molécules VIP de la composante 2. Colonne Nucleodur Pyramid 150 x 2 mm, 1,8 μ m ; phase mobile constituée d'un mélange d'eau et d'acétonitrile avec 0,1% et 0,08% d'acide formique respectivement en gradient ; T° four à 60°C ; débit de 0,6 mL/min ; volume d'injection de 3 μ L.

Concernant les deux autres molécules (apimaysin et 3'-méthoxymaysin) décrites comme ayant une activité inhibitrice de la croissance des larves, comme la molécule maysin, elles ont bien été retrouvées dans les variétés. Mais elles ne sont pas ressorties comme VIP pour une des variétés. En effet, La molécule apimaysin est présente dans toutes les variétés à des teneurs similaires. La molécule 3'-méthoxymaysin est en proportion aussi abondante dans la variété Java que dans la variété Nova. Cette molécule est presque aussi abondante que la molécule maysin. Ces molécules ne sont donc pas sous ou surexprimées dans une variété en particulière. Donc bien que ces molécules présentent des activités inhibitrices de la croissance des larves, elles ne suffisent pas à elles seules à donner une activité inhibitrice aux variétés étudiées qui les contiennent.

Afin de confirmer et d'approfondir la caractérisation des composés d'intérêt des soies de maïs dont les VIP par une approche complémentaire, un réseau moléculaire a été élaboré.

2.7. Identification par réseaux moléculaires

Les réseaux moléculaires décrits dans le chapitre 1 (*cf* paragraphe 5.2.3) sont une technique émergente de plus en plus utilisés, notamment en France dans le domaine de la phytochimie. En effet, ils cartographient les composés en les rassemblant en cluster en fonction de leur analogie de spectres de fragmentation MS^2 . Cette technique permet de visualiser rapidement les composés ayant le même schéma de fragmentation et donc appartenant à la même famille moléculaire. Ainsi les composés inconnus ou les VIP peuvent se retrouver dans un cluster de composés connus, et en faciliter l'identification. Lors d'analyse de dérégulation, les nouveaux composés peuvent être rapidement identifiés. Et lors de la comparaison de plusieurs espèces ou d'organes, les molécules spécifiques à une espèce ou à un organe sont rapidement repérées.

Nous avons souhaité évaluer cette approche dans la caractérisation des soies de maïs, l'utilisation d'un réseau moléculaire peut permettre de confirmer nos identifications préalables et de compléter les identifications d'autres composés ou des VIP.

2.7.1. Réseaux moléculaires dans les deux modes d'ionisation

Dans le chapitre 1, il a été mis en avant que suivant le mode d'ionisation des flavonoïdes, la fragmentation différait. Ainsi davantage de fragments peuvent être obtenus en mode d'ionisation positif que négatif dans le cas de C-glycosylation. Comme la construction du réseau se fait en fonction du nombre de fragments communs entre deux molécules, la comparaison des réseaux moléculaires dans les deux modes d'ionisation peut apporter des informations complémentaires. Le nombre de fragments communs minimum a été fixé à 3, d'après l'étude de la fragmentation des flavonoïdes, comme tous les composés ne donnent pas beaucoup de fragments, la valeur fixée est faible. Ces réseaux moléculaires ont été obtenus à partir des analyses effectuées par UHPLC/HRMS-MS de l'échantillon constitué du mélange des extraits de soies de maïs et ayant servi à l'identification des flavonoïdes glycosylés (*cf* paragraphe 3.4.5).

Les réseaux moléculaires présentés Figure III-38 permettent de mettre en évidence huit clusters principaux que cela soit en mode d'ionisation négatif (A), ou en mode d'ionisation positif (B). Lors de la comparaison des deux réseaux, certains clusters diffèrent. En effet, en mode d'ionisation positif, il

apparaît un cluster (cluster 9 rose) de cinq molécules de masse supérieure à 700 Da, qui n'est pas présent en mode d'ionisation négatif. De même en mode d'ionisation négatif, il y a un cluster de neuf molécules (cluster 1 vert - orange) qui n'est pas présent en mode d'ionisation positif. Ces différences entre les deux réseaux sont dues soit à des composés qui s'ionisent dans un mode et pas dans l'autre, soit à une meilleure fragmentation dans un mode qui génère suffisamment de fragments communs pour associer les molécules dans un même cluster.

La plupart des clusters sont associés par gamme de masse. Certains composés du cluster 1 ont un rapport m/z pair et leurs formules brutes contiennent un atome d'azote (Fig. III-22). Donc ce cluster correspondrait à des molécules de types acides hydroxamiques. Afin d'annoter les différents clusters, les composés déjà identifiés ont été positionnés sur les réseaux, bien que tous les composés identifiés (cf Tab III-7 à 10) n'aient pas été retrouvés dans ces réseaux. Ainsi quel que soit le mode d'ionisation les flavonoïdes ont été clustérisés par sous familles suivant le type de liaisons avec les sucres. En mode d'ionisation négatif, les mono *O*-glycosylés sont rassemblés (cluster 3), qu'ils aient un ou deux sucres, alors que les mono-*C*-glycosylés (cluster 8) sont séparés des di-*C*-glycosylés (cluster 7). Concernant les regroupements des trois flavonoïdes mono-*C*-glycosylés-*O*-glycosylés (cluster 2), il s'est fait, car les molécules ont des pertes de sucres identiques, mais des génines différentes. Alors que les trois flavonoïdes di-*C*-glycosylés (cluster 7) ont la même génine (apigénine) et une perte commune. Ainsi la molécule de rapport m/z 593 est liée à la molécule de rapport m/z 563, car elles possèdent toutes les deux un sucre hexose, alors que la molécule de rapport m/z 547 est liée à la molécule de rapport m/z 563, car elles ont en commun un sucre pentose.

De plus, les VIP ont également été recherchés dans les clusters, mais seuls deux VIP (maysin avec m/z 575 - Tr = 4,27 min et apigénine di-*C*-hexose avec m/z 593 - Tr = 2,54 min) ont été retrouvés respectivement dans les clusters 2 et 7. Ces composés se retrouvent avec deux autres molécules identifiées respectivement comme flavonoïdes mono-*C*-glycosylés-*O*-glycosylés et flavonoïdes di-*C*-glycosylés. Le réseau moléculaire confirme l'identification donnée pour ces deux composés VIP. L'identification des deux autres VIP n'a pas pu être complétée car les ions correspondants à ces molécules n'apparaissent pas dans les réseaux générés. Il est difficile de comprendre pourquoi les VIP ne sont pas tous retrouvés dans le réseau, alors que leur identification a pu être réalisée. Les VIP non repérés ont soit un ion moléculaire trop petit qui n'a pas été sélectionné lors de la fragmentation, soit les paramètres des réseaux doivent être modifiés pour obtenir plus de composés.

Quelques masses de molécules décrites dans la littérature (Tableau III-1) ont été retrouvées dans le réseau moléculaire, dont les trois flavonoïdes les plus décrits (maysin m/z 575, apimaysin m/z 559 et 3'-methoxymaysin m/z 589), qui se retrouvent dans le cluster 2. Ainsi que deux des composés *O*-glycosylés (m/z 461 et m/z 593), qui sont dans le cluster 3.

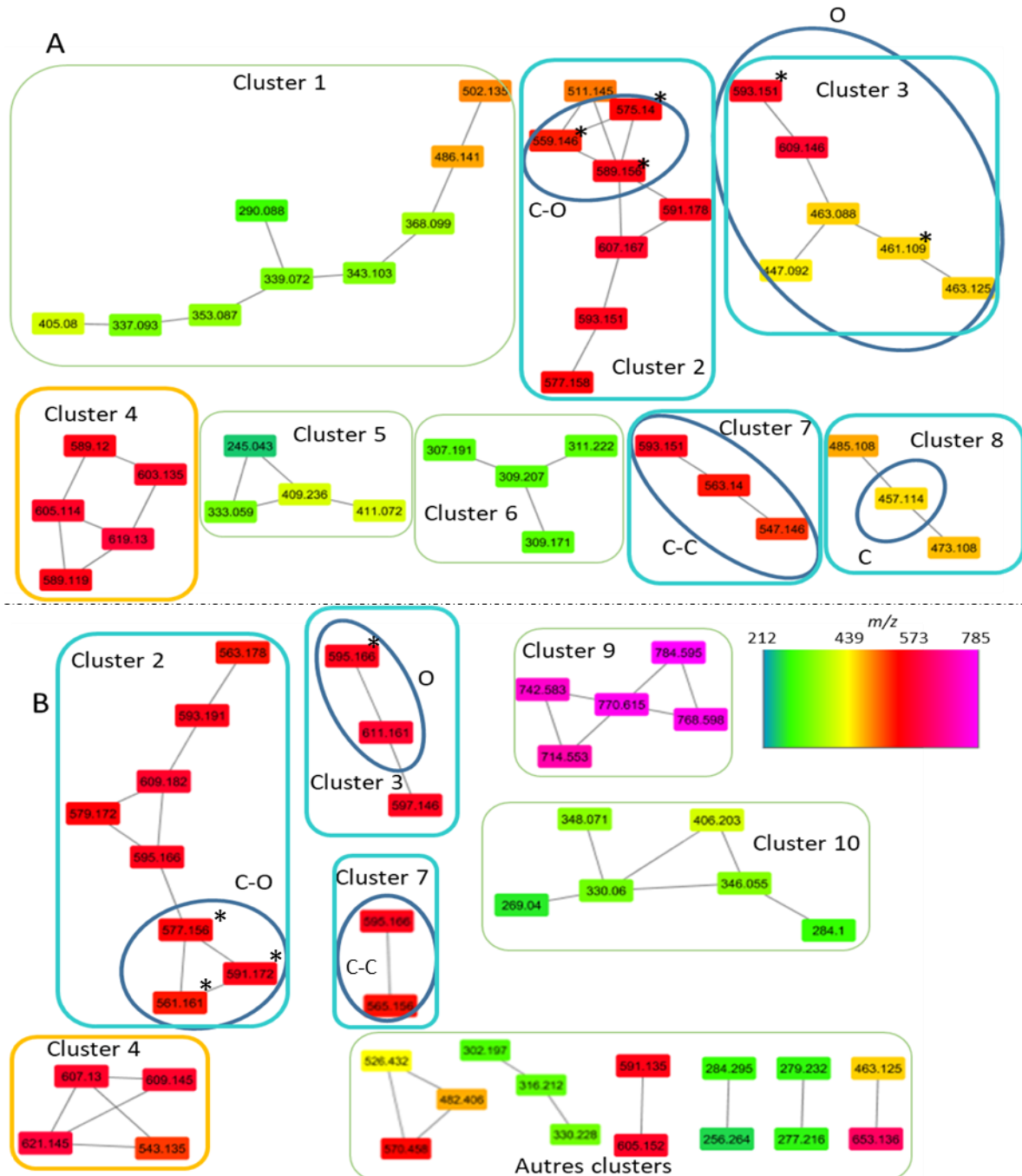


Figure III-38 : Réseaux moléculaires en mode d'ionisation négatif (A) et en mode d'ionisation positif (B). Les couleurs sont en fonction d'un gradient de rapport m/z (rouge : m/z les plus élevés et vert : m/z les plus faibles). Les cercles bleus correspondent aux flavonoïdes identifiés en fonction de leurs liaisons glycosylés (C : flavonoïdes mono-C-glycosylés ; C-C : flavonoïdes di-C-glycosylés ; O : flavonoïdes O-glycosylés ; C-O : flavonoïdes mono-C-glycosylés-O-glycosylés) Tolérance en masse de l'ion précurseur 0,02 Da ; tolérance en masse des ions fragments 0,1 Da. Paramètres du réseau : Cosinus score 0,7 ; nombre minimum d'ions fragments communs partagés par deux spectres MS/MS : 3 ; nombre maximum de nœuds voisins : 10 ; taille maximale des nœuds autorisés dans un seul réseau : 100 ; Run MS Cluster : Yes, accepte qu'il fusionne des spectres identiques.

Afin d'obtenir d'avantages d'informations sur les clusters n'ayant pas de molécules identifiées (Clusters 1, 4, 5 et 6), les formules brutes des masses de ces clusters, ainsi que ceux où des flavonoïdes ont été identifiés, ont été utilisées pour tracer la masse moléculaire en fonction du nombre équivalent de double liaison (Fig. III-39A) et ont également été reportées sur un diagramme de Van Krevelen (Fig. III-39B).

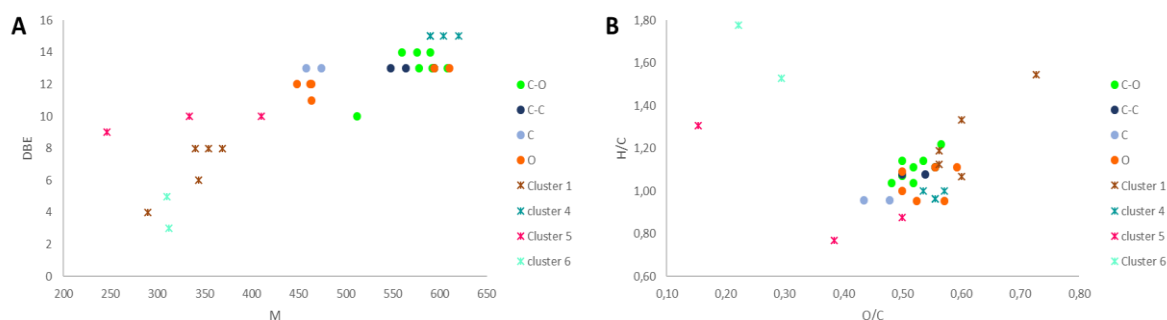


Figure III-39 : Représentations graphiques : (A) Nombre équivalent de double liaison en fonction de la masse moléculaire et (B) Van Krevelen, ratio H/C en fonction du ratio O/C. Ces graphiques ont été construits à partir de 35 composés et leur formule brute.

Trois des clusters (1 ; 5 et 6) non identifiés ont des masses plus faibles et beaucoup moins d'insaturations que les quatre autres clusters (2 ; 7 ; 8 et 3) correspondant aux molécules identifiées comme des flavonoïdes glycosylés. Le cluster 1 contiendrait des molécules de types acides hydroxamiques, qui ont des petits temps de rétention (< 2 min), comme les molécules qui constituent le cluster 5. Une des molécules du cluster 5 a peut-être une mauvaise attribution de formule brute, car elle a des valeurs de rapport H/C et O/C très éloignées des autres (Fig. III-39B). Le cluster 6 renferme des composés élués à des temps de rétention (> 6 min) plus élevés que les flavonoïdes et ces composés sont positionnés dans le diagramme de Van Krevelen dans la zone des lipides (Fig. III-22). Le cluster 4 est proche en termes de masse et d'insaturation des molécules flavonoïdes di-glycosylés (Fig. III-39A). Cela se confirme dans le diagramme de Van Krevelen, où le cluster 4 se retrouve confondu avec les clusters des flavonoïdes (Fig. III-39B).

Le cluster 4 étant proche des flavonoïdes, la valeur du cosinus score est modifiée afin de mettre en évidence si ces clusters peuvent se lier.

2.7.2. Modification du cosinus score

La valeur seuil du cosinus score correspond à la ressemblance qu'il existe entre deux spectres de fragmentation. Quand cette valeur est égale 1, cela correspond à deux spectres de fragmentation identiques, quand elle est proche de zéro, les deux spectres de masse sont totalement différents. Cette valeur a été abaissée à 0,6 pour mettre en évidence si des clusters en mode d'ionisation négatif (Fig. III-40) se regroupaient ce qui impliquerait une similitude dans les schémas de fragmentation des molécules appartenant à ces clusters.

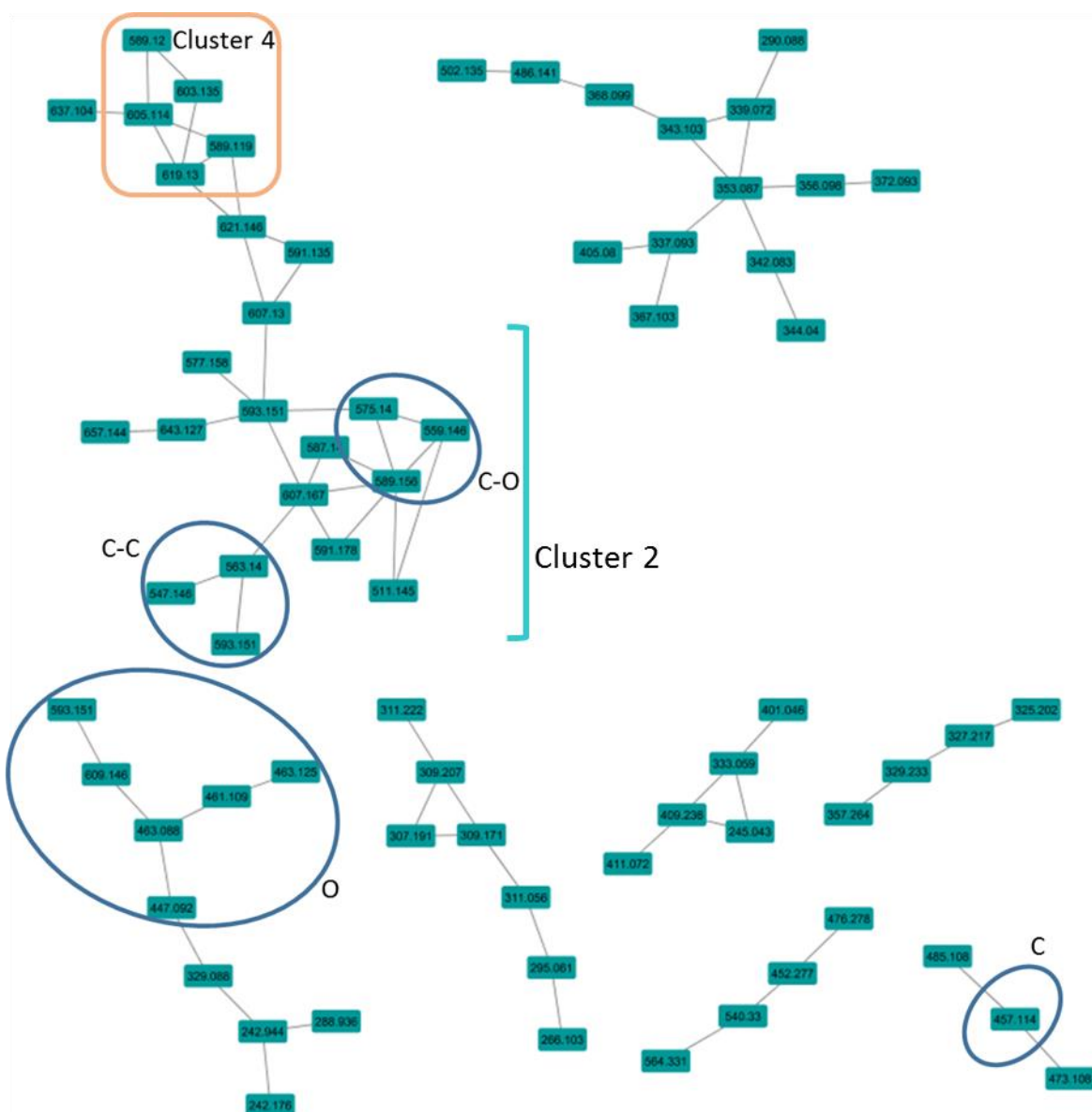


Figure III-40 : Réseau moléculaire en mode d'ionisation négatif avec cosinus score à 0,6.

Les clusters obtenus avec un cosinus score de 0,6 (Fig. III-40) sont plus grands. Certains clusters obtenus dans la figure III-38 sont maintenant regroupés. Ainsi, le cluster contenant les composés mono-*C*-glycosylés-*O*-glycosylés englobe également maintenant les di-*C*-glycosylés. En effet, une partie de la fragmentation de ces deux types de molécules peuvent se ressembler, car ils ont en commun le sucre lié sur un carbone, qui peut donner la même fragmentation. De plus, un autre cluster (4) qui n'avait pas de molécules identifiées s'est regroupé également à celui du cluster mono-*C*-glycosylé-*O*-glycosylé (cluster 2). Donc une partie de sa fragmentation ressemble à ce type de molécules, mais avec quelques différences.

Deux composés inconnus du cluster 4 ont une masse de 589 Da, comme une des molécules du cluster (2) mono-*C*-glycosylé-*O*-glycosylé, mais avec une insaturation en plus (Fig. III-39A). La formule brute de ces deux composés inconnus est en $C_{27}H_{26}O_{15}$, alors que le mono-*C*-glycosylé-*O*-glycosylé est en $C_{28}H_{30}O_{14}$, et correspond au chrysoériol C-(6-deoxy-xylo-hexos-4-ulopyranosyl)-2''-*O*-désoxyhexose. La comparaison des spectres de fragmentation de ces trois composés en mode d'ionisation négatif (Fig.

III-41) indique que les deux composés inconnus (1 et 2) ont une perte commune de 146 Da. Cette perte correspondrait à un motif *O*-désoxyhexose. Donc la différence entre ces deux composés est dans la structure du deuxième sucre (respectivement perte de 118 et de 134) et comme les ions moléculaires sont identiques, les génines sont donc aussi différentes.

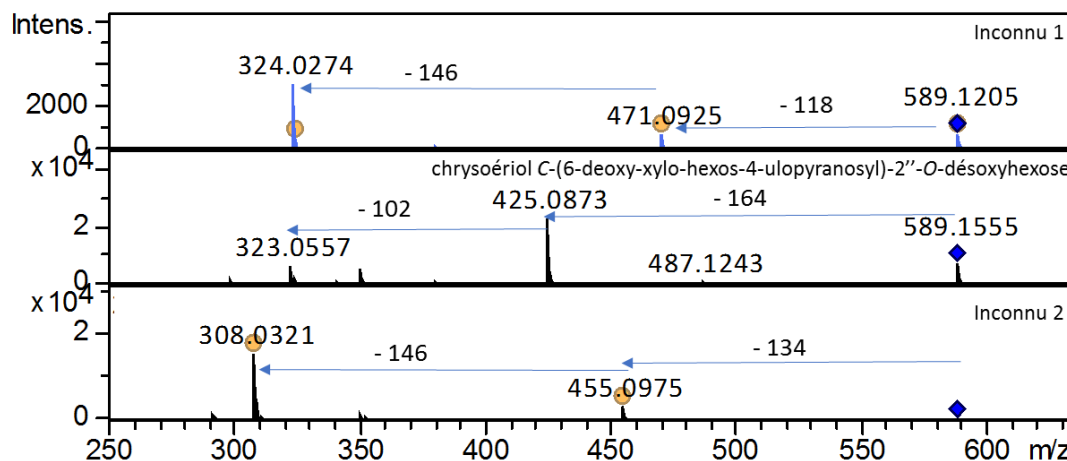


Figure III-41 : Spectres de fragmentation en mode d'ionisation négatif en UHPLC/HRMS-MS de trois molécules correspondant à une masse de 589 Da. Inconnus 1 et 2 dans le cluster 4 et chrysoériol C-(6-deoxy-xylo-hexos-4-ulopyranosyl)-2''-O-désoxyhexose dans le cluster 2.

La formule brute du fragment 471,0925 du composé inconnu 1 est $C_{23}H_{19}O_{11}$. La perte correspond à une formule brute de $C_4H_6O_4$. Cette perte pourrait correspondre à une coupure en $^{0,2}X$ dans un sucre de type C-(hexos-ulopyranosyl). La génine du composé 1 aurait une formule brute de type $C_{15}H_8O_6$, qui correspondrait à une génine de type anthraquinone avec une fonction acide ou dioxole et des fonctions alcool ou de type benzofuro-chromenone avec des fonctions alcool, par exemple.

La formule brute du fragment 455,0975 du composé inconnu 2 est $C_{26}H_{19}O_{10}$. La perte correspond à une formule brute de $C_4H_6O_5$. Cette perte pourrait correspondre à une coupure en $^{0,2}X$ dans un sucre de type acide glucuronique. La génine du composé 2 aurait une formule brute de type $C_{15}H_8O_5$, donc avec un oxygène en moins par rapport au composé 1. Donc cette génine pourrait correspondre au même type de génine que le composé 1, mais avec une fonction alcool en moins.

Dans les deux cas, ces composés ont une fragmentation similaire à celle du composé mono-C-glycosylé-O-glycosylé, et leur différence serait la nature de leur génine qui a une insaturation en plus, c'est pour cela qu'en fonction de la valeur du cosinus score, le cluster 4, dont font partie ces derniers composés, se lie ou non au cluster des mono-C-glycosylé-O-glycosylé.

Il apparaît ainsi bien complémentaire de construire le réseau moléculaire dans les deux modes d'ionisation et ce, d'autant plus que pour l'instant la librairie du GNPS référence plus de composés étudiés en mode d'ionisation positif. L'établissement des clusters a permis de mettre en évidence la présence de composés pour l'instant pas encore identifiés, mais appartenant à des clusters connus, ce qui permettrait de faciliter leur future identification. La variation de la valeur du cosinus a permis d'identifier un autre groupe de composés proche, qui pourrait également être identifié par la suite, dans le cadre d'une caractérisation globale de l'extrait.

3. Conclusion et perspectives

L'étude bibliographique du maïs et notamment de ses soies a mis en évidence qu'il contenait des familles moléculaires diversifiées dont des dérivés de 1,4-benzooxazine-3-one, des stéroïdes, des saponines et des composés phénoliques (acides phénoliques, flavonoïdes, comme anthocyanes et flavones). Elle a également montré que des flavones C-glycosylés permettaient d'inhiber la croissance des larves *H. zea*. Ainsi une étude métabolomique ciblée des flavonoïdes a été menée afin d'évaluer si la composition en flavonoïdes d'extraits de soie de maïs de différentes variétés pouvait être corrélée à l'efficacité de la variété sur l'inhibition de la croissance des larves du ravageur *H. zea*.

Dans un premier temps, l'empreinte phytochimique d'un extrait représentatif d'un mélange de 5 variétés différentes de soies de maïs a été réalisée par UHPLC/HRMS et a permis une caractérisation plus approfondie du contenu en flavonoïdes ce qui n'avait jamais été réalisé jusqu'à présent sans passer par l'isolement préalable des molécules. La quantification relative en UHPLC/MRM des flavonoïdes-glycosylés a permis de montrer que la variété de maïs Java se distinguait des quatre autres variétés. De plus, parmi les molécules qui permettent cette différenciation se retrouve bien la maysin, molécule déjà décrite comme inhibitrice de la croissance des larves.

L'approche métabolomique que nous avons mise en place s'avère intéressante en première intention pour aider à la sélection variétale et permettre un premier criblage de variétés potentiellement efficaces comme plante piège sans faire d'études lourdes et longues de terrain.

L'activité des molécules identifiées comme spécifiques de la variété Java pourrait par la suite être évaluée et ces molécules utilisées pour lutter contre le ravageur *H. zea*. Un criblage d'autres variétés pourrait également être mené pour sélectionner d'autres variétés pouvant aider dans le piégeage de ces ravageurs. Enfin, une combinaison de l'étude des volatils avec cette étude sur les flavonoïdes pourrait être réalisée pour mettre en évidence de façon plus approfondie les variétés qui seraient efficaces comme plante piège, non seulement en sélectionnant les variétés évitant la prolifération des *H. zea*, mais également en réduisant le choix à celles ayant également une bonne capacité à attirer les ravageurs *H. zea* et leurs ennemis naturels.

Références bibliographiques

- B. Abad-Garcia, S. Garmón-Lobato, LA. Berrueta, B. Gallo, F. Vicente**, New features on the fragmentation and differentiation of C-glycosidic flavone isomers by positive electrospray ionization and triple quadrupole mass spectrometry, *Rapid communication mass spectrometry*, 2008, 22, 1834-1842
- B. Abad-Garcia, LA. Berrueta, S. Garmón-Lobato, B. Gallo, F. Vicente**, A general strategy for the characterization of phenolic compounds in fruit juices by high-performance liquid chromatography with diode array detection coupled to electrospray ionization and triple quadrupole mass spectrometry, *Journal of chromatography A*, 2009, 1216, 5398-5415
- E. Allerdings, J. Ralph, H. Steinhart, M. Bunzel**, Isolation and structural identification of complex feruloylated hetrooxylan side-chains from maize bran, *Phytochemistry*, 2006, 67, 1276-1286
- T. Bhaigiyabati, T. Kirithika, J. Ramya, K. Usha**, Phytochemical constituents and antioxidant activity of various extract of corn silk (*Zea mays*. L), *Research journal of pharmaceutical, biological and chemical sciences*, 2011, 2(4), 986- 993
- V. Cambier, T. Hance, E. de Hoffmann**, Variation of DIMBOA and related compounds content in relation to the age and plant organ in maize, *Phytochemistry*, 2000, 53, 223-229
- T. Cajka and O. Fiehn**, Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics, *Analytical chemistry*, 2016, 88, 524-545
- O. Ceska and E. D. Styles**, Flavonoids from Zea Mays pollen, *Phytochemistry*, 1984, 23, 1822-1823
- D. Classen, J. T. Arnason, J. A. Serratos, J. D. H. Lambert, C. Nozzolillo, B. J. R. Philogène**, Correlation of phenolic acid content of maize to resistance to *Sitophilus zeamais*, the maize weevil, in CIMMYT's collections, *Journal of chemical ecology*, 1990, 16 (2), 301-315
- S. V. Copaja, B. N. Barria, H. M. Niemeyer**, Hydroxamic acid content of perennial triticeae, *Phytochemistry*, 1991, 30, 1531-1534
- F. Cuyckens, Y. L. Ma, G. Pocsfalvi, M. Claeysi**, Tandem mass spectral strategies for the structural characterization of flavonoid glycosides, *Analisis*, 2000, 28, 888-895
- T. Degen, C. Dillmann, F. Marion-Poll, T. C. J. Turlings**, High genetic variability of Herbivore-induced volatile emission within a broad range of maize inbred lines, *Plant physiology*, 2004, 135, 1928-1938
- J. Dong, L. Cai, X. Zhu, X. Huang, T. Yin, H. Fang, Z. Ding**, Antioxidant activities and phenolic compounds of cornhusk, corncob and stigma maydis, *Journal of the Brazilian Chemical Society*, 2014, 25 (11), 1956-1964
- A. El-Ghorab, KF. El-Massry, T. Shibamoto**, Chemical composition of the volatile extract and antioxidant activities of the volatile and nonvolatile extracts of Egyptian corn silk (*Zea mays* L.), *Journal of agricultural and food chemistry*, 2007, 55, 9124-9127
- C. A. Elliger, B. G. Chan, A. C. Waiss Jr., R. E. Lundin, W. F. Haddon**, C-glycosylflavones from ze mays that inhibit insect development, *Phytochemistry*, 1980a, 19, 293-297
- C. A. Elliger, B. C. Chan, A. C. Waiss Jr.**, Flavonoids as larval growth inhibitors, *Naturwissenschaften*, 1980b, 67, 358-360
- F. Ferreres, A. Gil-Izquierdo, P. B. Andrade, P. Valentão, F. A. Tomás-Barberán**, Characterization of C-glycosyl flavones O-glycosylated by liquid chromatography tandem mass spectrometry, *Journal of Chromatography A*, 2007, 1161, 214-223
- F. Ferreres, PB. Andrade, P. Valentão, A. Gil-Izquierdo**, Further knowledge on barley leaves O-glycosyl-C-glycosyl flavones by liquid chromatography-UV diode-array detection-electrospray ionization mass spectrometry, *Journal of Chromatography A*, 2008, 1182, 56-64

- A. Figueirinha, A. Paranhos, J. J. Pérez-Alonso, C. Santos-Buelga, M. T. Batista**, Cymbopogon citratus leaves: characterization of flavonoids by HPLC-PDA-ESI/MS/MS and an approach to their potential as source of bioactive polyphenols, *Food chemistry*, 2008, 110, 718-728
- T. Fossen, R. Slimestad, OM. Andersen**, Anthocyanins from Maize (*Zea mays*) and Reed Canarygrass (*Phalaris arundinacea*), *Journal of agricultural and food chemistry*, 2001, 49, 2318-2321
- G. Gattuso, C. Caristi, C. Gargiulli, E. Bellocco, G. Toscano, U. Leuzzi**, Flavonoid glycosides in bergamot juice, *Journal of agricultural and food chemistry*, 2006, 54, 3929-3935
- G. Glauser, G. Marti, N. Villard, GA. Doyen, JL. Wolfender, TC. Turlings, M. Erb**, Induction and detoxification of maize 1,4-benzoazin-3-ones by insect herbivores, *The plant journal*, 2011, 68, 901-911
- S. Gouinguéné, T. Degen, T. C. J. Turlings**, Variability in herbivore-induced odour emissions among maize cultivars and their wild ancestors (teosinte), *Chemoecology*, 2001, 11, 9-16
- E. Grata, D. Guilleme, G. Glauser, J. Boccard, PA. Carrupt, JL. Veuthey, S. Rudaz, JL. Wolfender**, Metabolite profiling of plant extracts by ultra-high-pressure liquid chromatography at elevated temperature coupled to time-of-flight mass spectrometry, *Journal of chromatography A*, 2009, 1216, 5660-5668
- H. Halbwirth, S. Martens, U. Wienand, G. Forkmann, K. Stich**, Biochemical formation of anthocyanins in silk tissue of *Zea mays*, *Plant science*, 2003, 164, 489-495
- K. Hasanudin, P. Hashim, S. Mustafa**, Corn silk (*Stigma Maydis*) in healthcare: A phytochemical and pharmacological review, *Molecules*, 2012, 17, 9697-9715
- <http://casmi-contest.org/2017/index.shtml>
- YJ. Jung, JH. Park, JG. Cho, KH. Seo, DS. Lee, YC. Kim, HC. Kang, MC. Song, NI. Baek**, Lignan and flavonoids from the stems of *Zea mays* and their anti-inflammatory and neuroprotective activities, *Archives of Pharmacal Research*, 2015, 38(2), 178-185
- K. A. Kim, S.-K. Choi, H.-S. Choi**, Corn silk induces nitric oxide synthase in murine macrophages, *Experimental and molecular medicine*. 2004, 36(6), 545-550
- M. J. Kim, K. R. Im, K. S. Yoon, S. W. Choi**, The inhibitory effect of hydroxycinnamic acid derivatives from corn (*Zea may* L.) bran on melanogenesis, *Journal of the society of cosmetic scientists of Korea*, 2009, 35 (2), 143-149
- SL. Kim, M. Kim, YY. Lee, GH. Jung, BY. Son, JS. Lee, YU. Kwon, YI. Park**, Isolation and identification of flavonoids from corn silk, *Korean journal of crop science* 2014, 59, 435-444
- T. Kind and O. Fiehn**, Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm, *BMC bioinformatics*, 2006, 7:234
- TG. Köllner, C. Schnee, J. Gershenzon, J. Degenhardt**, The sesquiterpene hydrocarbons of maize (*Zea mays*) form five groups with distinct developmental and organ-specific distributions, *Phytochemistry*, 2004, 65, 1895-1902
- AC. Kurilich and JA. Juvik**, Quantification of carotenoid and tocopherol antioxidants in *Zea mays*, *Journal of agricultural and food chemistry*, 1999, 47, 1948-1955
- K. H. Liland**, Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis, *Trends in analytical chemistry*, 2011, 30 (6), 827-841
- J. Liu, C. Wang, Z. Wang, C. Zhang, S. Lu, J. Liu**, The antioxidant and free-radical scavenging activities of extract and fraction from corn silk and related flavone glycosides, *Food Chemistry*, 2011, 126, 261-269
- Y.-Y. Liu, X.-H. Yang, D. Chen, J.-Z. Yang**, Isolation and structure identification of flavonoids from leaves of *Zea mays* L., *Journal of Jilin university medicine edition*, 2012, 38 (1), 67-69

- Z. A. Maksimović and N. Kovačević**, Preliminary assay on the antioxidative activity of *Maydis stigma* extracts, *Fitoterapia*, 2003, 74, 144-147
- Z. Maksimović, D. Malencić, N. Kovacević**, Polyphenol contents and antioxidant activity of *Maydis stigma* extracts, *Bioresource technology*, 2005, 96, 873-877
- G. Marti, M. Erb, J. Boccard, G. Glauser, GR. Doyen, N. Villard, CA. Robert, TC. Turlings, S. Rudaz, JL. Wolfender**, Metabolomics reveals herbivore-induced metabolites resistance and susceptibility in maize leaves and roots, *Plant, cell and environment*, 2013, 36, 621-639
- T. Michel, E. Destandau, C. Elfakir**, Evaluation of a simple and promising method for extraction of antioxidants from sea buckthorn (*Hippophaë rhamnoides* L.) berries: Pressurised solvent-free microwave assisted extraction, *Food chemistry*, 2011, 126, 1380-1386
- P. Milind and D. Isha**, Zea maize: a modern craze, *International research journal of pharmacy*, 2013, 4(6), 39-43
- T. Namba, H. Xu, S. Kadota, M. Hattori, T. Takachi, Y. Kojima**, Inhibition of IgE formation in mice by glycoproteins from corn silk, *Phytotherapy research*, 1993, 7, 227-230
- F. Nessa, Z. Ismail, N. Mohamed**, Antimicrobial activities of extracts and flavonoid glycosides of corn silk (*Zea mays* L), *International Journal of Biotechnology for Wellness Industries*, 2012, 1, 115-121
- O. Rau, M. Wurglics, T. Dingermann, M. Abdel-Tawab, M. Schubert-Zsilavecz**, Screening of herbal extracts for activation of the human peroxisome proliferator-activated receptor, *Pharmazie*, 2006, 61, 952-956
- SC. Ren, Z. Liu, X. Ding**, Isolation and identification of two novel flavone glycosides from corn silk, *Journal of medicinal plants research*, 2009, 3(12), 1009-1015
- SC. Ren, QQ. Qiao, XL. Ding**, Antioxidative activity of five flavones glycosides from corn silks, *Czech journal of food science*, 2013, 31, 148-155
- B. Rhino**, Thèse "Gestion par plante piège de la noctuelle ravageuse de la tomate *Helicoverpa zea* – Utilisation de la diversité intraspécifique du maïs », 2015, Montpellier
- B. Rhino, A. Verchere, C. Thibaut, A. Ratnadass**, Field evaluation of sweet corn varieties for their potential as a trap crop for *Helicoverpa zea* under tropical conditions, *International journal of Pest Management* 2016, 62, 3-10
- L. Riffault, E. Destandau, L. Pasquier, P. André, C. Elfakir**, Phytochemical analysis of Rosa hybrid cv. 'Jardin de Granville' by HPTLC, HPLC-DAD and HPLC-ESI-HRMS: Polyphenolic fingerprints of six plant organs, *Phytochemistry*, 2014, 99, 127-134
- E. Sarepoua, R. Tangwongchai, B. Suriharn, K. Lertrat**, Influence of variety and harvest maturity on phytochemical content in corn silk, *Food chemistry*, 2015, 169, 424-429
- AM. Shelton and FR. Badenes-Perez**, Concepts and application of trap cropping in pest management, *Annual review entomology*, 2006, 51, 285-308
- M. E. Snook, N. W. Widstrom, B. R. Wiseman, P. F. Byrne, J. S. Harwood, C. E. Costello**, New C4' hydroxy derivatives of maysin and 3' methoxymaysin isolated from corn silks, *Journal of agricultural and food chemistry*, 1995, 43, 2740-2745
- J. Stockel, M. Bar, J. N. Boidron, G. Bourgeois**, Methodological approach to identify chemical oviposition stimulants from maize for European corn borer, *Journal of chemical ecology*, 1987, 13 (3), 557-567
- LW. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W.-M. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reilly, J. J. Thaden, M. R. Viant**, Proposed minimum reporting standards for chemical analysis, *Metabolomics*, 2007, 3, 211-221

- R. Suzuki, Y. Okada, T. Okuyama**, A new flavone C glycoside from the zea mays with glycation inhibitory activity, *Chem. Pharm. Bull.*, 2003, 51(10), 1186-1188
- R. Suzuki, M. Lijima, Y. Okada, T. Okuyama**, Chemical constituent of the style of *Zea mays* L. with glycation inhibitory activity, *Chemical and pharmaceutical bulletin*, 2007, 55(1), 153-155
- J. Tian, H. Chen, S. Chen, L. Xing, Y. Wang, J. Wang**, Comparative studies on the constituents, antioxidant and anticancer activities of extracts from different varieties of corn silk, *Food & Function*, 2013, 4, 1526-1534
- PG. Tillman and B. G. Mullinix Jr.**, Grain Sorghum as a trap crop for corn earworm (Lepidoptera: Noctuidae) in cotton, *Environmental entomology*, 2004, 33 (5), 1371-1380
- M. Ueda and R. S. Bandurski**, A quantitative estimation of alkali-labile indole-3-acetic acid compounds in dormant and germinating maize kernels, *Plant physiology*, 1969, 44, 1175-1181
- V. Vukics and A. Guttman**, Structural characterization of flavonoid glycosides by multi stage mass spectrometry, *Mass spectrometry reviews*, 2010, 29(1), 1-16
- T. H. Walter and R. W. Andrews**, Recent innovations in UHPLC columns and instrumentation, *Trends in analytical chemistry*, 2014, 63, 14-20
- NW. Widstrom, and M. E. Snook**, A gene controlling biosynthesis of isoorientin, a compound in corn silks antibiotic to the corn earworm, *Entomologia experimentalis et applicata*, 1998, 89, 119-124
- J.-L. Wolfender, P. Waridel, K. Ndjoko, K. R. Hobby, H. J. Major, K. Hostettmann**, Evaluation of Q-ToF-MS/MS and multiple stage IT-MSⁿ for the dereplication of flavonoids and related compounds in crude plant extracts, *Analisis*, 2000, 28 (10), 895-906
- S. G. Wynn and B. J. Fougère**, Veterinary herbal medicine, Elsevier, St Louis, Missouri, 2007
- J. Xia and DS. Wishart**, Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis, *Current protocols in bioinformatics*, 2016, 55, 14.10.1-14.10.91
- J. Xiao, E. Capanoglu, AR. Jassbi, A. Miron**, Advance on the flavonoid C-glycosides and health benefits, *Critical reviews in food science and nutrition*, 2016, 56, S29-S45
- RA. Yost, and C. G. Enke**, Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation, *Analytical chemistry*, 1979, 51 (12), 1251-1264
- HJ. Zeringue Jr.**, Identification and effects of maize silk volatiles on cultures of *Aspergillus flavus*, *Journal of agricultural and food chemistry*, 2000, 48, 921-925
- Y. Zhang, L. Wu, Z. Ma, J. Cheng, J. Liu**, Anti-diabetic, anti-oxidant and anti-hyperlipidemic activities of flavonoids from corn silk on STZ-induced diabetic mice, *Molecules*, 2016, 21 (7), 1-11
- S. Žilić, A. Serpen, G. Akillioğlu, V. Gökmen, J. Vančetović**, Phenolic compounds, carotenoids, anthocyanins, and antioxidant capacity of colored maize (*Zea mays* L.) kernels, *Journal of agricultural and food chemistry*, 2012, 60, 1224-1231
- S. Žilić, M. Janković, Z. Basić, J. Vančetović, V. Maksimović**, Antioxidant activity, phenolic profile, chlorophyll and mineral matter content of corn silk (*Zea mays* L): comparison with medicinal herbs, *Journal of cereal science*, 2016, 69, 363-370

Conclusions – Perspectives

L'objectif de ces travaux était de développer de nouvelles méthodologies et de nouveaux outils pour obtenir des signatures moléculaires de milieux complexes et les interpréter plus facilement. En effet, que les échantillons proviennent de plantes, de milieux biologiques relevant du domaine environnemental, agro-alimentaire, cosmétique, pharmaceutique ou médical, ils sont souvent pourvus d'une grande chimiodiversité avec différentes familles moléculaires, elles-mêmes englobant un grand nombre de molécules de structures variées. Pour identifier le plus rapidement possible les composés déjà décrits dans la littérature et mettre en évidence des composés d'intérêt inconnus ou spécifiques, différentes méthodologies se développent et peuvent être sélectionnées et utilisées en fonction de la nature des échantillons, des molécules et de la problématique d'intérêt. Certaines de ces techniques ont donc été mise en œuvre au cours de cette thèse pour répondre à des études diverses (analyses ciblées d'extraits, comparaison de variétés ou d'organes de plantes, caractérisation phytochimique de plantes ou de composés spécifiques).

Une vue d'ensemble des outils et des stratégies analytiques utilisables pour la caractérisation de mélanges complexes décrits dans la littérature a d'abord été effectuée. La combinaison d'une technique séparative, d'une détection par spectrométrie de masse et d'un traitement de données adapté permet d'obtenir différents types d'empreintes moléculaires de milieux complexes. De plus en plus, les développements vont vers l'obtention d'outils de dérégulation permettant d'aider à repérer facilement, dans un mélange, les composés, mais également vers des outils de métabolomique pour pouvoir comparer de nombreux échantillons composés d'un grand nombre de molécules.

La première partie de ce travail (Chapitre 2) a été axée sur le développement de méthodes d'analyse non ciblées et ciblées en chromatographie sur couche mince (TLC). Une comparaison des profils phytochimiques TLC de différents organes de colza (*Brassica napus* L.) par des traitements statistiques (analyses en composantes principales) a été effectuée. L'originalité de cette démarche est de pouvoir réaliser des analyses statistiques directement depuis les images des plaques TLC, sans passer par un retraitement manuel. Ainsi des regroupements d'échantillons peuvent être visualisés par similitude de profil chromatographique TLC et il est possible de mettre en évidence les taches et donc les molécules qui permettent de distinguer les groupes. Afin d'améliorer l'identification des molécules d'intérêt, des couplages TLC/MS ont été mis en œuvre. Deux types de couplages ont été développés, un premier couplage avec une source sous vide, la source MALDI et un couplage à pression atmosphérique avec une source électrospray (ESI) et une source à ionisation chimique (APCI). Ainsi les couplages TLC/MS permettent de caractériser un large panel de composés et peuvent être utilisés pour des approches d'identification, de comparaison d'extraits avec l'appui de traitements statistiques et aller jusqu'à la caractérisation ciblée de spot d'intérêt.

Dans la seconde partie du travail de thèse (Chapitre 3), une étude métabolomique ciblée sur les flavonoïdes présents dans les soies de maïs, a été menée en complément du travail d'observation sur le terrain effectué par le CIRAD en vue de sélectionner la variété de maïs la plus efficace en tant que plante piège protectrice de plants de tomates. Afin d'étudier sélectivement les flavonoïdes extraits de soies de maïs, une première empreinte moléculaire obtenue par UHPLC/HRMS, a été cartographiée par un diagramme de Van Krevelen, ce qui a permis de mettre en évidence les différentes familles moléculaires présentes dans l'extrait et de sélectionner la plage de rétention du groupe de composés d'intérêt. Les analyses par UHPLC/HRMS-MS ont permis ensuite d'identifier une trentaine de flavonoïdes dans l'extrait par l'étude de leur schéma de fragmentation. Ces composés ont été analysés par UHPLC/MS pour obtenir leur quantification relative dans les 108 échantillons analysés et les données obtenues ont été analysées avec l'appui de traitements statistiques. Parmi les cinq variétés de soies de maïs étudiées, nos résultats permettent de distinguer deux variétés (la variété Java et la

variété Nova) des autres variétés. Or sur le terrain, le CIRAD a montré que la variété Java était la variété montrant le plus d'efficacité dans l'inhibition de la croissance des larves des insectes ravageurs. Quatre composés d'intérêt influent majoritairement cette séparation pour la variété Java. Parmi ces composés, seule la molécule de maysin a été décrite dans la littérature comme molécule inhibitrice de la croissance des larves. Les autres composés devront être testés afin de vérifier leur activité sur les larves, après avoir été formellement caractérisés. Une approche d'identification des composés discriminants a également été amorcée avec la construction des réseaux moléculaires. La cartographie du réseau moléculaire a permis de mettre en évidence des clusters formés en fonction de la structure des flavonoïdes (O ou C-glycosylés, Mono- et/ou di-glycosylés) et des clusters contenant des composés proches des flavonoïdes, des anthraquinones, avec des substituants glycosylés du même type. Pour aller plus loin dans l'étude, des standards d'autres familles moléculaires devraient être incorporés à ce réseau pour en faciliter son interprétation.

En conclusion, ces travaux ont conduit au développement de couplages TLC/MS et UHPLC/MS qui permettent d'obtenir des empreintes moléculaires d'extraits, de comparer plusieurs échantillons d'un point de vue moléculaire, d'obtenir des informations pertinentes et de les caractériser, ce qui nous permet d'élargir notre panel d'outils analytiques et de champ d'investigation.

Le savoir-faire acquis dans le cadre de ces premières études métabolomiques nous permet de les appliquer et de les transposer directement à d'autres problématiques dans des projets en cours, comme la compréhension des mécanismes de défense de la tomate, lors d'une collaboration avec les Instituts de Chimie de Nice et de Sophia Agrobiotech : les profils métabolomiques d'échantillons de plant de tomate sains et infestés par différents types de ravageurs (nématodes, noctuelles ...) ou par des combinaisons d'agressions modulées sont comparés. Dans le cadre du projet régional Chêne et Vin (APR-IR), dans la recherche de marqueurs moléculaires des différentes espèces de chêne, les études métabolomiques peuvent permettre de mieux sélectionner les espèces de chêne et de mieux comprendre et maîtriser l'impact de l'espèce sur les interactions entre le chêne et le vin dans l'élevage de vins de Loire. Ces développements seront également indispensables aux projets collaboratifs en cours de construction concernant l'étude des interactions plante-insecte, plante-environnement ou plante-plante, et aider ainsi à la compréhension de la prolifération des plantes envahissantes et invasives notamment dans les zones humides de la région Centre-Val de Loire.

Dans le cadre de la recherche de nouvelles molécules avec un potentiel d'activité, des études de phylogénie pourraient aussi être envisagées avec ces approches métabolomiques, en comparant le contenu moléculaire de plantes déjà connues pour leur potentiel actif et de plantes de même parenté. Cette étude pourrait également aider dans la compréhension des voies de biosynthèses des molécules dans les plantes et la répartition de ces molécules dans les organes de la plante.

Les approches qui ont été développées dans ce travail de thèse m'apportent une expertise qui pourra être mise au service de projets de recherche dans ces thématiques, mais également pour répondre aux problématiques de mes collègues de l'ICOA en mettant par exemple en place ces types de couplage pour aider à l'identification de produits lors de nouvelles synthèses. De plus, cette expertise est complémentaire à celle que j'ai en parallèle acquise dans le cadre de mon poste d'ingénieur d'étude. En effet, ces autres activités ont conduit principalement au développement d'une méthode d'analyse quantitative en HPLC/MRM MS, qui a ensuite été transposée en UHPLC/MRM MS pour réaliser, en collaboration avec l'Institut des Sciences et de la Terre d'Orléans (ISTO), le suivi de médicaments et de drogues dans les eaux usées et dans des carottes sédimentaires d'un bassin de décantation, afin

d'aider au suivi des molécules et à la datation de l'Anthropocène, mais aussi pour réaliser le monitoring des flux de polluants et aider dans la mise en place de procédés de traitements des eaux. Une autre étude a porté, en collaboration avec le laboratoire Interfaces, Confinement, Matériaux et Nanostructures (ICMN, CNRS Orléans) sur l'identification par UHPLC/MS de produits de dégradation du paracétamol et du diclofénac par ozonation. Enfin, des compétences en analyse quantitative et en validation de méthodes m'ont été apportées dans le cadre d'une collaboration avec les Hôpitaux Universitaires de Genève et avec l'association Niemann-Pick en développant une méthode UHPLC/MS pour déterminer la teneur de méthyl- β -cyclodextrine dans différents organes et dans le sérum de souris après injection intrapéritonéale, avec un suivi cinétique.

ANNEXES

ANNEXES Chapitre 2

Annexe 1 : Récolte et extraction du matériel végétal

La récolte du colza a été effectuée dans un champ du Loiret et conservée tout de suite au froid. Ensuite les différents organes ont été séparés à l'ICOA.

Les baies d'argousier proviennent de la Moldavie Roumaine et ont été conservées au congélateur jusqu'à leur extraction.

La poudre de fruit de mangoustan nous a été fournie par LVMH recherche.

Le jus de pommes enrichi en dérivés de quercétine, dès la fabrication, provient de la société Touraines Jus de Pomme

Le vin rouge est un côte du Rhône.

L'extraction assistée par micro-ondes des végétaux a été réalisée dans un four micro-ondes MycroSynth (Milestone, Sorisole, Italie) piloté par une console « easyCONTROL » en suivant les conditions opératoires du tableau A1.

Tableau A1 : Conditions opératoires d'extraction assistée par micro-ondes.

Plantes	Masse (g)	Solvant	Volume (mL)	Puissance (W)	Temps (s)	Nbre cycle
Colza <i>Brassica napus</i> L.	0,5	EtOH	5	1000	30	3
Argousier <i>Hippophae rhamnoides</i>	1	EtOH	7	1000	30	3
Mangoustan <i>Garcinia mangostana</i>	0,5	EtOH	7	800	30	3

Les extraits de colza, d'argousier et de mangoustan ont été centrifugés pendant 15 min à une vitesse de 9 000 rpm et le surnageant a ensuite été récupéré. Le surnageant a été filtré sur filtre seringue à 0.45 µm.

Le jus de pommes (entreprise Touraine jus de pommes) était sous forme lyophilisé. 1 g de poudre lyophilisé a juste été solubilisé dans 7 mL d'éthanol.

Aucune préparation d'échantillon n'a été réalisée sur le vin rouge (Côtes du Rhône).

Annexe 2 : Analyses TLC

Dépôts :

Avant les dépôts, les plaques ont été éluées avec du méthanol pour les laver, puis séchées dans une étuve à 100°C. Les dépôts ont été réalisés avec un déposeur semi-automatique Linomat IV (Camag, Muttenz, Suisse) ou un déposeur automatique TLC sampler 4. Ils ont permis d'effectuer des dépôts sous forme de spray à l'aide d'un courant d'azote. Les dépôts ont été réalisés à 10 mm du bord inférieur de la plaque. Les conditions de dépôts sont récapitulées dans le tableau A2 suivant l'échantillon.

Tableau A2 : Conditions de dépôts.

Echantillons	Volume (μL)	Start position (mm)	Largeur des bandes (mm)	Espace entre les bandes (mm)	Vitesse de dépôt (sec/μL)	Plaque gel de silice Si ₆₀ F ₂₅₄	Taille de la plaque (cm*cm)
Colza <i>Brassica napus</i> L.	5	10	8	5	10	TLC verre	20x10
Vin	10	10	8	7	7	TLC alu	5x7,5
Argousier <i>Hippophae rhamnoides</i>	10	8	5	5	7	TLC alu	5x7,5
Jus de pommes	10	10	8	7	7	TLC alu	5x7,5
Mangoustan <i>Garcinia mangostana</i>	10	6	5	5	7	TLC alu	5x7,5
Acides organiques	5	8	8	5	7	TLC verre	20x10
Stéroïdes	5	16	10	14	7	HPTLC verre	20x10

Conditions de séparation :

Les séparations ont été effectuées dans une chambre d'élution verticale sur 7 cm, avec une saturation de la cuve de 10 min par le solvant d'élution.

Colza :

Plaque composés polyphénoliques : Phase éluante : AcOET/H₂O/AF/AcOH (100/27/11/11 ; v/v/v/v)

Plaque composés lipidiques : Phase éluante : CHCl₃/n-heptane/MeOH (2/2/0.5 ; v/v/v)

Vin, Argousier, Jus de pommes :

Plaque composés polyphénoliques : Phase éluante : AcOET/MeOH/AF (90/10/1 ; v/v/v)

Mangoustan :

Plaque composés xanthones : Phase éluante : CHCl₃/n-heptane/MeOH (70/20/10 ; v/v/v)

Standards acides organiques :

Phase éluante : NH₄OH/EtOH/H₂O (12.5/75.5/12 ; v/v/v)

Standards stéroïdes et oxystérols :

Phase éluante : Hexane/Iso-octane/Isopropanol (69/24/7 ; v/v/v) + 1% AcOH

Révélation :

Pour la révélation, la pulvérisation des réactifs s'est faite soit avec un pulvérisateur de spray manuel, soit avec le derivatizer (Camag), qui permet de n'utiliser que 2 à 4 µL de réactif pour révéler une plaque.

Après la révélation, les plaques ont été prises en photos. Pour celles correspondant aux extraits de colza, cela a été réalisé avec l'appareil TLC visualizeur (Camag) pour une standardisation des images.

Flavonoïdes : NEU + PEG

La révélation est réalisée par pulvérisation du réactif NEU, puis après séchage du NEU au sèche-cheveux par pulvérisation de la solution de PEG 4000. La visualisation et la photo se font à une longueur d'onde de 366 nm.

Neu :

Le réactif est préparé en solubilisant 1 g de β-aminoethyl diphenylborate dans 100 mL de méthanol.

PEG :

5g de polyéthylèneglycol (PEG) 4000 a été dissous dans 100 mL d'éthanol.

Primuline :

5 mg de primuline est dissous dans 100 mL d'un mélange de solvant acétone/eau (8/2).

La visualisation et la prise de photo de la plaque révélée se font à une longueur d'onde de 366 nm.

La révélation étant non destructive, une autre révélation peut se faire sur la plaque révélée à la primuline.

p-anisaldéhyde :

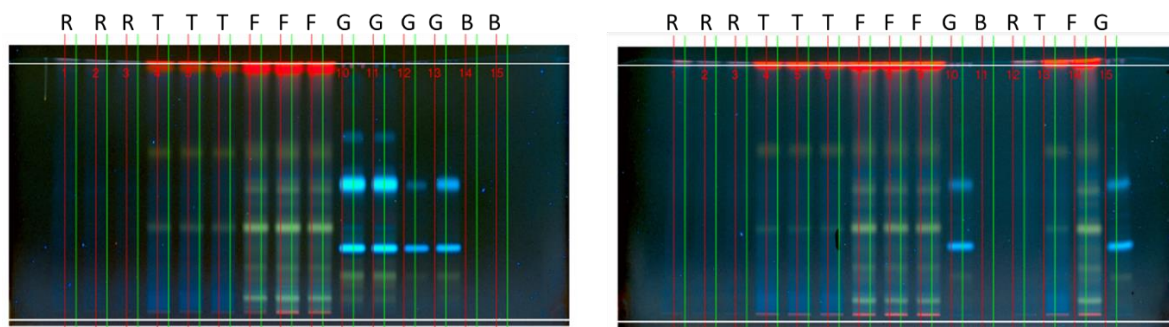
85 mL de méthanol est mélangé avec 10 mL d'acide acétique glaciale, 5 mL d'acide sulfurique et 0.5 mL d'anisaldéhyde.

Après la pulvérisation de cette solution, la plaque est chauffée à 100°C pendant 5 à 10 minutes. La détection des composés se fait dans le visible.

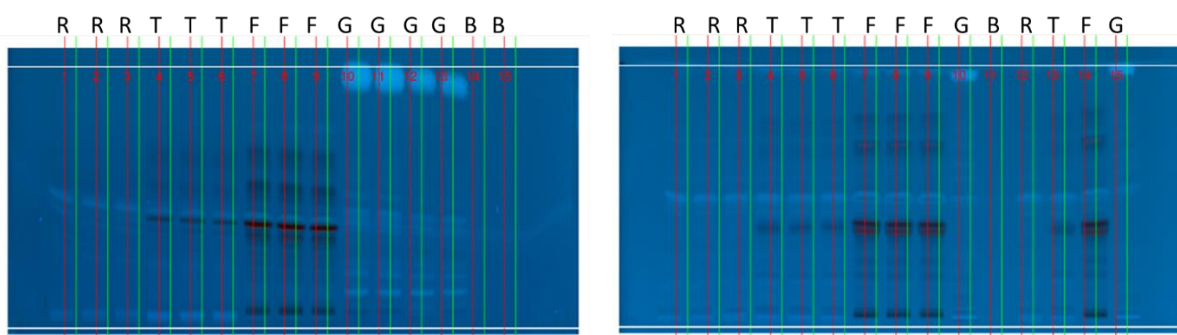
Annexe 3 : Plaques TLC de colza pour traitement avec le logiciel rTLC

Deux séries de plaque ont été soumises au logiciel rTLC pour les traitements statistiques par système d'élution et révélation.

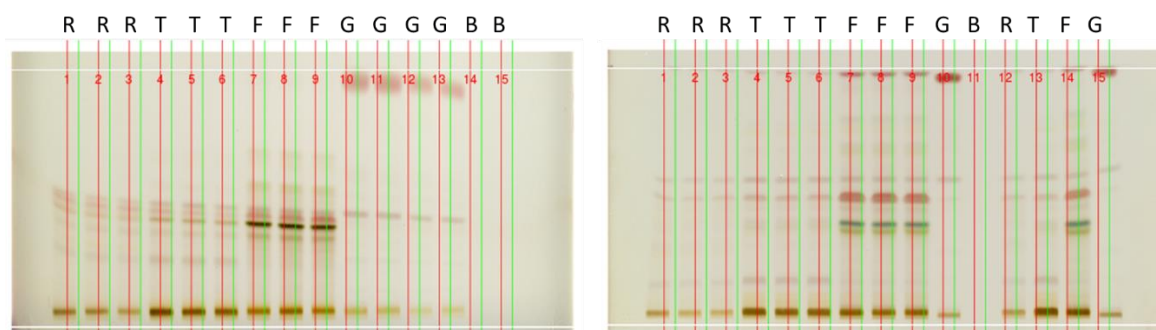
- Plaques avec phase d'élution AcOET/H₂O/AF/AcOH (100/27/11/11 ; v/v/v/v) avec révélation NEU-PEG à 366 nm (R : racine ; T : tige ; G : graine ; F : feuille) :



- Plaques avec phase d'élution CHCl₃/n-heptane/MeOH (2/2/0.5 ; v/v/v) avec révélation primuline à 366 nm (R : racine ; T : tige ; G : graine ; F : feuille) :



- Plaques avec phase d'élution CHCl₃/n-heptane/MeOH (2/2/0.5 ; v/v/v) avec révélation p-anisaldéhyde dans le visible (R : racine ; T : tige ; G : graine ; F : feuille) :



Annexe 4 : Matrices organiques couramment utilisées en MALDI-Tof-MS

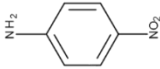
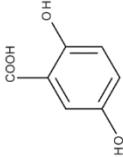
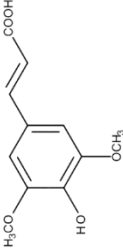
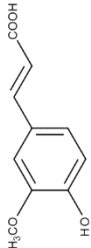
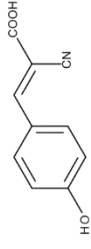
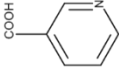
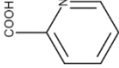
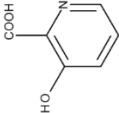
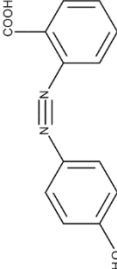
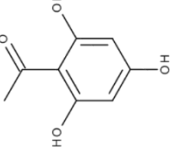
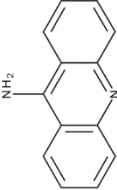
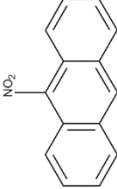
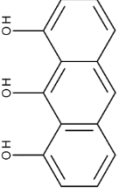
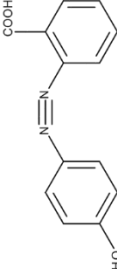
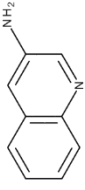
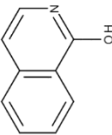
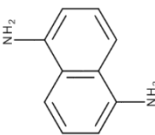
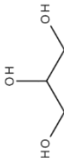
p-Nitroaniline (PNA)	Acide 2,5-dihydroxybenzoïque (DHB)	Acide sinapnique	Acide férulique (FA)	Acide α -cyano-4-hydroxycinnamique (CHCA)
				
Acide nicotinique	Acide picolinique	Acide 3-hydroxypicolinique (3HPA)	Acide 2-(4-hydroxyphénylazo)-benzoïque (HABA)	2,4,6-Trihydroxycétophénone (THAP)
				
9-Aminoacridine	9-Nitroanthracène	Dithranol	Acide 2-(4-hydroxyphénylazo)-benzoïque (HABA)	Acide succinique
				
3-Aminoquinoline	1-Isoquinolinol	1,5-Diaminonaphthalène	Glycerol	
				

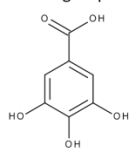
Tableau : Liste des principales matrices et utilisation

Composé	Abréviation	Utilisation
p-Nitroaniline	PNA	Peptides et protéines, Matrice liquide
Acide 2,5-dihydroxybenzoïque	DHB	Polymères synthétique polaire Molécules organique Hydrates de carbone
Acide sinapique	SA	Protéines >5kDa Dendrimères Fullerènes
Acide férulique	FA	Peptides et protéines, Laser UV
Acide α -cyano-4-hydroxycinnamique	CHCA	Peptides <5kDa Sucres
Acide nicotinique		Oligonucléotides >3,5kDa
Acide picolinique		Oligonucléotides
Acide 3-hydroxypicolinique	3HPA	Oligonucléotides >3,5kDa
2,5-dihydroxyacétophénone	DHAP	Protéines >10kDa Glycoprotéines
2,4,6-Trihydroxyacétophénone	THAP	Oligonucléotides <3,5kDa Hydrates de carbone acide Peptides phosphorylé
9-Aminoacridine	9-AA	Lipides Métabolites
9-nitroanthracène	9-NA	Fullerènes
Dithranol	DIT	Polymères synthétique non polaire Lipides Dendrimères
Acide 2-(4-hydroxyphénylazo)-benzoïque	HABA	Protéines >10kDa
Acide succinique		Peptides et protéines, Laser IR Polymères synthétique, Laser IR
3-Aminoquinoline		Hydrates de carbone
1-Isoquinolinol		Oligosaccharides
1,5-Diaminonaphthalène	1,5-DAN	Fragmentation de peptides et protéines Imagerie lipidique
Glycérol		Peptides et protéines, Matrice liquide

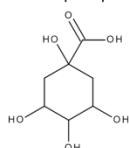
Annexe 5 : Structures des molécules discutées dans le chapitre

Polyphénols

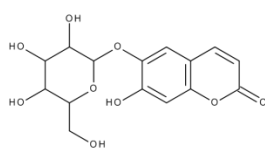
Acide gallique



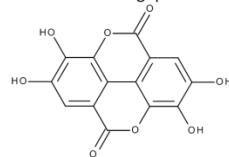
Acide quinique



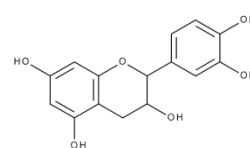
esculine



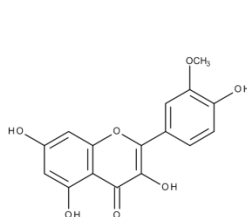
Acide ellagique



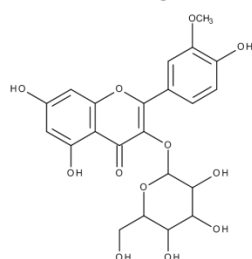
Catéchine



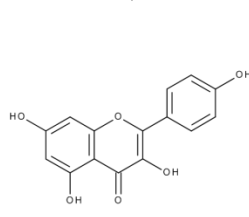
Isorhamnétine



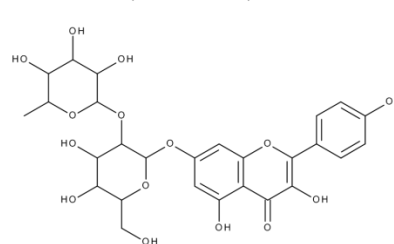
Isorhamnétine 3-glucoside



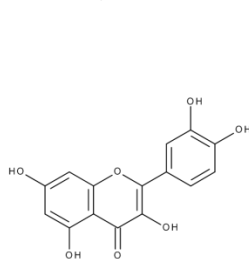
Kaempférol



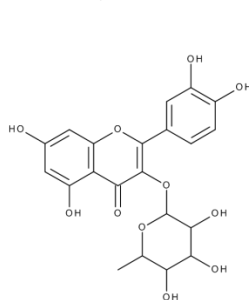
Kaempférol 7-neohesperidoside



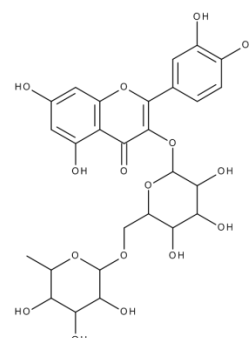
Quercétine



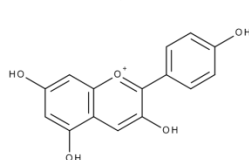
Quercitrine



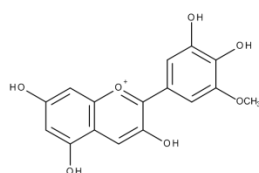
Rutine



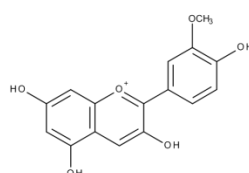
Pélgargonidine



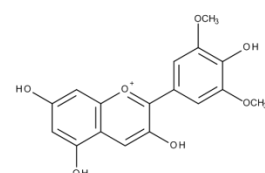
Pétunidine



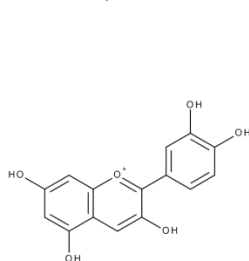
Péonidine



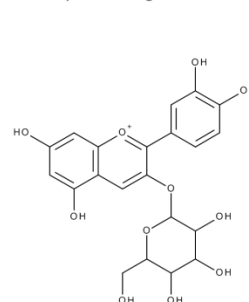
Malvidine



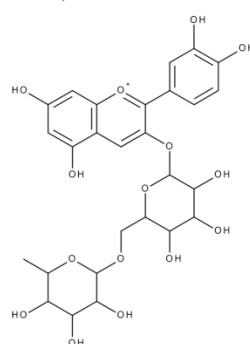
Cyanidine



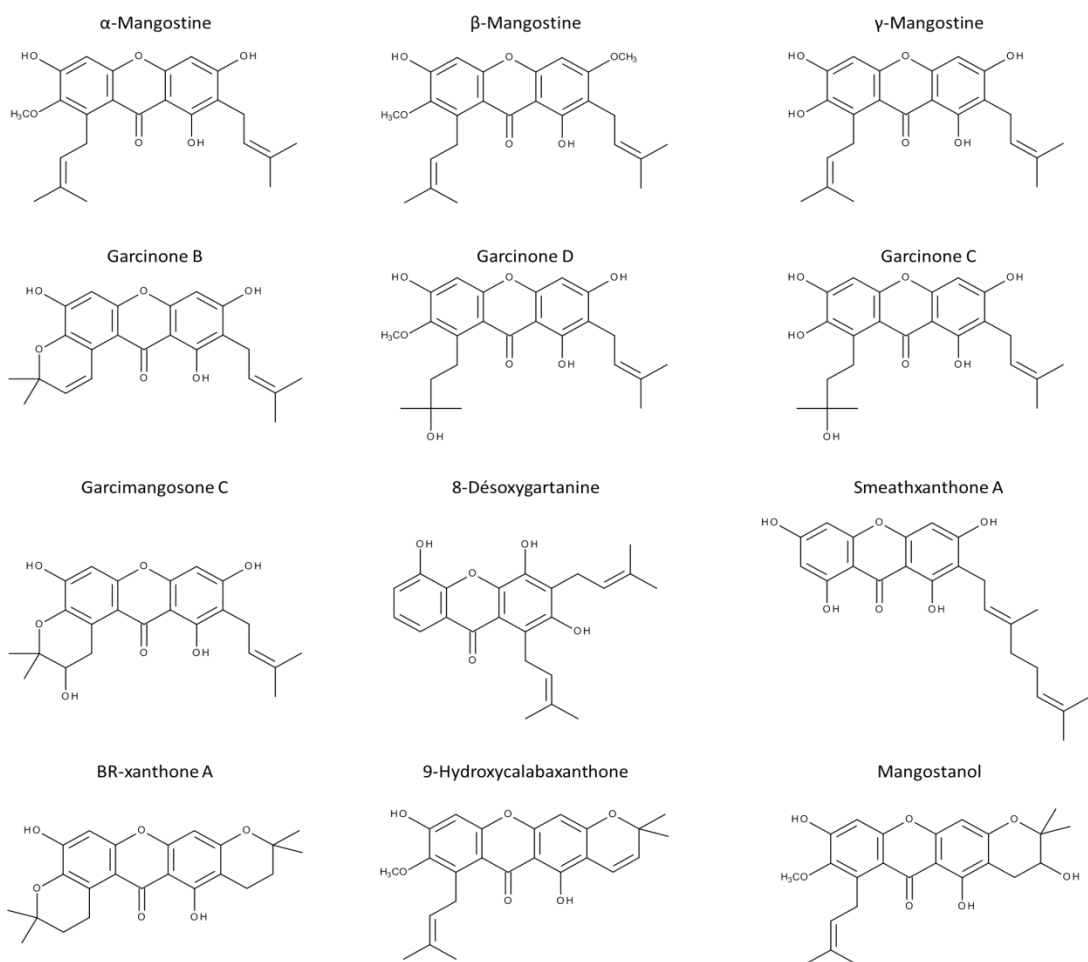
Cyanidine 3-glucoside



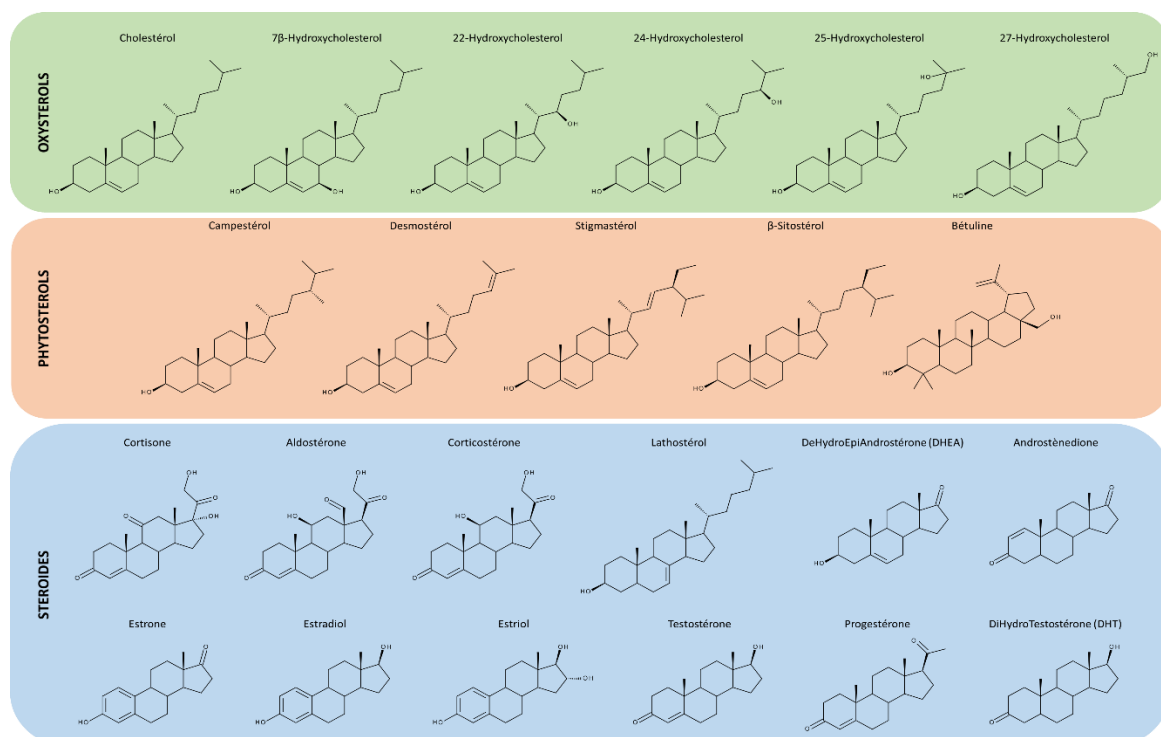
Cyanidine 3-rutinoside



Xanthones



Oxystérols, phytostérols, stéroïdes



Annexe 6 : Analyses TLC/MALDI-ToF MS

Pour réaliser le couplage TLC/MALDI-ToF MS, un adaptateur de plaque TLC aluminium a été utilisé. Les analyses ont été réalisées avec un spectromètre de masse MALDI-ToF MS, Autoflex de l'entreprise Bruker. Les conditions de source étaient : Ion source 1 à 19 kV ; Ion source 2 à 16,15 kV ; Lens à 8,25 kV ; Reflector à 8,25 kV ; Pulsed ion extracted à 20 ns pour le mode d'ionisation négatif et 40 ns pour le mode d'ionisation positif ; la fréquence d'acquisition à 6,7 Hz. L'appareil est équipé d'un laser à 337 nm. L'acquisition se fait par le logiciel FlexControl (v3.3) et le retraitement se fait via le logiciel FlexAnalysis (v3.3).

La matrice DHB utilisée a été préparée à 200 mg.mL⁻¹ pour l'immersion et 10 mg.mL⁻¹ pour le mode pulvérisation ciblée dans du méthanol.

La matrice inorganique est une matrice magnétique de nanoparticules d'oxyde de fer revêtues de silice à noyau-enveloppe, elle a été préparée selon la thèse de Justine Ferey (2017) à 10 mg.mL⁻¹ dans l'éthanol.

Annexe 7 : Analyses TLC/API TQ MS

L'analyse TLC/ESI TQ MS pour les standards d'acides organiques a été optimisée avec un mélange de solvants Eau/ACN (50/50) avec 0.1% d'acide formique comme solvant de désorption des taches à un débit de 0.2 mL/min. L'analyse en spectrométrie de masse a été effectuée sur un triple quadripôle TSQ Endura de chez Thermofisher. Les paramètres de source ESI étaient un « sheath gas » à 27, un « aux gas » à 9, une température de la ligne de transfert à 312 °C et de vaporisation à 210 °C, une tension de l'électrode de capillaire à 3500 V en mode d'ionisation positive et de -2500 V en mode d'ionisation négative. L'acquisition s'est fait en mode full scan sur le premier quadripôle avec le logiciel Thermo Xcalibur (v3.0 SP2).

L'analyse TLC/APCI TQ MS pour les molécules d'oxystérols, phytostérols et stéroïdes a été optimisée avec l'acétonitrile comme solvant de désorption des taches à un débit de 0.2 mL/min. L'analyse en spectrométrie de masse a été effectuée sur un triple quadripôle TSQ Endura de chez Thermofisher. Les paramètres de source APCI étaient un « sheath gas » à 30, un « aux gas » à 10, une température de la ligne de transfert à 325 °C et de vaporisation à 450 °C, un courant de décharge à 6 en mode d'ionisation positive. L'acquisition s'est faite en mode full scan sur le premier quadripôle avec le logiciel Thermo Xcalibur (v3.0 SP2).

ANNEXES Chapitre 3

Annexe 1 : Extraction du matériel végétal

L'extraction des composés polaires et moyennement polaires a été effectuée par une extraction assistée par micro-ondes dans un four MycroSynth (Milestone, Sorisole, Italie), muni d'un carrousel de 12 réacteurs, piloté par une console « easyCONTROL ». 200 mg de soie de maïs a été extrait avec 10 mL d'un mélange Eau/EtOH (50/50), avec une puissance de 700 W en 3 cycles de 30 sec.

Annexe 2 : Analyse UHPLC/HRMS et UHPLC/TQ

Les travaux présentés en UHPLC/MS ont été effectués avec deux colonnes Nucleodur C18 de chez Macherey-Nagel, nommée Sphynx RP et Pyramid de dimension 150*2 mm avec une taille de particules de 1,8 μm . Les analyses ont été effectuées à une température de four de colonne de 60°C. L'élution est un gradient composé d'eau et d'acétonitrile avec 0,1% d'acide formique à un débit de 0,6 mL/min. L'acétonitrile passe de 5 à 20% en 1,5 min et reste à ce pourcentage pendant 1 min. Puis il augmente à 35% en 1,5 min et à 60% en 1 min. L'acétonitrile reste à 60% pendant 1 min, pour passer ensuite à 100% en 1,5 min et reste à 100% pendant 1,5 min, avant de revenir aux conditions initiales d'équilibre qui sont maintenues pendant 5 min.

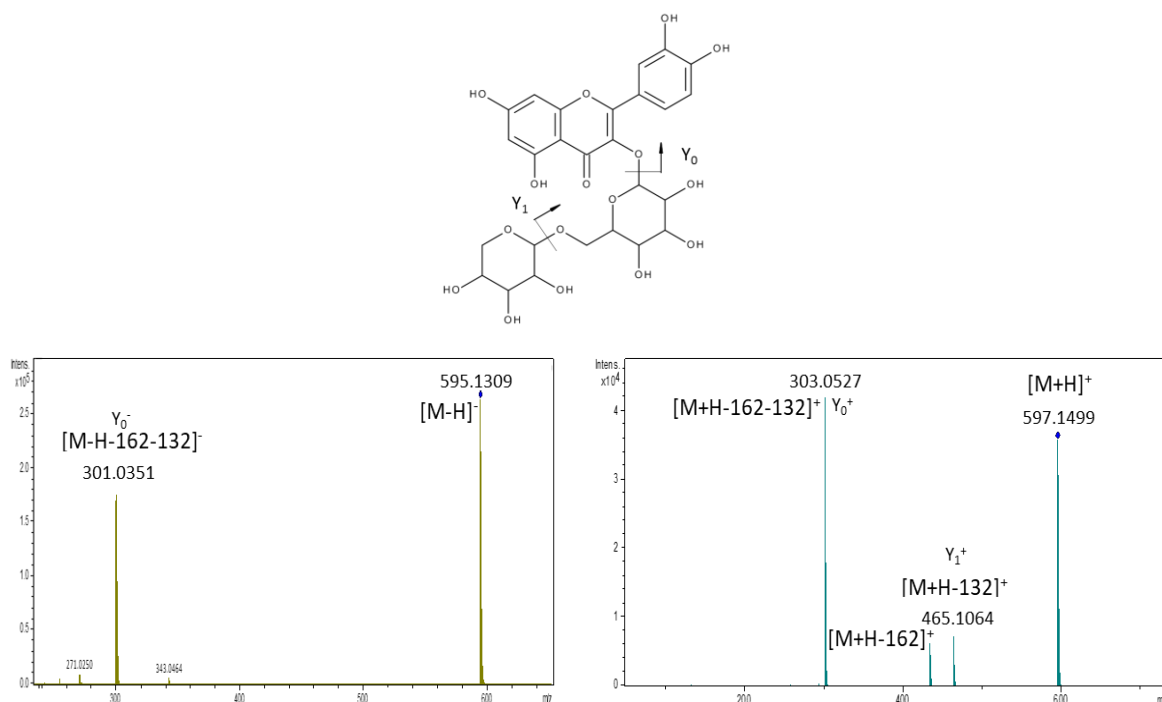
La chromatographie liquide a été couplée à un spectromètre de masse haute résolution (Maxis de Bruker), avec une source d'ionisation électrospray. L'analyse a été réalisée en mode d'ionisation négatif sur une gamme de masse de m/z de 50-1650. La tension du capillaire était de - 4 000V, une pression de gaz d'azote à 2 bars, une température de source à 200°C et un gaz de séchage à 10,5 L/min. En mode d'ionisation positif, les paramètres étaient une tension de capillaire de 4 000 V, un gaz de séchage au débit de 9,5 L/min, tous les autres réglages de paramètres restant identiques à ceux fixés pour le mode d'ionisation négatif. Une calibration externe a été réalisée avec un tuning mix et une correction interne avec deux masses 666,019886 et 1265,981563. L'analyse de fragmentation a été réalisée avec une énergie de 30 eV. Le spectre de masse est acquis pendant 400 ms, pendant ce scan, les deux ions les plus intenses sont alors fragmentés, s'ils sont au-dessus de 1 000 en intensité pendant 500 ms. La fenêtre d'isolation est de 4 unités pour un m/z de 120 et de 7 unités pour un m/z de 1 200. Ces deux masses ne sont plus utilisées pendant 7 sec.

Pour réaliser les analyses de quantification, la chromatographie liquide a été couplée avec un analyseur triple quadripôle (TSQ Endura de chez ThermoFisher), avec une source d'ionisation électrospray. La tension du capillaire était à - 2 500V, avec un gaz de nébulisation à 30 (unité arbitraire), un gaz auxiliaire à 5 (unité arbitraire) et un gaz d'entraînement à 3 (unité arbitraire). La température de la source était à 420°C et celle de la ligne de transfert à 356°C. Le gaz de collision est de l'argon, à une pression de 1,5 mTorr.

Les transitions MRM et les conditions sont données dans le tableau de l'annexe 4.

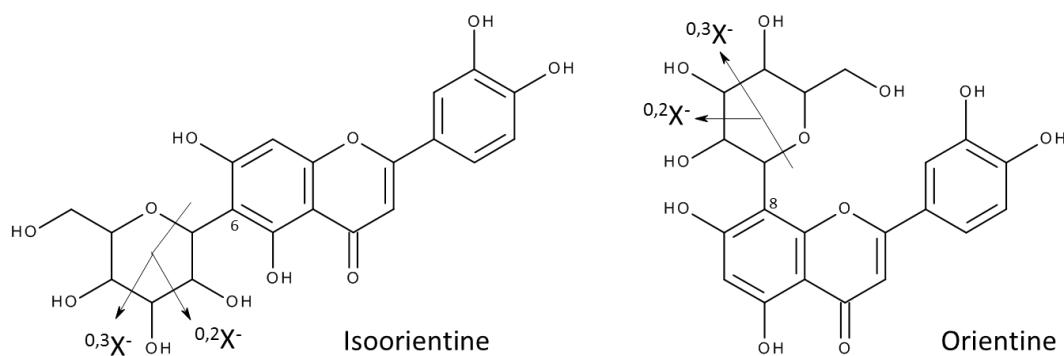
Annexe 3 : Spectres de fragmentation des standards et schémas de fragmentation

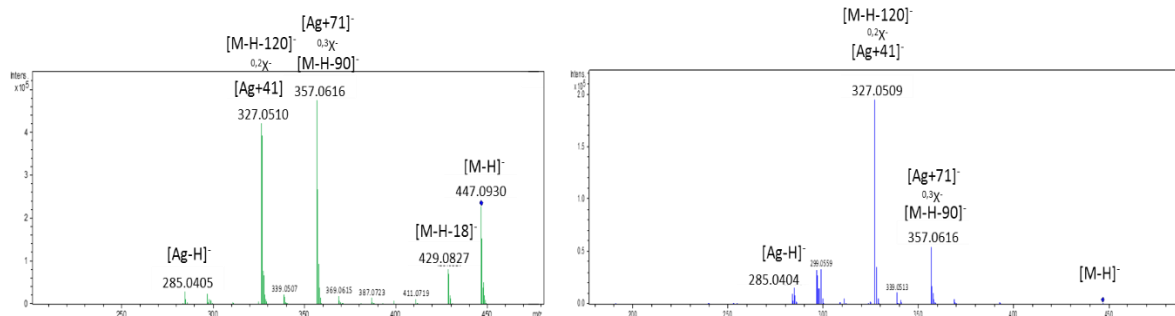
- Flavonoïde mono-*O*-glycosylé : Peltatoside



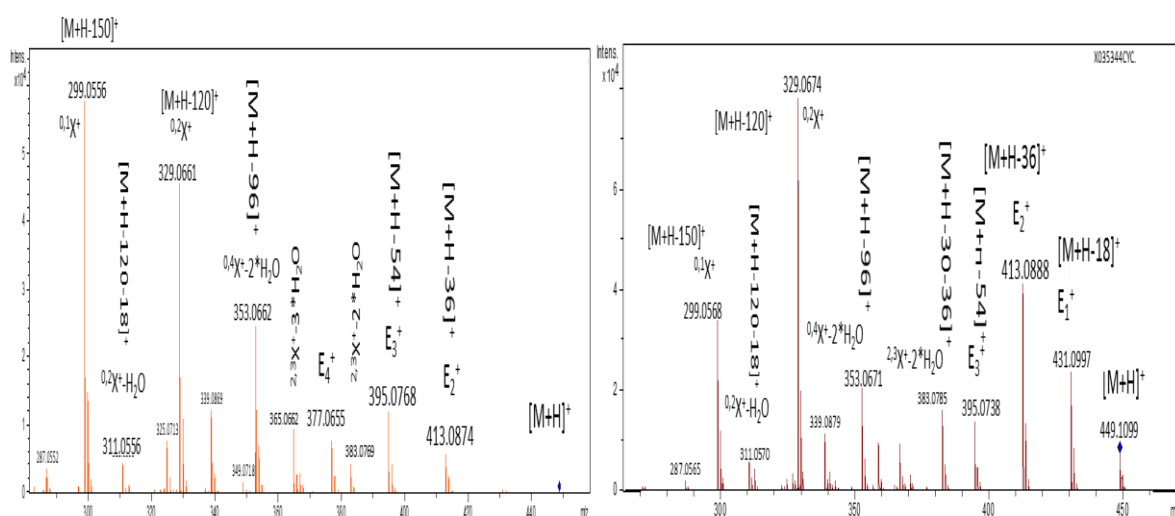
Spectres de fragmentation de la molécule peltatoside en mode d'ionisation négatif et positif, exemple de flavonoïde-*O*-glycosylé, obtenus sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV.

- Flavonoïde mono-*C*-glycosylé : Isoorientine et orientine





Structures et spectres de fragmentation des deux molécules isomères : lutéoline 6-C-glucoside (isoorientine) et lutéoline 8-C-glucoside (orientine) en mode d'ionisation négatif, obtenus sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV.



Spectres de fragmentation des molécules isomères lutéoline 6-C-glucoside (isoorientine) et lutéoline 8-C-glucoside (orientine) en mode d'ionisation positif, obtenus sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV.

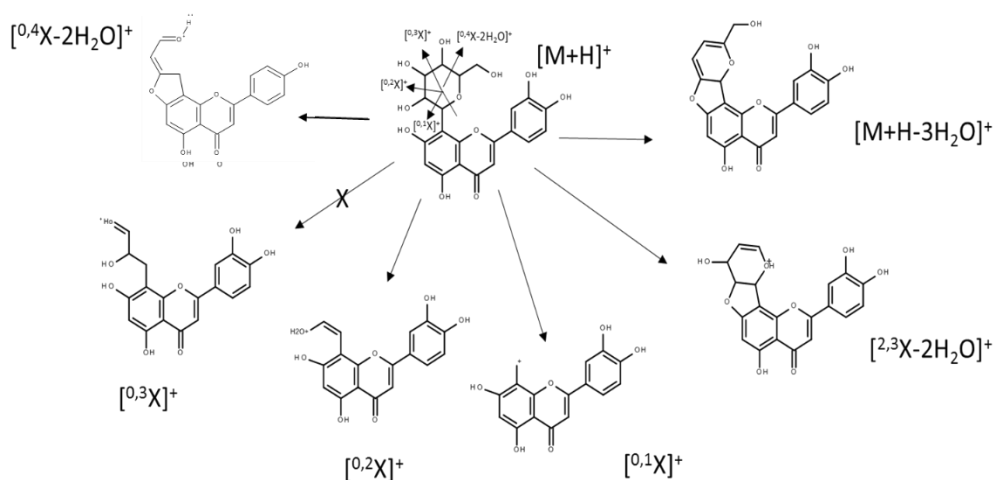
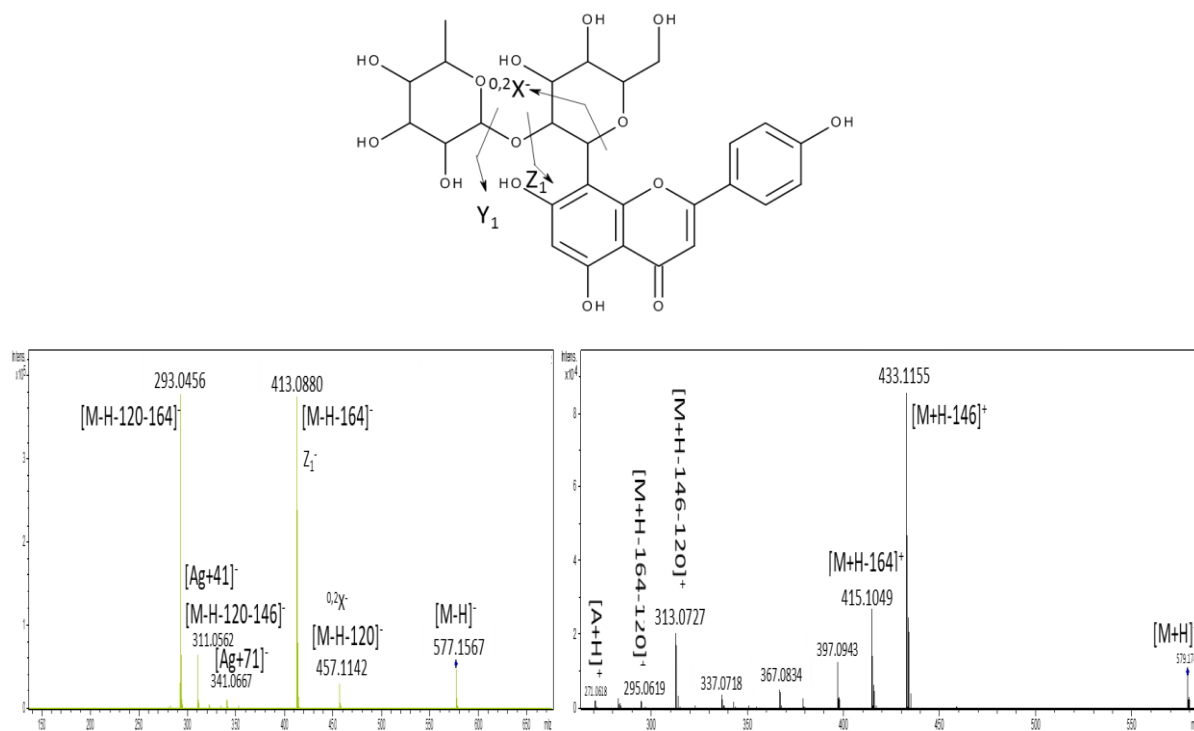


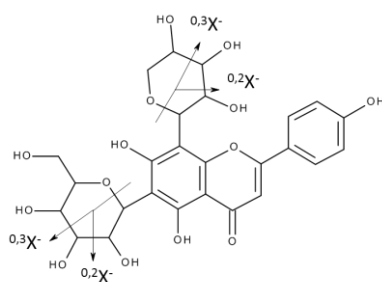
Schéma de fragmentation de l'orientine en mode d'ionisation positif.

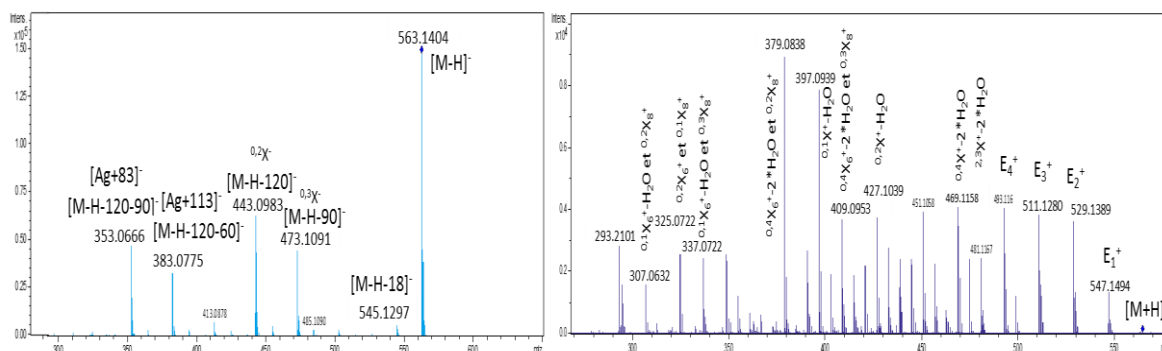
- Flavonoïde mono-C-glycosylé-O-glycosylé : Vitexine-2''-O-rhamnoside



Spectres de fragmentation de la molécule Vitexine-2''-O-rhamnoside en mode d'ionisation négatif et positif, exemple de flavonoïde mono-C-glycosylé-O-glycosylé, obtenus sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV.

- Flavonoïde di-C-glycosylé : Schaftoside





Spectres de fragmentation de la molécule schaftoside en mode d'ionisation négatif et positif, exemple de flavonoïde di-C-glycosylé, obtenus sur un appareil qTOF (Maxis de Bruker), avec une énergie de collision de 30 eV.

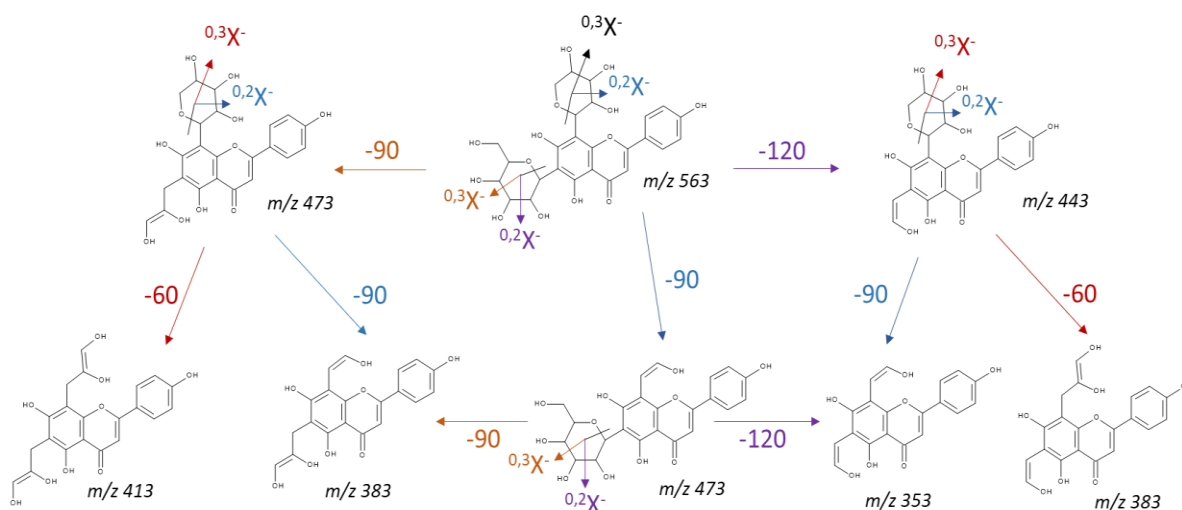
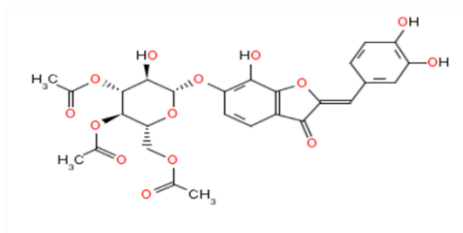


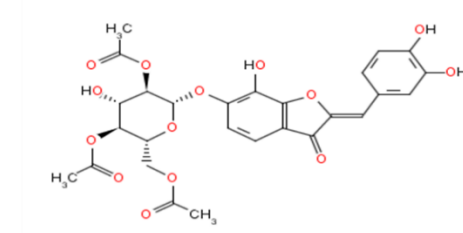
Schéma de fragmentation de la molécule schaftoside en mode d'ionisation négatif.

Annexe 4 : Propositions de molécules pour la formule brute $C_{27}H_{26}O_{14}$ de CompoundCrawler 3.0 obtenu via ChempSpider

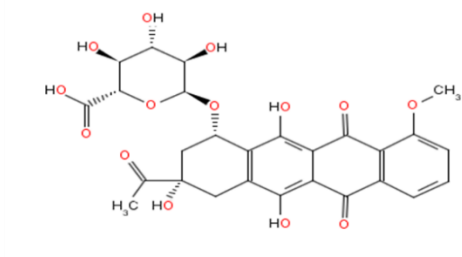
ID 8778584 : (2Z)-2-(3,4-Dihydroxybenzylidene)-7-hydroxy-3-oxo-2,3-dihydro-1-benzofuran-6-yl 3,4,6-tri-O-acetyl-beta-D-glucopyranoside



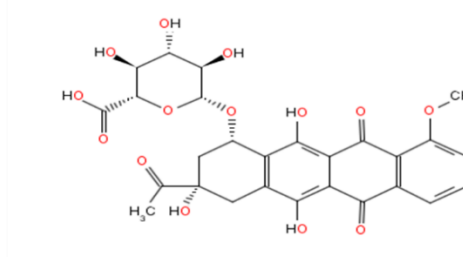
ID 8849987 : (2Z)-2-(3,4-Dihydroxybenzylidene)-7-hydroxy-3-oxo-2,3-dihydro-1-benzofuran-6-yl 2,4,6-tri-O-acetyl-beta-D-glucopyranoside



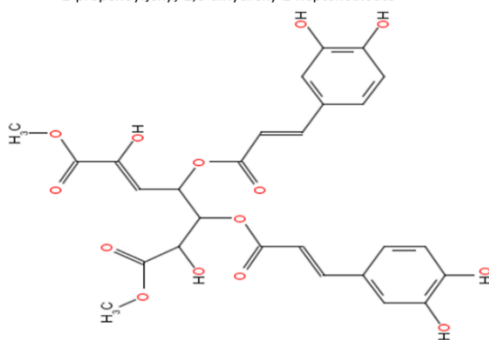
ID 9355752 : (1S,3S)-3-Acetyl-3,5,12-trihydroxy-10-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-1-tetracenyl alpha-D-glucopyranosiduronic acid



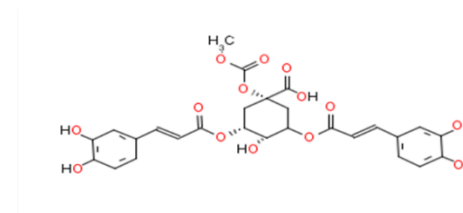
ID 9390350 : (1S,3S)-3-Acetyl-3,5,12-trihydroxy-10-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-1-tetracenyl beta-D-glucopyranosiduronic acid



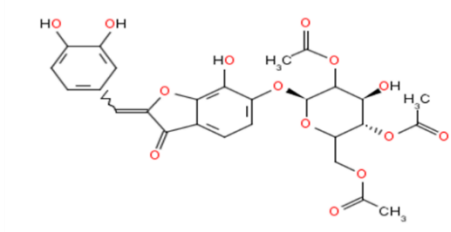
ID 20567074 : Dimethyl (2Z)-4,5-bis{[(2E)-3-(3,4-dihydroxyphenyl)-2-propenyl]oxy}-2,6-dihydroxy-2-heptenedioate



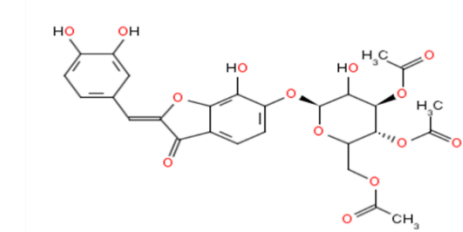
ID 23221032 : (3R)-3,5-Bis{[(2E)-3-(3,4-dihydroxyphenyl)-2-propenyl]oxy}-4-hydroxy-1-[(methoxycarbonyl)oxy]cyclohexanecarboxylic acid



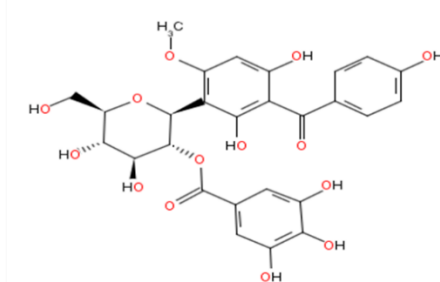
ID 24846159 : (2Z)-2-(3,4-Dihydroxybenzylidene)-7-hydroxy-3-oxo-2,3-dihydro-1-benzofuran-6-yl 2,4,6-tri-O-acetyl-beta-D-threo-hexopyranoside



ID 24846165 : (2Z)-2-(3,4-Dihydroxybenzylidene)-7-hydroxy-3-oxo-2,3-dihydro-1-benzofuran-6-yl 3,4,6-tri-O-acetyl-beta-D-threo-hexopyranoside



ID 34237713 : (1S)-1,5-Anhydro-1-[2,4-dihydroxy-3-(4-hydroxybenzoyl)-6-methoxyphenyl]-2-O-(3,4,5-trihydroxybenzoyl)-D-glucitol



Annexe 5 : Tableau des transitions MRM et paramètres d'acquisition pour le dosage quantitatif en UHPLC/MRM MS base résolution des 67 composés et de l'étalon externe

Tr	Compound	Polarity	Precursor (m/z)	Product (m/z)	Collision Energy (V)	Dwell Time (ms)	RF Lens (V)	Q1 Res. (FWHM)
0,57	533	Negative	533	191	25	5	200	2
0,59	405	Negative	405	191	25	5	160	0.7
1,1	401	Negative	401	153	25	100	160	0.7
1,13	558	Negative	558	346	25	10	200	0.7
1,59	431	Negative	431	137	25	75	160	0.7
2,54	593b	Negative	593	429	30	50	250	2
2,65	565	Negative	565	323	25	100	200	0.7
2,65	623	Negative	623	383	30	5	230	2
2,66	593da	Negative	593	429	30	50	250	2
2,76	Shaftoside	Negative	563	353	25	1	180	2
2,82	IsoOrientin	Negative	447	327	25	100	160	0.7
3	587b	Negative	587	425	25	100	200	0.7
3,05	Vitexine-2''-O-Rhamn	Negative	577	293	25	1	180	0.7
3,07	547a	Negative	547	473	20	1	230	2
3,12	Rutin	Negative	609	301	30	1	250	0.7
3,29	463a	Negative	463	301	20	5	200	0.7
3,42	607a	Negative	607	323	20	1	200	2
3,62	637a	Negative	637	299	25	1	250	2
3,67	593a	Negative	593	285	20	100	250	0.7
3,69	547ab	Negative	547	473	20	1	230	2
3,71	651	Negative	651	351	20	1	180	2
3,72	449a	Negative	449	151	20	75	160	0.7
3,88	461	Negative	461	299	20	1	180	2
3,89	447a	Negative	447	284	20	5	160	0.7
3,96	637b	Negative	637	299	25	1	250	2
4	547ac	Negative	547	473	20	1	230	2
4,03	589b	Negative	589	471	20	100	200	0.7
4,07	449b	Negative	449	151	20	75	160	0.7
4,08	463b	Negative	463	301	20	5	200	0.7
4,13	463c	Negative	463	301	20	5	200	0.7
4,24	591a	Negative	591	429	25	5	250	0.7
4,27	575	Negative	575	411	20	1	200	0.7
4,37	593c	Negative	593	411	20	10	250	2
4,39	473a	Negative	473	327	20	75	180	0.7
4,39	447ab	Negative	447	284	20	5	160	0.7
4,41	593db	Negative	593	429	30	50	250	2
4,46	577b	Negative	577	473	25	5	200	0.7
4,47	573a	Negative	573	324	20	75	180	0.7

4,5	511	Negative	511	273	20	50	180	2
4,57	561a	Negative	561	457	20	75	250	0.7
4,65	475a	Negative	475	299	25	75	180	0.7
4,68	603	Negative	603	395	20	75	180	0.7
4,73	559	Negative	559	395	20	10	200	0.7
4,74	593dc	Negative	593	429	30	50	250	2
4,74	547ad	Negative	547	473	20	1	230	2
4,8	473b	Negative	473	327	20	75	180	0.7
4,8	475b	Negative	475	299	25	75	180	0.7
4,82	457aa	Negative	457	311	20	5	200	0.7
4,83	573b	Negative	573	324	20	75	180	0.7
4,84	561b	Negative	561	457	20	75	250	0.7
4,91	589a	Negative	589	425	20	1	200	0.7
4,94	607b	Negative	607	425	20	75	200	0.7
4,95	657	Negative	657	425	20	100	200	0.7
4,95	487a	Negative	487	341	20	1	200	2
4,97	621	Negative	621	487	20	5	200	0.7
4,99	591b	Negative	591	429	25	5	250	0.7
5	587a	Negative	587	323	20	5	180	0.7
5,05	593dd	Negative	593	429	30	50	250	2
5,12	457ab	Negative	457	311	20	5	200	0.7
5,24	487b	Negative	487	341	20	1	200	2
5,29	515	Negative	515	369	20	10	180	0.7
5,32	603b	Negative	603	395	20	75	180	0.7
5,46	605	Negative	605	471	20	75	200	0.7
6,15	619	Negative	619	485	20	5	140	0.7
6,16	485	Negative	485	338	25	1	200	0.7
6,41	601a	Negative	601	187	20	5	200	0.7
8,25	417	Negative	417	401	25	250	160	2

Annexe 6 : Tableau résultats des répétitions extraction et injection

Ind	Variétés	593-353	623-383	563-473	593-429	447-327	593-413	447-327	595-301	593-413	577-413	607-443	651-351	547-383	461-299	447-284	547-383
Gar1-05MO1_2EI	Gar1-05	3788	39	1817	311	592	0	0	36949	0	3615	1269	34826	1103	1260	0	257
Gar1-05MO2_2EI	Gar1-05	5967	38	2685	305	311	0	0	29774	0	5786	1870	52975	2243	1327	0	469
Gar1-05MO3_2EI	Gar1-05	10719	58	5464	711	773	0	0	21502	0	15438	3224	78947	3153	2015	0	1086
Jav1-04MO1_2EI	Jav1-04	5308	3475	8066	4746	7770	4302	16520	29639	4596	1625	14194	718867	2515	120907	0	7426
Jav1-04MO2_2EI	Jav1-04	7601	1796	10631	4630	7856	2014	6979	48061	2181	3040	11714	775248	2942	68438	0	7504
Jav1-04MO3_2EI	Jav1-04	5934	2354	8949	8049	19805	3220	15804	15033	3485	2657	13924	756936	2924	93644	0	6841
Jav5-05MO1_2EI	Jav5-05	19611	8324	18782	19283	31310	4878	7161	30495	3969	48952	14408	788194	17875	67266	0	9481
Jav5-05MO2_2EI	Jav5-05	7345	7911	8917	13069	19250	5644	5303	46394	3479	13772	11520	599631	8572	72077	0	7701
Jav5-05MO3_2EI	Jav5-05	10773	9393	11524	13111	19837	5865	7426	20883	5268	20170	12256	722872	11012	77072	0	10982
Gar1-03_2EI	Gar1-03	9545	51	4538	711	0	0	0	24973	0	9817	2509	73012	522	1005	2479	1334
Gar1-03_2EI_160520121230	Gar1-03	8407	101	4744	499	643	0	0	23485	0	9186	2242	71135	739	1364	2687	1350
Gar1-03_2EI_160520200324	Gar1-03	8923	18	4308	714	470	0	0	23900	0	8343	2005	67667	964	1325	2274	1211
Gar1-04_2EI	Gar1-04	6701	63	3668	586	190	0	0	26360	0	5312	2260	70855	707	2013	7153	1314
Gar1-04_2EI_160520102237	Gar1-04	5878	100	3565	671	925	0	94	26392	0	5232	1882	69480	609	2122	9396	1555
Gar1-04_2EI_160520181330	Gar1-04	6354	27	2787	648	1198	0	0	26414	0	4628	1805	68035	1017	2422	9559	1277
Gar1-05_2EI	Gar1-05	4060	58	1716	475	327	0	56	40760	0	3703	1422	40832	1385	724	0	257
Gar1-05_2EI_160520140223	Gar1-05	3925	20	1833	377	279	0	0	39268	0	3614	1238	39804	888	982	0	263
Gar1-05_2EI_160520215317	Gar1-05	3469	41	1910	476	234	0	0	39143	0	3609	904	34870	1304	800	0	369
Jav1-04_2EI	Jav1-04	5522	3596	8062	6274	10017	5192	11766	32164	4725	1595	14993	845133	2377	126282	0	7450
Jav1-04_2EI_160520122812	Jav1-04	5254	3692	7161	6758	9438	4477	16456	30116	5070	1806	14721	816492	2097	121803	0	7046
Jav1-04_2EI_160520201906	Jav1-04	5371	3574	7102	5087	9142	4447	16025	29696	4325	1548	14565	784460	2846	126027	0	7004
Jav1-05_2EI	Jav1-05	21848	6153	27047	14381	6896	7153	2458	25986	8641	27094	70732	756468	9952	40603	3979	2877
Jav1-05_2EI_160520103819	Jav1-05	20343	6077	25127	13068	7536	7985	3460	27455	8449	27076	68510	749344	9238	40771	4066	2710
Jav1-05_2EI_160520182912	Jav1-05	20175	5611	26347	11626	7752	7138	3550	24805	8520	26211	65080	728652	8572	44017	4092	3758
Jav5-05_2EI	Jav5-05	21542	8822	22205	23109	38960	5725	4674	36728	4312	56800	18661	861883	18766	65751	0	10642
Jav5-05_2EI_160520141805	Jav5-05	20349	9062	20502	20980	41109	5292	2339	35294	4788	53882	16785	840440	21107	71492	0	12332
Jav5-05_2EI_160520220859	Jav5-05	19609	8165	19606	20084	36955	4824	3948	31625	3842	50585	15918	788280	20792	68977	0	10316

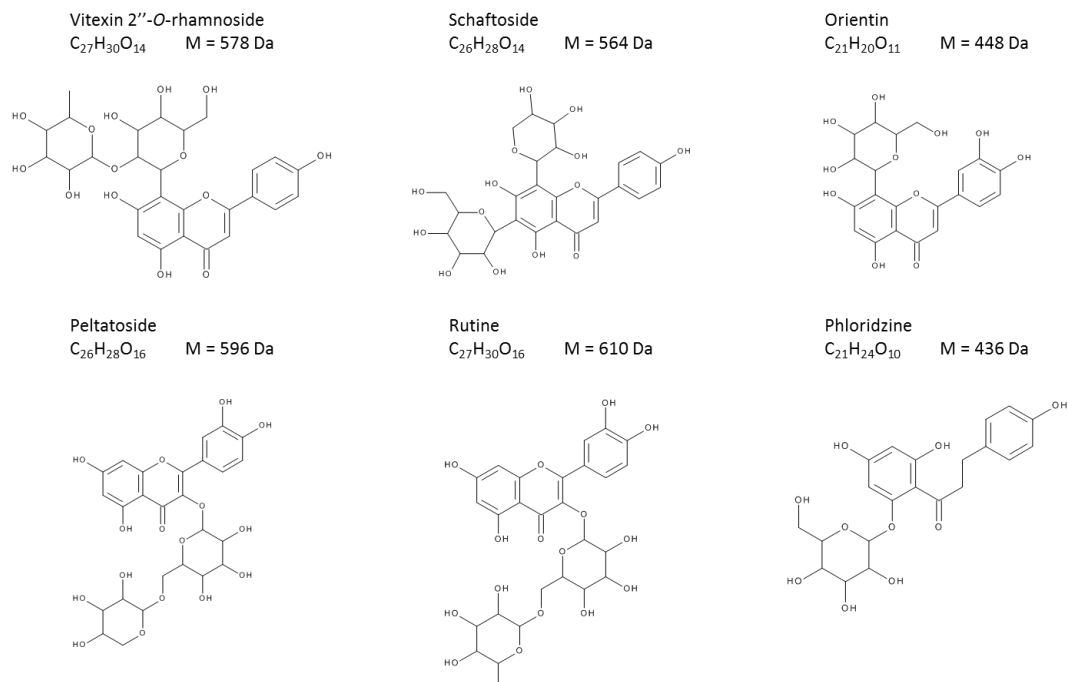
Annexes

Individus	Variétés	593-429	463-301	435-273	547-383	591-487	593-473	575-411	593-411	591-427	591-427	559-395	487-341	591-427	589-425	241-197	559-395
Gar1-05MO1_2EI	Gar1-05	235	396	207085	802	0	3499	33456	7426	204	1646	36659	0	2886	48956	9572	0
Gar1-05MO2_2EI	Gar1-05	299	141	150981	1330	0	3920	39059	9641	280	2092	65950	0	3694	66339	11764	0
Gar1-05MO3_2EI	Gar1-05	339	313	177152	2422	0	8163	83226	17016	683	2605	134236	0	6674	110548	15552	0
Jav1-04MO1_2EI	Jav1-04	20430	51543	82136	6711	2558	318928	3180957	821006	39275	5158	212223	39766	131946	2644828	6917	4438
Jav1-04MO2_2EI	Jav1-04	13978	20927	222133	9662	7553	197708	2235078	595278	40265	6603	243785	61634	95400	1924262	18620	4529
Jav1-04MO3_2EI	Jav1-04	23034	55098	156305	6547	3525	353517	3809513	957854	53463	5439	257131	58607	110834	2280271	9034	4795
Jav5-05MO1_2EI	Jav5-05	25604	80962	123153	12279	660	400168	4303986	1038566	57044	10679	1058944	83143	98715	1992897	2834	5129
Jav5-05MO2_2EI	Jav5-05	21607	65462	82473	6625	740	362470	3641189	864712	43554	9033	430999	42064	89121	1849385	2821	4937
Jav5-05MO3_2EI	Jav5-05	25005	66398	183636	7796	563	397490	4091135	995734	48547	10725	604795	48760	100878	2074358	4334	4043
Gar1-03_2EI	Gar1-03	323	0	218359	1089	0	5983	55809	12744	866	880	75490	0	4538	75984	15751	795
Gar1-03_2EI_160520121230	Gar1-03	186	488	213045	1029	0	3824	51550	9341	369	703	75172	0	4465	74021	15662	922
Gar1-03_2EI_160520200324	Gar1-03	360	320	208141	1226	0	5179	52230	11988	603	672	71218	0	5063	67349	15459	747
Gar1-04_2EI	Gar1-04	522	819	263135	877	127	7183	64806	13720	471	1247	63235	0	5979	75714	28337	1157
Gar1-04_2EI_160520102237	Gar1-04	323	938	263786	1079	216	6482	66962	14069	960	739	60895	0	5341	73325	28911	632
Gar1-04_2EI_160520181330	Gar1-04	411	748	246619	645	146	5939	59213	12924	636	718	58937	0	5137	69467	29026	338
Gar1-05_2EI	Gar1-05	196	410	221302	928	0	4605	38680	9123	434	1493	40600	0	2909	52685	9311	52
Gar1-05_2EI_160520140223	Gar1-05	290	305	210668	1114	0	4035	36349	9237	561	1648	40197	0	2609	50035	9493	423
Gar1-05_2EI_160520215317	Gar1-05	223	274	201594	643	0	4127	36748	8777	252	1224	36740	0	3481	49042	8894	442
Jav1-04_2EI	Jav1-04	22556	54032	88871	7196	3343	373151	3555226	934726	35378	7306	237207	39054	144719	2901053	5942	4953
Jav1-04_2EI_160520122812	Jav1-04	22566	54443	89894	6714	3087	362125	3352311	907417	36819	6952	222617	41452	136392	2732980	5921	3138
Jav1-04_2EI_160520201906	Jav1-04	19132	52851	83624	6100	3181	322849	3135893	838176	35369	6390	210288	48362	131133	2618024	6177	3809
Jav1-05_2EI	Jav1-05	9174	69948	221557	7362	3206	136053	1445144	348696	13351	34821	609700	135121	170535	3415435	5416	13512
Jav1-05_2EI_160520103819	Jav1-05	8862	69721	220784	8073	3016	127079	1412116	324947	16442	33447	579957	147443	168876	3276343	5635	15013
Jav1-05_2EI_160520182912	Jav1-05	8851	65490	212817	7720	2329	124291	1332961	312610	15506	30432	569046	169146	158693	3176375	5059	13425
Jav5-05_2EI	Jav5-05	31194	84556	130691	13005	1001	465593	4768516	1181100	49465	11599	1184588	75082	109422	2172976	2553	4668
Jav5-05_2EI_160520141805	Jav5-05	30424	83775	127827	13118	1084	433126	4556046	1121512	56351	11762	1113646	86598	105042	2122834	3047	3199
Jav5-05_2EI_160520220859	Jav5-05	24963	82669	120988	11454	1250	402082	4372575	1062181	51479	10768	1063153	83632	101390	1974482	2699	4612

Annexes

Individus	Variétés	485-381	487-341	487-341	301-151	329-229	589-455	619-485	343-284	309-291	309-291	311-293	311-293	309-291
Gar1-05MO1_2EI	Gar1-05	0	142	151	143	16300	149	262	0	107729	11892	242158	36773	11100
Gar1-05MO2_2EI	Gar1-05	0	467	487	99	22878	609	684	0	145440	15883	334312	48833	10064
Gar1-05MO3_2EI	Gar1-05	0	1846	1892	419	29282	2617	2956	0	163447	12375	359078	64585	16196
Jav1-04MO1_2EI	Jav1-04	1444	28298	27015	4628	14858	11607	232364	921	56207	8062	135778	13003	6191
Jav1-04MO2_2EI	Jav1-04	1585	37703	34940	4889	24250	20841	233802	1325	48974	6918	114965	17610	2836
Jav1-04MO3_2EI	Jav1-04	1311	34353	31526	4432	14314	19502	271002	1034	45944	5438	115415	18780	5080
Jav5-05MO1_2EI	Jav5-05	2870	54222	51030	3219	8018	162653	483620	1448	37814	3011	43316	7949	2537
Jav5-05MO2_2EI	Jav5-05	1904	41003	38458	3681	9208	51420	323850	1177	53252	6125	50789	9178	3003
Jav5-05MO3_2EI	Jav5-05	2448	35340	35384	2353	8324	72736	372845	1207	61532	3354	52289	7333	2737
Gar1-03_2EI	Gar1-03	0	2170	1890	311	56951	4550	6367	0	105838	14091	393593	64324	16624
Gar1-03_2EI_160520121230	Gar1-03	0	2129	2197	236	55984	4692	6459	0	110280	17774	388148	65867	16923
Gar1-03_2EI_160520200324	Gar1-03	0	1973	1769	249	53065	3972	6259	0	102976	13415	378533	59407	15981
Gar1-04_2EI	Gar1-04	0	3791	3268	377	99253	4317	7320	0	87416	15267	446952	93838	33615
Gar1-04_2EI_160520102237	Gar1-04	0	3388	3119	380	104638	4001	5921	0	86875	16646	440409	82777	33816
Gar1-04_2EI_160520181330	Gar1-04	0	3080	2791	380	97009	3808	6195	0	79555	16670	423776	96160	31998
Gar1-05_2EI	Gar1-05	0	286	275	212	16818	390	901	0	122131	10348	259191	22091	11950
Gar1-05_2EI_160520140223	Gar1-05	0	226	211	217	17048	382	540	0	113082	14494	247285	41521	13562
Gar1-05_2EI_160520215317	Gar1-05	0	229	232	249	15862	254	0	0	110772	8408	234863	35826	13050
Jav1-04_2EI	Jav1-04	1240	32725	29750	5062	16563	14212	266120	870	61346	7915	146484	15845	7270
Jav1-04_2EI_160520122812	Jav1-04	1657	31038	28837	4939	15297	12589	258506	1112	58287	9748	144155	16377	6451
Jav1-04_2EI_160520201906	Jav1-04	1553	28402	27710	4635	16365	13109	239906	902	59039	8961	133210	18813	7138
Jav1-05_2EI	Jav1-05	6720	218793	200545	4887	32696	162548	1277601	3232	34800	9401	150590	23407	6002
Jav1-05_2EI_160520103819	Jav1-05	7814	222604	201595	5031	31177	159055	1251553	3010	36146	10476	154616	24194	5888
Jav1-05_2EI_160520182912	Jav1-05	7193	200966	194281	4931	29793	146193	1161039	3434	33793	11385	141257	18910	5379
Jav5-05_2EI	Jav5-05	3281	62209	55138	2797	8772	193757	541017	1506	40220	3931	44370	7230	2438
Jav5-05_2EI_160520141805	Jav5-05	3110	57457	52356	3659	8244	188391	504336	1901	41262	3289	45854	7749	2356
Jav5-05_2EI_160520220859	Jav5-05	3006	54477	55097	2783	8214	178237	485726	1233	37587	3121	44025	7412	2372

Annexe 7 : Structures des standards utilisés



Annexe 8 : Séquences du passage des échantillons

Ces deux séquences ont été répétées trois fois en entier, en enlevant juste la gamme en commun.

Séquence 1

File Name	Path	Position	Inj Vol
Blanc	20170202	BA1	3.000
QC-Mix36soies	20170202	BA2	3.000
0ppm	20170202	BA3	3.000
0.001ppm	20170202	BA4	3.000
0.01ppm	20170202	BA5	3.000
0.02ppm	20170202	BA6	3.000
0.1ppm	20170202	BA7	3.000
0.2ppm	20170202	BA8	3.000
0.5ppm	20170202	BB1	3.000
1ppm	20170202	BB2	3.000
2ppm	20170202	BB3	3.000
10ppm	20170202	BB4	3.000
20ppm	20170202	BB5	3.000
Blanc	20170202	BB6	3.000
Nov1-02-2EI	20170202	BB7	3.000
Jav1-10-2EI	20170202	BB8	3.000

Séquence 2

File Name	Path	Position	Inj Vol
Blanc	20170221	BA1	3.000
QC-Mix36soies	20170221	BA2	3.000
0ppm	20170221	BA3	3.000
0.001ppm	20170221	BA4	3.000
0.01ppm	20170221	BA5	3.000
0.02ppm	20170221	BA6	3.000
0.1ppm	20170221	BA7	3.000
0.2ppm	20170221	BA8	3.000
0.5ppm	20170221	BB1	3.000
1ppm	20170221	BB2	3.000
2ppm	20170221	BB3	3.000
10ppm	20170221	BB4	3.000
20ppm	20170221	BB5	3.000
Blanc	20170221	BB6	3.000
Gar1-02-2EI	20170221	BE1	3.000
Nov1-04-2EI	20170221	BE2	3.000

Annexes

Gb1-04-2EI	20170202	BC1	3.000	Gar1-03-2EI	20170221	BD1	3.000
Gar1-08-2EI	20170202	BC2	3.000	SJ1-10-2EI	20170221	BD2	3.000
SJ1-05-2EI	20170202	BC3	3.000	Blanc	20170221	BB6	3.000
Nov1-08-2EI	20170202	BC4	3.000	Nov1-01-2EI	20170221	BD3	3.000
Jav1-06-2EI	20170202	BC5	3.000	Jav1-04-2EI	20170221	BD4	3.000
Blanc	20170202	BA1	3.000	Gb5-01-2EI	20170221	BD5	3.000
QC-Mix36soies	20170202	BA2	3.000	Gar1-05-2EI	20170221	BD6	3.000
GB04-2EI	20170202	BC6	3.000	QC-Mix36soies	20170221	BA2	3.000
Gar1-09-2EI	20170202	BC7	3.000	0ppm	20170221	BA3	3.000
SJ1-09-2EI	20170202	BC8	3.000	0.001ppm	20170221	BA4	3.000
Nov1-07-2EI	20170202	BD1	3.000	0.01ppm	20170221	BA5	3.000
Jav5-02-2EI	20170202	BD2	3.000	0.02ppm	20170221	BA6	3.000
Gb02-2EI	20170202	BD3	3.000	0.1ppm	20170221	BA7	3.000
Blanc	20170202	BA1	3.000	0.2ppm	20170221	BA8	3.000
Gar1-01-2EI	20170202	BD4	3.000	0.5ppm	20170221	BB1	3.000
SJ1-01-2EI	20170202	BD5	3.000	1ppm	20170221	BB2	3.000
Nov1-09-2EI	20170202	BD6	3.000	2ppm	20170221	BB3	3.000
Jav5-01-2EI	20170202	BD7	3.000	10ppm	20170221	BB4	3.000
Gb1-02-2EI	20170202	BD8	3.000	20ppm	20170221	BB5	3.000
Gar1-02-2EI	20170202	BE1	0.001	Blanc	20170221	BB6	3.000
Nov1-04-2EI	20170202	BE2	0.001	Gar1-02-2EI	20170221	BE1	3.000
QC-Mix36soies	20170202	BA2	3.000	Nov1-04-2EI	20170221	BE2	3.000
0ppm	20170202	BA3	3.000	Gar1-03-2EI	20170221	BD1	3.000
0.001ppm	20170202	BA4	3.000	SJ1-10-2EI	20170221	BD2	3.000
0.01ppm	20170202	BA5	3.000	Blanc	20170221	BB6	3.000
0.02ppm	20170202	BA6	3.000	Nov1-01-2EI	20170221	BD3	3.000
0.1ppm	20170202	BA7	3.000	Jav1-04-2EI	20170221	BD4	3.000
0.2ppm	20170202	BA8	3.000	Gb5-01-2EI	20170221	BD5	3.000
0.5ppm	20170202	BB1	3.000	Gar1-05-2EI	20170221	BD6	3.000
1ppm	20170202	BB2	3.000	QC-Mix36soies	20170221	BA2	3.000
2ppm	20170202	BB3	3.000	0ppm	20170221	BA3	3.000
10ppm	20170202	BB4	3.000	0.001ppm	20170221	BA4	3.000
20ppm	20170202	BB5	3.000	0.01ppm	20170221	BA5	3.000
Blanc	20170202	BB6	3.000	0.02ppm	20170221	BA6	3.000
QC-Mix36soies	20170202	BA2	3.000	0.1ppm	20170221	BA7	3.000
Jav1-07-2EI	20170202	BE3	3.000	0.2ppm	20170221	BA8	3.000
BlancSD-2EI	20170202	BE4	3.000	0.5ppm	20170221	BB1	3.000
Racinerie-2EI	20170202	BE5	3.000	1ppm	20170221	BB2	3.000
JauneACX-2EI	20170202	BE6	3.000	2ppm	20170221	BB3	3.000
Nov1-05-2EI	20170202	BE7	3.000	10ppm	20170221	BB4	3.000
Jav5-05-2EI	20170202	BE8	3.000	20ppm	20170221	BB5	3.000
Gb5-03-2EI	20170202	GA1	3.000	Blanc	20170221	BB6	3.000
Blanc	20170202	GA2	3.000	Gar1-02-2EI	20170221	BE1	3.000
QC-Mix36soies	20170202	BA2	3.000	Nov1-04-2EI	20170221	BE2	3.000
Gar1-03-2EI	20170202	GA3	0.001	Gar1-03-2EI	20170221	BD1	3.000
SJ1-10-2EI	20170202	GA4	0.001	SJ1-10-2EI	20170221	BD2	3.000

Annexes

Nov1-01-2EI	20170202	GA5	0.001	Blanc	20170221	BB6	3.000
Jav1-04-2EI	20170202	GA6	0.001	Nov1-01-2EI	20170221	BD3	3.000
Gb5-01-2EI	20170202	GA7	0.001	Jav1-04-2EI	20170221	BD4	3.000
Gar1-05-2EI	20170202	GA8	0.001	Gb5-01-2EI	20170221	BD5	3.000
Blanc	20170202	BA1	3.000	Gar1-05-2EI	20170221	BD6	3.000
SJ1-08-2EI	20170202	GB1	3.000	QC-Mix36soies	20170221	BA2	3.000
Nov1-06-2EI	20170202	GB2	3.000	0ppm	20170221	BA3	3.000
Jav1-05-2EI	20170202	GB3	3.000	0.001ppm	20170221	BA4	3.000
Gb5-05-2EI	20170202	GB4	3.000	0.01ppm	20170221	BA5	3.000
SJ1-04-2EI	20170202	GB5	3.000	0.02ppm	20170221	BA6	3.000
Gar1-04-2EI	20170202	GB6	3.000	0.1ppm	20170221	BA7	3.000
QC-Mix36soies	20170202	BA2	3.000	0.2ppm	20170221	BA8	3.000
0ppm	20170202	BA3	3.000	0.5ppm	20170221	BB1	3.000
0.001ppm	20170202	BA4	3.000	1ppm	20170221	BB2	3.000
0.01ppm	20170202	BA5	3.000	2ppm	20170221	BB3	3.000
0.02ppm	20170202	BA6	3.000	10ppm	20170221	BB4	3.000
0.1ppm	20170202	BA7	3.000	20ppm	20170221	BB5	3.000
0.2ppm	20170202	BA8	3.000	Blanc	20170221	BB6	3.000
0.5ppm	20170202	BB1	3.000				
1ppm	20170202	BB2	3.000				
2ppm	20170202	BB3	3.000				
10ppm	20170202	BB4	3.000				
20ppm	20170202	BB5	3.000				
Blanc	20170202	BB6	3.000				
1ppm	20170202	BB2	3.000				
Blanc	20170202	BB6	3.000				

**Signature moléculaire de milieux complexes :
Stratégies de couplage à la spectrométrie de masse et d'interprétation des données**

Résumé :

La composition en métabolites des organismes vivants évolue lorsqu'ils sont soumis à des modifications environnementales et diffère d'une espèce à l'autre. L'objectif de cette thèse a été de déployer des stratégies de couplages chromatographiques avec la spectrométrie de masse et de traitement des données MS pour répondre aux différentes problématiques rencontrées en termes de caractérisation, quantification et comparaison du contenu moléculaire de milieux complexes.

Ces études ont permis en premier lieu, de développer des méthodes de déréplication par TLC/MS pour identifier rapidement les molécules connues et évaluer la composition des milieux d'intérêt. Afin d'améliorer et de standardiser la comparaison de différents échantillons des traitements d'images des plaques TLC ont été effectués et des études statistiques ont été appliquées. De plus, des couplages de la TLC avec la MS ont été développés pour pouvoir identifier différentes familles moléculaires de faible masse (flavonoïdes, stéroïdes). Dans un second temps, une méthodologie a été mise en place en UHPLC/HRMS pour caractériser la présence de 30 flavonoïdes C et O-glycosylés dans des échantillons de soies de maïs de différentes variétés et aider à la sélection de variétés efficaces pour lutter contre l'infestation de ravageurs. Des approches de représentation visuelle du contenu moléculaire ont été utilisées pour cibler les molécules d'intérêt, puis elles ont été quantifiées par UHPLC/MRM MS. Les différentes empreintes chromatographiques ont été comparées en utilisant des traitements statistiques afin de mettre en évidence les différences entre ces variétés et identifier les molécules permettant cette distinction.

Mots clés : TLC/MS, UHPLC/HRMS, UHPLC/MRM MS, flavonoïdes, signatures moléculaires, analyses statistiques, réseaux moléculaires

**Molecular fingerprint of complex samples:
Strategies of mass spectrometry hyphenation and data interpretation**

Abstract:

The metabolite composition of living organisms changes when they are subjected to environmental modification and differs from one species to another. The aim of this PhD thesis was to deploy different chromatographic coupling strategies with mass spectrometry and MS data processing in order to respond in the most relevant way to the various problems encountered in terms of characterization, quantification and comparison of the molecular content of complex samples.

Firstly, these studies led to the development of TLC/MS dereplication methods to rapidly identify known molecules and to evaluate the composition of samples. In order to improve and standardize the comparison of different samples, image processing of TLC plates were performed and statistical studies were applied. In addition, the hyphenation TLC/MS have been developed to identify different low molecular mass families (flavonoids, steroids). In a second way, a methodology was implemented in UHPLC/HRMS to characterize the presence of C and O-glycosyl flavonoids in corn silk samples of five different varieties and help in the varietal selection to fight against pest infestation. Visual representation approaches of molecular content were used to target the molecules of interest which were they were quantified by UHPLC/MRM MS. The different chromatographic fingerprints were then compared using statistical treatments in order to highlight the differences between these five varieties and thus to identify the molecules allowing this distinction.

Keywords: TLC/MS, UHPLC/HRMS, UHPLS/MRM MS, flavonoids, molecular fingerprint, statistical analysis, molecular networking



Institut de Chimie Organique et Analytique
Université d'Orléans, CNRS UMR 7311, BP 6759
45067 Orléans Cedex 2, France

