Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Amandine HUEBER

Le 6 décembre 2021

Identification et caractérisation de nouveaux lipopeptides bactériens par spectrométrie de masse

---

Ecole doctorale : **BSB - Biologie, Santé, Biotechnologies**

Spécialité : **PHYSIOPATHOLOGIE**

Unité de recherche :
IRSD - Institut de Recherche en Santé Digestive

Thèse dirigée par
Nicolas CENAC et Justine BERTRAND-MICHEL

Jury

M. François FENAILLE, Rapporteur
Mme Claire CHERBUY, Rapporteure
Mme Sandra ALVES, Rapporteure
M. Jean-Charles PORTAIS, Examinateur
M. Thierry DURAND, Examinateur
M. Gilles DIETRICH, Examinateur
M. Nicolas CENAC, Directeur de thèse
Mme Justine BERTRAND-MICHEL, Co-directrice de thèse

Remerciements

Ce travail a été réalisé au sein de la plateforme MetaToul lipidomique de Toulouse et l'Institut de Recherche en Santé Digestive de l'université Paul Sabatier de Toulouse. Il a été financé dans le cadre d'une ANR par la région Occitanie et par l'institut national de la santé et de la recherche médicale.

Je tiens tout d'abord à remercier chacun des membres de mon jury pour avoir accepté de prendre le temps d'évaluer le travail issu de ces trois années de thèse. Merci au Docteur François Fenaille, au Docteure Sandra Alves et au Docteure Claire Cherbuy d'avoir accepté d'en être les rapporteurs. Merci au Professeur Jean-Charles Portais, au Docteur Thierry Durand et au Docteur Gilles Dietrich d'en avoir été les examinateurs.

Je voudrais remercier mon directeur de thèse Nicolas Cenac pour la confiance qu'il a placé en moi, pour ses nombreux conseils et surtout pour sa patience à toute épreuve. Son encadrement m'a permis d'évoluer, de progresser et de m'épanouir scientifiquement, cela a été un réel plaisir de travailler à ses côtés.

Je remercie également Justine Bertrand-Michel, ma co-directrice de thèse, pour m'avoir confié ce sujet de recherche au sein de son plateau technologique. Merci de m'avoir accompagnée pendant ces trois années et de m'avoir donné l'opportunité de travailler de manière autonome sur un grand nombre de spectromètres de masse.

Je tiens tout particulièrement à remercier le Professeur Jean-Claude Tabet pour m'avoir tant fait progresser et pour m'avoir transmis autant. Un grand merci pour m'avoir constamment conseillé tant sur le plan scientifique que personnel, d'avoir cru en mes capacités et de m'avoir poussée à donner le meilleur de moi-même. Je me souviens encore de nos premières journées fragmentation où (je dois bien avouer) je ne comprenais vraiment pas grand-chose, surtout concernant le sens des électrons ! Aujourd'hui je peux enfin « parler votre langage » et j'ai enfin compris comment placer ces flèches sur un mécanisme de fragmentation (même si ce n'est qu'un début) ! Je vous remercie sincèrement pour chacune de nos discussions, pour votre altruisme, votre humilité à toute épreuve et pour votre capacité d'écoute. J'espère avoir été à la hauteur de votre confiance.

Je voudrais remercier Pauline Le Faouder, pour ses conseils méthodologiques et instrumentaux qui m'ont permis à la fois de mener à bien mes travaux et d'élargir mes connaissances scientifiques. Merci d'avoir été présente lors des moments moins faciles.

Un grand merci à mes voisines de bureau qui ont rendu ces journées plus joviales : Jessica et Aurélie. Les filles, merci pour toutes les fois ou j'avais besoin de réconfort et que vous étiez présentes. Merci pour toutes les discussions et pour tous les bons moments passés ensemble et pour vos conseils, surtout ces derniers mois. Cette thèse n'aurait pas été la même sans vous.

Je remercie Nancy, merci d'avoir été si souvent à mon écoute. Merci pour toutes les fois ou tu m'as ramené un petit pain (que dis-je ? une chocolatine) pour me réconforter. Tu m'as beaucoup aidé pendant cette dernière année de thèse, nos conversations m'ont souvent apaisé, je t'en remercie. Je te souhaite tout le meilleur pour la suite.

Je souhaite remercier Camille ce fût un réel plaisir de travailler avec toi pendant ces trois années, merci d'avoir été mon unique « usine de production » de bactérie ! Merci pour tes nombreux conseils et pour ton aide. Je te souhaite une très bonne continuation, et pleins de bonnes choses pour la suite !

I would like to thank Martin Green, Jakub Ujma and Keith Richardson from Waters compagny for their help concerning the ion mobility experiments.

Je souhaite vivement remercier Geoffrey pour toutes les synthèses de lipopeptides ! Sans ces standards ce travail n'aurait jamais pu aboutir, je te souhaite pleins de bonnes choses (dans cette belle ville qu'est Strasbourg !)

Je souhaite remercier Florian, Hanna, Laurent, Emilien, Carla, et Sylvie de la plateforme MétaToul pour les nombreuses fois ou nous avons travaillé ensemble, et pour leurs conseils. Une mention toute particulière pour Hanna, un grand merci pour ta patience, pour avoir été si réactive quand j'avais besoin d'un prêt de spectro ! Merci de m'avoir fait confiance en ce qui concerne ton parc analytique.

Je souhaite tout particulièrement remercier les professeurs Dimitri Heintz, Pascaline Ullmann et Yanis François pour m'avoir transmis cette passion pour la spectrométrie de masse, sans eux je n'en serai pas là aujourd'hui.

Enfin, sur un plan plus personnel, je souhaite en premier lieu remercier mes amis, d'avoir été mon soutien sans faille pendant ces années : Antonin, Eugénie, Maud, Romain, Enzo, Clément et ceux que j'oublie. Merci pour ces week-ends autour d'une tarte flambée (et souvent d'une bière). Bref, merci pour toutes vos preuves d'amitié et pour m'avoir souvent aidé à souffler et à recharger les batteries.

Je souhaite évidemment à adresser un remerciement tout particulier à Antoine, mon meilleur ami depuis de longues années. Merci d'avoir toujours répondu présent quand j'en avais besoin, d'avoir toujours dis oui pour une bière (fortement réconfortante), ou pour ces moments particuliers à la table 12, merci pour ton écoute. Merci pour ton soutien indéfectible, et d'avoir été le carburant du moteur pendant ces trois longues années, tu le sais : cette réussite c'est aussi là tienne. T'es le plus chic des amis, et surement le meilleur conducteur d'autocar qu'il soit !

Merci à ma famille, à mes parents, et plus particulièrement à ma maman, pour m'avoir encouragée depuis le début. Merci de m'avoir tant apporté, et de toujours avoir cru en moi. Rien ne me comble plus que de voir cette fierté lorsque tu me regarde dans les yeux, merci indéfiniment.

Et enfin, Guillaume, je pense que ces trois années n'ont pas dû être faciles pour toi non plus, surtout la dernière. Je te remercie sincèrement pour ta patience et pour ta bienveillance. Merci d'avoir supporté mes moments de doutes pour finalement avoir traversé tout ça avec moi. Merci, de partager ma vie et simplement merci d'être la merveilleuse personne que tu es.

Résumé

De nombreuses études ont démontré l'implication du microbiote intestinal dans la physiologie et la physiopathologie de l'hôte. Parmi les bactéries probiotiques, *Escherichia coli* Nissle 1917 a été décrite pour ses propriétés analgésiques. L'étude d'extraits lipidiques de cette bactérie par spectrométrie de masse à haute résolution a permis d'identifier le C12AsnGABA, un lipopeptide qu'elle synthétise *via* son ilot pks. In *vivo*, le C12AsnGABA est capable d'inhiber l'hypersensibilité viscérale. Ce lipopeptide bactérien possède donc des propriétés analgésiques. Afin de déterminer si le C12AsnGABA appartient à une nouvelle famille de lipides, nous avons étendu l'analyse à plusieurs espèces bactériennes et mis en place un workflow analytique de chromatographie liquide couplée à la spectrométrie de masse en tandem à haute résolution spécifique des lipopeptides. Celui-ci, comprend plusieurs étapes interconnectées entres-elles et indispensables ayant été optimisées séparément. Vingt-trois nouveaux lipides bactériens ont ainsi pu être caractérisés. Classiquement, l'élucidation structurale de ces composés est réalisée par électrospray à l'aide de la spectrométrie de masse en tandem après de la fragmentation par dissociation induite par collisions. Lors de ma thèse, nous avons combiné les modes d'activation résonnant et non résonnant afin de mettre en évidence les pertes consécutives et les pertes directes pour permettre une meilleure compréhension des mécanismes de fragmentations impliqués. Parmi les résultats obtenus, des ions fragments produits inattendus ont été découverts et ont fait l'objet d'une seconde étude. Les courbes d'évolutions des profils en fonction de l'énergie ont été réalisées et ont permis de mettre en évidence pour la première fois la présence de courbes Gaussiennes élargies et de courbes composites. Ce comportement, jamais décrit auparavant, questionne sur l'origine de ces processus qui semblent être indépendants de l'instrumentation. Nous avons considéré la possibilité que deux chemins réactionnels puissent mener à la formation d'un même ion produit fragment dans le cas des courbes Gaussiennes élargies. Dans le cas des courbes composites, nous avons émis l'hypothèse que deux chemins réactionnels, le premier à basse énergie et le second à plus haute énergie menait chacun à la formation d'un ion de structures différentes mais de masse identique. Par des analyses de mobilité ionique cyclique couplée à de la spectrométrie de masse et des calculs quantiques, nous avons pu confirmer nos hypothèses.

Ces travaux de thèse ont permis d'identifier 23 nouveaux composés produits par le microbiote intestinal. Leur caractérisation structurale nous a permis d'identifier des comportements de fragmentation inattendus générant des ions fragments d'acides gras jamais décrits. Nous avons démontré que deux chemins réactionnels pouvaient exister pour générer deux ions produits de même rapport masse sur charge mais de structures différentes, se traduisant par des courbes composites.

Abstract

Numerous studies have highlighted the involvement of intestinal microbiota in the physiology and pathophysiology of the host. Among probiotic bacteria, Escherichia coli Nissle 1917 (EcN 1917) has been described for its analgesic properties. The study of EcN 1917 lipids extract by high-resolution mass spectrometry permit to identify C12AsnGABA, a lipopeptide that its synthesizes by bacteria pks island. *In vivo*, this compound is able to inhibit visceral hypersensitivity: this bacterial lipid shows analgesic properties. In order to determine if it belongs to a new lipid family, we extended the analysis to several bacterial species and develop an analytical workflow. This one combine a specific lipopeptides extraction and a specific lipopeptides analysis in liquid chromatography coupled with high-resolution tandem mass spectrometry and is constituting by several interconnected and essential steps having been optimized separately. Twenty-three new bacterial lipids have been identified, and characterized. Usually, the structural elucidation of new compounds is carried out by electro-nebulization combined with tandem mass spectrometry after fragmentation by collision-induced dissociation. During my thesis, resonant and non-resonant activation modes were combined to distinguish consecutive losses from competitive losses for a better understanding of their fragmentation mechanisms. Among the obtains results, unexpected fragment ions were discovered and were study. The evolution of the relative abundance of the ions from the CID spectrum in non-resonant mode as a function of the collision energy (ERMS curves) shows for the first time Gaussian profiles which can be broad or composites which seem to be independent of the mass spectrometer. This behavior, that has never been describe, bring up questions about the origin of the fragmentation mechanisms of these product ions. We considered the possibility that two reaction paths could permit the formation of the same fragment product ion for the board Gaussian curves. For the composite curves, we assume that two fragmentation reactions could occur: the first at low energy and the second at higher energy, each leading the formation of two isomers ions. These mechanisms of fragmentation have been confirmed by quantum calculations, and experimental evidences of isomeric ions formation is provided by cyclic ion mobility analysis coupled with mass spectrometry.

This work allows the identification of twenty-three new lipids, produced by the intestinal microbiota. The structural characterization allows us to identify unexpected fragmentation behaviors as the regeneration of fatty acids that has never been describe before. We have proposed two fragmentation mechanisms that could explain the formation of two isomers for the shapes of the ERMS curves, for the composites curves.

Liste des abréviations

AABG : Advanced active beam guide

ACC : Acétyl-CoA carboxylase

AcOEt : Acétate d'étyle

AcN : Acétonitrile

AEA : Anandamide N-araciodonoylehanolamide

AG : Acide gras

AGC : Automatic gain control

AGCC : Acide gras à courte chaîne

AGLC : Acide gras à longue chaîne

AGI : Acide gras insaturé

APCI : Ionisation chimique à pression atmosphérique

API : Atmospheric pression onisation

BUME : butanol-méthanol

C18-3OH : acide 3-hydroxy-octadécanoïque

CI : Ionisation chimique

CID : Collision induced dissociation

cIM : Cyclic ion mobility

Clb : Colibactine

COX : Cyclooxygenase

CV : Compensation voltage

CYP : Cytochrome

DART : Direct analysis in real time

DESI : Desorption electrospray

DEDL : détecteur évaporatif à diffusion de lumière

DFT : Density functional theory

DG : diacylglycéride

DT-IMS : Drift time ion mobility spectrometry

EC : Energie de collision

Ec : Energie cinétique

ECD : Electron capture dissociation

EcN : *Escherichia coli* Nissle 1917

E_{COM} : Energie au centre de masse

EDD : Electron detachment dissociation

EESI : Extractive electrospray ionisation

EET : acide epoxyeicosatriénoïque

EID : Electron induced dissociation

EIEIO : Electron impact excitation of ion from organics

E_{LAB} : Energie laboratoire

ELL : Extraction liquide liquide

EqC : nombre équivalent carbone

ER-IMS : Energy resolved ion mobility mass spectrometry

ERMS : Energy resolved mass spectrometry

ESI : Electrospray ionisation

ESSI : Extractive electrospray ionisation

FAIMs : Field asymmetric waveform ion mobility mass spectrometry

FAS : Fatty acid synthase

FAS-MS : Fast atom bombardment mass spectrometry

FID : Flamme ionization detector

FT : Transformée de fourier

FT-ICR : Spectrométrie de masse à résonnance cyclotronique des ions (Transformée de fourier)

GABA : Acide gamma-aminobutyrique

GC : Chromatographie gazeuse

GERLI : Groupe d'étude et de recherche en lipidomique

GL : Glycolipides

GPA : Glycérophosphate

GPE : Glycérophosphatidylethanolamine

GPI : Glycérophosphophoinositol

$GPIP_2$: Glycérophosphoionositol-biphopshate

GPL : Glycérophospholipide

GPG : Glycéroprostaglandine

HCD : Higher energy C-Trap dissociation

HEPT : Hauteur équivalente à un plateau théorique

HETE : Hydroeicosatetranoate

HILIC : Chromatographie liquide à interaction hydrophile

HR : Haute résolution

HYA : Acide 10-hydroxy-cis-12-octadecenoic

HYB : Acide 10-hydroxy-octadecanoic

HYC : Acide 10-hydroxy-trans-11-octadecenoic

IBS : Irritable bowel syndrome

IE : impact électronique

ICR : Ion cyclotron resonance

IKE : Ion kinetic energy

IRMPD : Infrared multiphoton dissociation

ISTD : Internal standard

KetoA : Acide 10-oxo-cis-12-octadecenoic

KetoB : Acide 10-oxo-octadecanoic

KetoC : Acide 10-oxo-trans-11-octadecenoic

LC : Chromatogrpahie liquide

LC-HRMS : Chromatographie liquide couplée à la spectrométrie de masse à haute résolution

LC-LRMS : Chromatographie liquide couplée à la spectrométrie de masse à basse résolution

LLE : Liquid-liquide extraction

LMCO : Low mass cut off

LO : Lipogenase

LpAA : Lipoaminoacide

LPS : Lipopolysaccharide

LR : Low résolution (basse résolution)

LT : Leukotriène

LTQ : Linear trap quadrupole

MALDI : Matrix assisted laser desorption ionisation

MeOH : Methanol

MIKE : Mass-analaysed ion kinetic energy

MG : Monoacylglycéride

MS : Spectrométrie de masse

MRM: Multiple reaction monitoring

MtBE: Méthyl *tert*-butyl éther

NAT: N-acétyltransférase

NL : Neutral loss (perte de neutre)

NRPS : peptide synthase non ribosomique

QET : Théorie du quasi-équilibre

RR : Processus de rearrangement

RS : Processus de rupture simple

SFAHFA : Short chain fatty-acid ester of fatty acid

SII : Syndrome de l’intestin irritable

SIMS : Secondary ion mass spectrometry

SL: Sphingolipides

SM: Sphingomyéline

SMB: Supramolecular bombardment

SN: Stereospecific numbering

SPE: Solid phase extraction

SRM: Single reaction monitoring

PA : Affinité protonique

PIS : Precursor ion scan

PrIS : Product ion scan

P-LC : Chromatographie polaire

PG : Prostaglandine

QqQ : Triple quadripôle

RCPC : récepteur couplé aux protéines G

RF : Radiofréquence

RMN : Résonance magnétique nucléaire

ROS : Espèces réactives de l'oxygènes

RP-LC : Chromatographie liquide en phase inverse

TG : Triacylglécide

Tm : Temps mort

TOF : Time of flight

Tr : Temps de rétention

TWIMS : Traweling wave ion mobility mass spectrometry

Abréviation des acides aminés :

Ala : Alanine (A)

Arg : Arginine (R)

Asn : Asparagine (N)

Asp : Aspartate / Acide aspartique (D)

Cys : Cystéine (C)

Glu : Glutamate / Acide glutamique (Q)

Gly : Glycine (G)

His : Histidine (H)

Ile : Isoleucine (I)

Leu : Leucine (L)

Met : Méthionine (M)

Phe : Phenylalanine (F)

Pro : Proline (P)

Ser : Sérine (S)

Thr : Thréonine (T)

Trp : Tryptophane (W)

Tyr : Tyrosine (Y)

Val : Valine (V)

Liste des figures

Figure 1 : Unité isoprène, composition élémentaire C_5H_8

Figure 2 : Fonction carboxylique, avec R chaine alkyle

Figure 3 : Différence entre des composés *trans* et *cis*

Figure 4 : Acide palmitique $C_{16}H_{32}O_2$

Figure 5 : Structure de l'acide linoléique, C18:2 Δ(9,12), $C_{18}H_{32}O_2$

Figure 6 : Schéma de la biosynthhèse des acides gras

Figure 7 : Machinerie enzymatique des acides gras

Figure 8 : Structure de l'acide arachidonique (A) et eicosapeintaénoïque (B)

Figure 9 : Métabolisme oxydatif de l'acide arachidonique chez les mammifères

Figure 10 : Structure de la prostaglandine 1 (PGE1)

Figure 11 : Structure du leukotriène LTB44

Figure 12 : Structure du 8-HETE

Figure 13 : Structure de l'anandamide (N-arachiodonoylanolamide (AEA))

Figure 14 : Structure de l'acide arachidonique lié au glutamate

Figure 15 : Structure de l'acide oléique lié au GABA

Figure 16 : Structure du squelette glycérol

Figure 17 : Structure d'un monoacylglycérol, le MG18:0 (1-octadecanoyl-sn-glycerol)

Figure 18 : Structure d'un dyacylglycérol, le DG18:0/18:0, 1-octadecanoyl-2-hexadécanoyl-sn-glycerol

Figure 19 : Structure d'un triacylglycérol, le TG 18:0/18:0/18:0, 1-octadecanoyl-2-octadécanoyl-3- octadecanoyl-sn-glycerol

Figure 20 : Structure de base d'un glycérophospholipide

Figure 21 : Structure de base d'un glycérophosphate (GPA)

Figure 22 : Structure de base d'une glycérophosphocholine (GPC)

Figure 23 : Structure de base d'une glycérophosphoéthanolamine (GPE)

Figure 24 : Structure des GPE sous forme diacyl, alkylacyl et alcèylacyl (A, B, C)

Figure 25 : Structure d'un plasmalogène ou R^1 et R^2 sont des chaines d'acides gras

Figure 26 : Structure de base d'un glycérophosphoglycérol (GPG)

Figure 27 : Structure de base d'un glycérophosphoionisitol (GPI)

Figure 28 : Structure de base d'un glycérophosphoinositol-3,5-biphosphate (PIP_2)

Figure 29 : Structure de base d'une glycérophosphatidylsérine (GPS)

Figure 30 : Structure de la sphingosine la plus répandue chez les mammifères

Figure 31 : Structure de la sphinganine la plus répandue chez les mammifères

Figure 32 : Exemple de structure d'un céramide (où R_2 est une chaine alkyle)

Figure 33 : Structure d'une sphingomyéline (où R_2 est une chaine alkyle)

Figure 34 : Exemple de structure d'un galactosylcéramide, composé de la famille des glycosphingolipides

Figure 35 : Structure du squelette cyclopentanophénantrène des stérols

Figure 36 : Structure du cholestérol : cholest-5-en-3-ol

Figure 37 : Structure de différentes hormones stéroïdiennes : l'estradiol (A), la testostérone (B) et la progestérone (C)

Figure 38 : Structure du cholécalciférol (vitamine D_3)

Figure 39 : Structure du 7-α-hydroxycholestérol, précurseur des acides biliaires

Figure 40 : Structure de deux acides billiaires majoritaires : l'acide cholique $C_{24}H_{40}O_5$ (A) et de l'acide chénodésoxycholique $C_{24}H_{40}O_4$ (B), et de de l'acide α-muricholique $C_{24}H_{40}O_4$ (C)

Figure 41 : Structure de l'acide glycocholique (A) et de l'acide taurocholique (B)

Figure 42 : Structure de l'isopentényl-pyrophosphate et du diméthylallyl-pyrophosphate

Figure 43 : Structure d'un prénol, composé issu de la sous famille des isoprénoïdes : le β-caroténe

Figure 44 : Structure d'un prénol issus de la sous famille des quinones : la ménaquinone, ou vitamine K2

Figure 45 : Représentation schématique des grands phyla des bactéries

Figure 46 : Gradient de quantités en bactéries dans le tractus digestif humain

Figure 47 : Acide α-mycolique, acide gras retrouvé chez les mycobactéries

Figure 48 : Schéma des principales classes d'acides gras retrouvés chez les bactéries

Figure 49 : Structure d'un acide gras diméthyl acétal (ou R^1 est une chaîne alkyle de longueur variable)

Figure 50 : Dosage relatif des acides gras totaux dans les deux souches bactériennes utilisées dans notre étude

Figure 51 : Exemple de saccharolipide, le lipide A

Figure 52 : Structure du lanostérol

Figure 53 : Structure du squalène

Figure 54 : Structures de a) l'acide désoxycolique et de b) l'acide lithocholique, les deux acides biliaires secondaires majoritaires

Figure 55 : Structures de a) l'acide ursodésoxycholique et de b) l'acide chénodésoxycholique

Figure 56 : Structure des acides gras à courte chaine, acide formique CH_2O_2 (A), de l'acide acétique $C_2H_4O_2$ (B), de l'acide propionique $C_3H_6O_2$ (C), de l'acide isobutyrique $C_4H_8O_2$ (D), Structure de l'acide butyrique $C_4H_8O_2$ (E), de l'acide iso-valérique $C_5H_{10}O_2$ (F), et de l'acide valérique $C_5H_{10}O_2$ (G)

Figure 57 : Exemple d'un acide gras à courte chaine estérifié par un acide gras hydroxylé

Figure 58 : Structure de l'acide 3-hydroxyoctadecanoïque

Figure 59 : Structure de l'acide 10-oxo-octadecanoic, KetoB

Figure 60 : Structure d'un commendamide

Figure 61 : Structure d'un exemple de polycétide : l'érythromycine

Figure 62 : Structure de la surfactine ($C_{53}H_{93}N_7O_{13}$), lipopeptide produit par les bacilles

Figure 63 : Structure du C12AsnGaba ($C_{20}H_{37}N_3O_5$), lipopeptide produit par EcN 1917

Figure 64 : Structure du C14Asn ($C_{17}H_{31}N_2O_4$), lipoaminoacide produit par EcN 1917

Figure 65 : Gamme de polarité de certains lipides, et de quelques solvants utilisés en chromatographie liquide, placés sur l'échelle de polarité de Hildebrand

Figure 66 : Principe d'une extraction sur phase solide

Figure 67 : Principales types de phases adsorbantes retrouvés dans les protocoles de lipidomique

Figure 68 : Représentation schématique d'un système de chromatographie gazeuse

Figure 69 : Structure d'un particule core-shell

Figure 70 : Comparaison des pics chromatographiques obtenus avec deux colonnes, l'une constituée de la technologie coreshell, l'autre constituée de particules totalement poreuses

Figure 71 : Structure de d'un lipoaminoacide arginine sous forme A) canonique ; B) déprotonée ; C) protonée et enfin zwiterionique (sous forme négative en D, et positive en E)

Figure 72 : Représentation schématique des éléments constituant le spectromètre de masse

Figure 73 : A) Représentation schématique de l'ionisation verticale, selon le principe de Franck-Condon et B) le spectre photo-électronique du bromobenzène

Figure 74 : Représentation schématique du chemin reactionnel de Brauman

Figure 75 : Principe de l'électronébulisation et cône de Taylor d'après *Schröder, E. Massenspektrometrie - Begriffe und Definitionen. (1991).*

Figure 76 : Représentation schématique d'un quadripôle vue latérale[(a)] et vue frontale[(b)]

Figure 77 : A) Domaines de stabilité des ions m_1, m_2, m_3 et m_4 dans les diagrammes $U=f(V_{RF})$, et B) le spectre de masse associé.

Figure 78 : A) Domaine de stabilité ions de masse croissante de m_1 à m_4 sur l'axe V_{RF} sachant que la tension continue est nulle, et B) Position de ces ions sur l'axe q_z ou la profondeur des puits sont schématisés par des verres.

Figure 79 : Principaux composants du système orthogonal TOF (a) linéaire spectromètre de masse en mode basse résolution, (b) spectromètre de masse en mode haute résolution

Figure 80 : Coupe longitudinale permettant la visualisation des électrodes d'un détecteur Orbitrap

Figure 81 : Représentation des vagues successives produites par le champ électrique oscillant, et du mouvement des ions dans la cellule de mobilité ionique TWIMS

Figure 82 : Schéma de la cellule de mobilité cyclique (waters)

Figure 83 : Représentation schématique d'un analyseur en tandem triple quadripôle

Figure 84 : Représentation schématique du mode d'acquisition SRM, sur un analyseur triple quadripôle

Figure 85 : Représentation schématique du mode de *balayage « Precursor ion scan »*, sur un spectromètre de masse triple quadripôle

Figure 86 : Représentation schématique d'un Q-TOF Xevo G2-XS Waters et de sa cellule de collision segmentée XS

Figure 87 : Représentation schématique d'un spectromètre de masse Q-Exactive Plus Orbitrap

Figure 88 : Représentation schématique d'un spectromètre de masse VELOS couplé à un analyseur Orbitrap

Figure 89 : Nomenclature mise en place spécifiquement pour les lipoaminoacides

Figure 90 : Structure du C13Asn

Figure 91 : Structure de certains LpAA identifiés à l'aide du Workflow lipidomique développé

Figure 92 : Schéma récapitulatif du workflow analytique développé pour la recherche de nouveaux lipides bactériens, produits par le microbiote intestinal

Figure 93 : Annotation des ions fragments produits à partir du C16Glutamate

Figure 94 : Structure du C16OMeGlu

Figure 95 : Annotation des ions fragments produits à partir du C12Glutamate

Figure 96 : Courbes ERMS issus de l'anion précurseur [C12Glu-H]- en a), et des ions fragments m/z 310 en b), m/z 199 en c) et m/z 128 en d).

Figure 97 : Profils d'évolution de l'abondance relative de l'ion fragment m/z 128 issu du précurseur $[C12Glu-H]^-$ en mobilité ionique en fonction de l'énergie de collision

Figure 98 : Profil d'évolution de l'abondance relative de l'ion fragment m/z 199 issu du précurseur $[C12Glu-H]^-$ en mobilité ionique en fonction de l'énergie de collision

Figure 99 : Positions possibles de la charge sur les LpAA où la chaine latérale de l'aminoacide conjugué est non-polaire, menant soit à la formation d'un cétène, soit à la perte d'un groupement CO_2

Figure 100 : Structure des ions m/z 128 et m/z 255 et m/z 199 produits en compétitions à partir du $[C16Glu-H]^-$ et du $[C12Glu-H]^-$

Figure 101 : Profils d'évolutions de l'abondance des ions fragments m/z B) 102, C) m/z 310 et D) m/z 198 en fonction de l'énergie de collision à partir du précurseur $[C12Glu-H]^-$ m/z 328 sous activation résonante (QqTOF Waters Xevo G2-XS)

Figure 102 : Profils d'évolutions de l'abondance des ions fragments m/z 310, m/z 128, m/z 199 et m/z 284 fonction de l'énergie de collision à partir du précurseur $[C12Glu-H]^-$ m/z 328 sous activation non-résonante (LTQ, Thermo)

Figure 103 : Profils ERMS calculés en fonction de l'abondance absolue des ions (m/z 199 et m/z 128, A et C), en comparaison avec les courbes ERMS normalisés par le courant ionique total pour les ions m/z 199 et m/z 128 en B et D à partir du précurseur $[C12Glu-H]^-$ m/z 328 sous activation non-résonante (QqTOF Waters)

Liste des tableaux et des schémas

Tableau 1 : Structure de certains acides gras saturés

Schéma 1: Mécanismes possibles de fragmentation par pertes directes compétitives (a) très mineure de CO_2 et (b) majeure d'H_2O qui est suivie de pertes consécutives (c) mineures de (H_2O+CO_2) et (H_2O+CO_2+HCN)

Schéma 2 : Mécanisme régiosélectif proposé pour la formation des ions complémentaires ions [*Lp*b+O]$^-$ et [y-H_2O]$^-$

Schéma 3 : Mécanismes possibles de fragmentation par pertes directes compétitives A) majeure de CO_2 et B) mineure d'H_2O à partir de [C12AA-H]$^-$ (avec AA=Val, Leu et Ile)

Schéma 4 : Proposition de mécanisme de fragmentation particulier de [C14Phe-H]$^-$, m/z 356, menant aux ions fragments m/z 226 et m/z 147 complémentaires.

Schéma 5 : Proposition de mécanisme de fragmentation particulier de l'ion [C12Trp-H]$^-$, menant au fragment de m/z 256, où R1 correspond à une chaîne alkyle $C_{10}H_{21}$de la partie acide gras

Identification et caractérisation de nouveaux lipopeptides bactériens par spectrométrie de masse

INTRODUCTION

PARTIE 1 : LIPIDES ET MICROBIOTE 7

Chapitre I : Les différentes classes de lipides, et principales fonctions associées 9

Chapitre II : Le microbiote intestinal et lipides associés 39

PARTIE 2 : ANALYSES LIPIDOMIQUE ET SPECTROMETRIE DE MASSE 65

Chapitre I : Généralités sur la lipidomique et les sciences du « omics » 67

Chapitre II : L'analyse des lipides 75

Chapitre III : La spectrométrie de masse 85

Chapitre IV : La spectrométrie de masse en lipidomique et les différentes types d'analyses possibles 136

PRESENTATION DES TRAVAUX DE THESE

Objectifs et travaux de thèse 145

Résultats 147

DISCUSSION GENERALE : 213

CONCLUSION ET PERSPECTIVES 231

ANNEXE 233

VALORISATION DES TRAVAUX DE RECHERCHE 253

BIBLIOGRAPHIE 257

INTRODUCTION GENERALE

Les pathologies du tractus digestif représentent à ce jour un défi sanitaire et économique majeur. Un dysfonctionnement de la communication entre le microbiote intestinal et l'hôte peut en être l'origine. L'étude des relations entre hôte-microbiote permettrait d'identifier de nouvelles stratégies thérapeutiques. Le syndrome de l'intestin irritable (SII) est le trouble gastro-intestinal le plus répandu. C'est aujourd'hui la première cause de consultation chez le gastroentérologue, et cette pathologie touche environ 11% de la population mondiale. Malgré les difficultés rencontrées quotidiennement par les patients, il reste à ce jour difficile d'identifier les causes à l'origine du SII, ce qui représente un verrou non négligeable dans le développement de thérapies adaptées en vue d'un traitement médicamenteux efficace. Actuellement, les traitements proposés sont symptomatiques et non efficaces. L'élucidation des mécanismes cellulaires responsables du SII pourraient mener à l'identification de médicaments spécifiques, permettant ainsi de prodiguer un traitement adapté aux patients atteints de ce syndrome particulièrement délétère. C'est dans ce contexte que le sujet de mes travaux de thèse se situent.

Une dysbiose bactérienne a été décrite chez les patients atteints du SII, cependant à ce jour, le rôle qu'elle pourrait avoir en lien avec le SII n'a pas encore été clairement défini. Dans cette optique, des travaux antérieurs ont permis de démontrer que les bactéries *Escherichia coli* Nissle 1917, une souche commensale de bactéries probiotiques, sont capables de synthétiser du GABA lié à un acide aminé et à un acide gras. Ensemble, ces trois composés forment un lipopeptide comme par exemple le C12AsnGABA. La liaison peptidique entre le GABA et l'acide gras permet au GABA de pouvoir passer au travers de la barrière intestinale. Le GABA, qui est le principal neurotransmetteur inhibiteur du système nerveux, peut ensuite se fixer sur son récepteur afin de réguler l'activité des neurones sensitifs, à l'origine de la nociception diminuant ainsi la douleur ressentie par les patients atteints du SII. Cette action bénéfique pour le patient, ne semble pas s'accompagner d'effets néfastes sur la motricité intestinale, ni sur la perméabilité cellulaire.

E. coli n'étant pas la seule bactérie du microbiote intestinale capable de synthétiser du GABA, l'objectif principal de mes travaux de thèse, consistent à étendre l'identification, la caractérisation et la quantification de nouveaux lipopeptides produits par l'ensemble du microbiote intestinal et d'évaluer leurs potentiels thérapeutiques. Une stratégie expérimentale de chimie analytique par chromatographie liquide couplée à la spectrométrie de masse à haute résolution (LC-HRMS) visant à l'identification et la caractérisation de ces lipopeptides a été mise en place. J'introduirai dans un premier temps la très grande famille des lipides dont

font partis les lipopeptides, puis le microbiote intestinal, ses fonctions principales et les lipides qu'il produit. Dans une seconde partie j'introduirai la lipidomique, la spectrométrie de masse et les grands principes de la fragmentation. Une dernière partie sera consacrée à la présentation de mes résultats de thèse.

Sommaire général

INTRODUCTION GENERALE

PARTIE 1 : LIPIDES ET MICROBIOTE 7

Chapitre I : Les différentes classes de lipides, et principales fonctions associées 9

1. Les acides gras et leurs métabolites 10

1.1 Les acides gras : nomenclature et biosynthèse 11

1.2 Les lipides issus du métabolisme des acides gras : les oxylipines 16

2. Les endocannabinoïdes 19

3. Les acides aminés acylés 19

4. Les glycérolipides 21

4.1 Les monoacylglycérols 21

4.2 Les diacyglycérols 22

4.3 Les triacylglycérols 22

5. Les glycérophospholipides 23

5.1 Les glycérophosphates 24

5.2 Les glycérophosphocholines 25

5.3 Les glycérophosphoéthanolamines 25

5.4 Les glycérophosphoglycérols 27

5.5 Les glycérophosphoinositols 27

5.6 Les glycérophosphosérines 28

6. Les sphingolipides 29

6.1 Les céramides 30

6.2 Les sphingomyélines 31

6.3 Les glycosphingolipides 32

7. Les stérols 33

7.1 Les stérols libres 33

7.2 Les stéroïdes 34

7.3 Les vitamines D 35

7.4 Les acides biliaires 36

8. Les prénols 38

Chapitre II : Le microbiote intestinal et lipides associés 39
1. Définitions et notions essentielles 39
2. Origine et composition du microbiote intestinal 40
2.1 Chez l'enfant 40
2.2 Chez l'adulte 41
3. Bactéries utilisées pendant mes travaux de thèse 43
4. Les fonctions du microbiote intestinal 45
5. Des modifications de la composition du microbiote intestinal peuvent être associées à des pathologies ou à désordres métaboliques 47
6. Les lipides bactériens 49
6.1. Les lipides structuraux 49
6.1.1 Généralités 50
6.1.2 Les glycérophospholipides 52
6.1.3 Les lipides à longue chaine et les lipides complexes 52
6.2. Les lipides issus du métabolisme du microbiote intestinal 54
6.2.1 Les acides biliaires secondaires et les dérivés du cholestérol 55
6.2.2 Les acides gras à courtes chaines 57
6.2.3 Les acides gras à longues chaines 59
6.2.4 Les commendamides 60
6.2.5 Les polycétides 61
6.2.6 Les lipopeptides et lipoaminoacides 62

PARTIE 2 : ANALYSES LIPIDOMIQUE ET SPECTROMETRIE DE MASSE 65
Chapitre I : Généralités sur la lipidomique et les sciences du « omiques » 67
1. Préparation des échantillons 68
1.1. L'extraction des lipides 69
1.1.1 L'extraction liquide-liquide 70
1.1.2 L'extraction en phase solide 71
2. L'influence de la matrice 74

Chapitre II : L'analyse des lipides 75
1. Les méthodes de séparation des lipides 76
1.1 La chromatographie en phase gazeuse 76
1.2 La chromatographie liquide 78
1.2.1. La chromatographie liquide à polarité de phase normale 81
1.2.2. La chromatographie liquide à polarité de phase inverse 81

1.2.3. Développement des systèmes chromatographiques et dernières innovations en colonnes chromatographiques 83

Chapitre III : La spectrométrie de masse 85

1. Un historique de la spectrométrie de masse 86

2. Généralités sur la spectrométrie de masse 89

3. Généralités sur les principaux constituants d'un spectrométre de masse 93

4. Les sources d'ionisation 95

4.1. Ionisation en phase gazeuse sous vide 95

4.2.1. Ionisation à pression atmosphérique 101

4.2.2 Ionisation par désorption à pression atmosphérique 103

5. Les différents analyseurs en spectrométrie de masse 105

5.1. Analyseur basse résolution avec balayage de tension 107

5.1.1 Le quadripôle : analyseur de faisceaux d'ions de faible énergie cinétique 107

5.1.2 Le piège à ions linéaire : analyseur par stockage et éjection sélective d'ions 109

5.1.3 L'analyseur temps-de-vol : faisceaux d'ions de haute énergie cinétique, et analyses sans balayage de champs 111

5.1.4 L'analyseur Orbitrap 113

5.1.5 L'analyseur de mobilité ionique 113

5.1.5 L'analyseur FT-ICR 113

5.2. Les analyseurs en tandem 123

5.2.1 Le triple quadripôle 126

5.2.2 L'hybride quadripôle – temps-de-vol (QqTOF) : analyse en haute résolution des ions produits 130

5.2.3 Les analyseurs Orbitrap en tandem : le Q-Exactive Plus et l'Electrospray Linear Trap Quadrupole Orbitrap VELOS 132

5.2.4 Séquence de balayage pour les expériences MS^n à l'aide d'un piège à ions 134

Chapitre IV : La spectrométrie de masse en lipidomique et les différentes types d'analyses possibles 136

1. Les analyses globales 136

2. Les analyses semi-ciblées 1377

3. Les analyses ciblées 138

4. La quantification en lipidomique 138

5. Le niveau d'identification en lipidomique 139

6. Les règles de chimie des réactions ions-molécules appliquées pendant ces travaux de thèse et nomenclature 141

PRESENTATION DES TRAVAUX DE THESE

Objectifs et travaux de thèse **145**
Résultats **147**
Publication 1 147
Publication 2 169
Publication 3 191

DISCUSSION GENERALE : **213**
CONCLUSION ET PERSPECTIVES **231**
ANNEXE 1 : La théorie du quasi-équilibre **233**
ANNEXE 2 : Supplementary information publication 1 ..... **237**
ANNEXE 3 : Supplementary information publication 2 ..... **251**

VALORISATION DES TRAVAUX DE RECHERCHE **253**
CURRICULUM VITAE **255**
BIBLIOGRAPHIE **257**

INTRODUCTION
PARTIE 1 : LIPIDES ET MICROBIOTE

Chapitre I : Les différentes classes de lipides, et principales fonctions associées

Les lipides constituent une vaste et importante famille de composés chimiques et sont retrouvés chez tous les organismes vivants[1-2]. Le nombre de métabolites est estimé à plus de 100 000 espèces chez les mammifères, dont 1 000 appartiennent à la grande famille des lipides[1]. Ils sont généralement définis comme étant des espèces chimiques insolubles dans l'eau et solubles dans les solvants organiques[2]. Cependant, cette définition limitée aux propriétés de solubilité des lipides tend à exclure certains lipides tel que les glycolipides et les isoprénoïdes, qui sont des lipides composés d'une partie hydrophile[3,4]. Pour donner une définition permettant de prendre en compte toutes les familles de lipides, il faut considérer que ce sont des composés de petites tailles moléculaires, hydrophobes ou amphiphiles et produits via la condensation des carbanions de thioesters et/ou de la condensation d'unités isoprènes C_5H_8 (Figure 1). Selon qu'ils soient hydrophobes ou amphiphiles, les lipides peuvent être plus ou moins polaires.

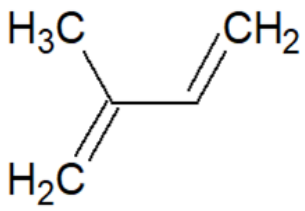


Figure 1 : Unité isoprène, composition élémentaire C_5H_8

Les lipides, constituent avec les glucides et les protéines les trois grandes familles de macronutriments, c'est-à-dire l'un des constituants des aliments contribuant à l'apport énergique nécessaire aux mammifères sous forme de calories. Les lipides jouent plusieurs rôles majeurs dans l'organisme, dans un premier temps ils permettent un stockage d'énergie, sous forme de triglycérides, notamment dans les tissus adipeux. Ils sont impliqués dans le transport de vitamines ou encore dans la transduction du signal. Ils jouent également le rôle de précurseurs pour les vitamines ou les prostaglandines par exemple.

Dans ce premier chapitre seuls les lipides non bactériens seront abordés, le prochain sera consacré aux lipides bactériens, bien qu'un grand nombre de lipides soient communs aux deux règnes.

La diversité chimique et structurale des lipides nécessite la mise en place d'une classification particulière[4]. Celle-ci permet d'inclure tous les lipides existants. Le consortium LIPID-Metabolites And Pathways Strategy (LIPID-MAPS®) définit et classe les lipides en fonction de leur biosynthèse.

Ils sont classés en 8 groupes sur la base de leurs origines biosynthétiques et donc de leurs différentes structures. On y retrouve : le groupe des acides gras, le groupe des glycérolipides, des glycérophospholipides, des sphingolipides, des stérols, des prénols, des saccharolipides et le groupe des polycétides. Ces 8 groupes de lipides peuvent être sous classés en deux types de lipides : les lipides polaires, comprenant les glycérophospholipides, les sphingolipides, les saccharolipides et les polycétides, et le groupe des lipides apolaires : les acides gras, les glycérolipides, les stérols et les prénols. L'ensemble des lipides est divisé en classes, sous-classes et espèces moléculaires[5].

1. Les acides gras et leurs métabolites

Les acides gras (AG) sont des acides carboxyliques (Figure 2) pouvant être sous forme libre, d'ester ou d'amide[4].

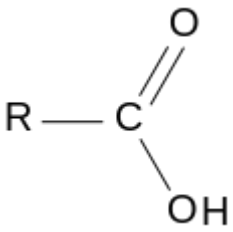


Figure 2 : Structure de la fonction carboxylique, avec R chaine alkyle

Le radical R de l'acide carboxylique est une chaîne aliphatique carbonée de longueur variable, généralement de quatre à plus de trente atomes de carbone. La chaîne carbonée d'un acide gras est considérée comme longue lorsqu'elle comporte entre 12 et 20 atomes de carbone, et très longue lorsqu'elle dépasse les 20 atomes de carbone[6]. Chez les mammifères, généralement, les acides gras sont composés d'une chaîne linéaire, avec un nombre pair d'atomes de carbone. Cette chaîne carbonée peut être saturée ou insaturée, elle peut contenir jusqu'à six doubles liaisons. Des modifications peuvent avoir lieu, comme une élongation de la chaine, une hydroxylation, ou encore une oxydation. Une nomenclature a été mise en place afin de décrire la structure de l'acide gras sous la forme Cx:y ou x correspond au nombre d'atomes de carbone et y au nombre d'insaturations.

La position des insaturations peut être notée selon deux écoles : la première en conformité avec la nomenclature internationale de chimie se note "$\Delta^{a, b, ..., n}$" où chaque lettre indique la position des insaturations par rapport à la fonction acide carboxylique. La seconde façon de noter les insaturation correspond à la nomenclature biochimique et se note "ω : n", où n correspond à la position de la première insaturation en partant du méthyl terminal.

En outre, la présence de doubles liaisons fait apparaître la notion d'isomères de configuration. L'isomérie *trans* et *cis*, est basée quant à elle sur l'encombrement stérique des groupements fonctionnels (Figure 3). La molécule est dite en configuration *cis* lorsque les groupements se trouvent du même côté de la double liaison. La molécule est dite *trans* lorsque ses substituants se trouvent de part et d'autre de la double liaison. La conformation *cis* est retrouvée chez les mamifères. Les acides gras insaturés présentent généralement des doubles liaisons de configuration *cis*. Cependant, il s'avère que celles-ci peuvent être de configuration *trans* dans de faibles proportions.



Composé *cis* Composé *trans*

Figure 3 : Différence entre des composé *trans* et *cis*

1.1.Les acides gras : nomenclature et biosynthèse

L'acide gras le plus représenté dans les cellules du règne animal est l'acide palmitique (C16:0 ; Figure 4), il est le précurseur d'autres acides gras, par différentes transformations (allongement de la chaine, hydrogénation de la chaine, ou β-oxydation)[7], représentés dans le **Tableau 1**[6,8].

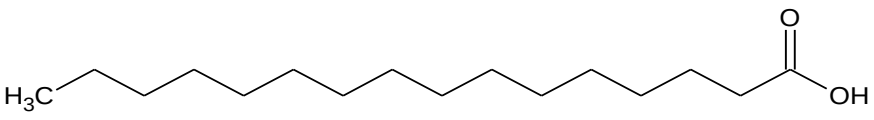


Figure 4 : Structure de l'acide palmitique $C_{16}H_{32}O_2$

La synthèse des acides gras monoinsaturés à moyennes et longues chaînes est initiée par la formation de l'acide palmitique, celui-ci est obtenu par l'action successive de deux enzymes l'acetyl-coA carboxylase (ACC) et la fatty acid synthase (FAS)[9]. L'ACC catalyse la première étape de la biosynthèse des acides gras. Elle mène à la conversion de l'acétyl-CoA en malonyl-CoA, qui est un intermédiare métabolique participant à la biosynthèse des acides gras. Le malonyl-CoA est pris en charge par la FAS. Cette enzyme joue divers rôles dans la synthèse des AG via des activités acétyl-CoA transferase, malonyl-CoA transferase, β-cétoacyl synthase / réductase, β-hydroxyacyl dehydratase, ect … Le nombre de cycles de la réaction enzymatique peut varier de deux à huit, ce qui permet d'aboutir à l'ensemble des acides gras intermédiaires entre les acides gras à courte chaine à partir de l'acide butyrique à l'acide stéarique[9].

Certains AG sont dits essentiels, comme l'acide linoléique, C18:2 $\Delta^{(9,12)}$ (Figure 5), et l'acide α-linolénique C18:3 $\Delta^{(9,12,15)}$ puisqu'ils ne peuvent être apportés que par l'alimentation. Ces acides gras sont principalement retrouvés dans les huiles végétales ou dans certains poissons[9].

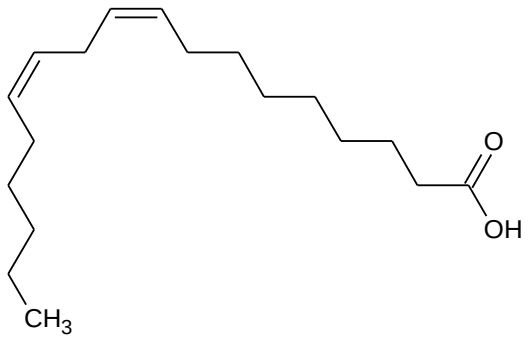


Figure 5 : Structure de l'acide linoléique, C18:2 $\Delta(^{9,12})$, $C_{18}H_{32}O_2$

Chez l'homme la synthèse des AG se déroule principalement dans les tissus lipogéniques, comme les tissus adipeux, les glandes mammaires, les reins et le foie. Dans certaines maladies, une lipogénèse accrue est observée dans les tissus touchés par la pathologie. La synthèse des acides gras insaturés à longues et à très longues chaînes est initiée par des précurseurs tels que les acétyl-CoA et le malonyl-CoA[4] (Figure 6).

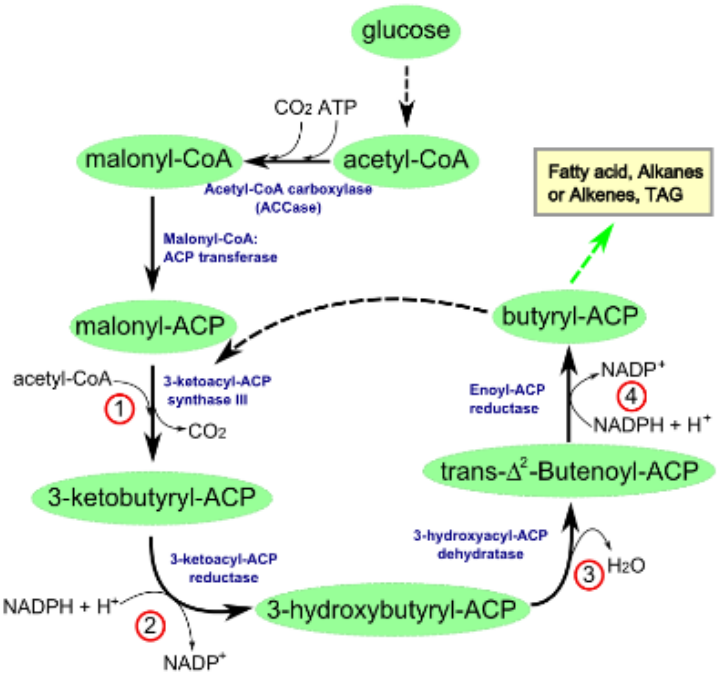


Figure 6 : Schéma de la biosynthèse des acides gras[10]

Puisque les mammifères ne sont pas capables de produire les acides α-linolénique et linonéique, la seule source de ces précurseurs est située dans l'alimentation. Ces AG sont alors pris en charge par la machinerie enzymatique[11] (Figure 7) permettant la production des différents AGPI de notre organisme[12].

Par voie enzymatique, l'acide linoléique (C18:2 n-6) peut être converti en acide γ-linolénique (C18:3 n-6) par une Δ6 désaturase[13]. L'élongation de cet acide par une élongase conduit à la formation de l'acide dihomo-linonélique (C20:3 n-6). L'acide néo-formé est modifié par une Δ5 désaturase conduisant à la formation de l'acide arachidonique (C20:4 n-6)[14].

L'acide α-linolénique (C18:3 n-3) peut être converti par une Δ6 désaturase en acide stéaridonique (C18:4 n-3) puis par une élongation à l'acide eicosatetraénoique (C20:4 n-3). Celui-ci, peut alors être substrat d'une Δ5 désaturase conduisant à l'acide eicosapentaénoique (C20:5 n-3), converti ensuite en acide docosapentaénoique (C22:5 n-3) sous l'action d'une élongase. Ce dernier est converti en acide tetracosapentaénoique (C24:5 n-3), qui peut alors subir une Δ6 désaturation pour former l'acide tetracosahexaénoique (C24:6 n-3)[15].

L'acide arachidonique et l'acide eicosapentanoïque sont les précurseurs dans l'organisme de la classe des eicosanoïdes, ayant de nombreuses fonctions notamment dans le processus d'inflammation, ou dans certaines maladies comme le diabète et les maladies cardio-vasculaires[16,17,18].

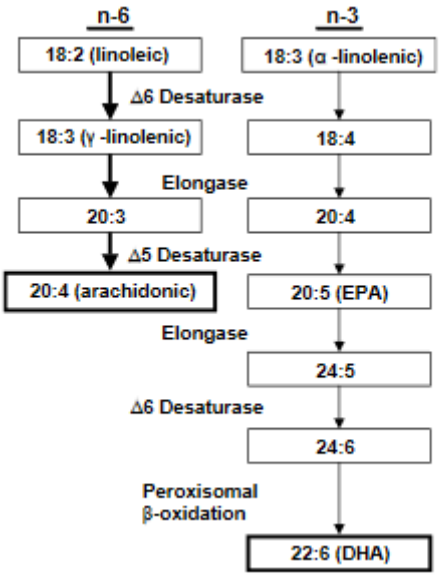
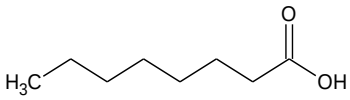
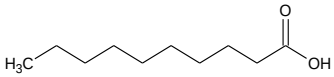
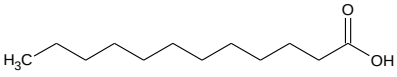
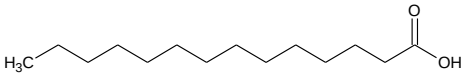
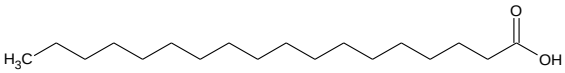
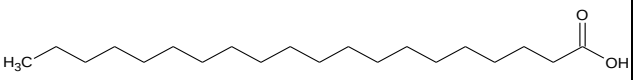
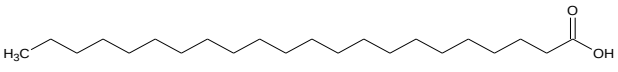


Figure 7 : Machinerie enzymatique menant aux différents acides gras par N. T. Nakamura et al[19]

Tableau 1 : Structure de certains acides gras saturés

Nom vernanculaire **Nom scientifique**	**Composition élémentaire** **Masse moléculaire**	**Structure de la molécule**
Acide caprylique Acide octanoïque	$C_8H_{16}O_2$ MW : 144.2114	
Acide caprique Acide décanoïque	$C_{10}H_{20}O_2$ MW : 172.2646	
Acide laurique Acide dodécanoïque	$C_{12}H_{24}O_2$ MW : 200.3178	
Acide myristique Acide tétradécanoïque	$C_{14}H_{28}O_2$ MW : 284.4772	
Acide stéarique Acide octadécanoïque	$C_{18}H_{36}O_2$ MW : 284.4772	
Acide arachidique Acide eicosanoïque	$C_{20}H_{40}O_2$ MW : 312.5304	
Acide behénique Acide docosanoïque	$C_{22}H_{44}O_2$ MW : 340.5836	

1.2 Les lipides issus du métabolisme des acides gras : les oxylipines

Les oxylipines représentent une très grande famille de composés, ce sont des lipides issus du métabolisme des acides gras polyinsaturés, et en particulier des acides arachidoniques et eicosapentaénoïques (Figure 8A et 8B). Ce sont des composés constitués de 18 à 20 atomes de carbone, issus de "l'arachidonate cascade" (Figure 9) [20,21], et résultent de différents processus d'oxydation. Cependant, chez les mammifères les oxylipines peuvent également être synthétisées à partir d'autres acides gras comme l'acide linoléique (C18:2, ω-6), l'acide-α linolénique (C18:3, ω-3), l'acide-γlinolénique (C18:3, ω-6), l'acide dihomo-γ linolénique ou eicosatriénoique (C20:3, ω-6), l'acide eicosapentaénoique (C20:5, ω-3), l'acide docosatétraénoique (C22:4, ω-6) ou l'acide docosahexaénoïque (C22:6, ω-3).

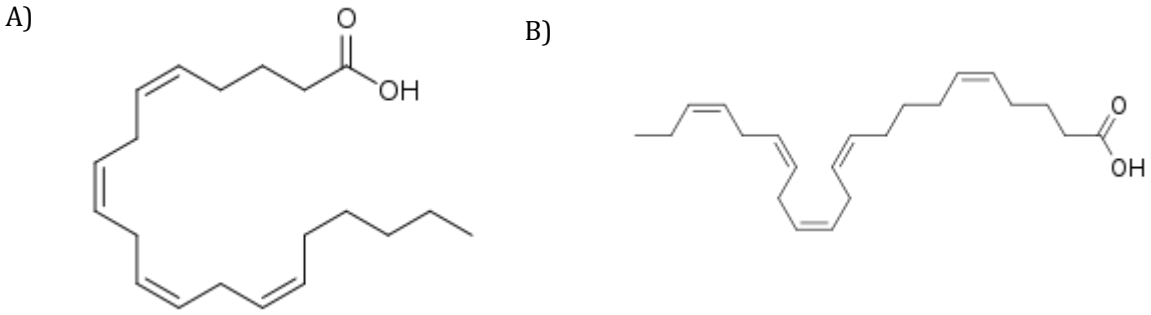


Figure 8: Structure des acides arachidonique (A) et eicosapentaénoïque (B)

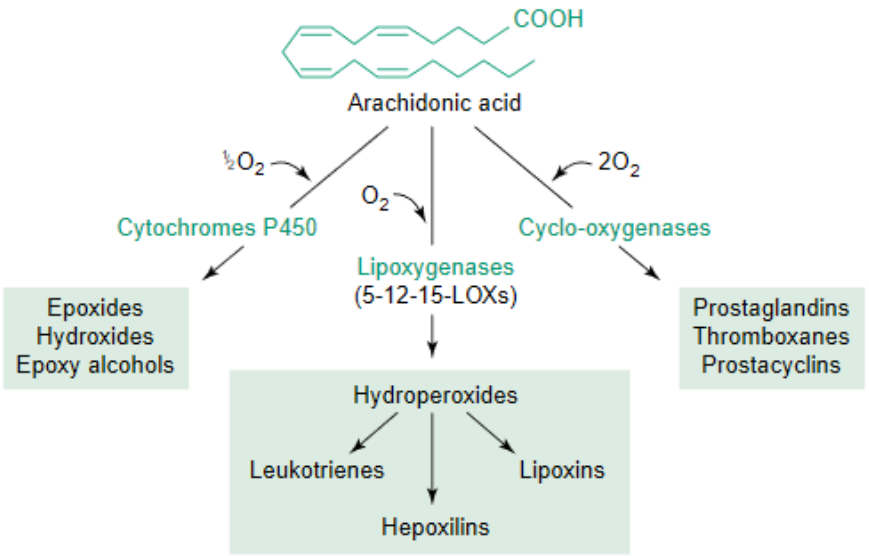


Figure 9 : Métabolisme oxydatif de l'acide arachidonique chez les mammifères, d'après Blée 2002[21]

Parmi les oxylipines se trouvent grand nombre de composés, qui sont divisés en 3 sous-classes en fonction de leur voie de biosynthèse enzymatique[20]: la voie des cyclooxygenase (COX), la voie des lipoxygenases (LO) et la voie des cytochromes P450 pouvant être époxygenase et/ou hydrolase[20]. Ces différentes sous-classes regroupent notamment les thromboxanes, les prostaglandines, les leucotriènes, les acides epoxyeicosatriénoïques (EET) et les hydroxyeicosatetraenoiques (HETE).

Les tromboxanes et les prostaglandines sont des eicosanoïdes issues de la voie de biosynthèse de la cyclo-oxygénase. Les COXs introduisent deux molécules de dioxygène sur l'acide arachidonique, conduisant à la synthèse des prostaglandines (Figure 10). Les prostaglandines, découvertes en 1935 par U. V. Euler, S. Bergstrom et B. Samuelson sont les précurseurs des tromboxanes. Les métabolites issus de ces enzymes COXs font partis de la famille des prostanoides. Structuralement, ce sont des métabolites proches de l'acide arachidonique : ce sont donc des composés à 20 atomes de carbone, possédant au minimum une insaturation ainsi qu'un cyclopentane entre les atomes 13 et 14 de carbone. On note également la presence d'un groupement hydroxyle en position C15 sur ces métabolites[20]. Les prostaglandines jouent des rôles physiologiques essentiels dans l'homéostasie : elles sont à l'origine de la vasoconstriction, de la vasodilatation, mais également de la migration et de la proliferation lymphocytaire.

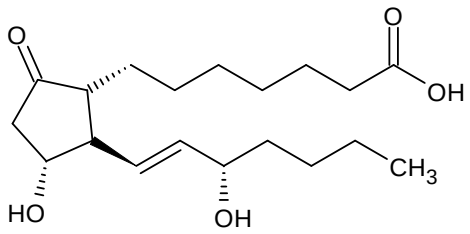


Figure 10 : Structure de la prostaglandine 1

Les leucotriènes sont issus de la voie enzymatique des lipoxygenases. Ils sont produits par l'action de la 5-lipoxygenase (5-LO), enzyme introduisant une molécule d'oxygène, formant ainsi les leukotriènes, les lipoxines ou les hepoxillines, comme par exemple la molécule de leukotriène B4 (LTB4, Figure 11), elle-même précurseur d'autres leucotriènes. Structuralement ce sont des métabolites à 20 atomes de carbone, comprenant trois doubles liaisons conjuguées, mais ne présentant pas de cycles. Les activités pharmacologiques des leucotriènes sont nombreuses, notamment en ce qui concerne la perméabilité vasculaire, l'activation de l'inflammation, notamment *via* la régulation des neutrophiles, ou encore la vasoconstriction[20].

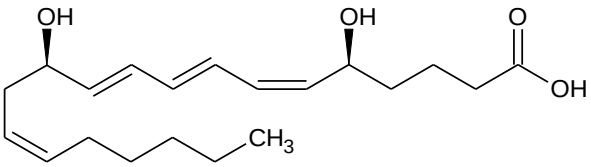


Figure 11 : Structure du leukotriène LTB4

Enfin, la voie des cytochromes P450 pouvant être époxygénase et/ou hydrolases, qui peut permettre la formation de ponts époxy entre deux carbone comme le 5,6-EET. Les HETE sont formés à partir de l'acide arachidonique. Structuralement, les HETE sont des métabolites à 20 atomes de carbone, composés généralement de 3 à 4 doubles liaisons carbone-carbone. Les métabolites issus de la voie des CYP les plus connus sont les 16-HETE, le 19-HETE et le 20-HETE (Figure 12). Ce sont des intermédiaires du métabolisme de l'acide arachidonique. Cette sous classe d'oxylipines va intervenir au niveau du transport de protéines membranaires[20,22].

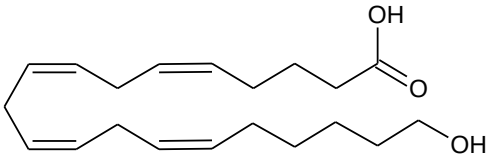


Figure 12 : Structure du 20-HETE

De manière générale, les oxylipines sont des métabolites comprenant 20 atomes de carbone. Ces molécules sont thermosensibles et très réactives, elles ont donc une demi-vie extrêmement courte, de l'ordre de quelques minutes[23].

Chez les mammifères, les oxylipines jouent un rôle essentiel dans de nombreux processus physiologiques comme dans la régulation du système cardiaque, rénal ou vasculaire. Ces composés sont également impliqués dans le développement maladies chroniques, lors du dérèglement de leurs voies de biosynthèse. Elles sont notamment impliquées dans l'obésité, l'inflammation, certains cancers et certaines maladies neurodégénératives et cardiaques. Ces oxylipines peuvent être utilisées comme biomarqueurs, permettant de diagnostiquer précocement des maladies[24,25,26] et leurs enzymes de biosynthèse constituent des cibles thérapeutiques. L'aspirine ou l'ibuprofène sont des médicaments inhibant la cyclooxygénase 2, permettant le bloquage de la production de PG lors de réactions inflammatoires[27].

2. Les endocannabinoïdes

Les endocannabinoïdes, comme le N-arachidonoyl-éthanolamine couramment appelés anandamide, est un neurotransmetteur lipidique synthétisé et libéré par les neurones au niveau du cerveau. Ce composé peut être généré par plusieurs voies de biosynthèse et de dégradation. L'anandamide (AEA, Figure 13) est le composé majoritaire de cette famille. La voie de synthèse fait intervenir la N-acyltransférase (NAT), qui permet un transfert de l'acide arachidonique (Tableau 1) sur une phosphatidyléthanolamine (PE, composé présenté au point 4.3). Ce transfert génère un N-arachidonoylphosphatidyléthanolamine, qui va être ensuite clivé *via* une phospholipase D, générant ainsi l'anandamide[28]. Ces composés agissent spécifiquement sur des récepteurs cannabinoïdes. Ils sont impliqués dans de nombreux processus physiologiques et pathologiques.

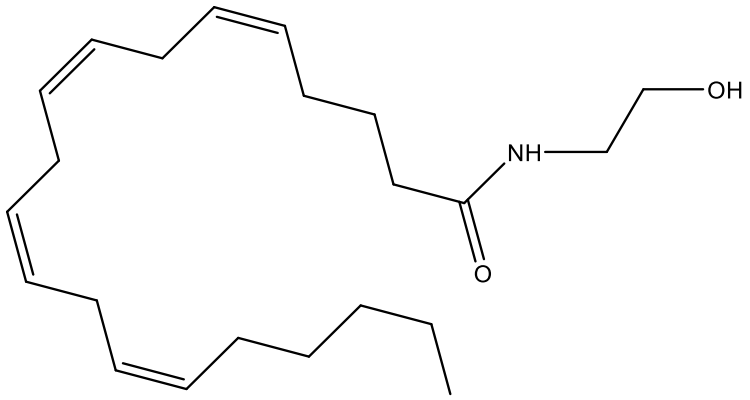


Figure 13 : Structure de l'anandamide (N-araciodonoylean olamide)

3. Les acides aminés acylés

Les acides aminés acylés aussi appelés lipoaminoacides (LpAA) sont des lipides composés de deux entités : une chaine d'acide gras et un acide aminé. Ces structures ont été décrites dans le cerveau de rats par l'équipe de J. Michael Walker[29]. Cinquante molécules ont été décrites, composées des acides aminés alanine, glutamique, histidine, glutamine, isoleucine, leucine, phenylalanine, méthionine, proline, serine, thréonine, tryptophane, tyrosine, valine, arginine ou asparagine.

Ces acides aminés peuvent être liés aux acides gras suivant : les acides oléique, palmitique, stéarique, arachidonique, docohexanoique. L'ensemble de ces acides gras sont couramment retrouvés dans le cerveau. Lorsque ces deux composantes sont associées, la molécule formée est un acide aminé acylé, comme l'acide arachidonique lié au glutamate (Figure 14).

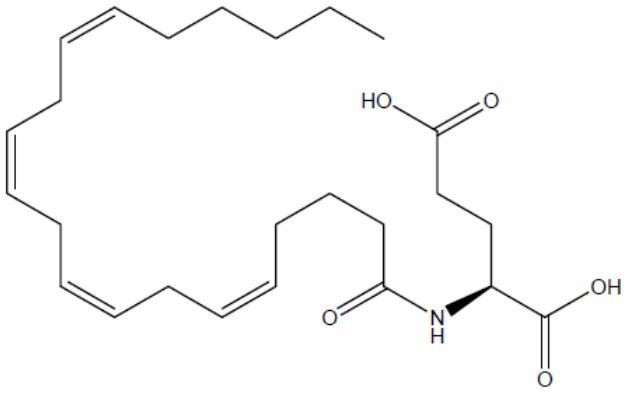


Figure 14 : Structure de l'acide arachidonique lié au glutamate

D'autres types de composés ont également été décrits par cette même équipe, ou un neurotransmetteur le GABA remplace l'acide aminé, permettant ainsi de former un l'acide oléique lié au GABA[29] (Figure 15).

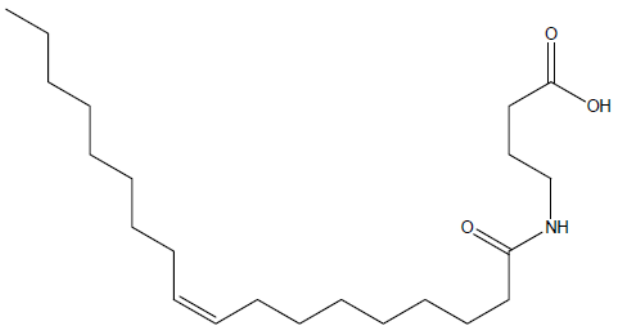


Figure 15 : Structure de l'acide oléique lié au GABA

4. Les glycérolipides

Les glycérolipides, aussi appelés acylglycérols ou glycérides sont des lipides non polaires. Ils sont composés d'un squelette glycérol (Figure 16) estérifié par une, deux ou trois chaînes d'acides gras formant respectivement les 3 classes de glycérolipides : les monoacylglycérols ou monoglycérides (MG), les diacylglycérols ou diglycérides (DG) et les triacylglycérols ou triglycérides (TG). Les structures de base des trois classes de glycérolipides sont présentées au cours des différentes sections de cette sous-partie[4,20].

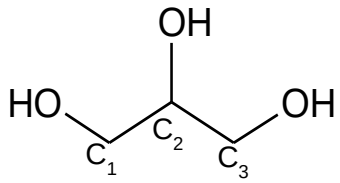


Figure 16 : Structure du squelette glycérol

La nomenclature de ces composés a été définie comme suit : MG(*sn-1/2*), DG(*sn-1/sn-2/3)* et TAG(*sn-1/sn-2/sn-3*), avec : *sn* « stereospecific numbering », qui représente la numérotation stéréospécifique au niveau du glycérol la position estérifiée du glycérol.

Chez les plantes, d'autres glycérolipides sont caractérisés par la présence de résidus de sucres liés par une liaison osidique située entre une des trois fonctions alcool du glycérol et une des fonctions alcool de l'ose. Ces glycérolipides constitués d'un galactose sont des galactoglycérolipides[20].

4.1 Les monoacylglycérols

Les monoacylglycérols sont des glycérolipides estérifiés par un seul acide gras, en position *sn*-1/2 (Figure 17). Ces lipides sont issus de l'hydrolyse des diacylglycérols via l'action d'une enzyme : la diacylglycérol lipase. L'action de cette enzyme va permettre la libération de l'acide gras en position *sn*-2[30]. Les MG contribuent à la réserve de glycérols, c'est donc une classe de lipides essentiels.

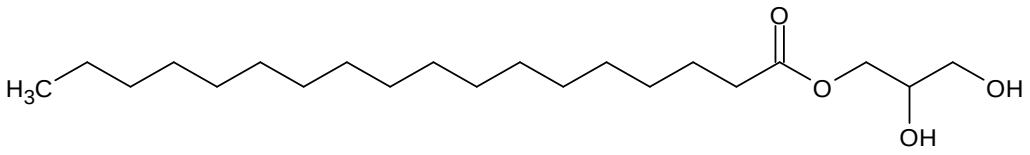


Figure 17 : Structure d'un monoacylglycérol, le MG18:0 (1-octadecanoyl-sn-glycerol)

4.2 Les diacyglycérols

Les diacylglycérols sont issus de l'estérification en position *sn*-1 et *sn*-2 d'un glycérol par deux acides gras (Figure 18). Les DG sont issus du réticulum endoplasmique par hydrolyse des triacylglycérides grâce à l'action d'une enzyme : la triacylglycérol lipase. Cela conduit à la formation d'un DG *via* l'hydrolyse d'un acide gras en position *sn*-3 [30,31]. Par ailleurs, les DG sont également issus des glycérophosphates et peuvent parfois intervenir dans la formation de certains glycérophospholipides tels que les glycérophosphocholines (GPC), les glycérophosphosérines (GPS) …[32]. Les MG et les DG ont des rôles importants dans l'homéostasie de l'hôte, dont certains rôles restent encore méconnus. Certaines études en lipidomique ont démontré l'accumulation de ces composés chez les personnes atteintes d'un déficit cognitif léger ainsi que chez les patients atteints de la maladie d'Alzheimer, où leurs rôles sont encore méconnus[33].

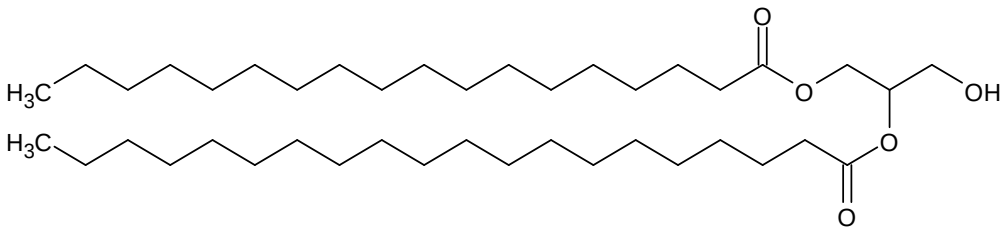


Figure 18 : Structure d'un diacylglycérol, le DG18:0/18:0, 1-octadecanoyl-2-hexadécanoyl-sn-glycerol

4.3 Les triacylglycérols

La synthèse d'un triacylglycéride (Figure 19) se produit en plusieurs étapes. En premier lieu, les glycérophophates dépourvus de leur groupement phosphate vont former des DG par l'action d'une acide phosphatidique phosphatase. Ces DG sont ensuite acylés par l'acyl transférase[30,31]. Les TG peuvent être hydrolysés en DG et en acides gras complémentaires. Dans le plasma, les TG sont les glycérolipides les plus abondants. Ces molécules représentent la réserve majeure d'énergie du corps humain et peuvent être stockées dans les vacuoles adipocytaires des tissus adipeux. Cette énergie stockée peut ensuite être redistribuée quand nécessaire, aux différentes cellules et organes du corps[34,35]. L'énergie produite par les TG est de l'ordre de 6 fois plus importante que celle produite par les glucides. De plus, les TG représentent une réserve d'énergie mobilisable sur plusieurs mois, on parle d'énergie utilisable à long terme via la formation de gouttelettes lipidiques. A l'inverse, les glucides constituent une réserve immédiate d'énergie, mobilisable pendant 24h uniquement. Le tissu adipeux sous-

cutané est composé de TG, qui agissent également comme isolant thermique chez les animaux à sang chauds[36,37].

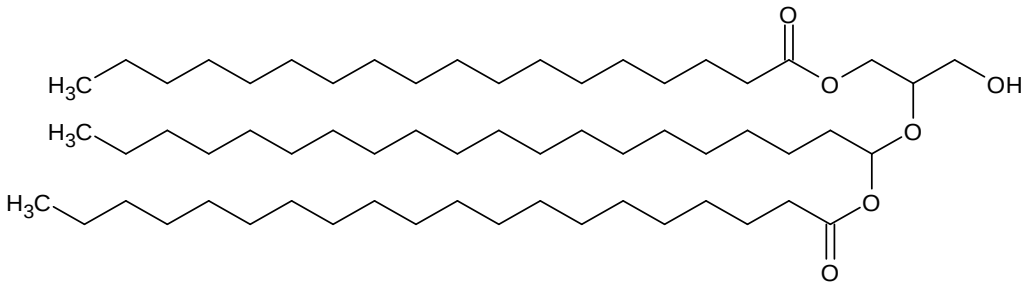


Figure 19 : Structure d'un triacylglycérol, le TG 18:0/18:0/18:0, 1-octadecanoyl-2-octadécanoyl-3- octadecanoyl-sn-glycerol

5. Les glycérophospholipides

Les glycérophospholipides (GPL) représentent la classe majoritaire des lipides retrouvés dans la nature. De par leurs propriétés amphiphiles, ces composés sont les constituants majeurs de la bicouche lipidique des membranes cellulaires. Ces GPL sont des sites de liaisons pour les protéines intra et inter cellulaires[4,20,38].

Les GPL sont des lipides impliqués dans certaines pathologies neurodégénératives comme la maladie d'Alzheimer[39]. En effet, ils sont des constituants des membranes neuronales, assurant le maintien d'une fluidité et d'une perméabilité. Chez les patients atteints de la maladie d'Alzheimer, le taux de GPL dans les membranes neuronales est fortement diminué[38]. Cette diminution s'accompagne d'une augmentation d'activités enzymatiques, comme c'est le cas par exemple pour les enzymes lipolytiques (enzymes visant à dégrader les lipides et en particulier les triglycérides en 3 AG). La diminution de la concentration des GPL peut entraîner une perturbation de la fluidité et de la perméabilité membranaire, ce qui peut avoir un impact sur l'homéostasie osmotique, et ainsi provoquer une inflammation et un stress oxydatif. L'ensemble de ces perturbations peuvent aboutir à un processus de neurodégénérescence[40]. Cette famille a pour structure commune le *sn*-glycérol 3 phosphate. La nomenclature pour cette classe de lipide est la suivante : GP(*sn-1/sn-2).* Où : GP indique la classe du glycérophospholipide, et *sn-1/sn-2* correspondent aux positions des deux acides gras sur le squelette glycérol. Les fonctions alcool en position *sn-1* et *sn-2* sont susceptibles d'être estérifiés par deux acides gras : ce sont des diacyl-phospholipides.

C’est la tête polaire qui permet de faire la distinction entre les différentes sous-classes de glycérophospholipides (Figure 20). Ainsi, 6 classes de phospholipides peuvent être différenciées : les glycérophosphates (GPA), les glycérophosphocholines (GPC), les glycérophosphoéthanolamines (GPE), les glycérophosphoglycérols (GPG), les glycérophosphoinositols (GPI) et les glycérophosphosérines (GPS).

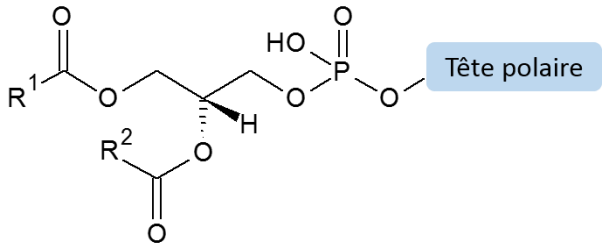


Figure 20 : Structure de base des glycérophospholipides

5.1 Les glycérophosphates

Les glycérophosphates sont les glycérophospholipides les plus simples et la classe de glycérolipide la plus minoritaire (Figure 21). A partir d’un glycérol et de deux acyl-transférases, ils sont produits dans la mitochondrie et le réticulum endoplasmique[1]. Cette sous classe de phospholipides est le précurseur d’autres glycérophoslipides et des glycérolipides comme les TG et les DG[31].

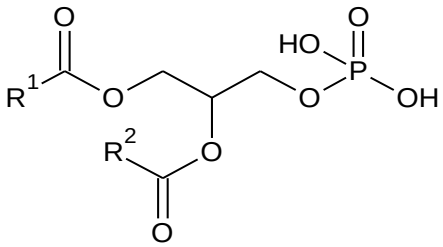


Figure 21 : Structure de base d’un glycérophosphate (GPA)

5.2 Les glycérophosphocholines

Les GPC (Figure 22), sont les lipides les plus abondants dans les membranes des cellules eucaryotes (tant bien animales et végétales), elles constituent jusqu'à 50% de la composition totale de la membrane[1]. Elles sont le principal composant des bicouches lipidiques, permettant d'assurer leur structure. Les PC sont caractérisées par le groupement choline qui les composent. Elles sont synthétisées au niveau du réticulum endoplasmique et dans l'appareil de Golgi à partir d'une choline phosphorylée et d'un DG[1]. Ces composés sont essentiels, puisque leur disparition s'accompagne généralement d'une inhibition de croissance et / ou de mort cellulaire. Hormis leur rôle majeur structural, les GPC sont impliqués dans la signalisation intracellulaire, ou encore dans la stabilisation de la chromatine. Les GPC sont les précurseurs des sphingomyélines.

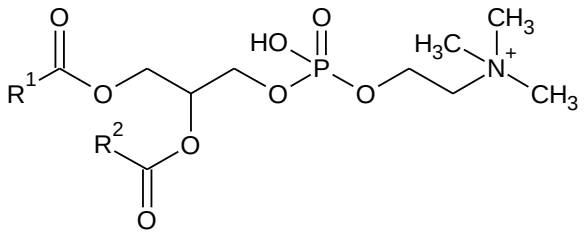


Figure 22 : Structure de base d'une Glycérophosphocholine (GPC)

5.3 Les glycérophosphoéthanolamines

Les glycérophosphoéthanolamines sont composés d'une tête polaire éthanolamine (Figure 23). Ce sont les lipides les plus abondants après les GPC chez les animaux et les végétaux. Dans les cellules animales, la GPE existe sous trois 3 formes : diacyl, alkylacyl ou bien alcènylacyl (Figure 24)[41]. Les GPE sont synthétisés au niveau de la mitochondrie et dans l'appareil de Golgi, à partir d'une éthanolamine phosphorylée et d'un diacylglycérol[42,1].Ces composés sont des médiateurs pro-homéostastiques, qui peuvent être administrés en tant que principe actif. Ils sont notamment utilisés dans les médicaments anti-inflammatoires et analgésiques comme le Normast et le Pelvilen.

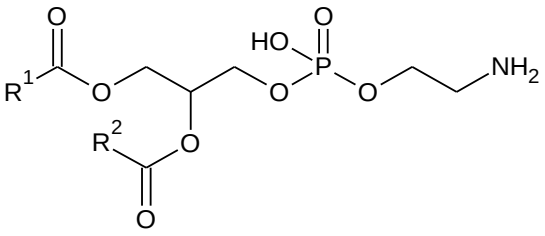


Figure 23 : Structure de base d'une Glycérophosphoéthanolamine (GPE)

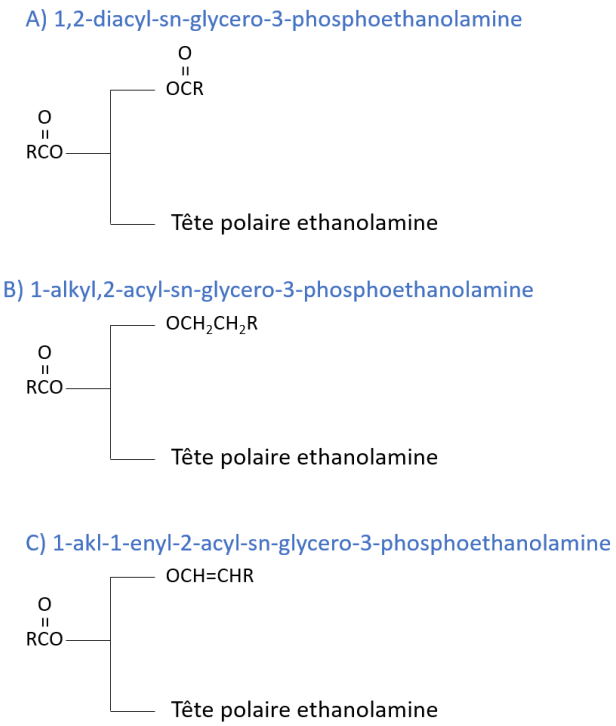


Figure 24 : Structure des GPE sous forme diacyl, alkylacyl et alcèylacyl (A,B,C)

Une sous classe de GPE existe, il s'agit des plasmalogènes (ou étherphospholipides, Figure 25). Ce sont des lipides complexes dont les formes plus abondantes sont les alcénylglycérophospholipides. La synthèse de ces composés nécessite des peroxysomes fonctionnels[43,44]. Un défaut de biosynthèse de ces composés est marqueur de certaines maladies peroxysomales humaines. Structuralement, chez les plasmalogènes, un acide gras est remplacé par un alcool gras insaturé lié au glycérol par une liaison éther[45]. Ces composés sont retrouvés dans la myéline.

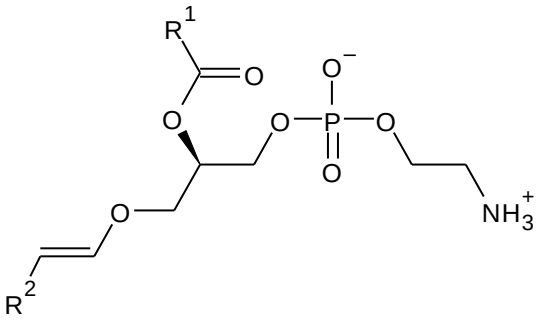


Figure 25 : Structure d'un plasmalogène ou R^1 et R^2 sont des chaines d'acides gras

5.4 Les glycérophosphoglycérols

Chez les animaux, les glycérophosphoglycérols représentent 1 à 2% des lipides totaux (Figure 26). Ils sont localisés dans les mitochondries, où ils servent de précurseurs pour la synthèse des cardiolipides. Ces lipides sont composés d'un glycérophosphate et d'une unité glycérol, formant la tête polaire. Chez les végétaux, les GPG sont retrouvés dans les plastes, où ils jouent un rôle essentiel dans la photosynthèse[46].

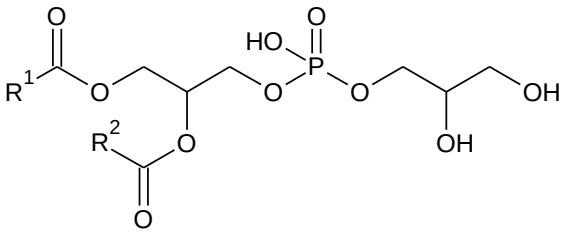


Figure 26 : Structure de base d'un Glycérophosphoglycérol (GPG)

5.5 Les glycérophosphoinositols

Les glycérophosphoionisitols (PI), sont présents dans le feuillet interne des membranes de tous les types cellulaires, et plus particulièrement dans le cerveau, où ils peuvent atteindre 10% de la totalité des GPL. Ils sont caractérisés par une tête polaire de type inositol, et possèdent un résidu alcool qui peut être phosphorylé en position 3, 4 ou 5. Une première phosphorylation conduit à la formation des glycérophosphoinositols phosphates (PIP, Figure 27), une seconde conduit à la formation des glycérophosphoinositols-4,5-biphosphates (PIP_2, Figure 28). Ces composés (PIP_2) peuvent être hydroxylés sur les positions 3-4, et également 3-5. Les PIP_2 peuvent être phosphorylés sur leurs troisièmes positions pour donner des PIP_3. Les PIP_2 peuvent être hydrolysés en diacylglycérol et en inositols triphosphate (IP_3)[47]. Ils sont très faiblement présents dans les membranes plasmiques (de l'ordre de 5%) mais, jouent un rôle essentiel dans le contrôle de la motilité cellulaire et dans le trafic membranaire[20,48,49].

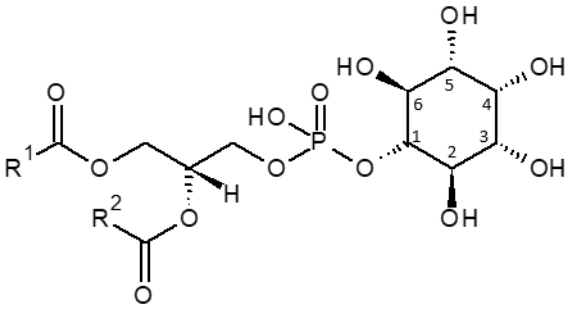


Figure 27 : Structure de base d'un Glycérophosphoionisitol (GPI)

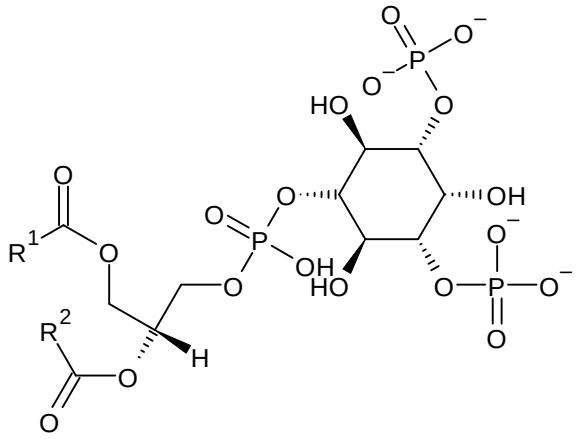


Figure 28 : Structure de base d'un Glycérophosphoinositol-3,5-biphosphate (PIP_2)

5.6 Les glycérophosphosérines

Les glycérophosphatidylsérines, représentant moins de 10 % du contenu lipidique total sont des glycérophospholipides (Figure 29) présentes au niveau des feuillets internes des membranes plasmiques et du reticulum endoplasmique des cellules végétales et animales, et des microorganismes. Les GPS sont constituées d'une tête polaire sérine, elles sont obtenues par la substitution de la choline ou de l'éthanolamine par une sérine. Etant des glycérophospholipides anioniques, les GPS sont capables de chélater les cations, et particulièrement le calcium. Cela va alors modifier la conformation de leurs têtes polaires entrainant un changement de fonction. Ces lipides se situent à la croisée de très nombreuses voies métaboliques. Elles sont notamment impliquées dans l'échange entre les têtes polaires de GPL (GPE et GPC), et constituent, des précurseurs pour leurs synthèses. Les glycérophosphosérines peuvent agir comme messagers de signalisation, via des interactions avec des protéines spécifiques[39,50].

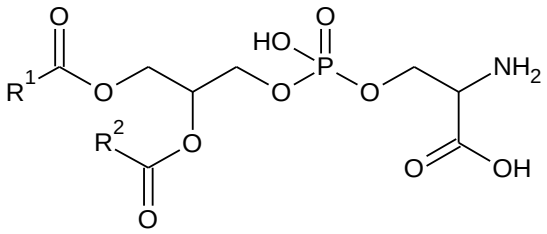


Figure 29 : Structure de base d'une glycérophosphatidylsérine (GPS)

6. Les sphingolipides

Les sphingolipides sont des composés présents dans les membranes cellulaires majoritairement au niveau du feuillet externe de la bicouche lipidique. Cette classe de lipides possède de nombreuses fonctions biologiques. Les sphingolipides sont impliqués dans un grand nombre de maladies génétiques, dans le phénomène d'apoptose, de différenciation et de reconnaissance cellulaire[51,52,53]. Les sphingolipides représentent une catégorie de lipides caractérisée par leur structure commune de type sphingoïde. Ces composés présentent une chaîne d'acide gras, deux fonctions hydroxyles et une fonction amine. La nomenclature de ces composés est la suivante : SL(d *sn-1/sn-2*). Où SL désigne la classe du sphingolipide, d désigne le groupement 1,3-dihydroxyle du motif sphingoïde, et *sn-1/sn-2* désignent respectivement le motif sphingoïde et le radical acyl relié par une liaison amide au groupement sphingoïde. La sous-classe des sphingolipides est déterminée par la nature de la tête (choline, monohexosyl, ou lactosyl par exemple), généralement polaire portée par le groupement hydroxyle primaire du squelette sphingoïde. Chez les mammifères, la base sphingoïde la plus généralement répandue est une sphingosine comportant 18 atomes de carbone et une double liaison (Figure 30). Toutefois il existe également une base sphinganine (Figure 31), comportant 18 atomes de carbone elle aussi, mais pas de double liaison. D'autres sphingolipides sont maintenant détectées, avec des longueurs de chaine d'acide gras variables, pouvant être saturées ou insaturées, méthylées ou oxydées (e.g. C16 :0, C16 :1).

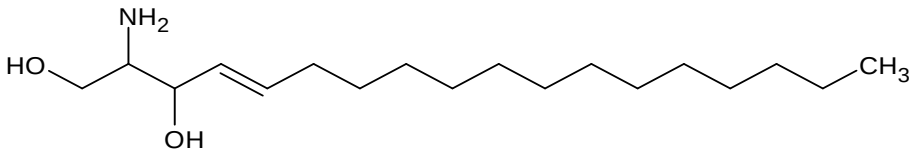


Figure 30 : Structure de la sphingosine la plus répandue chez les mammifères

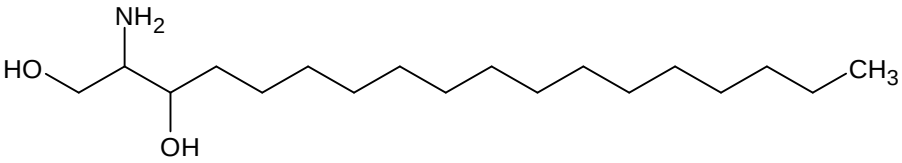


Figure 31 : Structure de la sphinganine la plus répendue chez les mammifères

De plus, il est possible de retrouver d'autres structures issues des spingolipides : les sulfatides. Ce sont des glycosphingolipides où l'un des deux résidus glucidiques a été estérifié par une molécule d'acide sulfurique[54]. Ces composés sont retrouvés dans les membranes des cellules du système nerveux.

6.1 Les céramides

Les plus simple des sphingolipides sont les céramides (Figure 32), ou les acylsphingosines. Ils constituent une famille lipidique structurellement homogène, et sont synthétisés dans le réticulum endoplasmique *via* une réaction entre une sphingosine et un acide gras, menant à la création d'une liaison amide[34]. Les acides gras impliqués dans cette réaction sont généralement composés d'une chaine alkyle entre 14 et 26 carbone, et peuvent être mono-insaturés[4].

Les céramides sont impliqués dans le métabolisme des sphingolipides, puisqu'ils servent de précurseurs aux cérébrosides, aux gangliosides, aux sphingophospholipides tels que la sphingomyéline, et aux acylcéramides. Ils sont également impliqués dans la régulation de l'apoptose.

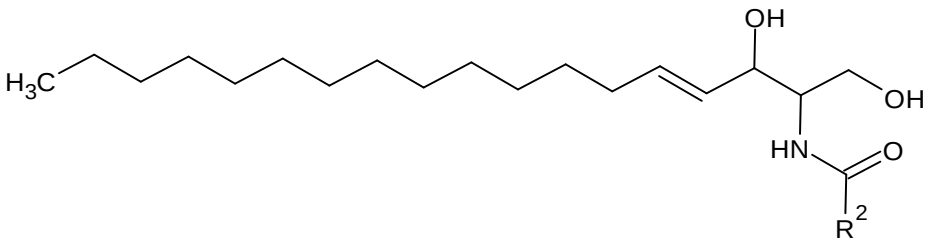


Figure 32 : Exemple d'un céramide (où R_2 est une chaine alkyle variable)

6.2 Les sphingomyélines

Les sphingomyélines, (Figure 33) sont les phosphosphingolipides les plus présents chez les mammifères, et sont des composants ubiquitaires des membranes des cellules animales. Ces composés sont particulièrement retrouvés au niveau du système nerveux, où ils entrent dans la composition des gaines de myélines. Elles sont synthétisées dans la membrane plasmique et dans l'appareil de Golgi, par réaction entre une céramide et une GPC[1,55]. Les chaines d'acides gras qui les composent sont généralement longues, et mono-insaturées. La chaine d'acide gras R^1 est de longueur plutôt fixe (chaine d'acide gras à 18 atomes de carbone), la chaine R^2 quant à elle présente une longueur variable. Les spingomyélines ne semblent pas exister chez les plantes et les microorganismes.

Les sphingomyélines se situent dans la couche externe de la membrane, où elles jouent un rôle essentiel de régulation des canaux ioniques. Elles sont également impliquées dans le fonctionnement des récepteurs couplés aux protéines G[55]. Elles sont capables de former des microdomaines plasmatiques ou radeau lipidique *via* leurs interactions avec le cholestérol[56].

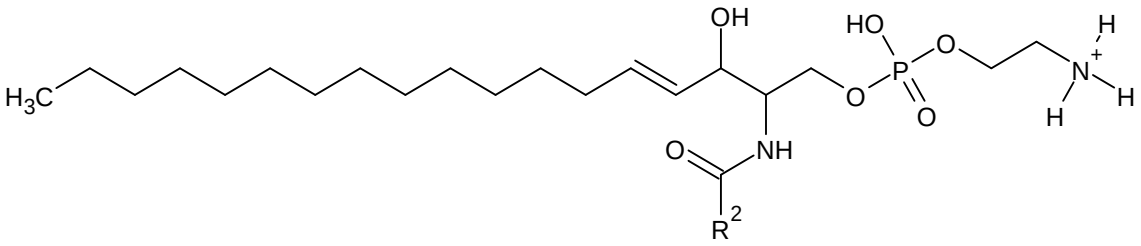


Figure 33 : Structure d'une sphingomyéline (où R^2 est une chaine alkyle variable)

6.3 Les glycosphingolipides

Les sphingolipides sont des lipides complexes, synthétisés dans l'appareil de Golgi. Ces lipides sont composés d'une unité céramide associée à un ou plusieurs sucres (Figure 34)[4,57]. Les deux principaux composés de la famille des sphingolipides sont les cérébrosides et les gangliosides[57]. Ces lipides sont généralement retrouvés dans le cerveau, ou dans les membranes des cellules ganglionnaires du système nerveux. Ils sont essentiels au développement et à la croissance des embryons, à la différenciation tissulaire ainsi qu'au fonctionnement du système nerveux[58].

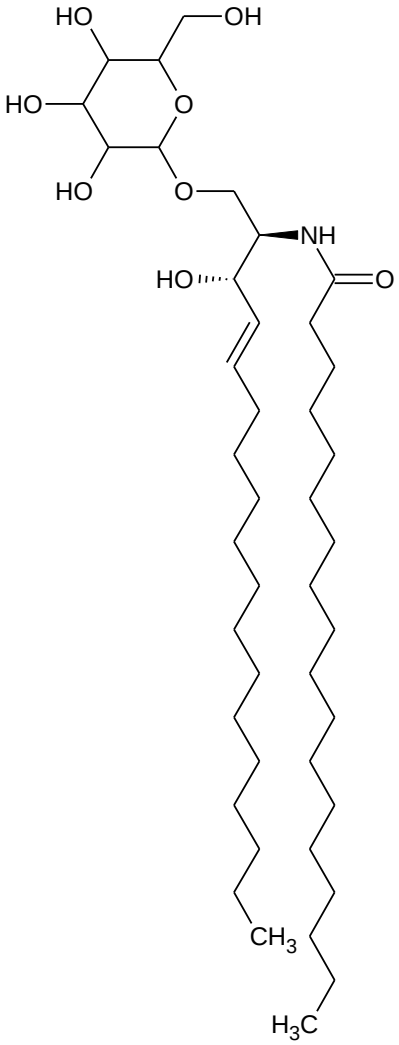


Figure 34 : Exemple d'un galactosylcéramide, composé de la famille des glycosphingolipides

7. Les stérols

Les stérols sont tous issus d'une structure commune : l'unité cyclopentanophénantrène (Figure 35). Il s'agit d'un tétracycle, responsable du caractère hydrophobe des stérols. Les stérols sont les principaux lipides non-polaires, qui constituent les membranes cellulaires. Parmi les stérols, plusieurs sous classes existent : les stérols libres, les stérols estérifiés, les oxystérols, les stéroïdes, les vitamines D et les acides biliaires[4].

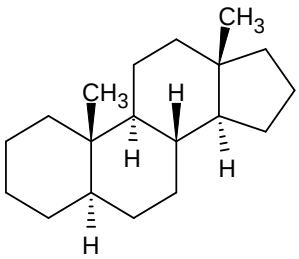


Figure 35 : Structure du squelette cyclopentanophénantrène des stérols

7.1 Les stérols libres

Les stérols les plus abondants chez les mammifères est le cholestérol (Figure 36). Comparé au squelette cyclopentanophénantrène, le cholestérol comporte deux fonctions méthyles en positions 10 et 13, une fonction hydroxyle en position 3 ainsi que des doubles liaisons en positions 5 et 6. Ils sont synthétisés dans le foie, et représentent des composants majeurs des membranes cellulaires assurant leur rigidité en s'intercalant entre les acides gras des phospholipides et des sphingomyélines. Les cholestérols sont les précurseurs des hormones stéroïdiennes, de la vitamine D et des acides biliaires. Leurs concentrations sont finement régulées, un dérèglement peut entrainer des perturbations plus ou moins importantes à l'origine de nombreuses pathologiques, comme par exemple l'athérosclérose[59].

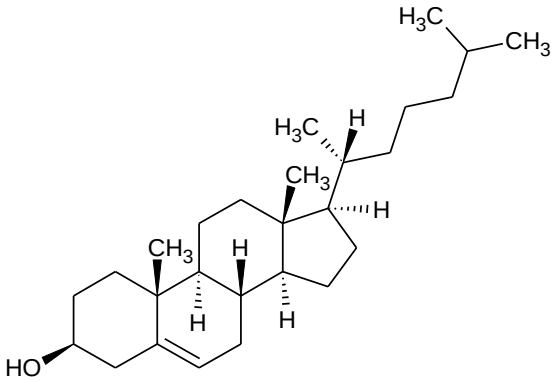


Figure 36 : Structure du cholestérol : cholest-5-en-3-ol

7.2 Les stéroïdes

Les stéroïdes sont des hormones, toutes synthétisées à partir d'un même précurseur : le cholestérol. Seuls les tissus dits stéroïdogéniques sont capables de synthétiser des stéroïdes. Ce sont principalement les glandes endocrines : les gonades les glandes surrénales, mais aussi le système nerveux central. La première étape de la synthèse de ces hormones à lieu dans la mitochondrie et dans le réticulum endoplasmique, ou la biosynthèse est catalysée par le cytochrome P450 et par des enzymes (hydroxystéroïdes déhydrogénases)[60]. Ces molécules sont impliquées dans de nombreuses fonctions physiologiques telles que le maintien de la pression artérielle, le développement des caractères sexuels secondaires, la reproduction, le mémoire, le métabolisme des glucides ou encore le développement osseux[61,62,63]. Les stéroïdes sont classés en 3 grandes sous familles. La première regroupe les stéroïdes composés d'un cycle oestrane à 18 atomes de carbone. Dans cette famille de stéroïdes se trouve notamment les oestrogènes (Figure 37 A). La seconde sous-famille de stéroïdes regroupe les composés ayant un cycle androstane à 19 atomes de carbone, notamment la testostérone (Figure 37 B). Enfin, la dernière classe de stéroïdes comporte un cycle prégnane à 21 atomes de carbone, comme par exemple la progestérone (Figure 37 C).

Il existe une autre famille de composés stérols très vaste, les oxystérols. Il s'agit de structures dérivées du noyau stérol sur lesquelles plusieurs oxydations ont eu lieu. Ces composés sont notamment retrouvés chez les plantes, les levures ou chez l'homme.

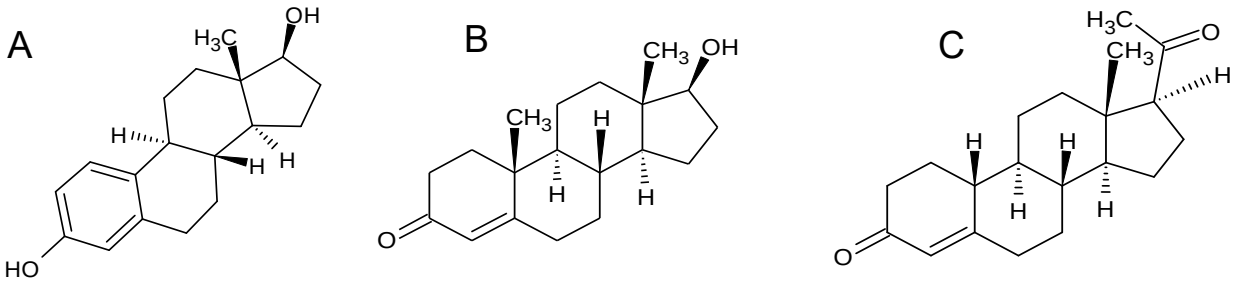


Figure 37 : Structure de différentes hormones stéroïdiennes : l'estradiol (A), la testostérone (B) et la progestérone (C)

7.3 Les vitamines D

Le terme vitamine D regroupe deux composés : l'ergocalciférol ou vitamine D_2, et le cholécalciférol ou vitamine D_3 (Figure 38). La première est issue de l'alimentation d'origine végétale, comme les céréales, alors que la seconde est produite par la peau, sous l'action des rayons UV. Elle peut également se retrouver dans l'alimentation d'origine végétale ou dans les poissons gras. Ces vitamines, considérées comme des hormones, facilitent l'absorption intestinale du calcium et du phosphate. Cette absorption est nécessaire à la minéralisation des tissus osseux, et donc à la croissance et au maintien de la structure osseuse[64,65]. Certaines études suggérent qu'un déficit en vitamine D pourrait entrainer un risque de développement cancéreux. Une carence en vitamine D entraine un rachitisme chez le jeune enfant en croissance[66]. Chez l'adulte, cette carence est associée à un trouble de la minéralisation osseuse, appelée ostéomalacie[67].

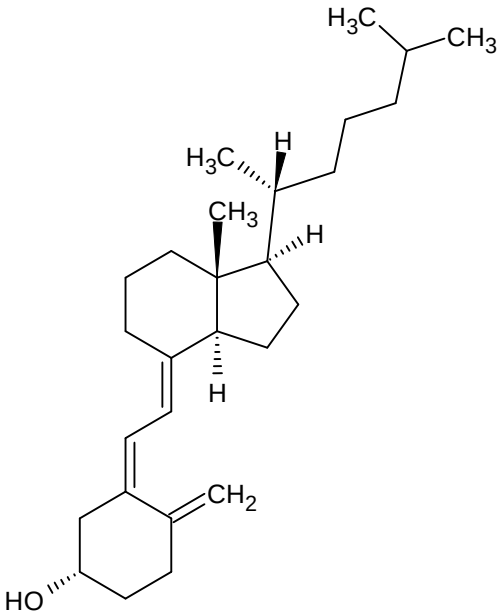


Figure 38 : Structure du cholécalciférol (vitamine D_3)

7.4 Les acides biliaires

Les acides biliaires sont une famille de stérols synthétisés dans le foie, à partir du cholestérol oxydé en position 7[68] : le 7-α-hydroxycholestérol (Figure 39). Ces molécules sont dérivées du cholestérol, mais ont en commun une structure stéroïdienne (le noyau stéroïdien) ainsi qu'une chaine latérale constituée au minimum de 4 atomes de carbone. Ce sont des molécules amphiphiles puisqu'elles sont composées d'une partie hydrophobe, par la structure stéroïdienne, et d'une partie hydrophile *via* la chaine latérale, et les groupements hydroxyls. Il existe une grande diversité d'acides biliaires au niveau structural définie par la position des groupements hydroxyls portés par les carbones ainsi que par la nature du groupement fixé à la chaine latérale. Ils peuvent se retrouver sous forme libre ou sous forme conjugués à la taurine ou à la glycine. Ces composés comportent une fonction acide. Les acides biliaires sont ensuite stockés dans la vésicule biliaire, sécrétés dans l'intestin, et retournent vers le foie via un cycle entéro-hépatique[69]. Les 4 fonctions principales des acides biliaires sont : 1) faciliter l'absorption intestinale des lipides, 2) faciliter l'excrétion biliaire de métabolites endogènes, 3) éliminer le cholestérol de l'organisme, et 4) participer à un maintien de l'homéostasie énergétique. Chez l'homme, les acides biliaires les plus répandus sont l'acide cholique $C_{24}H_{40}O_5$ (Figure 40A), l'acide chénodésoxycholique $C_{24}H_{40}O_4$ (Figure 40B) et l'acide muricholique $C_{24}H_{40}O_5$ (Figure 41C)[70]. L'acide cholique peut être conjugé à la glycine et à la taurine pour former l'acide glycocholique et l'acide taurocholique (Figure 41A et 41B respectivement)

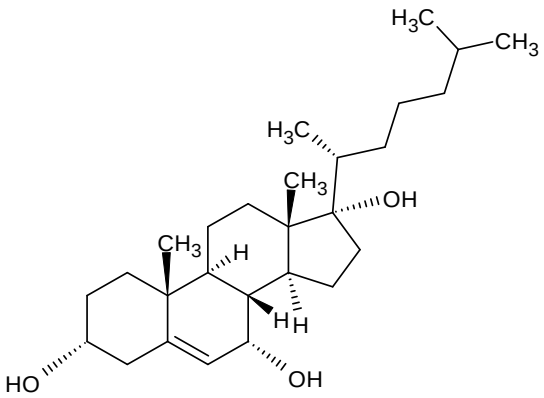


Figure 39 : Structure du 7-α-hydroxycholestérol, précurseur des acides biliaires

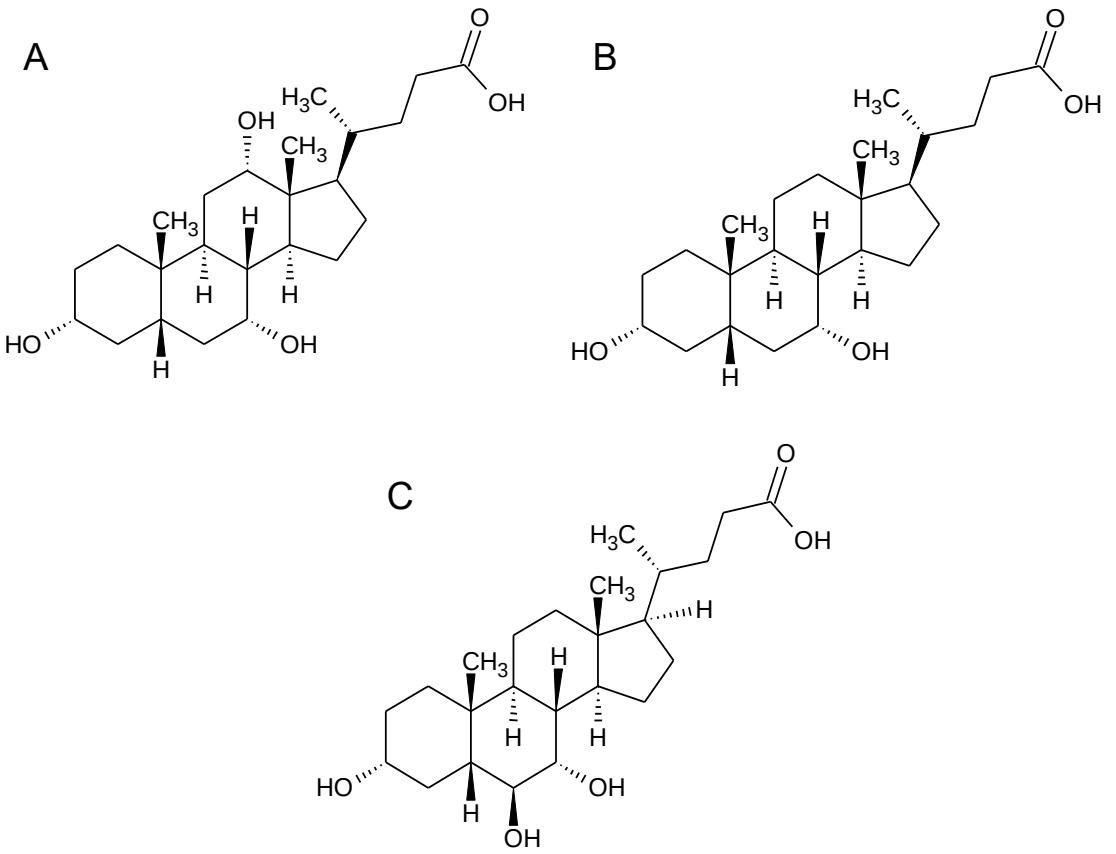


Figure 40 : Structure de deux acides billiaires majoritaires : l'acide cholique $C_{24}H_{40}O_5$ (A) et de l'acide chénodésoxycholique $C_{24}H_{40}O_4$ (B), et de de l'acide α-muricholique $C_{24}H_{40}O_4$ (C)

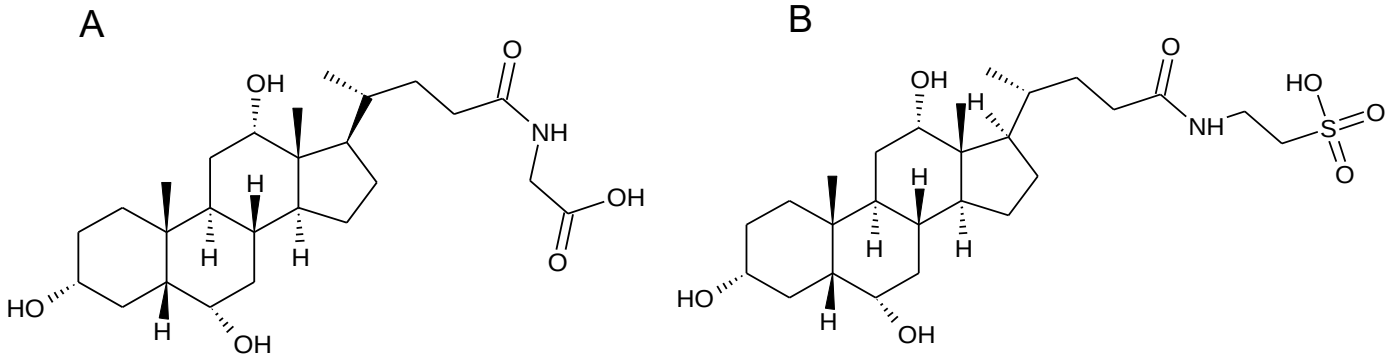


Figure 41 : Structure de l'acide glycocholique (A) et de l'acide taurocholique (B)

8. Les prénols

Les prénols sont synthétisés à partir de précurseurs constitués de 5 atomes de carbone : l'isopentényl-pyrophosphate et le diméthylallyl-pyrophosphate (Figure 42). Il existe différentes familles de prénols, classées en fonction de la longueur du squelette hydrocarboné et de leurs groupements fonctionnels. Par exemple les isoprénoïdes en C40 sont des caroténoïdes, comme par exemple le β-caroténe (Figure 43). Les quinones constituent une famille composée de dienes, les hydroquinones sont constitués de composés organiques aromatiques (Figure 44). Les composés prénolés jouent un rôle de protection contre le stress oxydatif, les maladies cardiovasculaires et certains cancers[35].



Figure 42 : Structure de l'isopentényl-pyrophosphate et du diméthylallyl-pyrophosphate

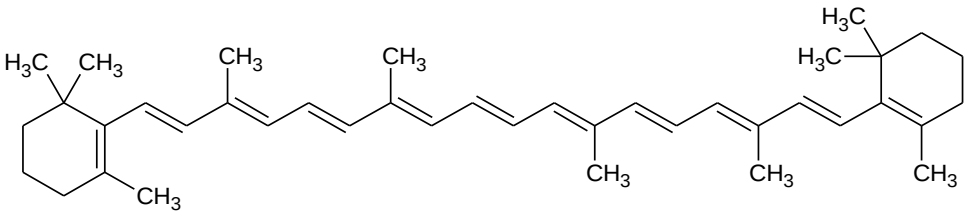


Figure 43 : Structure d'un prénol, composé issu de la sous famille des isoprénoïdes : le β-caroténe

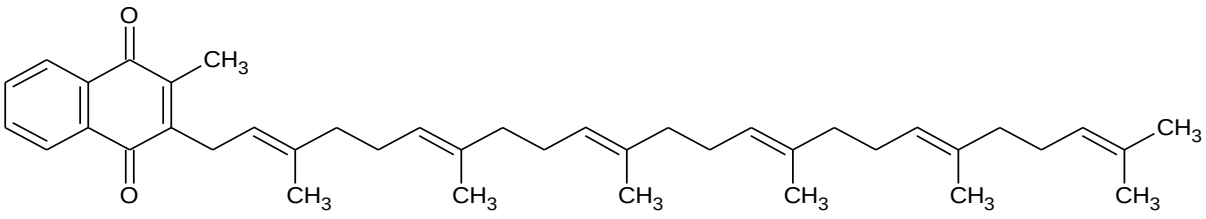


Figure 44 : Structure d'un prénol, de la sous famille des quinones : la ménaquinone, ou vitamine K2

Chapitre II : Le microbiote intestinal et lipides associés

Dans ce chapitre, les notions essentielles et indispensables concernant le microbiote intestinal seront présentées. Les notions telles que la de taxonomie, la dysbiose intestinale, la mise en place du microbiote intestinal chez l'enfant et l'adulte, et certaines pathologies liées au microbiote intestinal seront discutées. Dans un second temps les lipides du microbiote intestinal seront présentés en deux temps : les lipides structuraux puis les lipides issus du microbiote intestinal.

1. Définitions et notions essentielles

La notion de microbiote est introduit pour la première fois en 2001, définit comme un "système écologique de microorganismes commensaux symbiotiques et, parfois pathogènes qui résident dans le corps humain"[71]. Le microbiote n'est pas uniquement retrouvé chez les êtres humains, mais également chez les animaux et même les plantes. Depuis 2001 la définition à évoluée, le microbiote est une collection de microorganismes se regroupant dans une niche spécifique comme le microbiote intestinal, le microbiote pulmonaire ou le microbiote de la peau. Les microbiotes sont donc organes spécifiques puisque l'on ne retrouve pas les mêmes souches de micro-organismes en fonction de l'organe étudié. Cependant, le microbiote évolue au cours du temps dans une même niche comme au niveau du microbiote intestinal, objet de ce chapitre, où sa composition varie au cours de la vie de l'hôte.

Le microbiote intestinal, antérieurement appelé flore intestinale, est composé d'une population de bactéries et d'autres micro-organismes comme les archées, les virus et les champignons se trouvant dans le tractus gastro-intestinal. Il possède un rôle local en participant au maintien de l'homéostasie intestinale et un rôle systémique en étant impliqué dans plusieurs fonctions physiologiques vitales pour l'hôte, comme dans la maturation du système immunitaire. Le déséquilibre des populations composant le microbiote peut conduire à des pathologies digestives mais également à des pathologies non-intestinales comme l'obésité ou l'autisme[72,73]. Afin de restaurer cet équilibre, de nombreuses thérapies comme la prise de probiotiques ou les transplantations fécales de microbiote ont été développées. Des travaux plus récents sur l'activité de certains métabolites synthétisés par le microbiote intestinal ouvre la porte à des perspectives thérapeutiques prometteuses[74].

La caractérisation des métabolites du microbiote concerne essentiellement des molécules produites par les bactéries. Elles apportent à l'hôte entre 2 et 20 millions de gènes,

possédant, pour la majorité d'entre eux, des fonctions non redondantes. Taxonomiquement, les bactéries sont classées en phylas, classes, familles, genres et espèces. Le microbiote intestinal est composé de 60 genres, provenant de 4 phyla majoritaires : *Firmicutes, Bacteroidetes, Actinobactéries, Protéobactéries* (Figure 45). Pour ce chapitre, nous allons nous focaliser uniquement sur les bactéries.

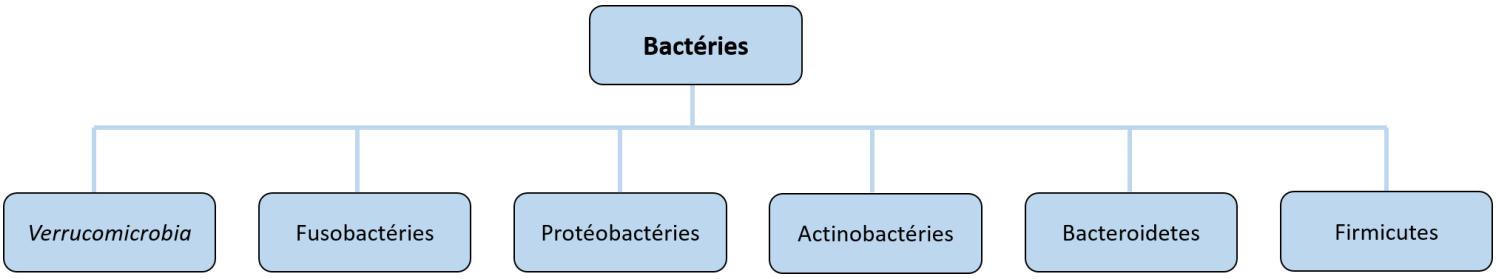


Figure 45 : Représentation schématique des grands phyla des bactéries

2. Origine et composition du microbiote intestinal

La composition du microbiote intestinal varie en fonction des espèces. Le microbiote intestinal de la souris, présente par exemple des différences lorsqu'il est comparé au microbiote intestinal humain. Au niveau des phylas de ces deux microbiotes, plus de 90% y sont communs, notamment par l'ultra dominance de deux phylas : les *Firmicutes* et les *Bacteroidetes.* La composition du microbiote intestinal varie également en fonction de l'âge de l'individu, cette partie est détaillée dans le point suivant.

2.1 Chez l'enfant

Le nouveau-né, se retrouve dès sa naissance en contact avec des bactéries, des archées, des virus et des organismes unicellulaires permettant la mise en place du microbiote intestinal. Lors d'un accouchement par voie basse, les bactéries du microbiote vaginal et fécal de la mère participent à la colonisation du tractus digestif fœtal.

Les premières bactéries colonisatrices retrouvées chez le nouveau-né sont aérobies et anaérobies facultatives et appartiennent à deux phyla : *Firmicutes* (*staphylocoques, enterocoques, …*) et *Proteobacter* (*Escherichia coli*). Elles vont modifier l'environnement pour

permettre l'implantation d'autres micro-organismes comme les bactéries anaérobies strictes (*Bacteroides, Clostridioides, Actinobacteries*). Le microbiote intestinal des nouveaux nés se caractérise par une faible abondance et une faible diversité bactérienne et par la dominance de deux phyla les *Proteobacteries* et les A*ctinobacteries*. Au cours de l'enfance, le microbiote du nouveau-né se diversifie, pour au final être dominé par deux autres phylas bactériens : les *Bacteroidetes* et les *Firmicutes*[75,76]*.*

Le microbiote intestinal du nouveau-né se complexifie de manière progressive au courant de sa vie. Dans un premier temps, certaines bactéries vont être apportées *via* le lait, et cela qu'il soit d'origine maternelle ou industrielle. Ensuite, ce sont d'autres bactéries, issues de l'environnement qui vont venir ajouter de la complexité au microbiote intestinal. Au bout d'un an, l'enfant possède un profil microbien similaire à celui d'un adulte en terme de diversité et de composition : il contiendra donc des espèces bactériennes appartenant aux phylas des *Firmicutes*, des *Bacteroidetes*, des *Proteobacteria* et des *Actinobacteria*[77,78,79]. C'est à l'âge de 3 ans que le microbiote intestinal de l'enfant commence à se stabiliser en terme de concentration bactérienne et d'espèces. Une fois que la population bactérienne du microbiote intestinale s'est stabilisée, elle reste la même tout au long de la vie de l'hôte en bonne santé.

2.2 Chez l'adulte

La quantité de bactéries composant le microbiote intestinal d'un adulte ne fait qu'augmenter en suivant un gradient de concentration de la bouche à l'anus (Figure 46)[80]. La quantité de bactérie atteint son maximum dans le côlon avec une concentration entre 10^9 à 10^{12} bactéries par gramme de tissu correspondant à 1000 voire 2000 espèces différentes[75]. Chez l'homme adulte, le microbiote intestinal représente environ 1 kilogramme de son poids total et est reconnu comme un organe à part entière[81]. Il y a dix fois plus de diversité chez les cellules bactériennes que chez les cellules humaines du corps[82]. L'intestin est l'organe le plus colonisé par le microbiote, il concentre environ 70% des bactéries[83,84]. Les *Firmicutes* (bactéries gram positives) et les *Bacteroidetes* (bactéries gram négative), constituent à eux deux près de 90% de la population totale du microbiote intestinal.

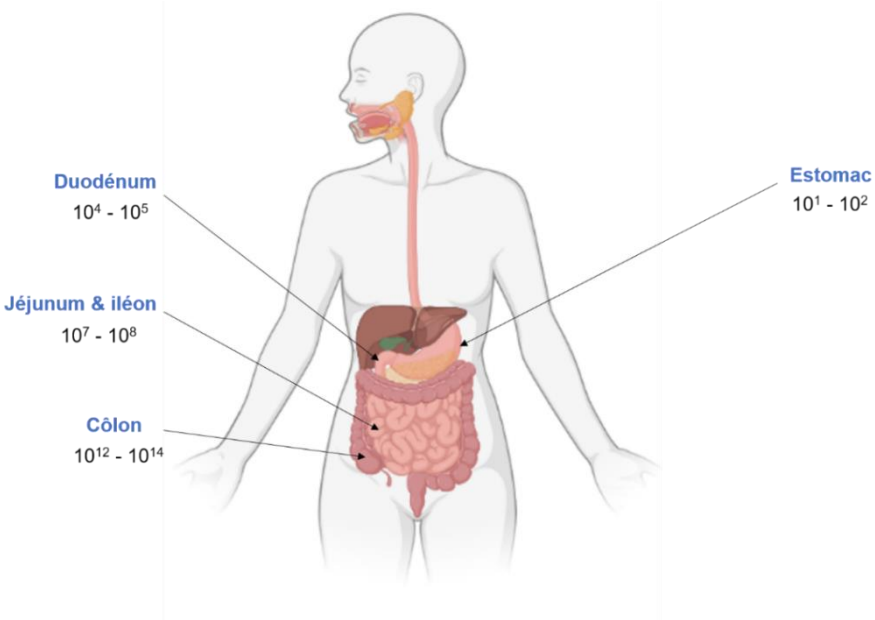


Figure 46 : Gradient de quantités en bactéries dans le tractus digestif humain

Les *Firmicutes*, qui sont plus abondantes, représentent entre 60 et 80% de la population bactérienne du tractus digestif. Ce phylum est majoritairement composé de bactéries appartenant à la classe des *Clostridia*. Les *Bacteroidetes,* représentant 30% de la population bactérienne du tractus digestif[85], est constitué par plusieurs espèces comme *Bactéroides fragilis.* Les deux autres phylas majoritaires de l'intestin humain sont les *Actinobacteria* et les *Proteobactaria*. Le phylum des *Actinobacteria* comprend les *Bifidobacteries* telles que *Bifidobacterium Longum.*

Bien que la composition du microbiote intestinal puisse varier d'un individu à l'autre, on retrouvera toujours le même profil d'espèces dominantes au cours du temps chez un même individu[86]. Cet équilibre du microbiote intestinal peut être perturbé par un changement brutal de régime alimentaire ou par un traitement antibiotique. La prise d'antibiotiques va induire la modification des espèces majoritaires composant le microbiote et/ou la disparition de certaines espèces initialement présentes. Ce n'est qu'un mois après l'arrêt du traitement antibiotique qu'un retour au profil initial des espèces bactériennes contenues dans le microbiote intestinal sera rétabli, c'est le phénomène de résilience[87,74].

3. Bactéries utilisées pendant mes travaux de thèse

Les probiotiques est un concept développé par Henry Tessier et Elie Metchnikov dans les années 1900, mais c'est en 1954 que Ferdinand Vergin utilise pour la première fois le terme probiote. Lors de la mise en place de cultures de bactéries anaérobies, Henry Tessier identifie que contrairement aux données de la littérature *Escherichia coli* n'est pas la bactérie majoritaire de la flore retrouvée chez l'enfant, mais que c'est une autre bactérie, méconnue qu'il nommera *Bacillus bifidus*. Il définit alors certaines bactéries comme « bonnes bactéries », abondantes de la flore infantile.

A ce jour, les probiotiques sont définis par l'OMS comme des micro-organismes vivants, qui confèrent un effet bénéfique à l'hôte lorsqu'ils sont présents en quantité adéquate. Pour qu'une bactérie soit considérée comme probiotique, celle-ci se doit de répondre à plusieurs obligations comme par exemple avoir une efficacité prouvée dans des études chez l'être humain. Plusieurs études cliniques cherchent à mettre en lien probiotiques et effet sur la douleur viscérale, ou plusieurs études montrent que certains probiotiques peuvent diminuer la douleur viscérale[88], alors que d'autres non[89]. Il apparait clairement que les effets de certains probiotiques ne sont pas généralisables à l'ensemble des probiotiques, de même espèce ou d'espèces différentes, et qu'un nombre limité de souches semblent avoir des effets bénéfiques pour l'organisme. Les probiotiques peuvent améliorer les fonctions de la barrière intestinale et possèdent des certains cas des fonctions anti-inflammatoires[90], anti-cancéreuses[91], anti-pathologiques[92] ou moduler le transit intestinal et la douleur viscérale.

La plupart des probiotiques qui sont utilisés à ce jour ont été caractérisées parmi les bactéries commensales de l'intestin humain, c'est le cas par exemple des deux bactéries probiotiques que nous avons utilisées lors de notre étude : *Escherichia coli Nissle 1917* (EcN) et *Lactobacillus animalis/murinus*.

Escherichia coli Nissle 1917 (EcN) a été la première souche de probiotique utilisée dans mes travaux puisqu'en 2017, notre équipe a identifié un nouveau lipide bactérien aux activitées analgésiques issu de cette bactérie : le C12asnGABA[93].

Lactobacillus animalis/murinus n'est pas une des bactéries les plus abondante du tractus digestif humain, cependant les *Lactobacillus* sont connues pour leurs effets bénéfiques sur l'homéostasie intestinale. De plus, notre équipe a démontré récemment que la présence de cette bactérie dans un modèle murin était corrèlement inversé à la diminution de la douleur viscérale.

Lactobacillus : Le genre des *Lactobacillus* décrit en 1901 par Beiierinck a été établi sur la base de la morphologie et des propriétés physiologiques des bactéries qui le composent[94]. Les *Lactobacillus* appartiennent au phylum des *Firmicutes* et à la famille des *Lactobacillaceae*, et sont caractérisées par leur forme en bâtonnet et leur production d'acide lactique. Ces bactéries gram+ sont immobiles, non flagellées, non sporogènes et anaérobies facultatives[94]. Les *Lactobacillus* ont été parmi les premières espèces de bactéries décrites pour leurs effets probiotiques. C'est en 1935 qu'est commercialisé pour la première fois *Lactobacillus casei shirota.* Les *Lactobacillus* sont des espèces fortement retrouvées au niveau des milieux riches en glucose comme les muqueuses intestinales, orales et vaginales où ils représentent 90 à 100% des bactéries. Le rôle probiotique des *Lactobacillus* lui est conféré en partie par sa production d'acides lactique et acétique, qui sont antibactériens, et son système de capture du fer efficace qui lui confère un avantage face aux pathogènes[95]. Elles participent également à la maturation du système immunitaire et de la barrière épithéliale intestinale[95].

Dans une majorité d'études, *L. murinus* et *L. animalis* sont considérées comme étant une seule et même espèce. Difficilement identifiables par les techniques classiques de caractérisation métabolique, elles se ressemblent également fortement d'un point de vue génétique puisque leurs ARN 16S présentent peu de différences. Il est donc suggéré de les identifier en tant que *L. murinus/animalis* et c'est ce que nous ferons dans notre étude[96].

Escherichia coli : Ces bactéries ont été décrites pour la première fois en 1885 par Theodor Escherich, lorsqu'il isole et caractérise ces bactéries dans des selles d'enfants[97]. Il nomme alors cette bactérie Bacterium coli commune. Le nom d'*Escherichia coli* sera reconu uniquement à partir de 1954. Les *Escherichia coli* appartiennent au phylum des *Proteobacteries* et à la famille des *Enterobacteriaceae*[98]. Ce sont des bactéries à gram -, anaérobies facultatives en structures de bâtonnets. Elles possèdent un flagelle pour se déplacer ou des pilis pour adhérer à la surface des cellules[99]. Environ 90 % de la population humaine possèdent des E. coli au sein de son tractus digestif[100]. Ces bactéries sont réparties en 7 groupes phylogénétiques (A, B1, B2, C, D, E et F) qui regroupent des bactéries commensales, pathogènes et probiotiques[101]. Le groupe phylogénétique B2 contient des *E. coli* possédant une capacité accrue de colonisation et d'implantation dans le tractus digestif humain. Dans la population française, 34% des *E. coli* du tractus digestif appartiennent à ce groupe[102]. Il est hautement hétérogène, et renferme à la fois des bactéries commensales et pathogènes, mais aussi probiotique comme *EcN*[101] *EcN* Nissle 1917 est une souche probiotique d'*Escherichia coli* décrite pour la première fois par Alfred Nissle. Pendant la

première guerre mondiale, il remarque son effet bénéfique en l'isolant du microbiote d'un de ses caporaux, inaffecté par la dysenterie[101].

E. coli est vendue en tant que probiotique dans certains pays sous le nom de Mutaflor® depuis 1917, son utilisation y est indiquée dans les cas de constipation chronique, dans les phases de rémission de la rectocolite hémorragique et dans le syndrome de l'intestin irritable. De plus, elle est souvent utilisée en tant que probiotique dans les études cliniques et pré-cliniques, ce qui a permis d'identifier certains des mécanismes *via* lesquels elle agît sur l'hôte.

4. Les fonctions du microbiote intestinal

Le microbiote intestinal est considéré comme un organe à part entière et joue un rôle sur l'homéostasie générale de l'hôte par ses fonctions métaboliques, immunologiques et endocrines[80,103]. L'ensemble des bactéries qui peuplent notre système digestif sont en contact avec les cellules de notre intestin. L'hôte et le microbiote intestinal ont établi une relation mutualiste et symbiotique : les deux partenaires ont un avantage à vivre ensemble. Les bactéries qui colonisent nos intestins sont impliquées dans de nombreuses fonctions physiologiques et participent également au maintien de notre santé. En échange, l'hôte procure un environnement stable aux bactéries ainsi que les nutriments, *via* l'alimentation, nécessaire à sa survie et à son développement. Différents facteurs, comme par exemple l'alimentation, sont impliqués dans le maintien de cette osmose entre les bactéries et hôte[86].

Bien que la composition taxonomique du microbiote intestinal puisse varier d'un individu à l'autre, les fonctions qu'il exerce sont majoritairement identiques[103]. Une des fonctions principales du microbiote intestinal est la fermentation des fibres non-digestibles. Ces fibres alimentaires, comme l'amidon ou les oligosaccharides, servent de substrat au microbiote. Elles vont être métabolisées en acides gras à courte chaine (AGCC). Cela implique de nombreuses activités métaboliques réalisées par le microbiote. La plupart des métabolites issus de la fermentation par le microbiote intestinal vont être utilisés par l'organisme et auront des répercussions sur la santé de l'hôte et le maintien de son homéostasie[86]. Le microbiote intestinal participe également à la fonction de barrière de la muqueuse intestinale en empêchant la colonisation du système digestif par d'autres bactéries qui pourraient être pathogènes pour l'hôte. Le microbiote intestinal participe à la production de xénobiotiques (métabolites non endogène à l'hôte) comme par exemple la synthèse de vitamines telles que les vitamines B2, B9, B12 ou K[104]. Ces vitamines d'origines bactériennes sont essentielles au bon développement de l'hôte et au maintien de son homéostasie.

Les bactéries assurent le développement et la maturation du système immunitaire et des cellules épithéliales du tube digestif[105]. Le microbiote intestinal participe aux effets trophiques de l'épithélium intestinal en favorisant le développement de microvillosités, la maturation des réponses immunitaires et adaptatives de l'hôte[106,107]. Les bactéries protègent également la muqueuse intestinale *via* la production de peptides antimicrobiens tels que les bactériocines. D'autres études ont permis de mettre en évidence que le microbiote intestinal permet d'améliorer considérablement la qualité du mucus[108].

A l'âge adulte, les souris axéniques, dépourvues de microbiote, présentent un grand nombre d'anomalies qui peuvent être corrigées en quelques semaines par l'inoculation d'un microbiote intestinal varié. L'absence de bactéries intestinales chez ces souris axéniques nécessite un apport calorique bien plus élevé ainsi qu'un régime riche en vitamines afin de maintenir un poids normal. Les souris axéniques sont plus sensibles aux infections car leur muqueuse intestinale est immature et la croissance épithéliale fortement ralentie[109]. Ces études permettent de mettre en lumière la nécessité et l'importance du microbiote intestinal. Les souris axéniques ont également été utilisées pour étudier le rôle du microbiote intestinal au cours du développement du système immunitaire du tractus digestif de l'hôte[110]. Des études ont démontré que l'exposition précoce à des bactéries est nécessaire pour le bon développement du système immunitaire et la mise en place de la tolérance du système immunitaire vis-à-vis du microbiote. L'absence à cette exposition bactérienne peut avoir des conséquences néfastes et irréversibles sur la santé de l'hôte. D'autres études se sont focalisées sur l'impact du microbiote sur les allergies. Certaines réactions allergiques déclenchées chez des souris axéniques ont été inhibées par la colonisation de l'intestin par le microbiote intestinal de souris sauvages[111].

5. Des modifications de la composition du microbiote intestinal peuvent être associées à des pathologies ou à désordres métaboliques

De nombreuses pathologies du tractus digestif et systémiques sont associées à une dysbiose du microbiote c'est-à-dire à une modification majeure de la composition taxonomique du microbiote intestinal. Certains troubles de l'immunité ou désordres métaboliques peuvent être déclenchés, ou aggravés par les bactéries intestinales de l'hôte. Au cours des dernières années, environ 25 pathologies ou troubles fonctionnels ont été associés à des modifications de la composition du microbiote intestinal[112]. Le microbiote aurait un impact direct sur l'activité des nerfs innervant le tractus digestif, sur les cellules épithéliales et sur l'activation des cellules immunitaires[113].

Parmi ces pathologies, le syndrome de l'intestin irritable (SII) est un trouble fonctionnel digestif associé à une dysbiose du microbiote chez 72% des patients. Il touche environ 5% de la population française, 11% de la population mondiale et principalement les femmes (sexe ratio : 2/1)[113]. C'est la première cause de consultation chez le gastroentérologue[114]. Bien que le syndrome de l'intestin irritable n'engage pas directement le pronostic vital, il altère profondément et de façon chronique la qualité de vie des malades. Les traitements étant dirigés contre les symptômes sont peu efficaces, c'est pour cela que le SII est un réel problème de santé publique. Cette pathologie se caractérise par des douleurs abdominales, des ballonnements (pouvant être associés à un métabolisme excessif du microbiote intestinal), et des troubles du transit intestinal[115]. Ce sont ces symptômes qui sont la cible des traitements. Le but des traitements est également de diminuer les conséquences psycho-sociales qui parfois peuvent être dramatiques. Les traitements de première intention sont médicamenteux, mais présentent des effets indésirables notoires. L'efficacité de ces traitements restent faibles et ne représente qu'un gain thérapeutique allant de 7 à 15%[116]. Il est donc urgent de développer de nouveaux traitements et de mieux caractériser la physiopathologie de ce syndrome.

De nombreuses études ont mis en évidence une dysbiose d'un point de vue taxonomique entre les patients atteints de SII et les patients contrôles. Néanmoins, les différentes études montrent de grandes divergences allant jusqu'à des contradictions dans la description des différences de composition bactérienne du microbiote de patients IBS et contrôle. Un microbiote est caractérisé par son répertoire de gènes exprimés et, plus important encore, par les protéines et les métabolites qu'il sécrète. De nombreuses études ont déterminé la composition taxonomique du microbiote, mais peu d'études ont caractérisé les gènes actifs, les protéines ou les métabolites. Les études portant sur les molécules produites par le

microbiote pourraient être plus pertinentes que celles visant à déterminer sa composition pour comprendre son rôle dans le SII. En effet, la composition taxonomique du microbiome humain varie énormément d'un individu à l'autre, tandis que sa composition génétique ou sa capacité fonctionnelle est très conservée[117]. La capacité fonctionnelle du microbiote est redondante et capitale pour maintenir la stabilité et la résilience du microbiome humain[117]. A la notion de redondance fonctionnelle du microbiote est associée l'idée que, jusqu'à un certain point, les espèces sont substituables d'un microbiote à l'autre en termes de fonctionnalité[117]. Ainsi, deux microbiotes différents dans leur composition peuvent posséder des profils de protéines et de métabolites aux fonctions similaires[117]. Il est donc majeur d'identifier les métabolites caractéristiques du microbiote intestinal des patients atteints du SII et de déterminer leurs effets sur l'hôte en mettant l'accent sur l'homéostasie intestinale. Compte tenu qu'aucune augmentation de la translocation bactérienne n'a pu être observée chez les patients SII, seules des molécules sécrétées et/ou transformées par les bactéries et capables de traverser la barrière intestinale pourraient avoir un impact sur la douleur viscérale. Parmi les métabolites bactériens plusieurs lipides comme les lipopeptides[93] et les acides gras à chaîne courtes[118] ou longues[119] sont capables de traverser la barrière épithéliale et de réguler l'activation des neurones sensitifs *in vivo*. L'identification et la caractérisation de nouveaux métabolites spécialisés produits par le microbiote intestinal semble aujourd'hui être une voie permettant de comprendre l'effet du microbiote intestinal sur certaines pathologies comme le SII.

6. Les lipides bactériens

Les lipides bactériens peuvent être de deux sortes :

- Les lipides dits de structure, qui sont les principaux constituants des membranes bactériennes. Selon les bactéries considérées les lipides structuraux qui les composent peuvent varier, tant en nombre, qu'au niveau des espèces lipidiques présentes. Cependant, les lipides structuraux majoritaires des bactéries, quel qu'en soit l'espèce considérée sont toujours des phospholipides et notamment des glycérophospholipides[120].

- Les lipides synthétisés ou transformés ; il s'agit de métabolites spécialisés (ou métabolites secondaires), que l'on retrouve chez les bactéries, les plantes et les champignons. Ces métabolites, à l'inverse des métabolites primaires n'assurent pas la croissance de l'organisme, et ne participent pas à l'assimilation des nutriments. Par exemple, ils peuvent avoir des rôles de de protection ou de défense vis-à-vis d'autres bactéries, c'est le cas des toxines. De manière générale, les bactéries peuvent produire et libérer des métabolites secondaires qui peuvent interagir avec leur environnement, et influencer les cellules de l'hôte[121]. Ces métabolites bactériens peuvent être des molécules lipidiques (acides gras à longue chaine, acides gras à courte chaine), des oligosaccharides, ou bien des métabolites issus d'acides aminés (comme le gaba, la sérotonine, l'histamine, les indoles, les catécholamines). Seuls les composés lipidiques seront abordés dans ce manuscrit.

Ces deux types de lipides bactériens seront décrits indépendamment, dans un premier temps nous nous focaliserons sur les lipides structuraux comme les glycérophospholipides et les saccharolipides, puis la seconde partie du chapitre sera dédiée aux lipides produits par le microbiote intestinal comme les acides biliaires ou les acides gras à courtes chaînes.

6.1. Les lipides structuraux

De manière générale, les données bibliométriques concernant les lipides structuraux bactériens sont très disparates parfois même contradictoires et sont peu abondantes rendant l'analyse bibliographique difficile. A l'exception des glycérophospholipides, les lipides majoritaires sont dépendant des bactéries étudiées ; il n'y a pas, à ce jour, de consensus sur les lipides composants les bactéries. Les concentrations et la nature des lipides structuraux varient également en fonction des facteurs environnementaux comme les conditions de culture (aérobie *vs.* anaérobie, liquide *vs.* solide, …), la température et la composition des milieux utilisés[120].

6.1.1. Généralités

Les acides gras bactériens présentent des longueurs de chaines carbonées variables, allant de 10 à 20 atomes de carbone avec une majorité de 15 à 19 atomes de carbone. Cependant, certaines données bibliométriques, notamment chez les mycobactéries mettent en évidence la présence d'acides gras avec des longueurs de chaine beaucoup plus grande : allant de 30 jusqu'à 90 atomes de carbone comme c'est le cas du dihydrophytyl alcool ou de l'acide α-mycolique (Figure 47). Ces acides gras à très longues chaines, sont retrouvés chez les *Actinobacteries* et les *Actinomycétales,* comme les mycobactéries par exemple [122].

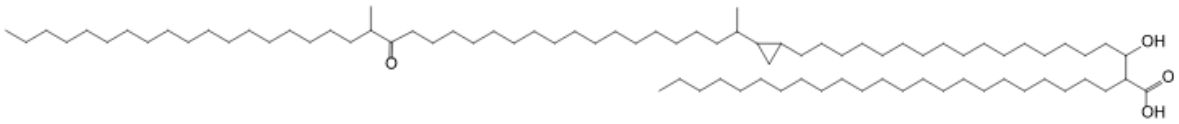


Figure 47 : Acide α-mycolique, acide gras retrouvé chez les mycobactéries d'après [123]

Il existe 4 groupes d'acides gras produits par les bactéries (Figure 48) :

- Les acides gras saturés et insaturés (AGI)
- Les acides gras branchés, qui peuvent être de conformation *iso* ou *antéiso*
- Les acides gras composés d'un groupement cyclopropane

Les acides gras branchés sont retrouvés parfois chez les bactéries gram négatives, ainsi que chez les *Lactobacilles,* mais sont absents chez les autres bactéries à gram positif.

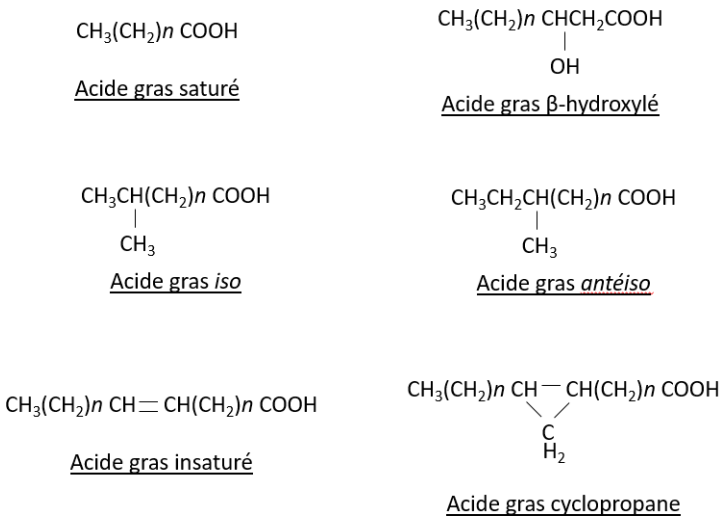


Figure 48 : Schéma des principales classes d'acides gras retrouvés chez les bactéries, adapté d'après[120]

Les acides gras saturés, peuvent être β-hydrolysés en positions 12, 14 16 ou 18. L'acide gras majoritairement β-hydrolysé est l'acide mycolique. Certaines études ont démontré que les acides gras insaturés bactériens possèdent généralement qu'une seule double liaison, se situant entre le carbone C_{11} et le carbone C_{12} et que les acides gras polyinsaturés étaient totalement absents chez les bactéries[120]. Les acides gras branchés majoritaires chez les bactéries sont les C19 :0 *antéiso* et les C19 :0 *iso*[122]. D'autres acides gras, plus rares peuvent être retrouvés chez les bactéries, il s'agit des acides gras aldéhydes et des acides gras diméthyl acétal (Figure 49).

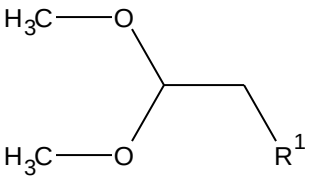


Figure 49 : Structure d'un acide gras diméthyl acétal (ou R^1 est une chaîne alkyle de longueur variable)

Au laboratoire, nous analysons couramment les acides gras totaux (AGT) par chromatographie gazeuse couplée à un détecteur à ionisation de flamme (GC-FID) après une étape d'extraction et d'hydrolyse. Nous avons pu ainsi caractériser les souches bactériennes utilisées durant ma thèse : *Escherichia Coli Nissle 1917* et *Lactobacillus animalis*. Les AG suivant ont été quantifiés : C16:0, C14:0, C18:0, C20:0, C16:1w7, C17:1w7, C18:1w9 cis, C18:1w7, C18:3w3, et C18:4w3 (Figure 50). Nos résultats ne sont pas en adéquation avec les études citées précédemment car des AG polyinsaturés, annoncés comme étant complètement absents du règne bactérien, ont été détectés.

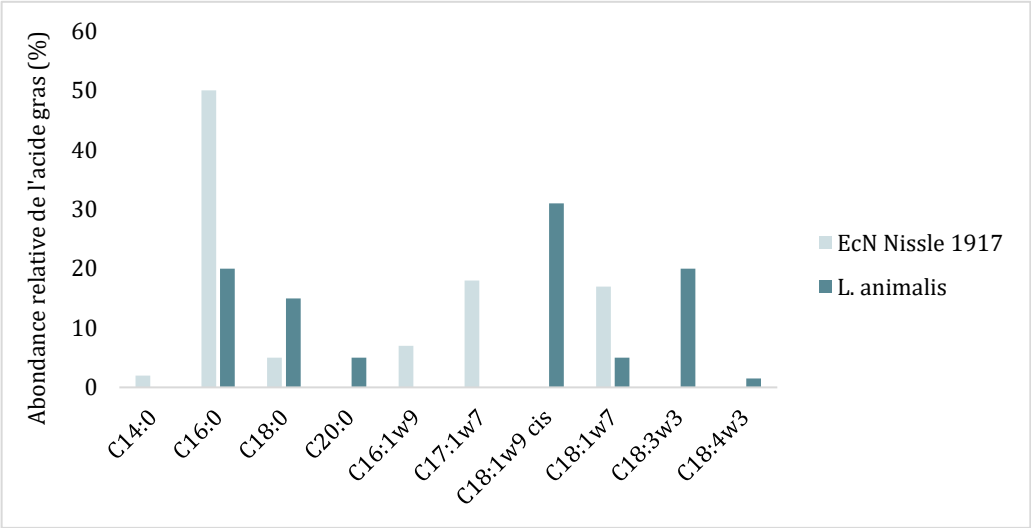


Figure 50 : Dosage relatif des acides gras totaux dans les deux souches bactériennes utilisées dans notre étude

6.1.2. Les glycérophospholipides

Afin de cataloguer l'ensemble des classes lipidiques présentes chez les bactéries, des expériences d'extraction des lipides totaux ont été réalisés, ce qui permet de mettre en évidence la présence de certaines classes lipidiques et de comparer leurs abondances relatives. Les GPG (glycérophosphoglycérols) et les GPE (glycérophosphatydyléthanolamine) sont les lipides les plus abondants quelles que soit les bactéries considérées, excepté pour les *Actinomycetales*, bactéries du phulym *Actinobacteria*. Les biglycophosphoglycérols sont également retrouvés mais en plus faible proportion. Il n'y a généralement pas de PC dans le règne bactérien, ou dans certains rares cas, en très faible quantité. Les PI non plus ne sont pas représentés, à l'exception des dérivés de PI, PI mannosides, chez les *mycobacterium,* les *propionibacterium,* les *corynebacterium* et les *arthrobacteria,* toutes des bactéries du phylum des *Actinobacteria*.

MacFarlane[124] réalise une étude similaire à celle précédente sur différentes bactéries et obtient des résultats semblables confirmant que les classes lipidiques majoritaires chez les bactéries sont les phosphoglycérols, les phosphatidyletanolamines, dans certains cas les phosphoinositols dérivés avec du mannose et que certaines classes de lipides sont rarement représentées dans le règne bactérien comme les phosphatidylcholines. A l'inverse, dans cette étude, les GPE sont retrouvés dans toutes les espèces de bactéries à gram négatif et les bacilles mais pas dans toutes les bactéries.

6.1.3. Les lipides à longue chaine et les lipides complexes

Les saccharolipides :

Les saccharolipides sont composés d'acides gras estérifiés directement liés sur des oses. Il y a peu de bibliographie sur cette classe de lipides, ce sont majoritairement des composés bactériens retrouvé également chez les végétaux. Le saccharolipide le plus connu est le C-acétylglucosamine, précurseur du lipide A (Figure 51), composé que l'on retrouve dans la membrane des bactéries gram négatives[125].

Le lipide A :

Le lipide A bactérien est constitué d'un squelette glucosamine disaccharide phosphorylé en positions 1 et 4 et acylé en positions 2 et 3, sur chacun des monosaccharides[126,127,]. La structure de ce lipide A est majoritairement bien conservée, mais dépend des espèces bactériennes. Habituellement, le lipide A est hexa-acylé, il est constitué de 6 chaines acylées, de longueurs variables, estérifiées avec le squelette disaccharides. Sa biosynthèse est réalisée dans le cytosol des bactéries, en 9 étapes successives, catalysées par des enzymes[128]. Ce sont les chaînes d'acide gras qui permettent la fixation à la membrane. Ce composé est extrêmement toxique et pathogène. Lors de la lyse des cellules bactériennes par le système immunitaire, le lipide A est libéré et peut provoquer des réactions immunitaire perturbant l'homéostasie de l'hôte[129]. Les réactions peuvent être variées : fièvre, diarrhée, et éventuellement choc septique.

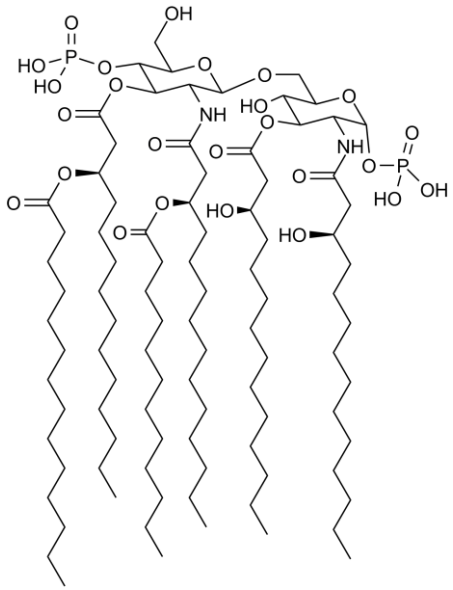


Figure 51 : Exemple de saccharolipide, le lipide A

Les lipopolysaccharides :

Les lipopolysaccharides (LPS) sont de grosses molécules constituées d'un lipide A et de saccharides[130]. On les retrouve dans les membranes des bactéries gram négatives. Ces molécules peuvent se lier au récepteur TLR4, ce qui permet la libération de cytokines (pro-inflammatoires). Dans les laboratoires de recherche, les LPS sont utilisées couramment afin de provoquer une inflammation, en tant qu'endotoxines[131].

Les stérols :

Les données concernant les stérols sont très contradictoires, certaines études affirment que les stérols ne sont pas retrouvés chez les bactéries[120] alors que d'autres ont décrits les voies de biosynthèse des stérols chez les bactéries[132]. Des stérols comme le lanostérol (Figure 52), ou le cycloartenol issus du squalène pourraient être retrouvés dans les bactéries (Figure 53).

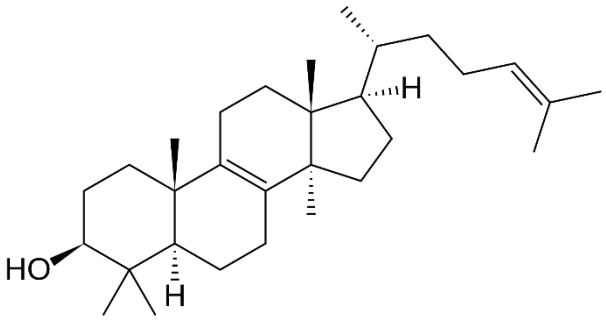


Figure 52 : Structure du lanostérol

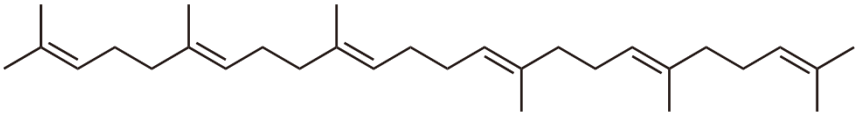


Figure 53 : Structure du squalène

6.2. Les lipides issus du métabolisme du microbiote intestinal

Le microbiote est considéré comme essentiel dans la physiologie de l'hôte, mais les mécanismes d'actions exercés par les bactéries sur l'hôte sont encore mal définit[133]. Les bactéries sont capables de produire et libérer des molécules (métabolites spécialisés) qui vont interagir avec l'environnement et ainsi influencer les cellules de l'hôte. Ces molécules peuvent provenir de la biotransformation de molécules, de la digestion de nutriments par les bactéries, comme par exemple les acides gras à courtes chaines, mais peuvent également être néo-synthétisées par les bactéries comme les commendamides. Il est estimé qu'environ 10% des molécules qui se retrouvent dans la circulation sanguine sont des métabolites issus du microbiote intestinal[121,134] Cette donnée renforce l'importance des métabolites bactériens dans la physiologie de l'hôte. Ces métabolites spécialisés produits par le microbiote seront décrits succinctement dans ce chapitre, cela concerne les lipides transformés par les bactéries dans

un premier temps, c'est-à-dire les acides biliaires et les acides gras à courtes chaînes. Seront ensuite décrits les lipides produits *de novo* par la bactérie, c'est-à-dire les lipides à longues chaînes (LCGA)[119], et les lipopeptides[93]. Les oligosaccharides ou les métabolites issus d'acides aminés comme la sérotonine, l'histamine, ou les catécholamines, bien que synthétisés de novo par les bactéries, ne seront pas décrits.

6.2.1. Les acides biliaires secondaires et les dérivés du cholestérol

Dans le côlon, le microbiote intestinal peut transformer des acides gras n'ayant pas été absorbés dans l'intestin grêle via des réactions enzymatiques. Ces transformations peuvent être des réactions d'hydrolyses, d'oxydations, de réductions ou encore d'hydroxylations. Par exemple, l'acide cholique qui provient majoritairement de la bile mais parfois également de l'alimentation et de la desquamation des cellules épithéliales intestinales, est transformé en coprostanol par le microbiote intestinal[135,136]. Environ 70% du cholestérol est utilisé par l'hôte, les 30% restants se trouvent dans le côlon et sont transformés par le microbiote intestinal. Cependant, le coprostanol n'est pas absorbé, il se retrouve donc éliminé dans les fèces.

Les acides biliaires sont issus de la transformation du cholestérol par le foie (Chapitre 1). Ils sont excrétés par les voies biliaires au niveau du duodénum et sont absorbés au niveau de l'iléon. Ils sont transportés par la veine porte jusqu'au foie, où ils sont à nouveau excrétés dans la bile : c'est le cycle entérohépatique[137]. Chaque jour, environ 5% des acides biliaires produits par l'hôte (ce qui représente environ 0.2 à 0.5 g/jour) échappent au cycle entérohépatique et se retrouvent en contact avec le côlon. Ces 5% d'acides biliaires vont être métabolisés et transformés en acides biliaires secondaires par le microbiote intestinal, via des processus enzymatiques (déconjugaison, oxydation, désulfation, épimérisation)[138]. Il existe plus de 20 acides biliaires secondaires, mis en évidence dans les fèces humaines. Ce sont les bactéries du genre *Clostridium* (une bactérie du phylum des *Firmicuites)* qui possèdent les enzymes permettant ces transformations. Les acides biliaires secondaires qui les plus majoritaires sont l'acide désoxycholique et l'acide lithocholique (Figure 54)[135]. Ces deux acides biliaires secondaires sont produits par la 7-α-déshydroxylase des bactéries, à partir des précurseurs que sont l'acide cholique et l'acide chénodésoxycholique.

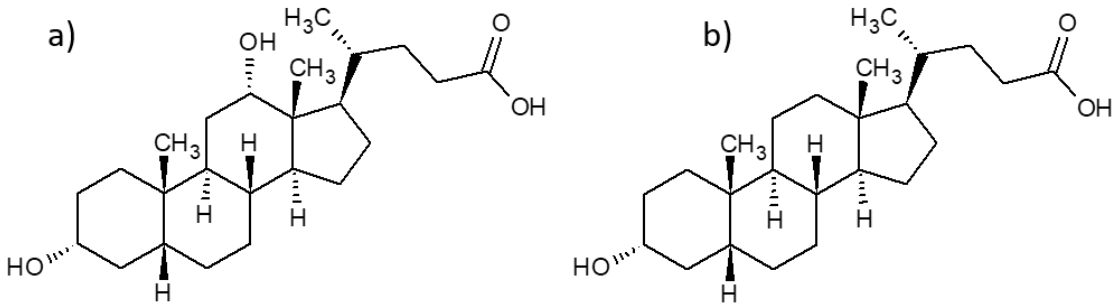


Figure 54 : Structures de a) l'acide désoxycholique et de b) l'acide lithocholique, les deux acides biliaires secondaires majoritaires

Ces acides biliaires secondaires influent sur l'état de santé de l'hôte. Par exemple, l'acide lithocholique est toxique pour les cellules hépatiques et une augmentation d'acide désoxycholique est associée à des risques accrus de cancers cancers coliques[139]. A l'inverse, l'acide ursodésoxycholique (Figure 55a), produit d'épimérisation de l'acide chénodésoxycholique (Figure 55b), présente des activités protectrices et est largement utilisé dans le traitement de maladies cholestatiques, notamment la lithiase ou la cirrhose biliaire[140].

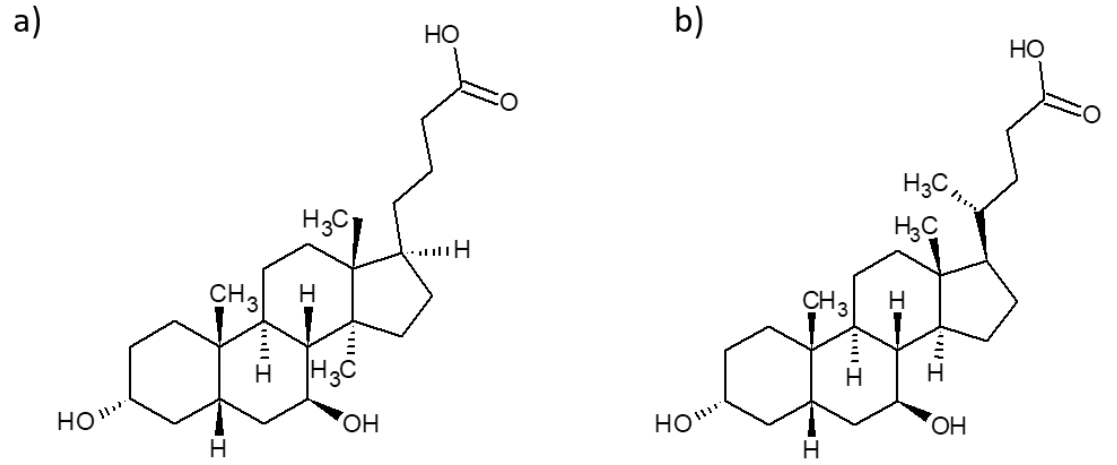


Figure 55 : Structures de a) l'acide ursodésoxycholique et de b) l'acide chénodésoxycholique

6.2.2. Les acides gras à courtes chaînes

Certaines fibres alimentaires sont constituées de polysaccharides, qui sont non digestibles par les enzymes des mammifères. Ces fibres non digestives progressent alors tout au long du tractus intestinal, tout en résistant aux différents processus liés à la dégradation et à la digestion. Au niveau de l'intestin elles se retrouvent alors au contact des bactéries du microbiote intestinal qui sont capables par fermentation[124,141,142] de digérer ces fibres alimentaires. Plus elles progressent vers le côlon, plus la concentration en bactéries est élevée, et donc plus leur fermentation est importante.

La fermentation bactérienne de ces fibres alimentaires va produire des acides gras à courte chaine. Ce sont principalement des bactéries anaérobies qui produisent les AGCC comme des bactéries des phyla *Bacteroides, Actinobactéries* et *Frimicutes* comme les Lactobacilles et les Bifidobactéries[124]. Les espèces majoritaires des acides gras à courte chaine sont les acides formique, acétique, butyrique, *iso*-butyrique, valérique, et *iso*-valérique (Figure 56). Les moins abondants sont les acides caprioïque, heptanoïque, octanoïque, nonanoïque et l'acide caprique. Les AGCC sont des « acides gras » solubles dans l'eau, ce sont donc, en réalité d'un point de vue chimique plutôt des acides organiques.

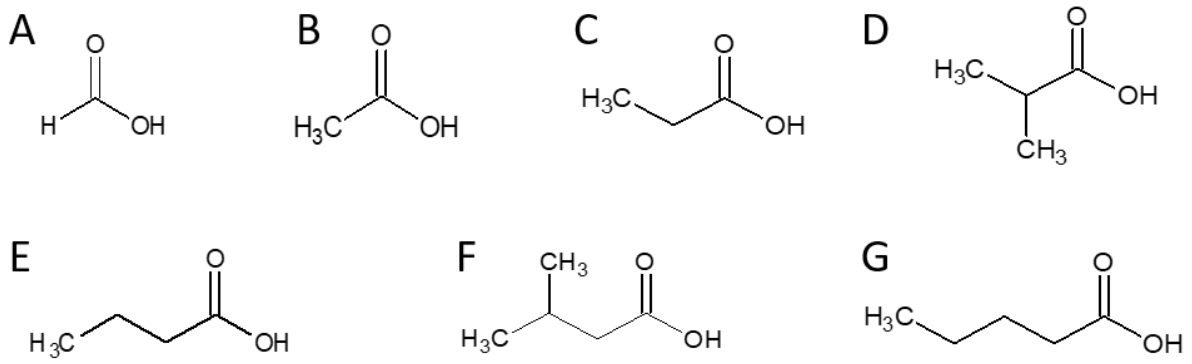


Figure 56 : Structure des acides gras à courte chaine, acide formique CH_2O_2 (A), de l'acide acétique $C_2H_4O_2$ (B), de l'acide propionique $C_3H_6O_2$ (C), de l'acide isobutyrique $C_4H_8O_2$ (D), Structure de l'acide butyrique $C_4H_8O_2$ (E), de l'acide iso-valérique $C_5H_{10}O_2$ (F), et de l'acide valérique $C_5H_{10}O_2$ (G)

Le microbiote intestinal va former trois acides gras à courtes chaines majoritaires, l'acide acétique, l'acide butyrique et l'acide propionique représentant 95% de la production des AGCC totale[143]. En réalité, il y a une disparité importante au niveau de la production de ces trois composés : 60% d'acide acétique, 20% d'acide butyrique et 20% d'acide propionique sont

produits par le microbiote intestinal[144]. Puisque ce sont les 3 acides gras majoritaires, leurs rôles seront décrits par la suite

Les AGCC sont rapidement absorbés par l'hôte *via* une diffusion non ionique à travers les membranes, et presque en totalité, puisque que 90% des AGCC produits vont être absorbés par les colonocytes et sont retrouvés dans la circulation sanguine[145,146]. Ils vont contribuer au métabolisme énergique de l'hôte, notamment au niveau des cellules épithéliales du côlon[147]. Les 10% de AGCC restants se retrouvent éliminés dans les fèces[148]. Cette production de métabolites secondaires peut varier selon la composition des fibres alimentaires, ou de la composition du microbiote intestinal. Ces AGCC constituent une source d'énergie importante pour les cellules du côlon, les muscles, les reins ou encore le cœur[149].

L'acide acétique ($C_2H_4O_2$) est produit majoritairement par les bactéries du phylum des *Bacteroidetes*. Il est synthétisé à partir du pyruvate *via* l'acétyl-CoA[150]. Au niveau intestinal, l'acide acétique favorise la différenciation des cellules lymphocytaires du système immunitaire (notamment les cellules T en cellules T régulatrices)[151]. L'acide acétique exerce donc un rôle majeur dans la tolérance du système immunitaire vis-à-vis du microbiote.

L'acide propionique ($C_3H_6O_2$), est également majoritairement produit par les bactéries du phylum des *Bacteroidetes*. Sa production par les bactéries intestinales est initiée par la conversion du succinate en méthylmalonyl-CoA[152]. Ce composé comporte les mêmes propriétés biologiques que l'acide acétique.

L'acide butyrique ($C_4H_8O_2$) est formé à partir de l'acétyl-CoA, et du butyryl-CoA. Ce composé possède des effets bénéfiques sur l'homéostasie intestinale, et des effets anticancéreux. En effet, le traitement de cellules tumorales par l'acide butyrique induit un processus de mort cellulaire[153,154]. Il est également capable d'inhiber la migration des cellules tumorales[155]. Cet AGCC a également des propriétés anti-oxydantes en diminuant la production de molécules réactives de l'oxygène (ROS). L'acide butyrique est capable d'augmenter l'expression d'enzymes permettant la conversion d'H_2O_2 en H_2O[156].

Etudier les AGCC permet par analogie l'étude de l'activité de fermentation du microbiote intestinal. C'est pour cette raison, que durant ma thèse, un développement analytique par chromatographie gazeuse couplée à un détecteur à ionisation par flamme, visant à doser ces composés a été réalisé. L'ensemble des AGCC sont disponibles commercialement, ce qui permet de réaliser des gammes étalons, et d'effectuer une quantification absolue de ces composés. Afin de valider la méthode, différentes matrices ont été extraites comme les féces de souris, ou encore le contenu cæcal par exemple. Ces travaux n'ont pas fait l'objet d'une publication.

Plus récemment, une nouvelle classe de lipides à courte chaine a été identifiée, il s'agit des AGCC estérifiés par un acide gras hydroxylé (SFAHFA, short-chain fatty acid esters of hydroxy fatty acids)[157]. Les AGCC qui sont estérifiés pour produire des SFAHFA sont les acides acétique et propionique, les acides gras hydroxylés sont constitués d'une chaine carbonée de moins de 20 atomes de carbone (Figure 57). Ces composés ont été retrouvés dans les colons et les contenus caecaux de rats, et sont synthétisés par des bactéries du phylum des *Bacteroidetes*. Pour l'heure, les activités biologiques de ces composés ne sont pas connues.

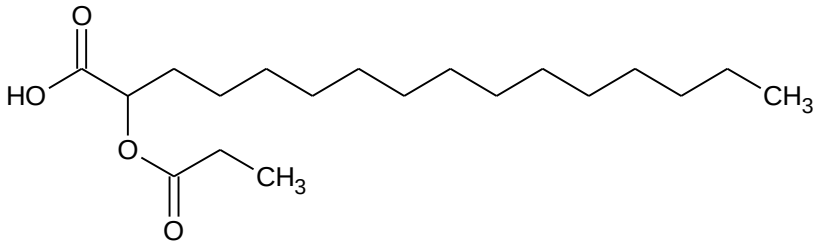


Figure 57 : Exemple d'un acide gras à courte chaine estérifié par un acide gras hydroxylé

6.2.3. Les acides gras à longues chaines

Les bactéries du microbiote intestinal sont également capables de produire des acides gras à longues chaines (AGLC), dont la production est très peu étudiée.

Des travaux antérieurs de l'équipe, portés sur la bactérie EcN Nissle 1917, ont permis de mettre en évidence que ces bactéries étaient capables de produire en grande quantité des d'AGLC hydroxylés en position 3, comme l'acide 3-hydroxyoctadecanoïque (C18-3OH), *via* une synthèse indépendante de polykétides synthases (Figure 58)[119].

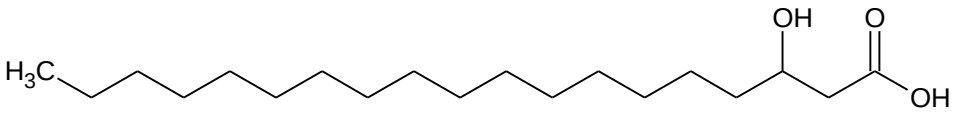


Figure 58 : Structure de l'acide 3-hydroxyoctadecanoïque

Les acides gras hydroxylés peuvent être saturés ou insaturés, avec un ou plusieurs groupement hydroxyles, ils peuvent être classés selon deux groupes : les acides gras mono-hydroxylés en position β ou α, ou les métabolites de l'acide linoléique [158]. Ils sont formés grâce aux enzymes de la famille des cytochromes P450, notamment grâce à la CYP1521A bactérienne[159]. Ces composés lipidiques peuvent s'accumuler dans les tissus, et sont capables de réguler l'expression génique via l'activation de récepteurs nucléaires. L'administration par voie orale à des souris souffrant de colites induit une diminution de l'intensité de l'inflammation via l'activation de PPAR-γ. Ce récepteur nucléaire localisé tout le long de l'intestin grêle et du côlon, possède une activité anti-inflammatoire[119].

Avec comme substrat l'acide linoléique, des bactéries comme les *Lactobacillus* ou les *Streptococcus* produisent *via* des hydratases des métabolites possédant des propriétés anti-inflammatoire. Pour l'instant, les métabolites caractérisés sont : 10-hydroxy-cis-12-octadecenoic acide (HYA), 10-hydroxy-trans-11-octadecenoique acide (HYC), 10-hydroxy-octadecanoicacide (HYB) ; 10-oxo-cis-12-octadecenoique (KetoA), 10-oxo-trans-11-octadecenoique acide (KetoC), 10-oxo-octadecanoique acide (KetoB, Figure 59).

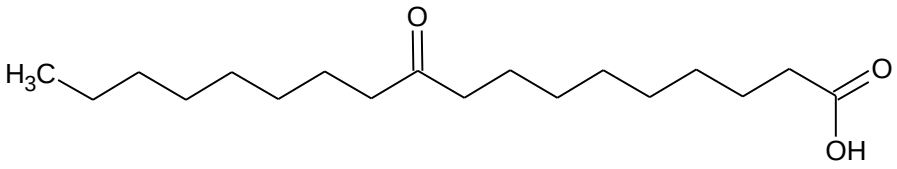


Figure 59 : Structure de l'acide 10-oxo-octadecanoique, KetoB

6.2.4. Les commendamides

Les bactéries du microbiote intestinal sont capables de produire une autre famille de lipides : les commendamides (Figure 60)[160]. Ces lipides ont été identifiés dans des bactéries du genre *Bacteroides*, plus particulièrement chez l'espèce Bacteroides vulgatus, retrouvés dans le tractus intestinal humain. Les commendamides composés d'un acide gras et d'une liaison peptidique sont fortement ressemblants aux anandamides et aux endocanabinoïdes. Les commendamides composés d'un acide gras et d'une liaison peptidique sont comparables aux anandamides et aux endocanabinoïdes des mammifères. Les commendamides sont des molécules de signalisation, activant des récepteurs couplés aux protéines G (RCPG). Potentiellement, les commendamides peuvent avoir un lien dans les maladies auto-immunes et l'athérosclérose[160].

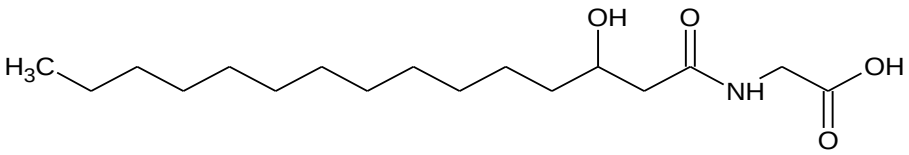


Figure 60 : Structure d'un commendamide

6.2.5. Les polycétides

Les polycétides sont obtenus par polymérisation entre les groupements propionyls et acétyls, sous l'action de la polycétide synthase. La grande diversité enzymatique des polycétides synthases entraine une très grande diversité structurale des polycétides. Les formes majoritaires sont cycliques et leurs structures peuvent être modifiées par glycosylation, méthylation, ou encore hydroxylation. Les polycétides sont majoritairement des métabolites spécialisés retrouvés dans le règle végétal (plantes, champignons) et bactériens. Ils sont connus pour leurs activités biologiques antiparasitaires, antibactériennes ou encore anticancéreuses. L'érythromycine (Figure 61), par exemple, est un polycétique, ayant une activité antibiotique à large spectre, utilisé principalement chez les personnes allergiques à la pénicilline[4].

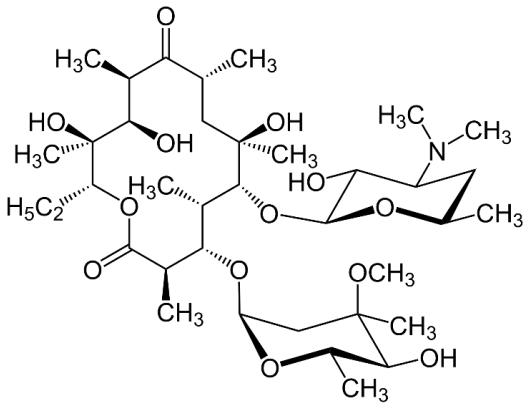


Figure 61 : Exemple de polycétide, l'érythromycine

6.2.6. Les lipopeptides et lipoaminolipides

Les bactéries du microbiote intestinal peuvent synthétiser des lipopeptides, qui sont des métabolites spécialisés : ce sont des lipides associés à des acides aminés, sous forme de peptides. Les lipopeptides sont des molécules peu décrites, qui font l'objet de peu d'études. Ces molécules sont notamment connues pour les propriétés antibiotiques, antivirales ou antifongiques dont elles bénéficient. Elles confèrent un avantage compétitif pour l'occupation des niches microbiennes lors de l'interaction avec d'autres microorganismes. *In vitro*, les lipopeptides des espèces bactériennes *Pseudomonas* et *Bacillus* comme l'amphisine ou la tolaasine ont des activités lytiques, elles inhibent la croissance d'un large panel de microorganismes comprenant les virus, les bactéries et les champignons[161]. Ces molécules sont formées par synthèse non ribosomale qui associe des synthases non robosomales (NRPS) et des polykétides synthases (PKS), qui sont impliquées dans la synthèse de plusieurs molécules, comme les polycétides ou encore de peptides.

Ils sont constitués d'une chaine alkyle liée par une liaison peptidique à un ou plusieurs acides aminés. Certaines espèces de bactéries, comme les *Bacillus* peuvent produire de la surfactine, lipopeptide de haut poids moléculaire ($C_{53}H_{93}N_7O_{13,}$ Figure 62)[162]. Les bactéries *Pseudomonas* peuvent produire de la syringomycine, composé de la sous classe des lipodepsinonapeptide[163,164], qui est également un composé de haut poids moléculaire ($C_{53}H_{85}lN_{14}O_{17}$). Les pseudomonas peuvent aussi synthétiser d'autres lipopeptides comme la tolassine (structure composée de 18 acides aminés, 1985 Da)[165], de la sclérostine ou l'arthrofactine[162].

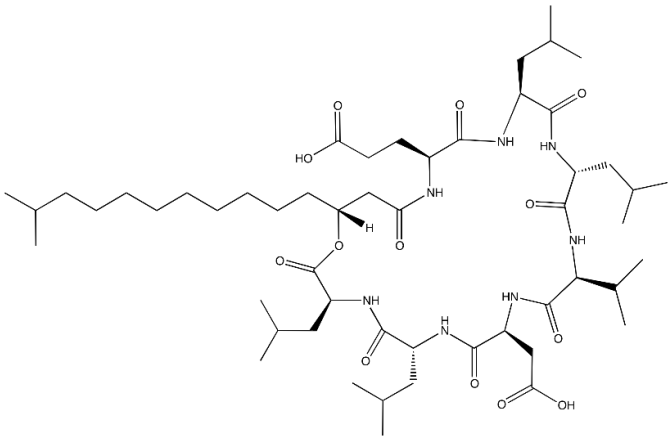


Figure 62 : Structure de la surfactine, ($C_{53}H_{93}N_7O_{13}$), lipopeptide produit par les bacilles

Plus récemment, notre équipe a identifié une nouvelle classe de lipopeptides synthétisés par les bactéries *E. coli*. Ce sont des lipopeptides de plus petite taille (aux alentours de 400 Da) composés d'une chaine alkyle, liée par une liaison peptidique à un seul acide aminé, celui-ci est lié via une liaison peptidique à un neurotransmetteur : l'acide γ-amino-butyrique (GABA) comme le C12AsnGABA ($C_{20}H_{37}N_3O_5$, Figure 63)[93]. Ce lipopeptide est capable de traverser la barrière épithéliale de l'intestin, à l'inverse du GABA seul. Le GABA, qui est l'un des neurotransmetteurs majoritaires du système nerveux central se retrouve alors au contact des terminaisons nerveuses et peut en activant le récepteur $GABA_B$ diminuer l'activation des neurones sensitifs, menant à une diminution de la douleur viscérale. Le traitement de souris ayant une hypersensibilité viscérale mécanique par le C12AsnGABAOH les rendait normosensibles.

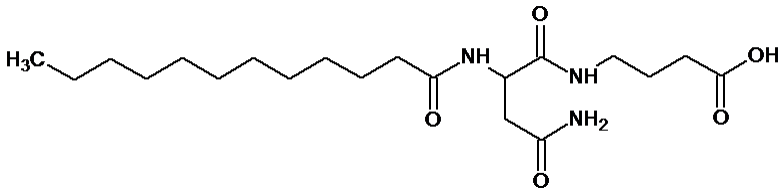


Figure 63 : Structure du C12AsnGaba ($C_{20}H_{37}N_3O_5$), lipopeptide produit par EcN 1917

La production de ces molécules est dépendante au minimum de 3 gènes de l'îlot *pks*. Cet îlot est un cluster de gènes, initialement identifié dans des isolats d'*E. coli IHE 3034*[166]*,* et contient 19 gènes, codant pour des pokyketides synthases (PKS), des synthases non ribosomales (NRPS) et des enzymes et protéines accessoires. Parmi ces gènes, ClbA, ClbB et ClbN sont nécessaires à la production de C12AsnGABA. Cet îlot pks peut également produire la colibactine.une génotoxine impliquée dans les cassures de l'ADN[160]. La colibactine produite par cette machinerie biosynthétique complexe implique l'action de deux protéines : ClbA et ClbS. ClbA est obligatoire dans la synthèse de la colibactine et d'autres métabolites bioactifs. Après son activation par ClbA, l'ajout de l'asparagine sur l'acide tétradécanoïque (acide myristique) est possible : cela mène à la formation du C14Asn (Figure 64). Ce composé est utilisé afin de produire de la précolibactine, *via* l'ajout de différents acides aminés par l'action de malonyol-CoA. La libération du C14Asn peut avoir lieu sous l'action de la peptidase ClbP qui va cliver la précolibactine en colibactine et en C14Asn. Le C14Asn, sous sa forme C14(D)Asn possèdent des activités antagonistes modérées aux récepteurs dopaminergiques et aux récepteurs à la 5-hydroxytryptamine[167] ($5\text{-}HT_7$).

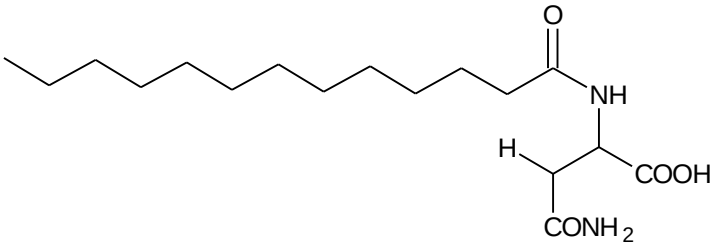


Figure 64 : Structure du C14Asn($C_{17}H_{31}N_2O_4$), lipoaminoacide produit par EcN 1917

Dans les probiotiques *E. coli*, trois autres métabolites issu de l'ilôt pks ont été identifiés : l'entérobactine, la salmochéline et la yersiniabactine. Ces molécules présentent des activités antibiotiques[168], antiparasitaires[168], immunusuppressives[169] et antitumorales[170].

Aujourd'hui, cet îlot pks a été identifié dans d'autres souches bactériennes, comme par exemples chez les enterobacters comme *Klebsiella pneumoniae* et *Citrobacter koseri*[171].

De manière générale, ces composés sont très peu décrits, il existe des lipopeptides cycliques, comme les gacamides, mais ce sont généralement de gros peptides de plus de 6 ou 7 acides aminés. En 2017 des lipoaminoacides ont été décrits, notamment dans le cerveau de rat, comme par exemple le C16Alanine ou encore le C16GABA[29]. En dehors du C14Asn et du C12AsnGABA décrit par notre équipe, les données bibliométriques sur les petits lipopeptides bactériens ne sont pas riches.

L'objectif de mon travail a été d'identifier et nouveaux composés lipidiques bactériens, l'outil de choix pour répondre à cette problématique a été la spectrométrie de masse. Le chapitre suivant de ce manuscrit sera consacré dans un premier temps à une étude bibliographique de la spectrométrie de masse et dans un second temps à la description des différentes techniques utilisées.

PARTIE 2 : ANALYSES LIPIDOMIQUE ET SPECTROMETRIE DE MASSE

Chapitre I : Généralités sur la lipidomique et les sciences du « omiques »

Depuis les années 1980, avec le développement exponentiel des techniques d'analyses, la mise en place d'analyses de plus en plus fines et complexes du vivant a été envisagée. En parallèle, le développement des techniques de bio-informatique a permis d'augmenter la rapidité du retraitement des données indispensables à ces nouvelles analyses. Par exemple, le séquençage de l'entièreté de l'ADN du génome humain, lancé en 1988[172] permet le développement de la « génomique », définie comme un domaine multidisciplinaire, visant à caractériser et à quantifier l'ensemble des gènes codant d'un organisme, à étudier leurs interactions, leurs relations et leurs influences sur l'organisme. Par analogie, d'autres disciplines émergent, visant à étudier un échantillon biologique dans sa totalité : la transcriptomique, qui étudie les ARN messagers, la protéomique qui étudie les protéines, la métabolomique qui s'intéresse à l'étude des métabolites (molécules de poids moléculaire < 1000 Da). La lipidomique qui est une sous partie de la métabolomique, concerne l'analyse des lipides[173].

La lipidomique correspond à l'analyse du lipidome et de ses fonctions[174]. La lipidomique est un terme introduit pour la première fois en 2001 par Kishimoto, et fait référence à l'analyse de la totalité des lipides. En 2003, le GERLI, qui est le groupe d'études et de recherches sur les lipides, définit la lipidomique comme la caractérisation complète de l'ensemble des molécules lipidiques contenues dans une cellule ou dans un tissu[175]. La lipidomique cherche à comprendre l'influence des lipides, sur les différents processus d'un organisme : signalisation cellulaire, composition membranaire, modulation des gènes ou interactions lipides-protéines. L'étape clé d'une analyse en lipidomique est la préparation des échantillons, cela regroupe plusieurs étapes successives comme l'homogénéisation ou la précipitation protéique. Les extraits lipidiques peuvent être analysés par spectrométrie de masse, ou bien par résonance magnétique nucléaire (RMN). Les analyses lipidomiques « single cell » qui émergent depuis peu, ont pour but de ne pas extraire et d'injecter directement une cellule en spectrométrie de masse[176].

1. Préparation des échantillons

La préparation des échantillons biologique est une étape déterminante et nécessaire pour l'analyse des molécules en lipidomique[3], elle comprend plusieurs étapes : le recueil des échantillons, l'arrêt du métabolisme, l'homogénéisation des échantillons et leur extraction. Malgré leur importance les deux premières étapes sont souvent négligées alors que l'extraction est quant à elle très bien documentée dans la littérature à l'aide de protocoles standards d'extractions spécifiques à chaque famille de lipides.

La réussite d'une expérience en lipidomique va dépendre non seulement des conditions d'analyse en chromatographie couplée/ou non à la spectrométrie de masse mais également de l'ensemble des étapes en amont portant sur la manipulation de l'échantillon et sur le choix de la matrice utilisée[177,178]. L'extraction consiste à extraire les molécules d'intérêts de la matrice biologique. Les échantillons peuvent être de différentes origines biologiques : humaine, animale, végétale ou encore bactérienne. Ils peuvent provenir de différentes matrices comme les tissus, les cellules ou encore les fluides biologiques. Toutes ces variations biologiques entraînent des différences de préparation de l'échantillon. Le protocole d'extraction est ajusté en fonction de la nature de l'échantillon étudié et du lipide à analyser. L'extraction se doit d'être la plus reproductive possible afin d'éliminer une source de variation inter-individuelle. Elle est généralement très chronophage et difficilement automatisable.

Dans certains cas une étape de pré-extraction, qui consiste en un broyage de l'échantillon, est nécessaire pour libérer les analytes de la matrice biologique et les rendre accessibles. Une étape d'homogénéisation est nécessaire pour l'extraction des lipides d'un tissu. Afin de suivre la totalité du processus de préparation d'un échantillon il est possible de suivre les rendements d'extraction et de préparation qui représentent le pourcentage d'extraction, propre à chaque composé. Pour cela, un standard interne (ISTD) est ajouté en début de préparation, il s'agit d'un composé absent initialement dans la matrice, et possédant des caractéristiques physico-chimiques identiques ou très proches de l'analyte. Dans le meilleur des cas, un standard interne identique à la molécule d'intérêt mais deutéré ou marqué à l'aide de ^{13}C est utilisé. Suivre le rendement d'extraction de ce composé permet de s'assurer de la qualité de la préparation des échantillons.

Le rendement d'extraction se calcule de la façon suivante :

$$Rendement\ d'extraction = {Aire\ ISTD\ extrait}/{Aire\ ISTD\ non\ extrait} * 100$$

Le rendement de préparation est calculé de la façon suivante :

$$Rendement\ de\ préparation = {Aire\ ISTD\ non\ extrait}/{Aire\ ISTD\ extrait\ avec\ matrice} * 100$$

Le rendement de préparation prend en compte l'effet matrice, qui sera détaillé plus tard dans ce chapitre.

1.1. L'extraction des lipides

La figure 65 représente la relation entre la polarité de certains solvants utilisés en chromatographie liquide et la polarité des classes de lipides, sur l'échelle de Hildebrand. Une étendue de polarité très importante est obsérvée dûe à l'écart de polarité entre les phospholipides possédant une tête polaire et les acides gras comme les triglycérides.

En lipidomique, il existe deux grands types d'analyse : l'analyse spécialisée d'une classe de lipide (par exemple, l'extraction des eicosanoïdes et leur analyse par LC-MS[179]), et les analyses dites globales[180]. Ces dernières ont pour but d'extraire la totalité du lipidome, et de l'analyser par LC-HRMS. Les analyses dites « molécules spécifiques » font appel à des techniques d'extraction spécifiques. A l'inverse pour les analyses dites « globales » l'extraction devra être à large spectre afin de recouvrir toute cette gamme de polarité.

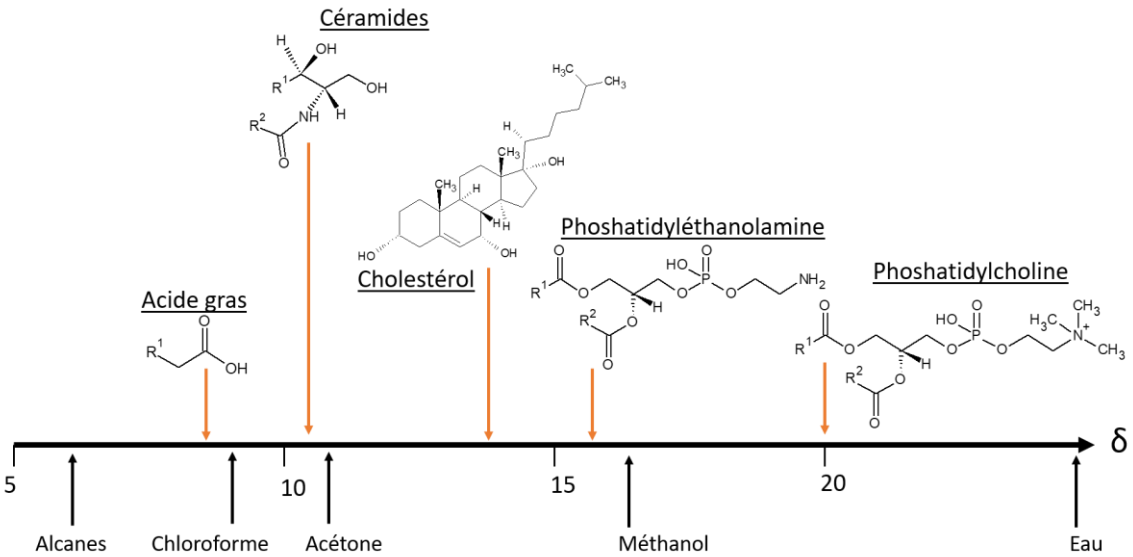


Figure 65 : Gamme de polarité de certains lipides, et de quelques solvants utilisés en chromatographie liquide placés sur l'échelle de polarité de Hildebrand[181]

En lipidomique deux grandes méthodes d'extraction sont couramment utilisées : l'extraction liquide-liquide[182], notamment le protocole de Bligh & Dyer[183], qui est utilisé dans les approches de lipidomique globale[184], et l'extraction sur phase solide (SPE).

Dans certains cas, il peut être nécessaire de réaliser une étape d'hydrolyse afin de libérer les lipides tels que les acides gras qui sont estérifiés aux glycérophospholipides ou aux triacylglycérides[185]. Dans d'autres cas, il est également nécessaire de cumuler les deux méthodes d'extractions (LLE et SPE), comme pour l'extraction des oxstérols[186].

1.1.1 L'extraction liquide-liquide

L'extraction liquide-liquide est une technique d'extraction chimique basée sur deux solvants non miscibles entre eux, générant un système biphasique. La séparation des métabolites est effectuée à l'interface des deux phases et dépend du coefficient de partage de chacun des analytes entre la phase aqueuse et la phase organique. Les lipides sont séparés en fonction de leurs affinités préférentielles pour l'une des deux phases. Le choix des solvants non miscibles entre eux est important et doit se faire en fonction des propriétés physico-chimiques des lipides à extraire et à séparer. Certains paramètres, comme les propriétés physico-chimiques des molécules à extraire ou le nombre d'extractions successives à réaliser pour épuiser totalement la matrice en analyte, sont à prendre en compte afin d'augmenter au maximum le rendement d'extraction.

En lipidomique, deux extractions liquide-liquide sont utilisées en routine, il s'agit des extractions selon les protocoles de Bligh & Dyer[183] et de Folch[187]. Pour l'extraction selon Bligh et Dyer, il s'agit d'un mélange de chloroforme, de méthanol et d'H_2O dans les propositions 8/4/3[183,187]. Le chloroforme faiblement polaire est utilisé en combinaison avec le méthanol, plus polaire, afin d'extraire l'ensemble du lipidome contenu dans un échantillon biologique. Le chloroforme et le méthanol forment la phase organique où se trouvent les lipides. L'eau va former la phase aqueuse, supérieure, contenant les sucres, les protéines, les sels et les autres composés hydrosolubles. Dans le protocole de Folch[187], la différence majeure est la présence de sels, qui modifie donc légèrement les lipides extraits de la matrice. Dans certains cas, le chloroforme est remplacé par le dichlorométhane, solvant moins toxique pour l'expérimentateur aux propriétés similaires[188].

D'autres mélanges de solvants existent et sont utilisés en lipidomique, comme l'extraction au butanol-méthanol (BUME)[189] ou l'extraction au méthyl tert-butyl éther (MtBE)[189], ou un mélange de trois solvants est utilisé, MtBE, MeOH et H_2O (10/3/2,5). Des rendements

d'extraction similaires à l'LLE du Bligh & Dyer ont été observés. Le MtBE étant un solvant de plus faible densité, la récupération de la phase organique est simplifiée puisqu'elle se situe sur le dessus. De manière générale, il existe de nombreuses méthodes d'extraction liquide-liquide[190], permettant d'extraire le lipidome et adaptées aux molécules souhaitées.

1.1.2 L'extraction en phase solide (SPE)

Avant d'effectuer une SPE, certaines étapes de prétraitement de l'échantillon sont parfois nécessaires afin de réduire les interactions entre la matrice et les lipides. Une étape de précipitation protéique permet d'éliminer les protéines pour éviter qu'elles entrent en interaction avec la phase adsorbante et détériorent la qualité de la séparation chromatographique et donc diminuent la sensibilité de détection (augmentation du bruit du fond). Les solvants organiques ou de fortes concentrations en sels induisent des altérations de la structure tridimensionnelle des protéines induisant leur précipitation, le MeOH est le solvant le plus efficace.

L'extraction sur phase solide repose sur l'adsorption des analytes (ou des contaminants selon les cas) sur une phase stationnaire solide (appelée adsorbant) et leur élution *via* une phase éluante liquide. La SPE peut se dérouler sur colonne ou sur plaque 96 puits. Cette extraction est divisée en plusieurs étapes. La première consiste à conditionner l'adsorbant, c'est-à-dire à l'activer avec un solvant organique, de l'eau, ou avec un mélange de solvants. Ensuite, l'échantillon est déposé sur l'adsorbant, plusieurs lavages successifs sont alors réalisés, permettant l'élimination des composés interférents. Enfin, l'élution des composés d'intérêts a lieu, *via* l'utilisation d'un solvant éluant adapté. L'affinité des composés restant sur la phase doit être plus grande envers le solvant pour pouvoir les décrocher (figure 66).

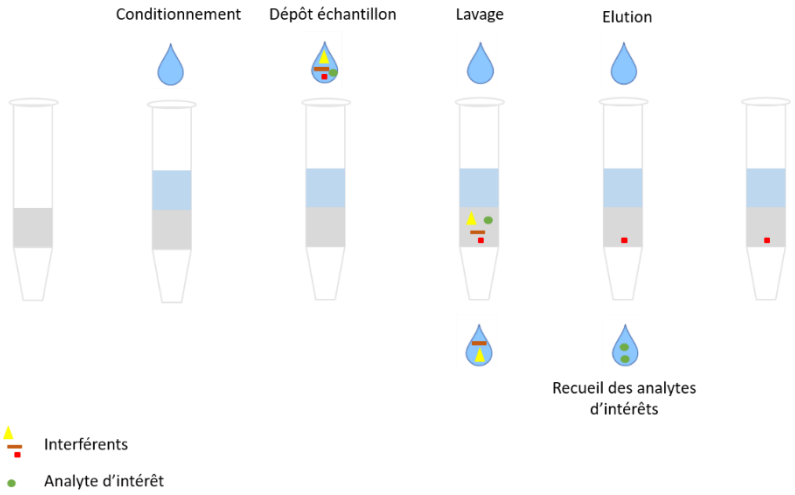


Figure 66 : Principe d'une extraction sur phase solide

Plusieurs types de phases adsorbantes composées de polymères ou de silice existent. Le choix de cette phase est très important dans le développement d'une méthode d'extraction en lipidomique et dépend des analytes à extraire, de leurs concentrations, du volume d'échantillon à déposer et du type d'échange recherché. Les phases polymériques offrent l'avantage d'être très stables et résistent à une large gamme de pH allant de 0 à 14. Les phases constituées de silice sont moins stables et supportent des variations de pH allant de 2 à 7.5 mais sont très sélectives.

Dans les adsorbants de silice, il existe plusieurs sous-type de phases : phase normale, phase inverse, silices échangeuses d'ions ou les phases mixtes. Les phases normales sont constituées de greffages DIOL, CN ou NH_2 et permettent l'extraction de molécules ayant des fonctions polaires. Les adsorbants en silice de phase inverse sont basés sur des interactions de type Van der Waals et permettent l'extraction des composés apolaires ou très faiblement polaires. Les phases échangeuses d'anions et de cations fonctionnent selon le même principe.

Les phases échangeuses d'anions fonctionnent selon un mécanisme de rétention lié à l'interaction ionique. L'extrémité du greffon de la phase (Figure 67) va créer une interaction très forte avec les composés de l'échantillon, portant des fonctions ionisables opposées. L'étape de lavage est réalisée *via* l'utilisation d'une base, puis d'un acide faible afin d'éliminer les cations et les anions faibles. L'interaction des phases échangeuses d'ions dépend essentiellement du pH et de la force ionique du contre-ion qui va permettre l'élution des

composés capturés par l'ion greffé de la phase. Il existe plusieurs types de greffons sur ces phases, comme par exemple les phases constituées d'amines quaternaires (SPE commerciales « SAX » interchim, SPE Oasis MAX, Waters). Au laboratoire, nous utilisons ce type d'interactions afin d'extraire les isoprostanes[191].

Les phases échangeuses de cations (de type commerciales « MCX » Waters, greffage représenté dans la Figure 67), permet l'extraction de composés basiques. Les plaques SPE de ce type sont recommandées pour le lavage de 99% des phospholipides[192].

Les phases polymériques, comme les phases « hydrophilic-lipophilic balance » (HLB) sont des phases autant hydrophiles qu'hydrophobes (greffage représenté sur la figure 67). Elles permettent l'élution de presque tous les composés (neutres, basiques, acides, pH 0 à 14). Au laboratoire, nous utilisons cette méthode d'extraction afin d'extraire les eicosanoïdes[179,193].

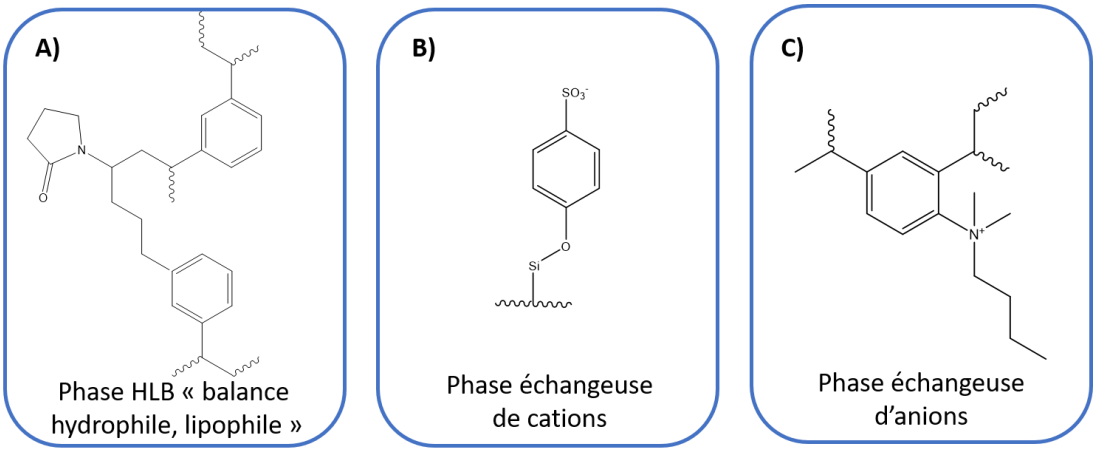


Figure 67 : Principales types de phases adsorbantes retrouvés dans les protocoles de lipidomique.

2. L'influence de la matrice

Lors de l'extraction des métabolites d'intérêt, la matrice biologique exerce un impact non négligeable appelé effet matrice. En effet, le rendement d'extraction va varier si l'extraction est réalisée à partir d'un fluide biologique (sérum, plasma), de cellules, ou de tissus (intestins, poumons, cœur…). Il s'agit de matrices complexes c'est-à-dire de matrices biologiques très riches, composées d'un nombre important de métabolites (primaires et secondaires), d'ADN, de protéines, … qui peuvent interférer avec les métabolites à extraire. Ces interférents peuvent être d'origine endogène, exogène, organique ou minérale. Lors de l'analyse en spectrométrie de masse, les interférents peuvent induire une suppression du signal. S'ils sont présents en trop forte quantité, ils vont absorber une partie conséquence des protons disponibles dans le milieu et empêcher les métabolites d'intérêts de s'ioniser et d'être détectés par l'instrument. L'influence de la matrice sur l'analyse peut se calculer de la manière suivante :

$$Effet\ matrice = {}^{ISTD\ non\ extrait}/_{Matrice\ extraite + ajout\ ISTD\ avant\ injection} * 100$$

Chapitre II : L'analyse des lipides

En raison de la complexité des extraits lipidiques, l'analyse par spectrométrie de masse est généralement associée à une technique séparative, mais peut être également effectuée en introduction directe dans le spectromètre de masse lors des analyses de shotgun lipidomique[194]. Les limites de cette approche sont nombreuses car une forte suppression du signal est observée, uniquement les lipides majoritaires seront détectés. Certains lipides ne sont visualisés en spectrométrie de masse uniquement si l'utilisateur de l'argent ou le lithium[195] pour générer des adduits, ce qui rend les analyses difficiles et augmente considérablement le nombre d'injections. Certains composés s'ionisent uniquement dans un mode, c'est par exemple le cas des anandamides ou des acyl-serotonine qui ne sont détectés qu'en ionisation positive alors que d'autres composés s'ionisent qu'en mode négatif : il est donc nécessaire d'injecter les échantillons dans les deux modes.

L'extrait lipidique peut également être analysé directement en RMN du ^{1}H ou du ^{31}P permettant une analyse globale quantitative mais la complexité des extraits lipidiques complique l'interprétation des spectres en raison d'un grand nombre de recouvrement des raies. Cependant, cette technique gourmande en matériel est non destructive et permet d'obtenir une empreinte lipidique quantitative de l'échantillon ce qui n'est pas le cas de la spectrométrie de masse.

Majoritairement, les lipides sont analysés en chromatographie liquide couplée à la spectrométrie de masse (LC-ESI/MS), à un DEDL ou à une détection corona, ou en chromatographie gazeuse couplée à un détecteur à ionisation de flamme (FID), en chromatographie gazeuse couplée à la spectrométrie de masse (GC-EI/MS).

Dans ces travaux nous nous sommes uniquement focalisés sur l'analyse par LC-MS, qui est de loin la technique la plus utilisée en lipidomique, nécessitant dans notre cas une extraction lipidique en amont.

1. Les méthodes de séparation des lipides

Historiquement c'est le botanique Mikhaïl Tswett qui s'intéressait à l'étude des pigments végétaux qui a développé et expliqué les techniques de séparation, ayant conduit aux bases de la chromatographie moderne dès les années 1900. Tswett a utilisé de l'éther de pétrole comme phase mobile, ainsi qu'une colonne remplie par de la craie (du carbonate de calcium) comme phase stationnaire. Il a réussi à détecter des zones de pigments différents, grâce à leurs différentes couleurs. A la fin des années 1970, la chromatographie moderne arrive dans les laboratoires et les industries *via* l'introduction de la chromatographie sur couche mince. Ces techniques se sont rapidement répandues dans toutes les industries chimiques : cosmétiques, para-chimiques ou encore pharmaceutiques. Aujourd'hui, les deux techniques séparatives majoritairement utilisées en lipidomique sont la chromatographie en phase gazeuse et la chromatographie liquide.

1.1 La chromatographie en phase gazeuse

La chromatographie en phase gazeuse (GC) est une méthode de séparation des molécules sous forme gazeuse, utilisée pour la première fois en 1941 par Martin et Synge. Les premières analyses de lipides, et plus particulièrement des acides gras libres, ont été réalisées en chromatographie en phase gazeuse, en 1952 par James et Martin[196]. En 1955, seulement 40 publications portent sur la GC, en 1966 ce chiffre passe à 10 000. Ces données permettent de mettre en évidence l'émergence d'une nouvelle technique analytique puissante.

En GC, la séparation des lipides est faite selon l'affinité qu'ils ont pour la phase stationnaire et leur température de vaporisation. La séparation des composés est effectuée en fonction de leur partition entre la phase mobile (gaz) et la phase stationnaire (colonne chromatographique). L'appareil de chromatographie est composé de 3 pièces principales. Pour commencer, il possède un injecteur permettant la volatilisation de l'échantillon contenant les analytes et entrainant les molécules en tête de colonne. Puis, il est composé d'une colonne de chromatographie comportant la phase stationnaire, celle-ci va séparer les analytes, elle est placée dans un four thermostaté (Figure 68). En GC, les colonnes de chromatographie sont des colonnes capillaires, de grande longueur (parfois jusque 100 mètres de longueur).

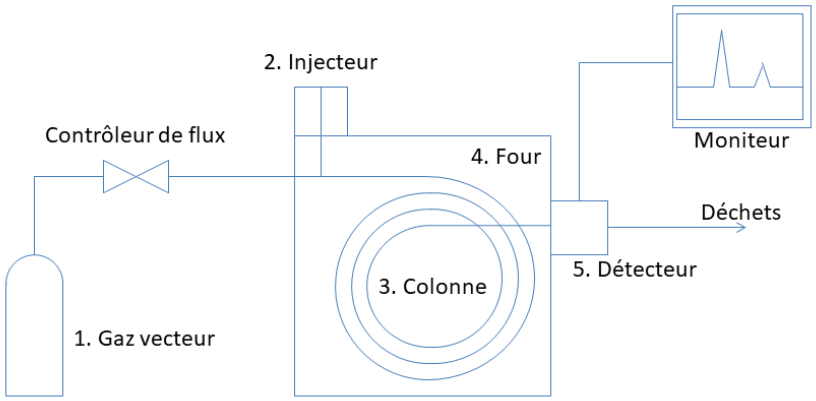


Figure 68: Représentation schématique d'un système de chromatographie gazeuse d'après[197]

Les colonnes capillaires de GC sont composées de deux éléments principaux : le tube et la phase stationnaire. La phase stationnaire est un film fin revêtu de polymère de haut poids moléculaire, thermiquement stable qui recouvre la paroi interne du tube. Son diamètre interne est compris entre 0.05 et 0.53 mm, l'épaisseur de ce film est proportionnellement liée à la rétention des analytes. Le pouvoir de résolution de la colonne est inversement proportionnel au diamètre de la colonne. Pour deux colonnes dont l'épaisseur du film est constante, diminuer son diamètre par 2 revient à doubler son efficacité.

La phase mobile est constituée d'un gaz inerte servant de gaz vecteur, qui va se déplacer à l'intérieur de ce tube, permettant la séparation des composés. Le gaz pousse les analytes, présents dans la colonne sous formes volatiles, vers le détecteur placé en sortie de colonne. Cette phase mobile est composée généralement d'hélium, d'azote ou de dihydrogène. Un système détendeur-régulateur va réguler les arrivées de gaz vecteur. La nature du gaz vecteur ne modifie pas de manière significative la séparation chromatographique. En revanche, la viscosité et la vitesse du gaz dans la colonne ont une influence sur la dispersion des composés dans la phase stationnaire. Les phases stationnaires les plus répandues sont des polymères siliconés dérivées du diméthyl polysiloxane. Elles sont greffées sur une colonne en silice *via* des liaisons O-Si-O. Des groupements CH_3 peuvent alors être greffés sur la silice. La polarité de la phase stationnaire va dépendre de ces groupements substitués et de leurs quantités relatives. Les colonnes non-polaires (composées à 100% de diméthyl polysiloxane, comme les colonnes commerciales « DB 1ms ou DB 5ms ») sont conseillées pour l'analyse des amines, des pesticides, des hydrocarbures, des composés soufrés, des stérols, ou des arômes et parfums. Les colonnes moyennement

polaires seront greffées de 90 à 65% de diméthylpolysiloxane, et de phényles, (colonnes commerciales « DB-35ms ») et sont conseillées pour l'analyse des médicaments, des glycols, des pesticides et des stéroïdes. Les colonnes polaires, sont composées de polyéthylène glycol, comme les DB-wax/ Stabilwax, et sont utilisées pour l'analyse des acides gras à longue chaîne, des acides organiques libres comme les SCFA. Enfin, la GC est également constituée d'un détecteur placé en fin de colonne dont l'objectif est la détection des analytes à la sortie de la colonne chromatographique.

L'analyse des lipides en GC est décrite notamment pour les acides biliaires, les stéroïdes, les acides gras à longues et courtes chaînes, hydroxylés ou non[198,199]. Cependant, tous les composés ne peuvent pas être analysés en GC, certains lipides thermosensibles ne peuvent pas être vaporisés ce qui rend leur analyse en GC impossible. Il est nécessaire que les analytes soient thermiquement stables et possèdent une pression de vapeur suffisante pour être volatilisés en GC. Ce n'est pas le cas de toutes les classes de lipides, parfois trop sensibles à la chaleur ou trop peu volatiles[200]. Le manque de volatilité peut être pallié par une étape de dérivatisation, consistant à diminuer le point de volatilité d'un composé. Par exemple les acides gras sont dérivés en esters méthyliques pour être analysés. Ces esters méthyliques sont préparés par transestérification des TG ou des PL à l'aide d'un catalyseur acide ou basique (HCl, BF3, méthanolate de sodium)[201]. Cette technique peut être couplée à deux types de détecteur : le FID ou bien, un spectromètre de masse à l'aide d'une ionisation à impact électronique ou chimique[202,203].

Cette technique de séparation des analytes est essentiellement utilisée, en lipidomique pour l'analyse des acides gras et des stérols. Elle sera utilisée uniquement pour l'analyse des acides gras à courte chaine, pendant mes travaux de thèse. L'ensemble de ces résultats ne seront pas présentés dans cette thèse.

1.2 La chromatographie liquide

La chromatographie liquide est une technique complexe, où plusieurs facteurs, listés ci-dessous peuvent influer.

- Le facteur de rétention ou de capacité k
- Le facteur de sélectivité α
- L'efficacité N
- La granulométrie ΔP
- La résolution R

Le facteur de capacité ou de rétention est un paramètre expérimental qui permet de décrire la vitesse de progression des solutés dans la colonne, il représente la mesure universelle de la rétention, déterminée à partir du temps de rétention (noté Tr) et du temps mort (noté Tm ou T0). Le temps de rétention est propre à chaque analyte, le temps mort est quant à lui, le temps nécessaire à la phase mobile pour traverser la colonne chromatographique. Ainsi, le facteur de capacité est défini par la formule suivante :

$$Facteur\ de\ capacité\ k = {(Tr - Tm)}/{Tm}$$

Plus le temps de rétention est élevé, plus le facteur de capacité augmente, ce qui signifie que la colonne retient beaucoup la molécule. Cette valeur de k est utilisée afin de comparer les temps de rétention sur les différents systèmes, puisqu'elle est indépendante de la longueur de la colonne et du débit appliqué. Lorsque l'on augmente la longueur de la colonne, on augmente proportionnellement les valeurs de Tr et de T0, et donc k reste inchangé. La valeur de k reflète la séparation chromatographique, lorsque k est situé entre 2 et 10, la séparation est optimale. A l'inverse, lorsque k est au-delà de 30, alors la séparation chromatographique est lente, et lorsque k est inférieur à 1, la séparation est mauvaise. En chromatographie liquide k peut être optimisé en agissant sur la composition de la phase mobile et de la phase stationnaire.

Le facteur de sélectivité α, représente la capacité de la colonne à séparer deux analytes voisins. Ce paramètre intervient dans la finesse des sommets des pics chromatographiques. Il est lié à la différence d'interaction de chaque composé avec les phases mobiles et stationnaires. Il se calcule selon la formule suivante, où α doit être supérieur à 1, pour que la séparation des analytes soit optimale.

$$\alpha = {Distribution\ du\ soluté\ le\ plus\ retenu}/{Distribution\ soluté\ le\ moins\ retenu}$$

L'efficacité N représente le nombre de plateaux théoriques que possède la colonne, et exprime la finesse des pics chromatographiques. La chromatographie est un phénomène continu et dynamique où chaque soluté se déplace progressivement en une suite d'étapes

distinctes. Ce processus se caractérise par un cycle d'adsorption et de désorption. Chaque cycle correspond à un nouvel état d'équilibre de toute la colonne. Un plateau théorique est défini comme une portion de la colonne, de longueur L, dans laquelle l'équilibre de répartition entre les deux phases est réalisé. Pour chaque plateau théorique N, la concentration en analyte est à l'équilibre entre les deux phases. A chaque nouvel équilibre, le soluté avance d'un plateau théorique dans la colonne. Pour une longueur L constante, l'efficacité d'une colonne augmente lorsque N augmente ou si H (hauteur du plateau) diminue. Cette efficacité est dépendante de la nature de la colonne et du type de phases qui la compose. Plus une colonne possède un nombre de plateaux élevé, plus elle est efficace, de la même manière, plus la hauteur du plateau H est faible, plus la colonne est efficace. L'efficacité N est calculée selon la relation suivante :

$$N = {}^{L}/_{H}$$

La hauteur du plateau (HEPT ou H) est essentiellement liée aux caractéristiques de la colonne : son diamètre, sa porosité, la taille de ses particules ou encore sa qualité de remplissage.

Lors du développement d'une méthode analytique, l'ensemble de ces paramètres, le type de colonnes et les phases mobiles sont à prendre en compte en fonction de l'analyte. Plusieurs types de techniques de chromatographie en phase liquide existent, elles peuvent être classées en fonction de la polarité de la phase stationnaire. Deux phases stationnaires sont couramment utilisées et permettent de distinguer les chromatographies liquides à phase polaire ou phase normale (P-LC) et celles à phase stationnaire apolaire ou de phase inverse (RP-LC). D'autres phases sont toutefois utilisées pour l'analyse de certaines classes de lipides, comme la chromatographie d'interactions hydrophiles (HILIC). L'analyse des lipides nécessite l'utilisation d'une gamme étendue de supports chromatographiques en lien avec la diversité de polarité et les propriétés physico-chimiques de ces molécules. Dans cette partie seront détaillé la sélectivité des colonnes, les tailles des particules et les principales colonnes utilisés pour les analyses en lipidomique.

La granulométrie correspond à la perte de charge ΔP, qui représente la différence de pression entre le haut et le bas de la colonne. Plus la granulométrie est fine, plus la perte de charge est importante. C'est pour cela qu'en diminuant la longueur de la colonne, le temps de rétention est diminué. Les plus petites particules offrent une efficacité élevée, cependant, elles entraînent davantage de pertes de charges.

La résolution R est le dernier paramètre à prendre en compte, celui-ci fait intervenir l'ensemble des paramètres décrits jusqu'à présent. La résolution mesure le degré de séparation à la base des pics. Il se calcule de la manière suivante :

$$R = {Tr2 - Tr1}/{\omega 2 + \omega 1}$$

Ou Tr est le temps de rétention et ω la largeur à la base du pic chromatographique

Ce rapport de résolution est optimal lorsqu'il est proche de 1, dans ces conditions les deux analytes 1 et 2 sont efficacement séparés. Lorsque ce rapport est en dessous de 1, alors la résolution est insuffisante, les pics correspondants aux analytes 1 et 2 se recouvrent.

1.2.1. La chromatographie liquide à polarité de phase normale

Lors d'une analyse par chromatographie liquide sur phase normale, la phase stationnaire est polaire et la séparation des analytes est basée sur l'échange d'interactions polaires. La phase stationnaire est, par exemple, composée de silice ou de groupement amine. La rétention des métabolites analysés en NP-LC dépend donc de la nature de leur tête polaire[204]. Dans certains cas, un sel comme l'acétate d'ammonium peut être ajouté aux phases mobiles, pour permettre une meilleure séparation chromatographique, des pics chromatographiques affinés et une meilleure détection des lipides. L'analyse par NP-LC permet de séparer les lipides selon leurs classes et polarité. Elle est utilisée, par exemple pour l'analyse des glycérophospholipides qui sont séparés par sous-familles en fonction de la nature de leurs têtes polaires.

1.2.2. La chromatographie liquide à polarité de phase inverse

Lors d'une analyse par RP-LC, la phase stationnaire est généralement composée de silice greffée de chaînes alkyles (de C4 à C30), de groupements cyano ou parfois de groupements phényls. La phase mobile utilisée est généralement composée d'un mélange hydro-organique. Généralement en début d'analyses, les conditions sont polaires et constituées d'un mélange aqueux/méthanol ou aqueux/acétonitrile. La force éluante va être augmentée au cours du gradient, par ajout de solvants moins polaires qui constituent la phase mobile. L'élution des composés est réalisée selon la théorie solvophobique qui considère que la rétention des analytes est liée au contact entre eux et les chaines alkyles de la phase stationnaire solvatée. Ainsi, ce sont les analytes les plus apolaires qui auront une plus grande interaction avec la phase, et seront alors retenus davantage que les composés polaires. De même, les lipides présentant une ou plusieurs insaturations sont élués plus tôt que les lipides saturés[20]. Il est possible de prévoir la rétention des composés en phase inverse en calculant le nombre d'équivalent carbone (EqC). EqC = CN-2n, où CN est le nombre de carbones des chaînes et n le nombre d'insaturation. L'EqC permet de prévoir les éventuelles co-élutions des lipides en chromatographie liquide lorsque l'on utilise une colonne phase inverse. L'addition d'acide formique à 0.1% permet d'améliorer la séparation chromatographique en affinant la largeur des pics. Ces colonnes sont utilisées pour l'analyse des eicosanoïdes[179], des isoprostanes[191], ou des aminolipides[93].

1.2.3. Développement des systèmes chromatographiques et dernières innovations en colonnes chromatographiques

Depuis une quarantaine d'années, de nouvelles technologies concernant les systèmes chromatographiques émergent. Dans un premier temps, les diamètres internes des tubulures (connectiques dans le système chromatographique) diminuent de plus en plus, (moins de 10 µm actuellement) pour diminuer les volumes morts. En parallèle, les systèmes d'injection permettant d'injecter des volumes d'échantillons de plus en plus restreint se développent, permettant d'atteindre une limite d'injection minimale de 0.1 µL pour les derniers appareils commerciaux. Avec le développement de colonnes chromatographiques de diamètre interne plus faible : microLC et nanoLC, pouvant atteindre 0.075 mm, les analyses LC-MS deviennent alors de plus en plus sensibles permettant de mieux détecter les composés (plus faible limite de détection) et de mieux les séparer. Lorsque des colonnes de diamètres internes plus faible sont utilisées, les débits de solvants chromatographiques sont diminués provoquant une meilleure focalisation du spray, un meilleur transfert de ce spray dans le spectromètre de masse et une meilleure ionisation des analytes. Il en résulte une meilleure détection. En conséquence, les analyses LC-MS deviennent miniaturisables, de plus en plus rapides et de ce fait de plus en plus efficaces. En lipidomique, ce sont les colonnes de microLC qui sont essentiellement utilisées afin de miniaturiser les analyses, les colonnes de nanoLC sont utilisées pour la protéomique.

Depuis 2007, les technologies dites « coreshell » se développent (Figure 69). Elles ont étés initiées par Horváth, puis par Kirkland. Ces supports de particules de silices sont superficiellement poreuses, et sont composées d'un centre non poreux « core », sur lequel se fixe une enveloppe « shell », de silice poreuse.

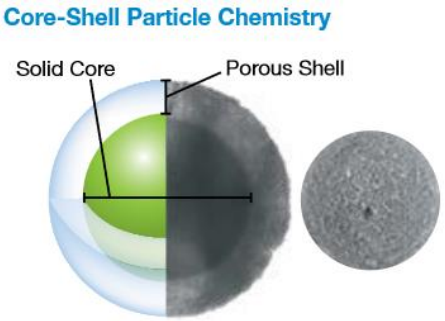


Figure 69 : Structure d'un particule coreshell[205]

Ces colonnes coreshell sont mieux remplies, leurs tailles sont plus homogènes et plus régulières comparé à celles constituées uniquement de particules poreuses. Ces particules poreuses entrainent le phénomène d'anisotropie d'écoulement, due à la non-homogénéité des tailles des particules, ou encore à un mauvais remplissage de la colonne. Ce manque d'uniformité entraîne la création de différents chemins de longueurs variables que la phase mobile peut traverser ayant pour conséquence un élargissement des pics chromatographiques (Figure 70).

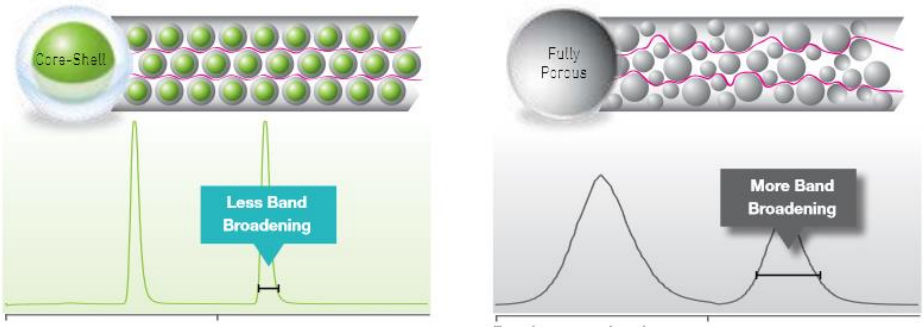


Figure 70 : Comparaison des pics chromatographiques obtenus avec deux colonnes, l'une constituée de la technologie coreshell, l'autre constituée de particules totalement poreuses[205]

Cette technologie Coreshell qui s'accompagne d'une diminution du diamètre des pores et des particules permet de limiter la diffusion coaxiale des analytes entrainant une meilleure séparation, une meilleure résolution et une meilleure reproductibilité des analyses par rapport à des analyses effectuées sur colonnes complètement poreuses

C'est l'ensemble de ces évolutions, qui permettent de développer des analyses LC-MS de plus en plus fines, permettant de détecter des analytes de plus en plus minoritaires et faiblement représentés dans l'échantillon. Ces dernières technologies permettent également des gains de sensibilité importants, (section développée en détails dans la partie 5 du manuscrit). Cependant, cette miniaturisation des analyses LC-MS s'accompagne de modifications des différents paramètres du spectromètre de masse, qui sont parfois chronophages.

Chapitre III : La spectrométrie de masse

Différentes techniques analytiques permettent la détection et l'identification de biomolécules. Les deux méthodes qui sont actuellement les plus utilisées sont la résonance magnétique nucléaire et la spectrométrie de masse. La résonance magnétique nucléaire a été la première méthode utilisée en métabolomique, et se trouve être la seule technique à donner une structure moléculaire de façon générale et cela le plus souvent sans ambiguïté. C'est une technique rapide, non destructrice, robuste, quantitative et applicable à toutes les biomolécules. En contrepartie, il est préférable d'étudier des composés purifiés, et en quantité suffisante (de l'ordre de plusieurs microgrammes)[206,207] pour des analyses RMN. En métabolomique et plus particulièrement en lipidomique, les mélanges étudiés sont généralement complexes, et les quantités dispersées de composés sont souvent représentés à l'échelle du nanogramme rendant l'analyse par RMN délicate[208]. Le couplage LC-RMN pourrait être une solution, cependant malgré le système développé par Brucker, il ne s'est pas popularisé comme l'ont été les couplages GC-MS ou LC-MS.

La spectrométrie de masse, est une méthode alternative, qui peut permettre dans certains cas de déterminer des structures moléculaires à partir de mélanges complexes grâce aux différents types de couplages possibles avec des méthodes séparatives comme la GC ou la LC. La détection et l'identification de nouvelles molécules est possible en routine dès que les composés sont connus et présents dans une banque de données. Toutefois quand il s'agit de composés *de novo* la détermination de leurs structures reste toujours difficile, et implique souvent des experts, les déterminations structurales peuvent être très longues et laborieuses. A l'inverse de la spectrométrie de masse, la RMN se base uniquement sur des couplages chimiques entre certains atomes (par exemple proton-proton) pour déterminer leur environnement et des déplacements chimiques basés sur des effets électroniques liés au groupements fonctionnels plus ou moins polaires, et donc prédictibles. La détermination structurale d'un composé en spectrométrie de masse implique son ionisation et la dissociation en phase gazeuse des espèces moléculaires chargées positivement ou négativement. La chimie de ces dissociations se base sur une chimie particulière où les ions sont désolvatés et dont leurs orientations se basent sur la thermochimie et la cinétique. En phase gazeuse, deux molécules homologues ionisées peuvent présenter des chemins différents de fragmentations. Par exemple, les acides aminés déprotonés $[Asp\text{-}H]^-$ et $[Glu\text{-}H]^-$ se dissocient différemment bien qu'ils soient homomogues *via* la présence d'un groupement CH_2 sur la chaine latérale. En effet, le premier perd NH_3 sous activation collisionnelle alors que le second perd plutôt H_2O[209]. Ainsi, un changement complet d'orientation peut avoir lieu selon des mécanismes de fragmentations spécifiques et donc une modification des ions fragments peut en résulter.

Seuls les calculs quantiques peuvent nous permettre d'interpréter et/ou de confirmer les mécanismes que nous pouvons proposer en se basant uniquement sur la chimie organique classique.

Bien que la spectrométrie de masse ait été moins utilisée pour l'analyse de petites molécules au cours des années 1980 et 1990, au profit des analyses de grosses molécules, elle s'est développée avec l'apparition des méthodes de désorption par électronébulisation (électrospray, ESI[210,211]) et des lasers UV assisté par matrices (MALDI[212]). Toutefois, le développement de la spectrométrie de masse en métabolomique est encore plus récent[213], et fréquemment employé dans le cadre d'analyses « omiques » en raison de sa grande sensibilité[208,207]. Aujourd'hui, dans ce manuscrit, dans la mesure du possible nous respecterons les règles de IUPAC dans la nomenclature[214].

1. Un historique de la spectrométrie de masse

Il est difficile de considérer la spectrométrie de masse d'aujourd'hui sans rappeler ses origines. Ce n'est que par ce biais, basé sur l'évolution des idées, parallèlement à celles d'autres sciences, que l'on peut comprendre la situation de la spectrométrie de masse d'aujourd'hui. En fait, les mathématiques, la physique, l'atomistique, la chimie-physique, la chimie …, ont vraiment avancé ensemble, l'un pouvant répondre au besoin de l'autre. Dans la première partie de ce chapitre l'historique de la spectrométrie de masse sera développé.

Les premiers pas de la spectrométrie de masse remontent à la fin du XIXème siècle, pour répondre aux premières questions de l'atomistique, c'est-à-dire la constitution de l'atome pour en faire un modèle. Ainsi Thomson en 1897 a montré les propriétés de l'électron vis-à-vis de champs magnétiques et de champs électriques[215,216]. En 1912 il construit son premier spectromètre de masse parabolique. A la suite de cela, Dempster introduit en 1918 un instrument à double focalisation, afin de répondre au besoin de la détermination des masses atomiques de certains éléments à l'état gazeux, et il génère le premier spectre de masse. L'année suivante, Aston mettait en évidence les premiers isotopes stables, notamment ceux du néon. A partir de là, le modèle atomique simpliste initié par Thomson (une sphère neutre constituée de charges positive et négatives en vibrations) commençait à évoluer et à se concrétiser par différents modèles de plus en plus organisés selon les avancées physiques. Dans un premier temps se développe le modèle planétaire de Rutherford-Bohr, où les électrons sont sur des orbites circulaires concentriques correspondantes à certaines raies émises de l'atome d'hydrogène, et décrites par la formule de Rydberg. Plus tard, Sommerfield introduit un nouveau modèle selon lequel certaines orbites sont elliptiques et non circulaires.

La mesure des rayonnements en spectroscopie devient alors de plus en plus précis notamment avec la démonstration du dédoublement des raies. C'est ainsi qu'a été introduit la notion de spin, ou l'électron tourne autour de lui-même, soit dans un sens, soit dans l'autre. L'ensemble de ces étapes constituent les prémisses du modèle moderne de l'atome. Concernant la spectrométrie de masse, elle a été dédiée principalement jusqu'aux années 1940 à l'étude des différents éléments du tableau périodique[217]. En 1938 Harold Washburn développe l'analyse par ionisation électronique en phase gazeuse des composés organiques volatils par spectrométrie de masse, qui trouve son application directe dès les années 1941 dans l'analyse des gasoils utilisés dans l'aviation pendant la Seconde Guerre mondiale, et l'analyse du caoutchouc retrouvé dans les pneumatiques des véhicules de guerre.

Au milieu des années 1950, le développement de la chromatographie gazeuse couplée à la spectrométrie de masse, notamment par Fred W. Mc Lafferty commence[218]. Dès les années 1960, l'analyse des produits naturels suffisamment volatils et des biomolécules de petites tailles débute. Les molécules moins volatiles sont analysées après leur dérivation chimique. Cela concerne les molécules comme les sucres (méthylation), les stéroïdes (la silylation des groupes hydroxyles), les acides gras (estérification) et les lipides en général. Ces analyses sont réalisées *via* une ionisation par électron, IE (initialement nommée impact électronique ou bien bombardement électronique, terminologie ensuite prohibée par IUPAC)[214].

Au milieu des années 1960, Munson et Field développent l'ionisation chimique permettant de déterminer les masses moléculaires des molécules à partir de molécules protonées (mode ions positifs) ou déprotonées (mode ions négatifs). Cependant, l'information sur les fragmentations est perdue car ces espèces sont souvent stables et possèdent une faible énergie interne selon le gaz réactif utilisé. Dix années plus tard, après l'accident chimique à Seveso, l'analyse des ions négatifs en ionisation chimique se développe très rapidement permettant la détection des molécules chlorées, comme les dioxines et les biphényles. En même temps, des techniques d'ionisation par champs et par désorption de champs permettant l'étude des molécules moins volatiles et plus fragiles se développent.

A la fin des années 1970, les techniques de désorption SIMS (Secondary-ion mass spectrometry) et surtout de FAB-MS (Fast atom bombardment mass spectrometry) se sont développées *via* l'utilisation de matrices comme le glycerol. Cette dernière méthode répond aux besoins d'analyses de molécules non volatiles et fragiles notamment pour les grosses molécules comme les polysaccharides, les peptides et les protéines. Le FAB permet une première avancée, puis au début des années 1980 apparaît la désorption laser infrarouge[219] qui était trop énergétique. Au milieu des années 1980, la désorption par laser UV avec matrice

(MALDI) apparait[220], en même temps que les méthodes à pression atmosphérique[221] (API). A cette période des méthodes par ionisation chimique à pression atmosphérique (APCI) émergent et permettent l'ionisation en phase gazeuse et l'électrospray (ESI) pour la désorption de plus grosses molécules, par Fenn[210] et en désorption laser par Tanaka[,222]. Cette dernière approche a été supplantée par le mode MALDI apparu en même temps.

Développement des techniques omiques : Dès lors, les techniques "omiques" les plus variées se développent. Les efforts ont surtout été réalisés en protéomique et en génomique, au détriment de l'étude de petites molécules rencontrées en métabolomique et en lipidomique. Les acquis concernant la décomposition des ions obtenus dans les années 1970-1980 ont été peu à peu oubliés car la fragmentation des peptides, impliquant principalement la rupture des liaisons peptidiques, reste plus simple à prédire et interpréter.

Aujourd'hui, ce passé sur l'interprétation des fragmentations des ions dérivant de petites molécules est pratiquement effacé, ayant pour conséquence un manque dans l'approche des déterminations structurales de petites molécules en métabolomique. Des programmes in-silico se développent, comme ils se basent sur des systématiques bien éloignées des règles de la chimie, ils ne permettent pas de déterminer sereinement les structures des molécules étudiées[223]. C'est notamment le cas de solutions informatiques développées actuellement par certains constructeurs de spectromètre de masse. Une partie y sera consacrée dans le chapitre présentant mes résultats de thèse.

2. Généralités sur la spectrométrie de masse

Dans cette partie, seront abordés certaines méthodes non utilisées pendant ma thèse, mais dont certains concepts sont nécessaires à la compréhension des résultats expérimentaux. Le mode d'ionisation par électrospray et les différents analyseurs utilisés pendant mes travaux seront décrits en détails.

En spectrométrie de masse, différentes étapes sont à considérer : comme la séparation des échantillons (en ligne où hors ligne), en amont de la source. Si elle se fait en ligne, le mode de séparation doit être cohérent avec le mode d'ionisation, le choix de solvants, co-solvants et sels ne doit pas seulement répondre aux exigences de la méthode séparative mais au bon fonctionnement de la source d'ionisation. Le mode d'ionisation doit être cohérent avec la molécule à analyser et à étudier. Concernant les molécules faiblement polaires, représentées par une forme canonique (charges partielles faibles, Figure 72A), plutôt volatiles, une ionisation en phase gazeuse peut être appliquée et donc, le couplage GC-MS est pertinent (ionisation par électrons ou ionisation chimique). Il est éventuellement possible d'utiliser les ionisations en photo-ionisation à pression atmosphérique ou en APCI, en couplage LC-MS. Dans ces conditions d'ionisation en phase gazeuse, des distributions de clusters non covalents mono-chargés (homo-multimères protonés ou déprotonés) sont formées où les interactions sont uniquement du type liaison hydrogène. Ces clusters réagissent par réaction ion/molécule avec la molécule à analyser afin de produire uniquement des agrégats simplement chargés. A l'inverse, les molécules polaires peuvent former de zwitterions, où les charges sont plus importantes que dans les formes protonées et déprotonées. (Un exemple de protomère de chacune des formes citées sont représentés dans la Figure 71D et 71E). Ces espèces ne peuvent pas être obtenues en phase gazeuse. En effet, pour produire ces zwitterions, les molécules polaires ont besoin d'être solvatées. Comme elles ne sont pas très volatiles, elles ne peuvent pas être ionisées en APCI. En conséquence, il est nécessaire d'utiliser une méthode de désorption permettant le passage en phase gazeuse de paquets de matières, conduisant à des agrégats multichargés (système hétérogène d'un mélange de la (ou des) molécule(s) à analyser, de solvant, de sels, de charges, etc).

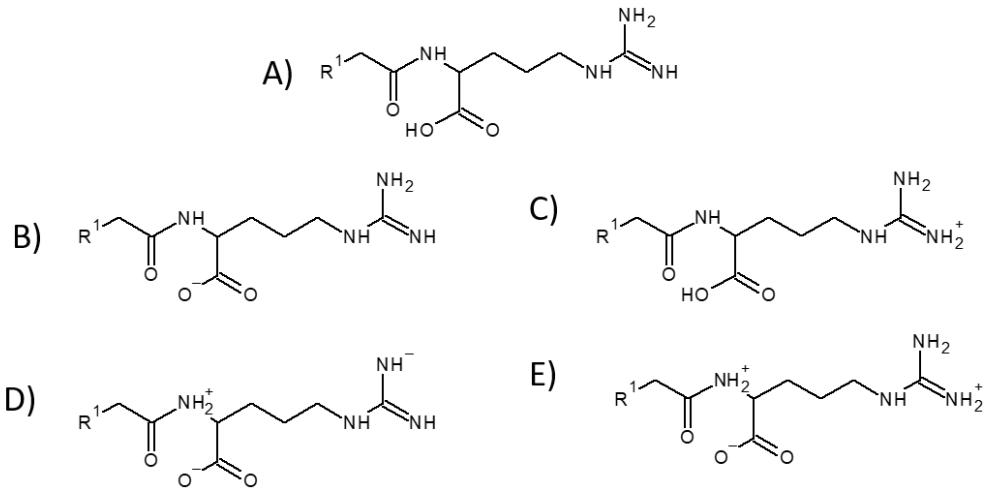


Figure 71 : Structure de d'un lipoaminoacide arginine sous forme A) canonique ; B) déprotonée ; C) protonée et enfin zwiterionique (sous forme négative en D, et positive en E)

Dans tous les cas de figure qu'il s'agit de molécules volatiles ou non, stables ou fragiles, il faut nécessairement la présence d'une charge positive sur la forme cationique (Figure 71A) et d'une charge supplémentaire aux deux charges opposées préexistante dans le zwitterion afin de les détecter en spectrométrie de masse.

Au début du XXème siècle, le rapport e/m était utilisé pour noter pour un élément. Puis de 1950 jusqu'au milieu des années 1970, les rapports m/e sont utilisés. Le terme e correspondait à un seul électron compte-tenu des modes d'ionisations en phase gazeuse utilisé. A la fin des années 1970 est apparu la notion du rapport m/z, permettant ainsi de se libérer de la valeur numérique de l'électron. Le rapport m/z est alors mesuré, où *z* représente le nombre de charges, un nombre pur ce qui fait que la dimension du rapport m/z est toujours une masse. Lors de l'ionisation en phase gazeuse sous vide z est généralement égal à 1 (sauf en impact électronique où dans certains cas, notamment pour le cas de composés aromatique, il est possible d'arracher 1, 2 ou 3 électrons). A pression atmosphérique, il est possible de détecter des ions radicaux mais exclusivement des composés mono-chargés. Au contraire, pour les méthodes de désorption où il est possible d'obtenir des ions multichargés, la masse peut être différente du rapport m/z. Le rapport m/z mesuré dans le spectre de masse doit être traduit. Quand une distribution d'état de charges est produite, un programme de déconvolution donnant une masse moyenne (au sens mathématique) devra être utilisé.

Ce rapport m/z peut être déterminé en basse résolution avec des spectromètres de masse équipés de filtres multipolaires (quadripole) ou avec les pièges à ions. Ces instruments permettent de mesurer des rapports m/z nominaux (nombres entiers qui ne tient pas compte

des défauts de masses des atomes). Le rapport m/z peut également être déterminé en haute résolution permettant de séparer des ions isobares et ainsi de déterminer de manière précise les rapports m/z, avec en général un rapport m/z donné à quatre décimales. Par ce biais, il est possible de remonter aux compositions élémentaires des ions. Cette approche reste précise pour des ions dont la masse n'excède pas 400-500 u (ou Dalton noté Da) et peut être réalisée sur des appareils équipés d'analyseurs de type temps-de-vol (time-of-flight TOF), à Transformée de Fourier (FT) ou à résonance cyclotronique des ions (FT-ICR).[1]* Compte tenu de la production d'espèces intactes, il est nécessaire d'avoir des informations sur la structure. Certains ions fragments caractéristiques correspondant à des pertes informatives et spécifiques (par exemple des pertes d'H_2O pour un alcool ou pour un acide carboxylique) doivent être mis en évidence. Ainsi, il est essentiel de donner de l'énergie sur l'ion d'intérêt, d'où l'introduction de la spectrométrie de masse en tandem.

Développement des techniques en tandem : Les tandems ont été dans un premier temps des instruments à double secteurs (ce sont des instruments à 2 champs, le premier magnétique et le second électrique) [224,225,226,227,228,229], qui depuis quelques décennies ont disparus des laboratoires d'analyse et de recherche, au profit d'instrumentation en tandem de type "Benchtop" introduits à la fin des années 1970 et au début des années 1980. Il s'agit de tandems de type triple quadripolaire conçus par Enke et Yost[230]. Dans ces spectromètres de masse, les ions sélectionnés dans le premier quadripole (un filtre de masse) sont soumis à des collisions produites dans le second quadripole (cellule de collision où seule la radiofréquence est appliquée, « RF only » quadripole). Le fonctionnement de cet instrument sera détaillé par la suite. En parallèle, les pièges à ions (ITD : ion trap detector)[231] où les ions sont formés dans le piège et sont analysés par éjection, se sont développé. Dans ce piège, il est possible de produire des séquences de fragmentations MS^2 ou MS^n. Cependant, la limite de ces instruments à champs quadripolaires était limitée en résolution, c'est à dire la capacité de l'appareil à faire la distinction entre deux ions m/Δm les plus proches possibles. De plus, de véritables réactions ions/molécules se produisaient dans cette cellule et le rapport m/z des ions produits pouvait être supérieur à celui de l'espèce moléculaire ionisée.

[1] * *Ici, seuls les instruments commerciaux seront décrits, afin de répondre à notre besoin. Les instruments expérimentaux de laboratoire auquels nous avons pas eu accès ne seront pas abordés*

Développement des appareils à temps-de-vol : Dans les années 1980, les instruments à analyseurs TOF se développent afin de répondre aux besoins de la désorption laser fonctionnant par impulsion comme les TOF/MS. Afin de minimiser les effets de dispersion temporelle des ions et de leurs dispersion en énergie cinétique (Ec) un réflectron (miroir électrostatique) est ajouté à l'analyseur[232]. Les ions peuvent se former à différents endroits de la source et ainsi subir un champ d'accélération légèrement différent.

Par la suite, l'introduction d'un quadrupole et d'une cellule de collision entre la source et le temps-de-vol, a permis d'obtenir un tandem hybride (noté Qq/TOF ou Q/TOF) de haute résolution (hybride à cause de l'énergie cinétique des ions qui est de plusieurs eV ou de dizaines d'eV dans les régions des quadripôles et de plusieurs keV à l'entrée du temps-de-vol)[233]. Les détails sur les fonctionnalités et le fonctionnement de cet appareil sera donné par la suite, puisque c'est sur un instrument de ce type qu'un développement analytique a été réalisé pendant ma thèse.

Les instruments hybrides : Il existe évidemment d'autres instruments hybrides, comme ceux couplés à un analyseur Orbitrap[234] de très haute résolution ou, moins courant, les Qq/FT-ICR. Contrairement à l'Orbitrap qui est utilisé uniquement comme détecteur, la cellule ICR peut permettre l'activation particulière de certains ions. Les instruments Qq-FT-ICR (ou avec multipoles Marshall[235]) possèdent divers modes d'activation en plus du mode CID. Il s'agit des activations :

- Par bombardement d'électron de plus de 10 eV (electron induced dissociation, EID[236] ou encore EIEIO) quel que soit la charge (multiple ou simple, positive ou négative) de l'ion.
- Par capture d'électron dissociative (electron captured dissociation, ECD) avec des électron de faibles énergies cinétique pour dissocier des ions positifs multichargés[237].
- Par dissociation induite par détachement d'électron à partir d'ion négatifs multichargés (electron detachment dissociation, EDD), utile surtout pour l'étude d'ions multichargés, ce qui ne fait pas parti de notre domaine d'intérêts[238].

Ces modes d'activations ne seront pas décrits par la suite car ils n'ont pas été utilisés durant ma thèse. La première méthode (EID), qui pourtant serait utile pour l'étude de lipides insaturés n'est pas encore installée sur les instrument commerciaux que nous utilisons[239].

3. Généralités sur les principaux constituants d'un spectromètre de masse

La première partie d'un spectromètre de masse est une source d'ionisation, avec en amont soit une sonde d'introduction directe, soit une méthode séparative (GC ou LC), (Figure 72).

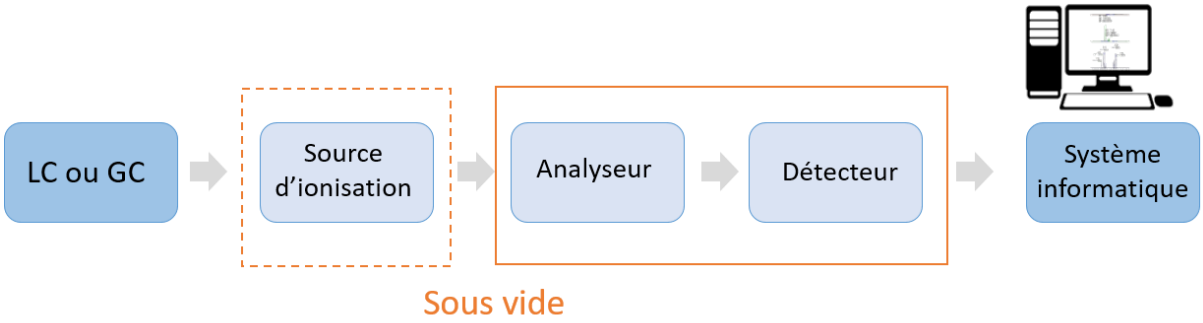


Figure 72 : Représentation schématique des éléments constituant le spectromètre de masse (la pression dans la source peut être la pression atmosphérique ou un vide plus ou moins poussé)

La source d'ionisation peut être sous vide lorsqu'il s'agit des modes d'ionisation en phase gazeuse comme l'EI (electron ionization, ionisation par électron) et la CI (chemical ionization, ionisation chimique), ou à pression atmosphérique pour les méthodes d'ionisations de type API. Dans ce dernier cas, la présence d'une interface [240] entre la source API et l'analyseur est nécessaire avec comme intermédiaire, une cellule à pression réduite pour la désolvation des agrégats chargés. La deuxième partie de l'instrument est l'analyseur, généralement sous vide, ou sous pression réduite (Figure 73). Les ions y sont transférés sous:

- Une faible énergie cinétique de l'ordre de 5 à 20 eV, sous une pression de 10^{-4} à 10^{-6} mbar. Cette pression permet d'éviter la dispersion des ions et leurs dissociations dans l'analyseur pouvant être un quadripole, un multipole ou un piège à ions. Le transfert du faisceau d'ions est facilité en faisant varier la "terre" (définie comme tension de référence, égale à 0) sur les électrodes. Dans cette zone, sous vide, des lois de balayages des tensions (continue et de l'amplitude de la tension RF) pour séparer les espèces de m/z différents sont appliquées, il s'agit ici d'un filtre de masse.

- Quelques keV (sous un vide de 10^{-10} à 10^{-11} mbar), dans un piège à potentiel quadro-logarithmique (Orbitrap)[234] constitué d'une électrode axiale, en forme de fuseau enfermée dans une enceinte sous forme de « tonneau » coupée en deux au milieu de l'axe de son plus grand diamètre. Sur cette paire d'électrodes symétriques, il apparaitra de manière alternative des charges induites par la présence du nuage d'ions en mouvement, sous forme d'hélices le

long de l'axe de l'électrode « fuseau », plus ou moins déphasé selon leurs rapports m/z. Un courant alternatif de fréquence liée au rapport de chacun des ions du nuage ionique entre ses électrodes est mesuré en produisant un signal transitoire (mélange de fréquences d'amplitudes différentes) dont la durée sera garante de la résolution. Dans ce cas, aucun balayage de tension ne se produit, le mouvement des ions est lié au potentiel de piégeage quadro-logarithmique.

- Une forte énergie cinétique de plus de 10 keV (sous un vide de 10^{-7} à 10^{-8} mbar), dans un tube de vol dont la longueur et la pression sont corrélées afin d'éviter toute collision entre les ions en vol et les molécules de gaz résiduelles (correspondant au libre parcours moyen des ions). Toute collision reviendrait à disperser les ions et à allonger leur temps de vol, ce qui entrainerait une perte de résolution ou une perte définitive des ions fragments par neutralisation de la charge ou par dissociation. Dans ce dernier cas de transmission des ions de la source vers l'analyseur, il n'y a pas de balayage, et seuls les temps-de-vol sont mesurés après une accélération des ions sous une haute tension. Cet instrument, contrairement au quadripole, fonctionne par impulsions.

La troisième partie de l'instrument est le détecteur, qui dépend de l'analyseur (Figure 73). Le détecteur est sous vide pour éviter les claquages électriques et sa détérioration *via* le dépôt de sels. Ce détecteur permet l'analyse des courants d'ions, de manière continue, avec le chaneltron (multiplicateur d'électrons), ou une galette multicanaux qui enregistre le passage des paquets d'ions.

Ces trois parties décrites succinctement (source, analyseur et détecteur) constituent le spectromètre de masse. Une dernière partie correspond au système informatique pour enregistrer les spectres de masses et piloter l'instrument. Ce système est totalement indépendant du spectromètre de masse, il reçoit uniquement les courants électriques qu'il traite.

4. Les sources d'ionisation

Le mode d'ionisation choisi va dépendre du caractère volatile ou non des molécules à étudier. Les molécules plus ou moins volatiles pourront subir après évaporation, soit une ionisation sous vide (EI et CI), soit une ionisation à pression atmosphérique (APCI, APPI, ou DART direct analysed in real time ; un mode classé dans l'ionisation ambiante dont on ne discutera pas dans ce manuscrit). Dans le cas des molécules non volatiles ou lorsqu'il s'agit de sels, l'évaporation de ces composés étant impossible, des processus de désorption seront utilisés, soit sous vide (MALDI) soit à pression atmosphérique comme l'ESI, l'AP/MALDI, DESI et l'EESI (extractive electrospray).

4.1. Ionisation en phase gazeuse sous vide

Ionisation par électrons : Ce mode d'ionisation implique la production d'ions moléculaires $M^{+\bullet}$ caractérisés par une forte énergie interne. L'énergie interne correspond à l'énergie vibrationnelle ou rotationnelle distribuée dans l'état fondamental et les états électroniques excités. Cette distribution d'énergie vibrationnelle s'étale sur 10 à 15 eV et dépend exclusivement des propriétés spectroscopiques de la molécule ionisée. Dans la source d'ionisation, les ions moléculaires sont des systèmes isolés, l'énergie interne des ions moléculaires ne peut donc pas être dispersée exception faite du rayonnement possible des ions se stabilisant. Le principe de Franck-Condon représente l'ionisation verticale des molécules allant de M en $M^{+\bullet}$, et leurs répartitions entre l'état fondamental et ses états électroniques excités, avant d'être redistribués sur des niveaux inférieurs, au moment où les surfaces de potentiels des états (représentés par des courbes) se croisent. Ce croisement permet aux molécules excitées de passer à un niveau d'énergie moindre (Figure 73).

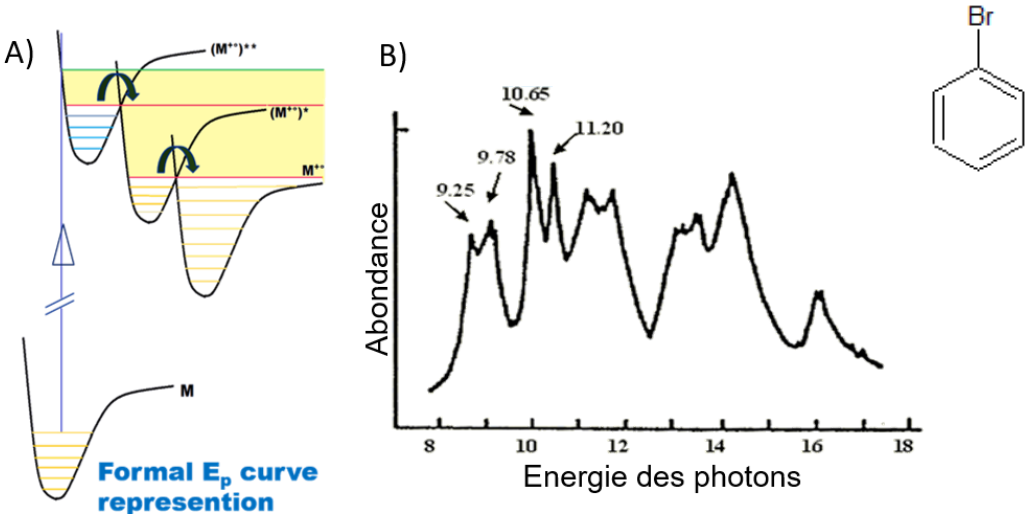


Figure 73 : A) Représentation schématique de l'ionisation verticale, selon le principe de Franck-Condon et B) le spectre photo-électronique du bromobenzène

Pour exemple, le spectre photoélectronique du bromobenzème (Figure 73) montre la distribution de l'énergie vibrationnelle dans l'état fondamental et les états excités. Cette distribution est propre à chaque composé, c'est une garantie de la reproductibilité des spectres dans leur fragmentation. Dans notre étude, nous considèrerons une distribution globale, sans séparer l'état fondamental des états excités, représentés par une courbe en cloche lors de nos expériences. Cette distribution très approximative est suffisante pour rendre compte qualitativement de l'orientation de processus de décompositions compétitifs soit dans la source soit dans d'autres régions du spectromètre de masse.

La fragmentation de ces ions est principalement due à la présence des états excités, qui sont souvent des états triplets (di-radicalaires) qui induisent la fragmentation alors qu'habituellement celle-ci est induite par la charge. En d'autres termes, les fragmentations sont induites par la présence de radicaux, même si la charge se trouve séparée du site radicalaire (c'est le cas pour les ions distoniques). Chimiquement, les ions radicaux sont des espèces instables, ce qui explique souvent l'absence de l'ion moléculaire dans les spectres de masse. Néanmoins, les spectres sont parfaitement reproductibles d'un instrument à l'autre, et c'est pourquoi les premières banques de données par ce mode d'ionisation ont été construites des 1975, et sont encore alimentées à ce jour et toujours utilisées. Les fragmentations obtenues peuvent s'interpréter à partir des règles de la chimie radicalaire, et ont été dans les années 1950 à l'origine de la construction de la théorie cinétique du quasi-équilibre (détaillée en Annexe 1). Même s'il est possible de remonter le puzzle de la molécule grâce aux ions fragments, l'absence de l'ion moléculaire peut être un fort inconvénient lors de

la détermination de structure. Pour l'élucidation structurale, il est nécessaire d'utiliser une méthode d'ionisation permettant de déterminer ces ions moléculaires.

L'ionisation Penning[241] où les ions sont formés par interactions molécules-atomes métastables pourrait être utilisée. Dans ce cas, l'énergie libérée lors de la désexcitation de l'état métastable est directement transférée à la molécule, qui peut alors être ionisée si cette énergie est suffisante pour que l'on se situe au-dessus du seuil d'ionisation de la molécule (énergie d'ionisation). Si cette énergie de désexcitation est proche de l'énergie d'ionisation de la molécule alors, l'ion moléculaire ne se décompose pas, et reste la seule espèce présente sur le spectre de masse. Cette méthode n'a jamais vraiment été développée commercialement, et n'est pas utilisée dans les laboratoires d'analyses. A l'inverse, l'ionisation chimique est commercialement viable, et facile à introduire sur les instruments. Cette méthode produit peu d'ions moléculaires par échanges d'électrons mais des molécules protonées ou déprotonées par échanges de protons.

Ionisation chimique : Inventée en 1971 par Mundson et Field, elle est basée sur des réactions qui ont lieu dans une source où la pression se situe entre 0,1 et 1 mbar. Ce mode d'ionisation repose sur la connaissance des réactions ion-molécule[242]. Afin que ces réactions puissent se faire, il est nécessaire que les ions aient une énergie cinétique très faible, de manière à ce qu'un des partenaires puisse capturer le second. Cette réaction de capture est déterminante pour les réactions ions-molécules à plusieurs étapes. En d'autres termes, les ions réactifs doivent être thermalisés (relaxés en énergie). Ces réactions ont été initiées dans un but de meilleure compréhension des réactions pouvant avoir lieu dans la stratosphère, ainsi que pour augmenter nos connaissances sur la composition des météorites[243], autour des planètes [244,245,246], des échantillons lunaires[247], de l'atmosphère planétaire, ou encore pour l'étude de l'environnement de Saturne [248,249].

Ces réactions ions-molécules peuvent être réalisées en plusieurs étapes, la première consiste en l'ionisation du gaz (noté G) *via* une ionisation par électron (noté e) (Eq. 1a).

$$\underline{\text{Eq. 1a}} : G + e \rightarrow G^{+\bullet} + 2^{e}$$

Ce gaz peut réagir sur lui-même, afin de former une molécule protonée réactive (Eq 1b) (comme par exemple les ions CH_5^+, H_3O^+, $(CH_3)_3C^+$ et NH_4^+).

$$\underline{\text{Eq. 1b}} : G + \langle e \rangle \rightarrow G^{-\bullet}$$

De la même manière, certaines molécules peuvent conduire aux ions réactions négatifs, comme Cl^-, OH^-, CH_3O^-, CH_2CN^- (en d'autres termes des solvants utilisés classiquement en LC-MS). Afin de produire ces ions réactifs (GH^+ ou $[G-H]^-$), la première étape consiste en une ionisation par électrons (Eq 2a et 2b).

$$\underline{\text{Eq. 2a}} : G^{+\bullet} + G \rightarrow GH^+ + [G-H]^\bullet$$

$$\underline{\text{Eq. 2b}} : G^{-\bullet} + G \rightarrow [G-H]^- + [G+H]^\bullet$$

Les ions radicaux produits sont très réactifs, et les électrons libérés sont thermalisés. Ainsi, dans la source un plasma constitué d'un mélange d'ions positifs et d'ions négatifs coexistants est produit.

Les gaz composés d'hétéroatomes (G) sont tous capables de s'autoprotoner, exception faite de l'acétone. Ces molécules de gaz protonées peuvent réagir sur la molécule d'intérêt, pour être solvatée et conduire à la formation d'un ion adduit (M,GH^+). Cette réaction est réversible et exothermique, puisqu'il s'agit de solvatation. A partir de cet ion adduit, le proton peut passer de GH^+ sur M, afin de produire (G,MH^+) selon un équilibre et ensuite se désolvater afin de produire $MH^+ + G$. Cette réaction est décrite selon le chemin réactionnel de Brauman décrivant la réaction ion-molécule par transfert de proton exothermique, où la réaction évolue vers les densités d'états les plus grandes (orientation thermo-cinétique). La conséquence de cette orientation peut se traduire par l'augmentation de l'énergie interne initiale, qui traduit une efficacité de réaction de plus en plus faible. C'est pour cela que la barrière centrale devient l'étape limitante (Figure 74).

Pour que cette réaction puisse avoir lieu, il est nécessaire que l'état final soit plus bas que l'état initial, c'est-à-dire qu'en allant de la gauche vers la droite, au fur et à mesure de l'avancement de la réaction, le nombre de fréquences infrarouges disponible augmente. C'est pour cela que naturellement la réaction se déroule de la gauche vers la droite, vers le sens exothermique. Ce chemin et les mêmes règles s'appliquent aux ions négatifs, lorsque l'on part de $[G-H]^- + M \text{ –()}\rightarrow G + (M-H)^-$.

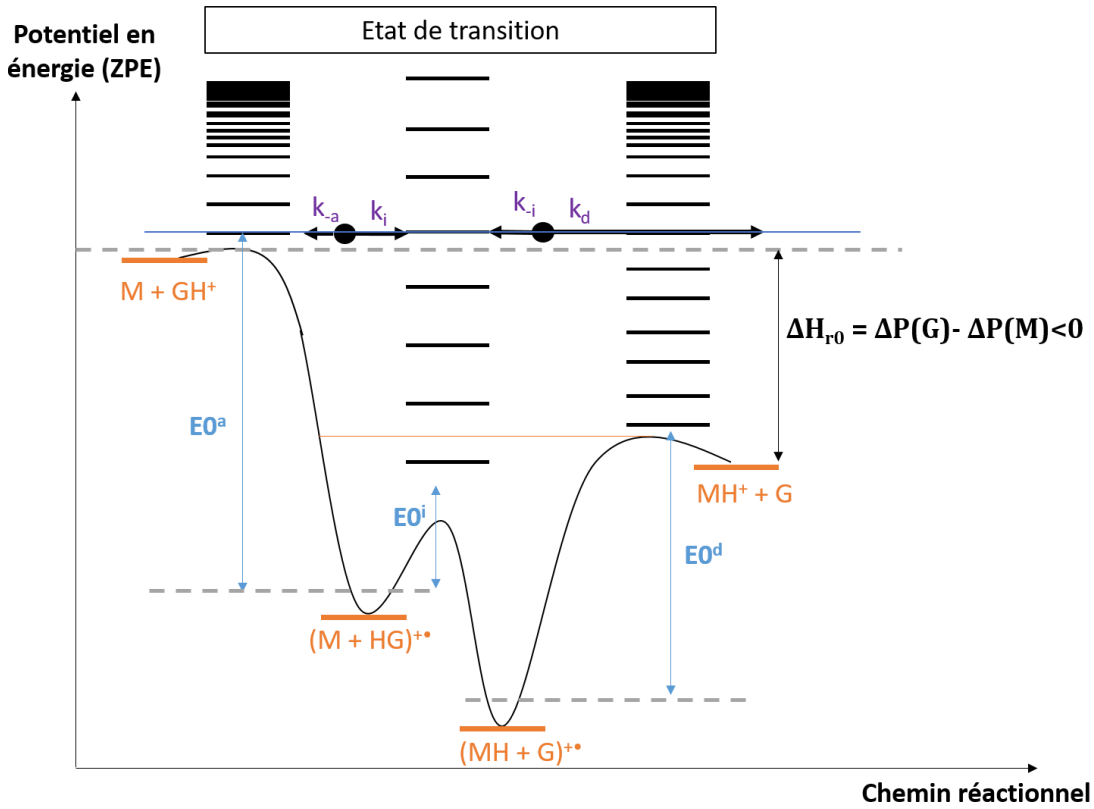


Figure 74 : Représentation schématique du chemin reactionnel de Brauman

Afin d'estimer si une réaction ion-molécule est exothermique ou non, il est nécessaire de définir deux fonctions d'états thermodynamiques (représentées par des réactions de pertes de protons fictives, mais calculables) :

- La première notion importante est la basicité (ΔG^0_{Bas}) ou l'affinité protonique, qui représente la capacité d'un composé à capter un proton (ΔH^0 = PA), cette fonction ne s'applique que pour les ions positifs.

$$MH^+ \rightarrow M+H^+$$

- La seconde, qui concerne les ions négatifs, s'apelle l'acidité, qui représente la capacité qu'à une molécule à perdre un proton. (ΔG^0_{Acide}) ou (ΔH^0_{Acide}).

$$M \rightarrow [M-H]^- + H^+$$

Ainsi, pour que GH^+ transfère son proton à M, il faut que $PA_M > PA_G$, c'est-à-dire que la molécule doit être plus acide que le gaz, et pour que $[G-H]^-$ arrache un proton à M, il faut que $\Delta H^0_{Acide}(G) > \Delta H^0_{Acide}(M)$.

Les ions négatifs sont des espèces beaucoup plus stables que les ions positifs, pour deux raisons. La première est chimique, une molécule protonée conduit à un grand nombre de fragmentations catalysée par le proton. En présence d'un anion, l'hétéroatome portant la

charge possède un nombre d'électrons du gaz rare qui le suit, et se trouve naturellement éléctronégatif, ce qui défavorise sa prompte décomposition. La seconde raison est liée à l'énergie interne emportée par les ions résultants de l'exothermicité de la réaction ion-molécule. Cette exothermicité s'accumule dans la liaison créée. Ainsi, pour les ions négatifs, la liaison créée est dans B ($H^+ + [B\text{-}H]^- \rightarrow B$) et donc les ions négatifs $[M\text{-}H]^-$ possèdent une faible énergie interne. Cette situation est contraire à celle des ions positifs où la liaison créée est dans MH^+ ($H^+ + M \rightarrow MH^+$) emportant une grande partie de l'exothermicité et ainsi se dissocie beaucoup plus rapidement.

Par exemple, une molécule d'ester (PA = 196 $kcal.mol^{-1}$) protonée par l'acétonitrile (PA = 186 $kcal.mol^{-1}$) va conduire à une énergie interne d'environ 10 $kcal.mol^{-1}$ (~0.45 eV ; avec 1 eV = 23,06 $kcal.mol^{-1}$) insuffisante pour conduire aux décompositions car 0,7-0,8 eV sont nécessaires pour atteindre un seuil de fragmentation possible. Cependant, la protonation de l'ester par le méthane (PA = 135 $kcal.mol^{-1}$), va conduire au maximum à 61 $kcal.mol^{-1}$(~2,6 eV). Cette fois, l'énergie est suffisante pour rompre une liaison covalente. Cette énergie est partiellement dispersée par collisions de très basse énergie cinétique avec le gaz réactif neutre (relaxation en énergie interne).

Les fragmentations des ions MH^+ ne sont pas nécessairement identiques à celles de $M^{+\bullet}$. Puisque dans un cas il s'agit d'un ion à nombre pair d'électrons, alors que l'autre possède un nombre impair d'électrons (ion radical). Dans ce dernier cas, très souvent l'ion radical s'isomérise en un intermédiaire distonique (où ce n'est pas la charge positive qui induit la fragmentation, mais le radical.

Cette partie a été développée surtout dans le but de dégager les conditions exigées pour qu'une réaction ion-molécule soit efficace (c'est-à-dire une réaction où à chaque collision un produit de réaction est formé). Cette efficacité ne dépend que de la première étape de la réaction ion-molécule, la réaction de capture, et cela que la réaction soit très fortement ou très faiblement exothermique : cela ne changera pas l'efficacité globale de la réaction. Ce mode d'ionisation n'a pas été utilisé pendant ma thèse, cependant le principe a été décrit puisqu'il est nécessaire pour la compréhension de l'ionisation en phase gazeuse à pression atmosphérique.

4.2.1. Ionisation à pression atmosphérique

Les techniques à pression atmosphérique permettant l'ionisation en phase gazeuse sont principalement l'ionisation chimique à pression atmosphérique, la photo ionisation à pression atmosphérique et le mode d'analyse direct en temps réel. Seule la première méthode d'ionisation, qui potentiellement, pourrait être utilisée pour certains composés parmi les plus volatils en métabolomique sera développée. Toutes ces méthodes (ainsi que les méthodes de désorption à pression atmosphérique) utilisent le même type d'interface, même si les configurations changent selon la société qui construit l'instrument[240]. Ces interfaces doivent permettre le transfert des ions de la pression atmosphérique à un pré-vide (pression réduite) où la désolvatation des agrégats chargés se produit afin d'être ensuite échantillonnés à l'aide de cônes successifs.

L'ionisation chimique à pression atmosphérique : Elle a été introduite en 1973 par *Horning et al*.[250], et s'effectue en phase gazeuse où l'échantillon est en solution dans la phase mobile. Les analytes sont introduits au travers d'un capillaire à pression atmosphérique, et sont nébulisés par du diazote hautement chauffé (de 300°C à 400°C). Cette étape permet l'évaporation rapide de la solution d'analytes (noté M) de manière à obtenir l'échantillon en phase gazeuse dans une ambiance saturée de gaz, constituée par les solvants évaporés (noté Solv)[251]. Les molécules de solvant, en phase gazeuse seront ionisées par décharge Corona, qui produit un nuage d'électrons, permettant à la fois l'ionisation par électron du solvant (formation de $Solv^{+\bullet}$), mais aussi à la thermalisation des électrons. Le solvant à l'état gazeux est protoné ($SolvH^+$) par «auto protonation» (sauf l'acétone)[252] comme en ionisation chimique. Cette fois-ci, à pression atmosphérique, les ions réactifs sont solvatés pour former en général des clusters chargés (homo-multimères protonés $Solv_nH^+$ ou déprotonés $Solv_{(m-1)}(Solv\text{-}H)^-$ (ou $[Solv_m\text{-}H]^-$) avec m toujours beaucoup plus petit que n, car les molécules qui sont déprotonées sont acides et par conséquent moins solvatées. Ces ions réactifs libres de molécules de solvant n'existent pas dans ces conditions, leur solvatation est immédiate par multiples collisions avec les molécules de solvants en phase gazeuse. Ces clusters vont réagir comme en ionisation chimique sur la molécule d'intérêt M par capture (selon l'équation Eq 3a). Il en est de même pour les ions négatifs (selon l'équation Eq 3b) pour donner des agrégats mono-chargés (cluster + molécule) :

$$\underline{\text{Eq. 3a}}: M + Solv_nH^+ \rightarrow [M, Solv_nH^+]$$

$$\underline{\text{Eq. 3b}}: M + (Solv_m\text{-}H)^- \rightarrow [M, Solv_m\text{-}H^-]$$

Les agrégats de tailles variables (selon les clusters ayant réagi) pourront coexister sous certaines conditions. Après leur formation, ceux-ci pourront perdre une ou plusieurs molécules de solvant, conduisant à un nouvel agrégat de taille plus faible mais moins réactif. Cela s'explique par exemple en mode positif, en comparant l'affinité protonique du cluster qui a perdu x molécules de solvant et celle de «l'analyte»: lorsque ces valeurs sont comparables ($PA_M \sim PA_{Solv(n-x)}$) alors l'agrégat protoné à plutôt une longue durée de vie[253,]. Pour les clusters deprotonés, ce raisonnement s'applique en comparant les acidités. Après libération des molécules de solvants, l'agrégat obtenu a une longue durée de vie dès que les acidités deviennent proches [$\Delta H^0_{Acide}(M) < \Delta H^0_{Acide}(Solv_{m-y})$]. Pour des systèmes polyfonctionnels, cette interprétation doit être faite pour chacune des fonctions chimiques. En conséquence, des agrégats avec des tailles différentes coexistent. Cela a été démontré à partir d'agrégats $M(H_2O)_xH^+$ (x=2 et 3) avec M étant l'ester méthylique de l'acide *p*-aminobenzoïque, étudiés par spectroscopie des ions[254]. Ces espèces qui survivent sont injectées dans la zone de désolvatation, au travers du système de l'interface (à pression atmosphérique vers la pression réduite)[255,256,257,258,259] pour conduire à des ions protomères dont la charge ne peut pas migrer d'un site à l'autre, comme c'est le cas pour les composés aromatiques.

Selon les conditions de désolvatation, c'est-à-dire *via* la modification des paramètres du « sampling cone » (voltage de cône), il apparaît à faible voltage principalement les espèces intactes. L'augmentation de cette valeur de sampling cone peut conduire à des fragmentations. Toutefois, il n'est pas toujours possible de travailler avec des voltages trop faibles, la diminution de cette valeur de sampling cone va défavoriser la transmission des ions de plus hauts rapports m/z. Il est donc nécessaire d'effectuer un compromis entre bonne transmission des plus hauts rapports m/z et fragmentation *via* ce paramètre.

Il n'est pas possible de voir des ions multichargés en APCI puisque les réactions ions-ions de même charge ne peuvent pas avoir lieu. De plus, on ne peut pas avoir de formation de sels protonés ou déprotonés, car ils ne sont pas volatils, dans le sens où ils sont nécessairement solvatés. Pour étudier ces espèces, il est nécessaire d'appliquer des processus de désorption. De plus, il ne peut pas y avoir de formation de molécules cationisées par des cations métalliques. Ils ne sont pas volatils car leur solvatation est très forte en raison de la présence des charges. Ces espèces qui pourront être désorbées intactes, peuvent être sous forme de cation solvaté ($[M+Na]^+$ où le cation est en interaction non covalente avec des doublets libres d'hétéroatomes), ou des formes tautomères des précédentes comme les sels protonés (i.e., $[(M-H+Na)+H]^+$) ou déprotonés (i.e., $[(M-H+Na)-H]^-$)[260].

4.2.2 Ionisation par désorption à pression atmosphérique

Parmi les nombreuses méthodes de désorption à pression atmosphérique comme l'électrospray, le nanospray[261], la désorption sous électrospray[262], l'«extraction» par électrospray (EESI)[263], l'évaporation par laser combiné à l'électrospray (LAESI)[264] ou la désorption laser assistée par matrice sous pression atmosphérique, nous ne développerons que l'ionisation par électro-nébulisation utilisée durant cette thèse.

Ionisation par électronébulisation : L'hydrodynamisme est connu depuis le XVIIIème siècle, et a été introduit par Bossut *via* le traité théorique et expérimental de l'hydrodynamisme[265]. Par la suite, l'électro-hydrodynamisme a surtout été développé par l'Abbé Nollet[266], souvent en discussion et en collaboration avec Benjamin Franklin. Ainsi, ils ont domestiqué les charges introduites sur des gouttelettes. L'électro nébulisation en tant que telle est utilisée pour la première fois par Dole *et al.* en 1968 afin de mettre en évidence des particules, *via* l'utilisation de la spectrométrie de masse. Fenn développe réellement cette technique en 1984[267] dans le cadre de détection de très gros polymères *via* l'utilisation d'analyseur quadripolaire.

L'ESI est une technique d'ionisation à pression atmosphérique, qui contrairement à l'APCI, ne permet pas d'évaporer très rapidement les gouttelettes de solvant. En effet en ESI les gouttelettes sont chargées, et donc leur évaporation devient quasiment impossible. Néanmoins, le fait d'évaporer partiellement ces gouttelettes multichargées conduit à la perte de l'équilibre entre le nombre de charge répulsives et la tension superficielle de la gouttelette conduisant sur le cône de Taylor lors de la limite de Rayleigh à des explosions de Coulomb (Figure 75). Ces explosions se font selon une cascade de plus en plus rapide afin d'arriver à la fin, à la désorption de gros agrégats porteurs de nombreuses charges[268,,251], il s'agit du modèle de l'évaporation ionique[269,270]. Ce sont ces agrégats qui seront partiellement désolvatés, dans le capillaire de transfert et surtout dans la zone de pression réduite où les molécules de solvants sont éliminées ainsi que certaines charges[271,272]. Si la molécule est petite, évidemment à la fin il ne restera plus qu'une seule charge, ainsi qu'une ou deux molécules de solvants. Aux concentrations les plus élevées, les formes protonées deviennent défavorisées au profit des formes cationisées. Cela va néanmoins dépendre des propriétés physico-chimiques de la molécule analysée. L'analyse de polymères, comme le polyéthylène glycol va conduire strictement aux formes cationisées par le sodium, alors que certains acides aminés acides refuseront de donner des formes cationisées, comme c'est le cas pour certains lipoaminoacides de notre étude.

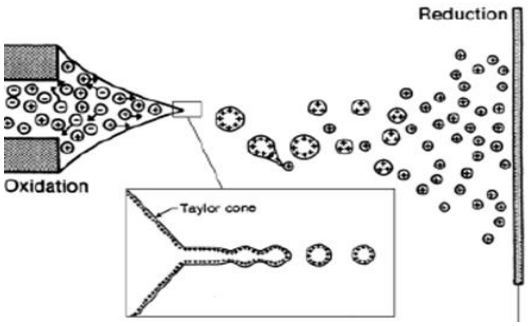


Figure 75 : Principe de l'électonébulisation, de la limite de Rayleigh et du cône de Taylor d'après *Schröder, E. Massenspektrometrie - Begriffe und Definitionen. (1991).*

Dans le cas de plus grosses molécules que celles rencontrées en lipidomique et en métabolomique, les ions libérés après désolvatation se retrouvent multichargés selon une distribution d'état de charge qui dépend des conditions de désolvatation et des propriétés physico-chimiques des solvants, additifs, et des groupements réactifs de l'analyte (la basicité ou l'acidité en phase gazeuse). Dans ce cas, le rayon de la gouttelette diminuera à son minimum, il n'y aura pas de désorption mais seulement une désolvatation continue par pertes consécutives de molécules de solvants jusqu'aux dernières selon un modèle établi par La Mora[273]: le modèle de la charge résiduelle. Les gros agrégats multichargés sont transférés dans la zone de pression réduite, afin d'être désolvatés et plus ou moins déchargés, selon leurs tailles et leurs potentiels de désolvatation. Il en résulte une distribution d'état de charges continue dépendant, généralement, de la conformation des macromolécules chargées.

C'est la méthode la plus populaire, puisqu'elle permet un couplage LC-MS[206]. Toutefois sa limite majeure se trouve au niveau de la gamme dynamique d'analyses quantitatives qui n'excèdent pas deux ordres de grandeur. Ce plus, avec cette technique d'ionisation, les effets de suppression d'ions à cause de la matrice sont importants, contrairement aux techniques à ionisation en phase gazeuse comme l'APCI[274].

5. Les différents analyseurs en spectrométrie de masse

En spectrométrie de masse, c'est l'analyseur qui permet de séparer les ions par leurs rapports m/z respectifs, pour cela, il mesure le rapport masse-sur-charge de la molécule (*m/z*). Plusieurs paramètres permettent de caractériser un analyseur : la gamme de rapport m/z qu'il peut mesurer, sa vitesse d'acquisition, sa sensibilité, sa résolution, sa précision en masse et sa compatibilité avec les différentes sources d'ionisations à l'aide d'interfaces adaptées[240].

La précision de mesure des rapports m/z des ions mono-isotopique et de leurs ions fragments dépend de la résolution de(s) l'analyseur(s) du spectromètre de masse. La résolution correspond à la capacité d'un spectromètre de masse à distinguer deux ions voisins de rapports masse-sur-charge m/z distincts d'une différence Δm/z[275]. La valeur de Δm/z peut se mesurer par la largeur à mi-hauteur d'un signal à un m/z donné (Full Width of the peak at Half its Maximum height, FWHM)[275,276]. Plus la résolution d'un analyseur est élevée, plus la mesure en masse de l'ion est précise à condition que l'étalonnage en rapport m/z soit constant dans le temps. Ces paramètres définissant la résolution et la mesure de masse, sont très importants et doivent être considérés pour toute analyse.

La précision en masse, caractérisée par la résolution, correspond à la différence entre la masse calculée (M_{cal}) de l'ion mono-isotopique et la masse mesurée expérimentalement (M_{exp}), et s'exprime en partie par million (ppm) et se calcule de la manière suivante :

$$Erreur\ de\ mesure\ en\ masse = {Mexp - Mcal}/{Mcal} * 10^6$$

Il existe une grande variété d'analyseurs caractérisés par des performances très différentes, deux grands sous-types d'analyseurs sont distingués : ceux qui travaillent en haute résolution et ceux en basse résolution. Par exemple, pour un ion m/z 400, l'analyseur utilisé est de basse résolution, si la valeur de Δm/z > 0,1 (l'unité est donnée en m/z). Pour une valeur de 0,01 < Δm/z < 0,1, la résolution est considérée comme moyenne et pour un Δm/z << 0,01 il s'agit de haute résolution.

En basse résolution les rapports m/z sont souvent exprimés en valeurs nominale. A l'inverse, les appareils de haute résolution pourront donner les valeurs de rapports m/z au millième d'unité. Les spectromètres de masses à transformée de Fourrier atteignent des valeurs de résolution de 1 000 000[277]. Ces appareils sont dits à très haute résolution, et sont constitués d'un analyseur par résonance cyclotronique des ions ou d'un analyseur Orbitrap.

Aujourd'hui les spectromètres de masse sont constitués de plusieurs analyseurs consécutifs séparés par de simple lentilles électrostatiques placées à la suite les unes des

autres. Ils impliquent des faisceaux d'ions de faible énergie cinétique (<<100 eV) comme pour les triples quadripôles (QqQ). Les rôles attribués à ces tandems sont d'identifier les molécules, d'élucider leurs structures, et de les quantifier en atteignant des limites de sensibilité les plus basses possibles. Ils permettent de réaliser des analyses MS/MS (MS^2) et certains avec les pièges à ions permettent d'effectuer des analyses séquentielles MS^n.

En 1965, K. Jennings a été le précurseur de la spectrométrie de masse en tandem avec l'activation collisionnelle[278]. Ensuite, F.W. McLafferty a introduit de nouveaux concepts comme la comparaison de structures communes pour des ions de même rapports m/z avec des instruments en tandem pour l'étude de la structure d'ions de très faibles tailles[279,280]. C'est en utilisant des instruments à doubles secteurs de géométrie Nier-Johnson, où les rapports mesurés sont les rapports E/B, que la forme des signaux et des spectres d'énergie cinétique des ions fragments (IKE: ion kinetic energy) sont alors explorés. Ils ont été ensuite utilisés afin de comparer les états de transitions des chemins menant aux ions fragments de structures communes ou isomères, notamment par Beynon et Cooks en 1969[281]. Finalement, Beynon [282] et Cooks[228] ont introduit des instruments à doubles secteurs à géométrie inversée (B/E), et les spectres MIKE (mass-analyzed ion kinetic energy). A la fin les années 1970 les premiers instruments commerciaux sont apparus (Varian devenu Finnigan/Thermo et VG devenu Waters). Ces instruments permettaient d'isoler des ions isobares de petites tailles[283] et d'en étudier leurs structures respectives en utilisant les processus d'activation collisionnels.

Au cours de mes travaux de thèse, j'ai pu travailler sur différents spectromètres de masse, ayant des spécificités différentes. L'ensemble des analyseurs avec les différentes théories de bases concernant leurs fonctionnements ne seront pas développés. Par contre, nous mettrons en valeur les paramètres modifiables importants. Dans un premier temps, nous nous consacrerons aux analyseurs simples, puis aux analyseurs en tandem.

5.1. Analyseur basse résolution avec balayage de tension

Aucun analyseur simple n'a été utilisé durant ma thèse, cependant les principes de ces analyseurs doivent être décrits, puisque ce sont les mêmes qui seront utilisés dans les instruments en tandem (QqQ, QqTOF, LTQ) qui eux ont été utilisés durant mes travaux de thèse.

5.1.1 Le quadripôle : analyseur de faisceaux d'ions de faible énergie cinétique

Les analyseurs de type quadripôlaires sont les plus répandus, ils ont été inventés par Wolfgang Paul en 1953. Un quadripôle est composé de quatre électrodes parfaitement parallèles (Figure 76) entre lesquelles des potentiels opposés sont appliqués entre deux électrodes consécutives[20]. Sur chacune des électrodes est appliqué un potentiel Φ_0 (défini selon l'équation 4):

Eq. 4 : $[\Phi_0 = U \pm V_{RF} \cos(\Omega t)]$

ou ± signifie seulement une possibilité de déphasage de 180° de la radiofréquence

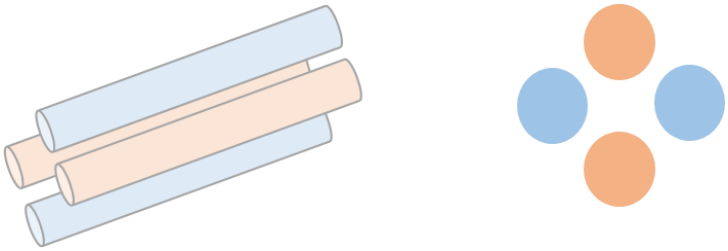


Figure 76 : Représentation schématique d'un quadripôle de vue latérale et frontale

Sur une première électrode est appliqué un potentiel Φ_0 qui est constitué de deux variables : une tension continue U et une tension alternative d'amplitude V_{RF}. Sur l'électrode consécutive, est appliquée un potentiel $-\Phi_0$. Ces tensions sont référencées à la terre prise comme 0. Cependant ensuite le référentiel sera changé (donc la valeur de 0), afin de permettre un transfert des ions vers le détecteur. Ainsi, les molécules chargées vont passer au travers du quadripôle en parcourant une trajectoire sinusoïdale, sous l'effet du champ quadripolaire. Lorsque l'amplitude des oscillations se stabilise, les ions entrent en résonnance, et peuvent sortir de l'analyseur quadripolaire pour atteindre le détecteur.

Dans ces conditions un champs quadripolaire est créé, où les ions peuvent être en équilibre, et traverser entièrement le quadripôle pour en émerger [284,285,286]. A l'inverse, d'autres ions ont des mouvements sinusoïdaux dont l'amplitude diminue ou augmente, entrainant une neutralisation de la charge sur les électrodes avec lesquelles ils entrent en contact, ces ions ne peuvent donc plus émerger à l'extrémité axiale de l'analyseur. La résolution des équations de mouvement d'un ion permet de définir les paramètres de stabilité établis à partir des équations de Mathieu (non présentées ici). Ces paramètres sont notés : a et q proportionnel chacun aux rapports zU/m et zV_{RF}/m respectivement. Ainsi, on peut définir des domaines où les ions m_1, m_2, m_3 et m_4 sont stables (Figure 77) et d'autres en dehors des leur frontières, où les ions sont prématurément éjectés. La stabilité des ions m_1, m_2, m_3 et m_4 peut être représentée sur un diagramme $U=f(V_{RF})$ où la droite de fonctionnement $U/V_{RF} = C^{te}$ définie une résolution donnée. Sur cette droite, les ions m_1, m_2, m_3 et m_4 peuvent émerger du quadripôle uniquement pour les zones de la droite de fonctionnement définies par les segments de la droite inclus dans chacun des domaines de stabilité des ions. Ainsi, le quadripôle agit en filtre de masse, il est alors possible de constituer un spectre de masse avec l'ensemble des ions présents dans le diagramme de stabilité qui peuvent émerger du filtre de masse. Ces signaux, sont comme des U retournés, dont la largeur est constante quel que soit le rapport m/z. Aujourd'hui, ils ont une forme quasi-Gaussienne grâce à la présence de pré- et post-filtres et de lentilles

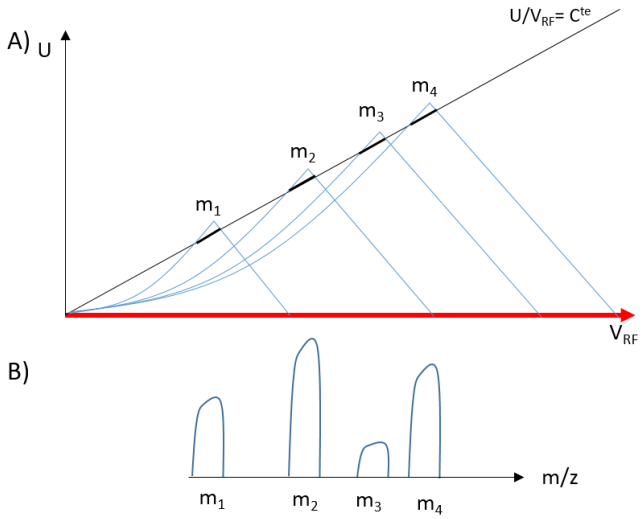


Figure 77 : A) Domaines de stabilité ddes ions m_1, m_2, m_3 et m_4 dans les diagrammes $U=f(V_{RF})$, et B) le spectre de masse associé

5.1.2 Le piège à ions linéaire : analyseur par stockage et éjection sélective d'ions

Le piège à ions linéaire a les mêmes caractéristiques qu'un piège 3D. Ils fonctionnent tous les deux uniquement avec une tension V_{RF} et une tension continue maintenue nulle. Néanmoins, celle-ci pourrait être employée pour compenser les effets de la charge d'espace due au nombre d'ions piégés. Notons que dans le LTQ (Linear Trap Quadrupole) les effets de charge d'espace sont moindres que dans le piège 3D. Le LTQ peut travailler en basse résolution avec une éjection résonante (soit quadripolaire soit hexapolaire) qui est radiale (axe *r*) alors que dans le piège 3D, elle l'est de façon axiale (axe *z*). Ainsi en basse résolution, on passera de la droite de fonctionnement q_z (pour le piège 3D) en q_r (pour le LTQ) sachant que q_z=-2q_r. Toutefois, cela ne change pas les règles de fonctionnement, pour simplifier les explications uniquement q_z sera utilisé. Par ailleurs, l'axe z est employé lors de l'éjection des ions hors du piège pour les analyser en haute résolution *via* une C-trap, ou pour les activer dans une autre cellule d'activation (cellule HCD higher energy collision dissociation).

Le domaine de stabilité de chacun des ions (m_1, m_2, m_3 et m_4) peut être représenté selon des segments de droite se superposant sur l'axe de la tension V_{RF} (Figure 78A), ou ils peuvent être placés sur la droite de fonctionnement q_z (Figure 78B) où les rapports m/z de ces ions se disposeront, pour une tension V_{RF} donnée, de manière croissante en q_z. La valeur de q_z=0.908 correspond à la limite de stabilité et celle-ci correspond à la résonance quadripolaire, position où les ions seront naturellement éjectés du piège. Avec le piège commercial LTQ, l'éjection résonante est hexapolaire et a lieu à q_z=0,78. Associée à une diminution de la radiofréquence Ω, l'éjection hexapolaire permet d'accroître la gamme en m/z. Ainsi pour éjecter par résonance hexapolaire, les ions m_i piégés à une valeur de $q_{z,mi}$ donnée (et inférieure à $q_{z,eject}$=0,78), il faut augmenter l'amplitude V_{RF} de manière à ce que la valeur de $q_{z,mi}$ augmente pour tendre vers q_z 0.78.

Cette approche permet de savoir si un ion est stable ou pas, mais pas de savoir s'il est stabilisé de la même manière quel que soit sa position sur son axe de stabilité q_z. Par exemple, pour un nombre constant dans le piège, un ion particulier sera-t-il stabilisé de la même manière sur une faible ou forte valeur de q_z par la force de rappel du champ de piégeage ? Ainsi, pour une tension V_{RF} donnée, les différents ions m_1, m_2, m_3 et m_4, ne seront pas piégés sur la même position en q_z et donc pas de la même manière. La profondeur du puits de pseudopotentiel (D_z liée à la force de piégeage) varie avec la valeur de q_z. Pour les faibles valeurs de q_z, la profondeur du puits de pseudopotentiel est très faible, entrainant une forte instabilité des ions s'y trouvant dès qu'il y a une charge d'espace. A l'inverse, pour des valeurs q_z plus élevée, le puits de pseudopotentiel est plus profond, les ions y seront donc

plus stables. Cependant, dès que les ions arrivent à $q_z = 0.908$ la profondeur du puits s'annule car c'est à $q_z = 0.908$ que la profondeur du puits de pseudopotentiel est égale à 0 et les ions sont éjectés. Schématiquement, la profondeur des puits de pseudopotentiels peut se traduire par des verres de tailles différentes où elle augmente de 0 à 0.908 avant de devenir nulle. Ainsi, en augmentant de manière régulière la haute tension, il est possible d'éjecter dans l'ordre des ions m_4, m_3, m_2 puis enfin, l'ion m_1 (Figure 78B).

Pour un nombre d'ions donné, les répulsions de Coulomb entre eux, mènent à la formation d'un nuage de taille bien précise et constante. Les effets de charge d'espace des ions dans le nuage seront d'autant plus dramatiques que le puits de pseudopotentiel est petit. Si un ion m_i/z lié à cette charge d'espace fait parti des ions les plus proches des électrodes du piège (par expansion répulsive de la taille du nuage d'ions), il peut alors être éjecté prématurément si sa position en q_z est à 0,78. Son éjection n'a plus lieu sur l'éjection quadripolaire pour être détecté à m_i/z mais est détecté à $(m_i/z)x(0,78/0,908)$, faisant apparaitre des espèces inattendues (pics fantômes). Pour éviter une telle situation le nombre d'ions dans le piège doit être régulé, en utilisant « l'automatic gain control » (AGC) qui permet de maintenir constant et pas trop élevé le nombre d'ion injecté dans le piège.

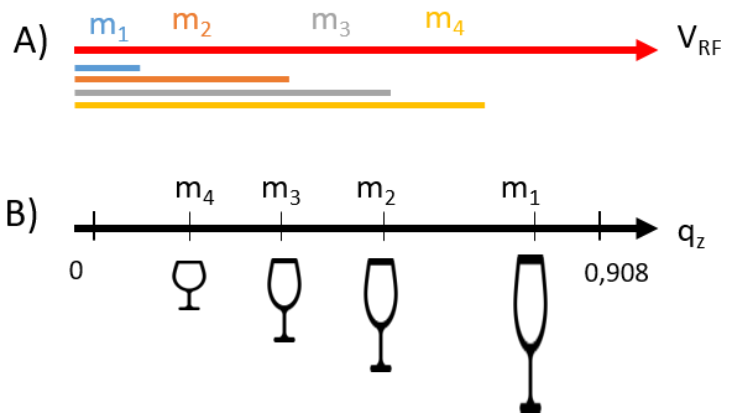


Figure 78: A) Domaine de stabilité ions de masse croissante de m_1 à m_4 sur l'axe V_{RF} sachant que la tension continue est nulle, et B) Position de ces ions sur l'axe q_z ou la profondeur des puits sont schématisés par des verres.

Avec ces analyseurs, lorsque l'on étudie des ions possédant de nombreux isotopes abondants, ils affectent la position de l'ion mono-isotopique d'intérêt. C'est-à-dire que son rapport m/z est déplacé légèrement vers les plus hautes valeurs de m/z. Afin de pallier à ces perturbations dues aux charges d'espaces, il est possible de réaliser un mode de balayage « inverse », nommé « reversed scan » ou « zoom scan » où le premier ion du massif isotopique éjecté n'est plus l'ion mono-isotopique mais un ion isotope naturel à M+n.

5.1.3 L'analyseur temps-de-vol : faisceaux d'ions de haute énergie cinétique, et analyses sans balayage de champs

L'analyseur à temps-de-vol (TOF) a été décrit en 1948 et exploité commercialement en 1955[287,288]. Dans un analyseur TOF, tous les ions m_i, quel que soit leurs rapports m/z sont accélérés par une haute tension V_o pulsée de plusieurs dizaines de kV à la sortie de la source. Ainsi, tous les ions ont la même énergie cinétique s'exprimant telle que : $zeV_o=m_iv_i^2/2$. Les ions fortement accélérés « volent » au travers du tube de vol selon un même chemin dans le vide. La longueur de ce tube est calculée par le constructeur, en tenant compte du vide existant, de manière à ce que la probabilité de collision avec le gaz résiduel soit nulle (que le libre parcours moyen des ions correspond à la distance entre la source et le détecteur). Dans ces conditions, ni la dispersion du faisceau d'ions, ni une neutralisation de la charge ou une dissociation induite par collision peut avoir lieu, ce qui se ferait au détriment de la résolution et de la sensibilité. Comme les ions m_i ont tous la même Ec, leur vitesse v_i dépend de la masse de l'ion considéré [$v_i=(2zeVo/m)^{1/2}$]. Ainsi tous les ions de vitesses différentes arriveront à des temps différents au détecteur. Les ions les plus légers arriveront ainsi les premiers au détecteur, et les ions les plus lourds arriveront plus tard.

L'inconvénient majeur des analyseurs de type TOF est que l'ionisation dans la source peut ne pas toujours se faire en même temps et au même endroit avant l'accélération des ions ou que, dès la sortie de la source, le transfert des ions conduit à leur dispersion spatiale qui peut être aggravée par la durée de l'impulsion de la haute tension. La conséquence est que potentiellement, des ions de rapports masse-sur-charge identiques peuvent présenter des Ec initiales différentes. Ceci entraine des vitesses différentes ou des distances parcourues différentes donc des temps de vols différents pour un même ion m_i, et donc une dispersion des mesures de rapports masse-sur-charge ayant pour conséquence une dégradation de la résolution. Une telle dispersion temporelle est rencontrée lorsque la source d'ionisation est externe et où les ions sont injectés dans une zone où s'exercera l'accélération pulsée. Afin de limiter ce phénomène, sur le tandem Qq/TOF que nous avons utilisé, le faisceau d'ions qui émerge de la cellule de collision et des lentilles qui la suivent, est injecté dans le tube de vol pour être dévié d'un angle $\theta < 90°$ où $\theta=tang^{-1}(v_{TOF}/v_{faisc})$ avec v_{TOF} et v_{faisc} les vitesses des ions dans le tube de vol et ceux incidents émergeant de la cellule de collision. Cet angle θ est aussi celui du tube de vol de Guilhaus[289]. Cette entrée presque orthogonale permet de limiter la dispersion spatiale des ions à une dispersion radiale définie par l'angle solide du faisceau d'ions émergeant des lentilles post-collision. L'avantage des collisions dans cette région permet une relaxation radiale de l'énergie cinétique des ions mais aussi une meilleure focalisation dans l'axe des lentilles d'où une amélioration de la résolution du tandem Qq/TOF. Parmi les différentes modifications instrumentales introduites pour améliorer la résolution de

ce type d'instrument, nous nous focaliserons sur celles qui est directement appliquée dans le cas de l'injection d'ions produits à partir d'une source externe comme l'ESI. Il y a par exemple, l'introduction d'un miroir électrostatique (nommé réflectron) qui avait été conçu par une équipe française (Lebeyec et al)[290] pour l'étude des décompositions métastables d'ions produits par désorption laser pour être ensuite appliqué au mode d'ionisation/désorption MALDI et plus tard pour les sources d'ionisation externe.

Le réflectron (Figure 79) est disposé à l'extrémité du tube de vol, à l'opposé de la source (ou de la zone d'injection des ions formés dans une source externe). Il est constitué de lentilles électrostatiques sur lesquelles un gradient de tension est appliqué. Dans le cas de notre instrument, l'axe du réflectron diffère de celui du temps de vol. Les ions de même rapport m/z entrent de manière plus ou moins simultanée (dû à la dispersion en Ec) en profondeur dans le réflectron. Les ions m_i de plus grande Ec entrent plus loin dans le réflectron, ou ils sont ensuite repoussés vers le chemin inverse sous un autre angle et une plus forte énergie cinétique, alors que ceux ayant une Ec plus faible entrent à peine dans le miroir électrostatique et en ressortent immédiatement avec une Ec encore plus faible. Ainsi les ions m_i ressortent avec une plus forte Ec mais en retard par rapport à ceux de plus faible Ec. Ces deux paquets d'ions peuvent alors après un temps de vol arriver en un point commun : une surface commune perpendiculaire à l'axe moyen du faisceau qui émergea du réflectron. Si en cet endroit est placé le détecteur, les ions m_i ayant subi une dispersion d'énergie cinétique sont alors détectés en même temps. Comme les ions fragments produits en vol par dissociation des ions m_i sont plus légers que leur précurseur ils vont être plus rapidement éjectés et auront un temps de vol différent de leurs précurseurs. Le réflectron permet un allongement de la distance parcourue par les ions dans l'analyseur, ce qui permet en plus de l'effet du miroir électrostatique, d'augmenter de façon considérable la résolution de 5 000 à 20 000. Ce gain en résolution se fait du détriment de la gamme de masse utilisable. En effet, les ions de rapports m/z supérieurs à 3000 m/z ne sont plus détectables par le TOF.

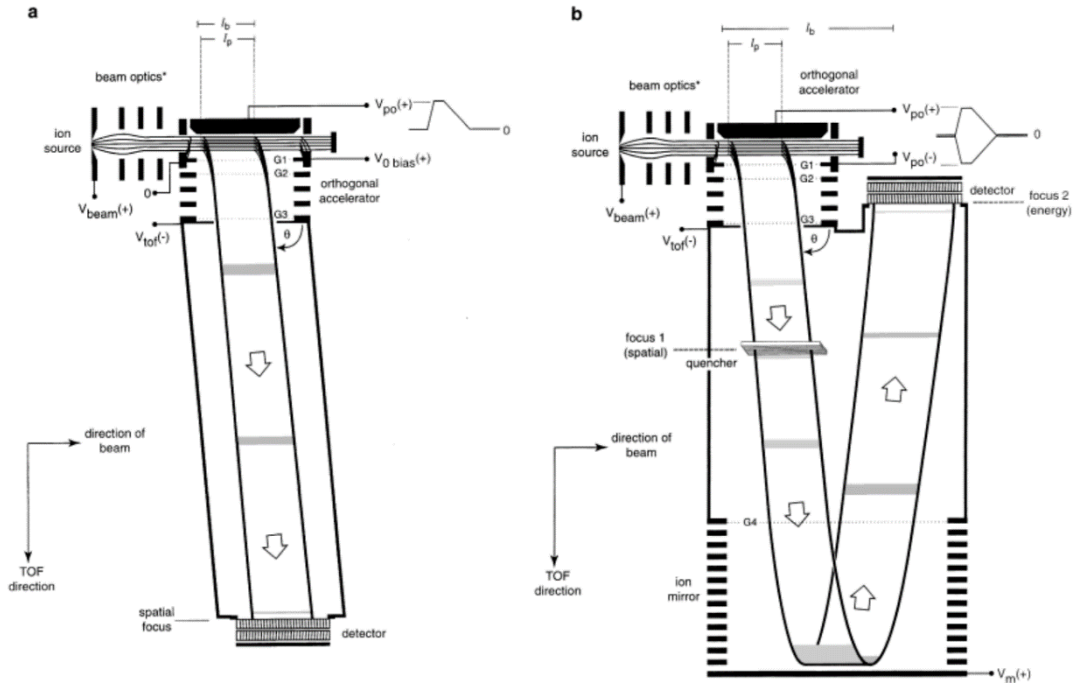


Figure 79 : Principaux composants du système orthogonal TOF (a) en basse résolution, (b) en mode haute résolution d'après [289]

5.1.4 L'analyseur Orbitrap

En 1923, Kingdon[291–293] a mis en place les prémices du futur analyseur Orbitrap sous la forme d'un cylindre, avec à son centre un filament permettant le piégeage des ions. Un premier champ quadripolaire est créé par le cylindre et un second par le filament central. Ce dernier champ est inversement proportionnel à sa distance par rapport au filament. Un tel champ est créé par un potentiel logarithmique dont la dérivée donne l'expression du champ. Compte tenu des deux composantes du potentiel, il sera considéré comme un potentiel quadro-logarithmique. En 1981, Knight réalise une modification essentielle sur son spectromètre de masse, afin d'étudier les ions produits par un laser. Il referme les extrémités du cylindre proposé par Kingdon pour permettre la mise en place d'un véritable piège à ions[294]. Par la suite, il ajoute un détecteur temps-de-vol au cylindre de Kingdon, pour détecter des ions éjectés par le piège et de leurs rapports m/z. Au milieu des années 1990, A. Makarov, a calculé les lignes de potentiel quadro-logarithmique du piège de Kingdon. Au sein de la société HD Technologies Ltd. (Atlas House, Simonsway, Manchester, M22 5PP, U.K.) il a construit un piège ayant les formes de ces lignes de potentiels. C'est ainsi qu'est né le piège électrostatique Orbitrap[295] caractérisé par l'absence de tension radiofréquence dont la première publication est parue en 2000[234]. Cette nouvelle technologie est constituée globalement de trois électrodes.

L'une est l'électrode centrale, en forme de fuseau imbriqué dans les deux autres, qui sont en forme de «tonneau creux coupé à mi-hauteur» (Figure 80)[234]. La forme de ces trois électrodes vont notamment modifier la façon dont il faut injecter les ions dans un piège sans radiofréquence de piégeage d'ions : l'Orbitrap[296,297,298,299,300].

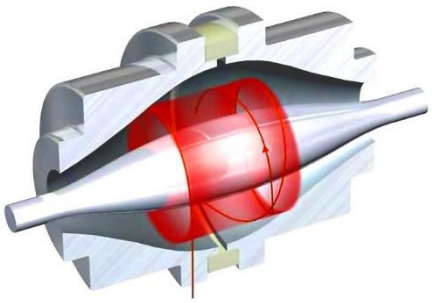


Figure 80 : Coupe longitudinale permettant la visualisation des électrodes d'un détecteur Orbitrap

Après injection grâce à une haute tension continue, un paquet d'ions va être piégé. Ils vont prendre des trajectoires particulières qui peuvent être décrites par deux composantes : une fréquence ω_r dont le mouvement est radial (r) et une fréquence ω_z dont le mouvement est axial (z). La fréquence axiale ω_z étant indépendante de l'énergie des ions et de leurs dispersions spatiales, c'est elle qui sera choisie pour mesurer le courant transitoire. Pour la mesure d'un tel courant lorsque les ions sont soit d'un côté soit de l'autre, il faut que l'électrode « tonneaux » soit bien coupée en deux dans son axe. C'est l'origine du choix de la configuration de ces deux électrodes symétriques. Les ions vont être séparés en fonction de leurs rapport m/z. Les ions les plus légers ont des forces de répulsions avec les électrodes supérieures et inférieures plus fortes, ce qui les entraînent au plus proche de l'électrode centrale et les obligent à parcourir un chemin le long de l'axe z plus important. A l'inverse, les ions de rapports m/z plus grands, présentent des forces de répulsions plus faibles avec l'électrode centrale, cela aura pour conséquence des fréquences de mouvements plus faibles pour ces ions. C'est par ces fréquences de mouvements différentes que le détecteur Orbitrap permet d'effectuer une séparation des ions en fonction de leurs rapports masse-sur-charge.

En 2005 le premier spectromètre de masse avec la technologie Orbitrap est commercialisé par l'entreprise ThermoFisher. L'Orbitrap est uniquement un détecteur qui fonctionne tel un piège à ions, sans radiofréquence, dans lequel les ions se mettent en orbites à une fréquence spécifique de leur rapport masse-sur-charge. Les ions sont injectés par paquets dans la cellule et de façon tangentielle à l'aide d'une haute tension. Comme les ions fragments sont envoyés par paquets, il est nécessaire de les stocker et compresser ce nuages d'ions en amont : c'est le rôle de la C-Trap. C'est grâce à un vide très poussée, de l'ordre de

10^{-11} hPa, et d'une injection rapide des ions depuis la C-Trap, que l'ensemble des ions restent stabilisés quelques centaines de millisecondes dans l'Orbitrap. Les fréquences axiales sont mesurées, elles sont obtenues par l'apparition alternatives des charges induites sur les deux électrodes qui composent le « tonneau ». L'ensemble des fréquences ainsi obtenues sont détectées sous forme d'un signal transitoire (de l'ordre de quelques centaines de millisecondes), qui est d'autant plus long que la pression est faible. Plus ce signal est long plus la résolution est grande. C'est *via* la transformée de Fourrier de l'ensemble de ces fréquences, que l'on obtiendra les rapports masse-sur-charge de chacun des ions constituant les paquets d'ions stockés dans la C-trap. C'est par ce traitement informatique que cet analyseur fera parti des instrument FT/MS.

Le critère de Nyquist est à l'origine de la haute résolution que peuvent atteindre les détecteurs Orbitrap. Il établit que plus le nombre de « points » effectués lors d'une analyse est grand, plus le calcul de la fréquence est « juste », et plus la précision en masse est augmentée. Cependant, l'augmentation du nombre de point doit s'accompagner d'une augmentation du temps de mesure. Généralement, la précision en masse sur ces spectromètres de masse atteint l'ordre du ppm. Ces détecteurs Orbitrap sont principalement couplés en tandem. Cependant, il existe également des spectromètres de masse Orbitrap couplés directement avec une source d'ionisation par électron ou même par électrospray (ce sont les technologies Exactive, ThermoFisher). Cet appareil, présent au laboratoire n'étant pas utilisé au courant de ma thèse, il ne sera pas détaillé davantage.

5.1.5. L'analyseur de mobilité ionique

Dans cette partie seront présentés un bref historique concernant le développement des techniques de mobilité ionique, les principes, les calculs permettant de comparer les sections efficaces de collisions expérimentales et théoriques, et enfin les techniques de couplages entre mobilité ionique et spectrométrie de masse. L'ensemble de cette étude bibliographique concernant la mobilité ionique ne sera pas développé comme je l'ai fait pour les autres techniques que nous avions utilisées sur notre plateforme. En effet, ne possédant pas l'instrumentation adéquate nous avons développé une collaboration, au sein de groupes d'experts, avec dans un premier temps M. Green, J. Ujma et K. Richardson (Waters, Wilmslow, Angleterre) et ensuite avec D. Ropartz et H. Rogniaux (Plateforme BIBS, INRAE, Nantes). Ainsi les notions essentielles à la compréhension des résultats obtenus avec ces nouveaux collaborateurs seront abordées.

Historique sur la mobilité ionique : Les fondements de la mobilité ionique sont apparus dès les années 1900, ou trois grands axes de recherche sont à considérer dans le développement de cette technique :

1) Au tout début des années 1900, Langevin élabore un travail théorique portant sur l'ionisation des gaz et sur la théorie cinétique des gaz[301]. Il met en évidence le rôle des forces attractives de partenaires sur la section efficace de collision. Il décrit pour la première fois les interactions ion-molécules et l'influence qu'elles peuvent exercer sur la mobilité.

2) En 1930 Mitchell et Ridler se consacrent à l'étude de l'influence du champ électrique et de la pression du gaz tampon sur la mobilité ionique en étudiant la diffusion de cations dans l'azote[302]. Leurs études sur la mobilité permettent de définir K la mobilité ionque en fonction de la relation suivante :

$$K = {}^{E}/_{N}$$

E : champ électrique, N : densité du gaz

Il est possible de définir différents types de régimes électriques : un régime de faible champ où K est une constante, et un régime de haut champ où K est dépendante du champ électrique. En parallèle, à la même période, les premières études portant sur les innovations techniques, concernant les optiques ioniques et l'amélioration de la détection des ions voient le jour. Les travaux de Cravath[303] et de Van de Graaf en 1926[304] sur l'injection des ions par paquets via l'utilisation de portes à ions (appelées « ion shutters ») améliorent grandement la technique. En 1935, Bradburry[305] développe une technique de mobilité où il utilise un champ uniforme (c'est-à-dire constant) entre deux plaques.

3) Le développement des instruments de mobilité ionique a lieu 20 ans plus tard, principalement initié par Mason et Mc Daniel[306], qui élaborent ensemble la théorie de diffusion des ions en phase gazeuse sous l'influence d'un champ électrique. Les premiers appareils commerciaux ont été développés à partir de leurs travaux. Dans les années 1960 Ferguson développe un instrument à « écoulement d'ions après ionisation » (« flowing afterglow ») utilisé dans le domaine de l'étude d'atmosphère variées en chimie-physique. Les développements qui suivent permettent le couplage de la mobilité ionique avec d'autres techniques de chimie analytique, notamment la GC, la LC et la spectrométrie de masse. Le premier couplage IMS-MS est réalisé en 1962 par Earl W. Mc Daniel, puis par les groupes de Ferguson et de Smith dans les années 1970, pour permettre l'étude des interactions ions-molécules en phase gazeuse[307,308]. A cette période, la mobilité

ionique est encore nommée « chromatographie plasma »[309] (semblable à la technique « flowing-afterglow »), et prend de l'ampleur, pour l'étude de composés organiques[310] comme l'héroïne et la cocaïne[311]. De plus, l'intérêt croissant des applications militaires pour cette instrumentation propulse cette technique en avant, notamment puisqu'elle permet la détection de nombreuses armes chimiques, explosifs, et drogues.

Principe de la mobilité ionique : La mobilité ionique représente la capacité d'un ion M à se déplacer dans un environnement gazeux, sous l'effet d'un champ électrique E[312]. L'ion est entrainé par le champ électrique et est ralenti par les collisions avec le gaz[313,314]. La mesure de la mobilité ionique K consiste à déterminer le temps nécessaire pour que l'ion M traverse un tube de distance L. Classiquement, les paquets d'ions voyagent au travers d'un tube de dérive, rempli de gaz, sous l'influence d'un champ électrique uniforme, c'est leurs arrivées respectives au détecteur qui est mesurée selon l'expression suivante :

$$v = KE$$

Dans cette équation, v est la vitesse de l'ion dans la cellule de mobilité ionique, K est la mobilité ionique exprimée en $cm^2.V^{-1}.\ s^{-1}$ et E est le champ électrique (eV) appliqué.

Les ions se déplacent à une vitesse constante dans le tube, leurs temps de diffusion respectifs vont dépendre de leurs sections efficaces de collision, c'est-à-dire de leurs conformations. Généralement, deux isomères peuvent être séparés en mobilité ionique. Le moins compact des deux auras *a priori* un temps de dérive plus élevé que le plus dense. Plus sa conformation est « ouverte » (volume de plus faible densité) plus sa surface efficace de collision avec le gaz (appelé gaz tampon) augmente et donc plus l'ion sera ralenti. Le gaz tampon va ralentir les ions en fonction de leur taille, leur forme, et leur charge. La différence de temps de dérive des ions peut s'expliquer par la mesure des surfaces de collision efficace (collision cross-sections (CCS ou Ω)) qui représentent la capacité de chaque ion à subir des collisions, grâce à la relation de Mason-Schamps[315] :

$$Mobilité\ ionique = \frac{3ze}{16N} x \sqrt{(\frac{1}{m} + \frac{1}{M})}\ x \sqrt{\left(\frac{2\pi}{kT}\right) x \frac{1}{\Omega}}$$

Où *z* correspond à l'état de charge de l'analyte ionisé, *e* la charge de l'électron, N est la densité du gaz tampon (calculée à partir de la mesure de pression réduite dans le tube de dérive), m la masse du gaz, M la masse moléculaire de l'ion, k est la constante de Boltzmann, et T la température du gaz tampon. Cette valeur de CCS (Ω) dépend aussi du gaz tampon utilisé mais pas de l'instrument[316].

Cette relation n'est valable que dans les cas où la masse de l'ion M est plus grande que celle du gaz[317]. La valeur de cette section efficace Ω peut être calculée pour chaque composé, il s'agit d'une propriété intrinsèque de la molécule, qui peut être obtenue à partir d'une structure théorique (ou validée) de l'ion. Il existe aujourd'hui différents instruments de mobilité ionique couplés à la spectrométrie de masse. Leurs configurations diffèrent d'un constructeur à l'autre, et dépendent notamment dans la façon dont le champ électrique et le gaz tampon sont implantés dans la cellule de mobilité. Tout comme en spectrométrie de masse, il y a quatre parties essentielles à prendre en compte lors d'une analyse de mobilité ionique : la source d'ionisation, l'injection des ions dans la zone de mesure de mobilité des ions, l'analyseur et le détecteur.

Outre l'ionisation à l'aide du ^{63}Ni[318] qui conduit à la formation d'ions radicaux et l'ionisation par transfert de proton, (PTR, proton transfer reaction)[319] qui produit de molécules protonées ou déprotonées (technique dérivant du mode «flowing-after-glow», les modes d'ionisation les plus courants, combinés aux analyseurs IMS, sont les processus de désorption par ESI et par MALDI.

L'injection des ions dans le tube de dérive constitue une étape clef dans une expérience de mobilité ionique. Il est nécessaire que le temps d'injection soit le plus court possible, afin d'obtenir la meilleure résolution possible du temps de dérive des ions. La première méthode développée par Brabury-Nielson et Tyndall pour l'injection des ions est la « porte » à ions. Il s'agit de deux plaques percées, parallèles, avec deux électrodes aux extrémitées, permettant le piégeage des ions. Lorsqu'un paquet d'ions est formé, le potentiel de la seconde électrode est annulé, le passage des ions vers le tube de dérive est alors possible. Cette technique est basée sur l'envoi de paquets d'ions par impulsion contrôlées dans le temps, selon un cycle pouvant se répéter n fois. Cependant, seule la faible quantité d'ions stockés entre les deux plaques est transférée vers le tube, les autres ions sont perdus, cela constitue la limite de la technique. En 1997, Smith[320] développe une technique permettant un gain de sensibilité, où les ions sont piégés dans un système de lentilles co-axiales dont l'intérieur peut-être décrit comme un entonnoir. Il s'agit d'une alternance d'électrodes cylindriques sur lesquelles des potentiels opposés sont appliqués, focalisant ainsi les ions sur leur axe commun, et les injectant dans le tube sans perte notable. En mobilité ionique, plusieurs types d'analyseurs coexistent, on retrouve notamment les tubes de mobilité ionique avec un champ électrique homogène (DTIMS Drift time ion mobility spectrometry), ceux avec un haut champ asymétrique (DMS, FAIMS Field Asymmetric Waveform), et ceux à champ oscillants (TWIMS Traveling wave IMS ou TWIG), notamment introduits par Waters.

Les analyseurs DTIMS sont composés d'une série d'électrodes cylindriques co-axiales, séparées par des anneaux isolants où est placée une résistance. Une différence de potentiel est appliquée sur les deux extrémités du tube de dérive ce qui donne naissance à un champ électrique faible mais homogène. Le DTIMS mesure un temps de dérive correspondant à la durée de migration au travers du gaz tampon dont l'ion a besoin pour atteindre le détecteur. Dans le tube de mobilité ionique la pression en gaz est constante, la vitesse des ions l'est aussi, c'est pour cela que le temps de dérive des ions peut être mesuré[321]. Il est possible de calculer la constante de vitesse des ions injectés ainsi que les sections efficaces de collision respective des ions vis-à-vis du gaz tampon.

Les analyseurs DMS et FAIMS se développent dans les années 1990 par Buryakov[322] ou les ions sont discriminés d'une manière comparable à un quadripôle. Cette technique fonctionne tel un filtre ionique, permettant la transmission de certains ions ciblés sous l'effets de paramètres contrôlés. La séparation des ions nécessite l'application sur une surface (considérée comme une électrode) d'une tension alternative (de l'ordre de 200 mHz) sous forme d'onde dissymétrique et la mise à la terre de l'autre électrode[323]. Le principe est d'appliquer de forts champs électriques oscillants et opposés sur deux électrodes perpendiculaires à la trajectoire des ions au travers du gaz tampon[313]. Par ce fait, les ions oscillent de haut en bas entre les deux électrodes. Le champ appliqué est composé de portions de champ élevé de courte durée et de portions de champ faible de plus longue durée. Lorsque le champ élevé est appliqué les ions suivent une trajectoire sinusoïdale, responsable d'un déplacement des ions d'une électrode vers l'autre. Pour que les ions n'aillent pas sur les plaques, une tension continue dit voltage de « compensation » CV (compensation voltage) est ajoutée[324]. En faisant varier cette tension CV, les ions peuvent, selon leur surface efficace de collision, émergés du système sans atteindre les électrodes. Une valeur de CV est appliquée spécifiquement pour une valeur de k donnée, ce qui permet aux ions ciblés de conserver une trajectoire rectiligne, et de sortir vers l'analyseur. Les ions sont détectés en fonction de leurs valeurs de CV. Sur ce type d'instrument, les valeurs de CCS ne sont pas disponibles. Cependant, l'appareil peut être utilisé selon deux modes : le premier consiste à balayer une gamme de CV afin d'obtenir une empreinte globale des ions présents, ou bien en sélectionnant un CV d'intérêt, permettant ainsi de réaliser une analyse ciblée[325].

Les technologies TWIMS (travelling wave ion mobility spectrometry) fonctionnent selon un principe similaire aux DTIMS. Le champ électrique va disperser les ions à différentes vitesses au travers du gaz tampon. Contrairement aux DTIMS, les TWIMS utilisent des champs oscillants et non des champs électriques uniformes[313]. Les champs oscillants du TWIMS peuvent être représentés comme des vagues successives[326] (Figure 81). Dans cette technologie, une série d'impulsions de voltage est utilisée afin de propulser les ions vers le gaz tampon et au travers de nombreuses collisions. Au fur et à mesure que les vagues passent le long de l'appareil, les ions « surfent » un certain temps, puis sont entrainés par la vague. Ce processus est répété, ce qui permet la séparation des ions en fonction de leur mobilité ionique. Les ions avec une mobilité ionique plus faible seront emportés plus vite que les ions à longue mobilité ionique, les ions sont détectés en fonction de leurs arrivées au détecteur. Ces technologies sont développées par la société Waters, le premier TWIMS commercialisé date de 2006, avec le spectromètre de masse Synapt HDMS (puis les Synapt G2 en 2011 et les Synapt G2-Si en 2013), et constituent la majorité des cellules de leurs spectromètres de masse à mobilité ionique[327].

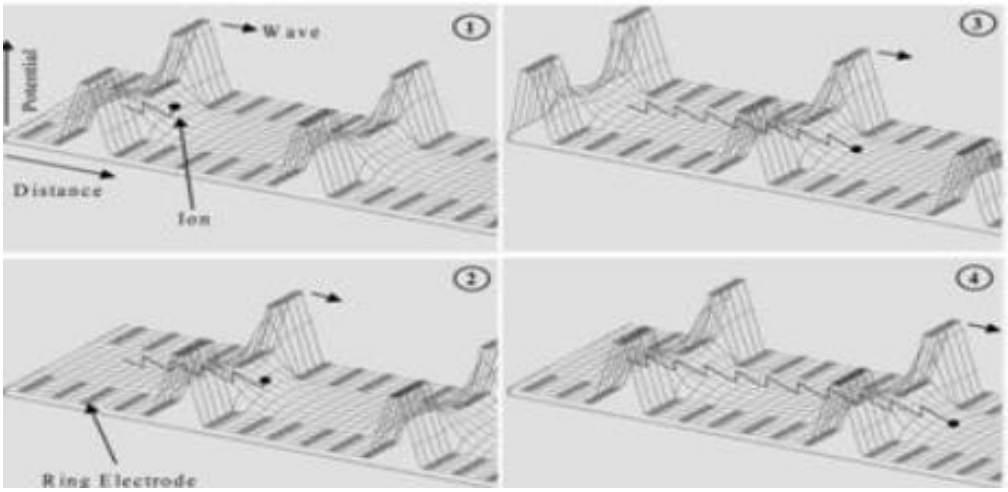


Figure 81 : Représentation des vagues successives produites par le champ électrique oscillant, et du mouvement des ions dans la cellule de mobilité ionique TWIMS[326]

Les technologies TWIMS peuvent être modifiées, car le TWIMS peut être remplacé par une cellule de mobilité cyclique (cIM), placée orthogonalement à l'axe principal de l'appareil, composé de deux zones (Figure 82). La première constitue le corps principal et est composée de 608 électrodes, la seconde constitue la zone où l'entrée et la sortie des ions est effectuée. L'avantage qu'offre cette technologie est un allongement du « tube de dérive », sans pour autant augmenter la taille de la cellule de mobilité, où les ions peuvent effectuer autant de tours que nécessaire. L'augmentation du trajet des ions, permet d'accroitre la résolution en temps de dérive, résultant en une séparation plus efficace de certains isomères[328]. Dans ces appareils la résolution augmente, puisqu'elle est proportionnellement liée à $\sqrt{n}$ ou n correspond aux nombres des tours dans la cellule de mobilité cyclique. C'est cette technologie qui a été utilisée durant mes travaux de thèse, en collaboration avec M. Green, K. Richardson et J. Ujma de la société Waters (Wimslow, UK).

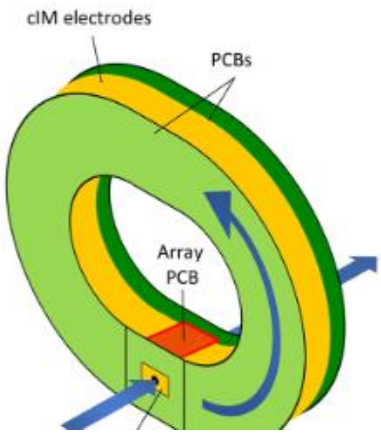


Figure 82 : Schéma de la cellule de mobilité cyclique d'après[328]

Au départ, le moyen de détection des ions, étaient souvent des plaques de faraday, où le courant induit par le flux d'ions en fonction du temps était mesuré. A l'heure actuelle, ces cellules de dérive sont souvent couplées à des analyseurs tels qu'ils ont été décrits précédemment (TOF, ion trap ou filtre de masse quadripolaires et maintenant Orbitrap)[322]. Les premiers couplages IM-MS permettent l'observation de différents conformères, notamment en ce qui concerne le cytochrome C. Celui-ci après désolvatation est 7 fois chargé et présente 4 conformères une fois désolvaté en phase gazeuse[329].

Les couplages IMS-MS permettent de donner une nouvelle dimension à la métabolomique. En effet, il était jusqu'à présent difficile d'identifier certains métabolites, à cause de l'existance de nombreux composés isobares ou isomères qui ne pouvaient pas être distingués malgré les couplages GC et LC[316]. Le couplage IMS-MS permet la séparation de ces molécules et leur identification. Pour l'analyse du lipidome, les techniques d'IMS-MS se

sont développées et permettent d'améliorer la fiabilité de l'identification des lipides[330], notamment pour les composés isomériques (isomères de positions), ou pour distinguer la position de doubles liaisons dans des échantillosn biologiques complexes. En effet, à l'aide des calculs de CCS et de la séparation réalisée en IMS-TOF, il est possible de distinguer des glycérophosphatydilcholines comme les PC 18:1 (6Z)/18:1(6Z), les PC 18:1 (9Z)/18:1(9Z) et les PC 18:1 (9E)/18:1(9E)[331].

5.1.6. L'analyseur FT-ICR

La spectrométrie de masse à transformée de Fourrier (FT-MS) est apparue au milieu des années 1970 avec la résonnance cyclotronique des ions (FT-ICR), initiée par Comisarow et Marshall[332,333,332]. Il s'agit d'une cellule cubique, installée dans un tube à l'intérieur d'un champ magnétique de quelques Tesla, initié à l'aide d'un système supraconducteur. Les ions sont produits *via* une ionisation par électrons à l'intérieur de cette cellule, et sont piégés par des plaques opposées. Dans de telles conditions, les ions tournent autour d'un axe dans la cellule. Dans les années 1985 cet instrument se développe surtout pour la recherche fondamentale. Dans les années 1990, les instruments ont été suffisamment modifiés pour être automatisés, et les rendre accessibles à la chimie analytique. Aujourd'hui, les champs des aimants ont fortement augmenté à 12 Tesla environ et les cellules sont devenues cylindriques. La résolution de cet appareil se situe au-delà du million pour des rapports masses-sur-charge aux alentours de 200. Ces instruments se sont démocratisés, les coûts sont amoindris avec des systèmes récupérateurs d'hélium produit par l'aimant supraconducteur. Ces appareils sont dédiés à l'analyse des biomolécules quel que soit leurs tailles.

5.2. Les analyseurs en tandem

Pour l'identification de nouvelles molécules et leur caractérisation structurale notamment en lipidomique, la mesure des rapports m/z des espèces moléculaires chargées ou de leurs ions fragments impliquent dans le tandem, un dernier analyseur haute résolution. La haute résolution permet l'attribution plus précise de la composition élémentaire des ions fragments ce qui facilite l'interprétation des spectres de masse ou des spectres de fragmentation.

Les ions fragments peuvent être produits à l'aide de plusieurs mode d'activation, nous nous focaliserons uniquement sur les modes collisionnels d'activation. Les modes comme la photo-dissociation (IRMPD), le bombardement électronique (EID) ou les autres modes d'activations plutôt dédiés à l'analyse d'ions multichargés ne seront pas abordés. L'activation collisionnelle des ions peut se produire de deux manières différentes : le mode resonant et le mode non resonant qui sont dépendant de l'instrumentation choisie. Il est possible d'activer les ions après une ou deux collisions (mode non resonant dit « chauffage rapide »), ou après plusieurs centaines de milliers de collisions (mode resonant dit « chauffage lent »)[334].

Activation des ions sous mode résonant : Le mode résonant est le mode d'activation retrouvé dans les techniques de piégage d'ions (piège à ions et cellule ICR). L'ion sélectionné est caractérisé par sa fréquence de mouvement propre, ω_M pour une valeur de $q_{z,act.}$ donnée. Cet ion est excité en lui appliquant ω_M dont l'amplitude doit être bien inférieure à la profondeur du puits de pseudo-potentiel afin d'éviter son ejection. Ces collisions sont faiblement énergétiques, mais plus grandes que les collisions utilisées lors de réactions ions-molécules. En général lors d'une collision activante, la suivante est souvent désactivante, cela signifie qu'elle amène l'ion à sa relaxation en Ec et en énergie interne. Cependant, il reste toujours un résiduel d'énergie vibrationnelle dans l'état fondamental. Au cours des collisions ce résidu d'énergie vibrationelle va s'accroître, jusqu'au moment où l'énergie interne atteint le seuil de dissociation de l'ion. La durée de cette opération de collisions multiples atteint une à plusieurs dizaines de millisecondes. L'excitation avec le mode resonant est donc appelée « chauffage lent ».

Exemple : Pour un ion précurseur m/z 400, si nous souhaitons détecter un de ces ions fragments de rapport m/z 85 suite à son activation, il est nécessaire que cet ion soit à une valeur $q_z < 0.908$ pour que ce dernier reste piégé. Dans les conditions d'activation recommandée par le constructeur ($q_{z;act} = 0.25$), l'ion produit, de rapport m/z minimum capable d'être détecté est défini par l'équation suivante : $m = (400/0.908)*0.25 = 111$, donc les ions de rapport m/z < 111 ne peuvent pas être piégé et donc détecté. Ce phénomène est appelé low

mass cut off (LMCO). Dans ces conditions, l'ion fragment d'intérêt m/z 85 aura une valeur $q_{z,m/z\ 85}$ de piégeage de 1.17 (i.e., $q_z = {}^{400}/_{85} * 0.25$), valeur bien supérieure à 0.908 où l'éjection naturelle des ions a lieu[335]. Afin de le détecter, il est nécessaire de le piéger dès sa formation, il faut alors que sa position sur le diagramme de stabilité soit à une valeur de $q_{z,quad}$ < 0.908. Pour cela la valeur de $q_{z,act.}$ de l'ion précurseur m/z 400 doit être ramenée à une valeur plus faible corrélée à celle ($q_{z,m/z\ 85}$) imposée pour la détection de l'ion fragment m/z 85. Si cet ion doit être formé à $q_{z,m/z\ 85}$ = 0,85 alors la valeur de $q_{z,act}$ de son ion précurseur m/z 400 devra être égale à 0,18 (résultant de la proportion ${}^{85}/_{400} * 0{,}85$). Cela signifie que l'ion m/z 400 est piégé dans un puits moins profond, il est donc nécessaire de réduire l'amplitude d'excitation de l'ion activé, sous peine de le voir éjecté du puits et de le perdre. En conclusion, si nous souhaitons voir cet ion fragment de faible rapport masse sur charge, il est nécessaire de lui donner beaucoup d'énergie interne, il faut donc augmenter son temps d'activation de 10 ms à 100 ms par exemple.

Avantages et inconvénients d'une excitation par mode résonant : Pour finir, il faut souligner qu'avec le mode résonant ce sont uniquement les processus compétitifs qui peuvent être mis en évidence. Avec ce type d'analyseur seuls les ions parents sont activés. Les ions fragments issus de ces ions ont une Ec qui tend vers 0 par relaxation, leurs fréquences de mouvements sont très différentes de celles de leurs précurseurs (la fréquence d'un ion est inversement proportionnelle au rapport m/z pour une valeur de q_z donnée). C'est pour cette raison qu'ils ne peuvent pas être excités par la fréquence d'activation des ions parents. Lorsque des ions fragments de décompositions consécutives apparaissent sur le spectre d'ion produit à partir d'un ion parent sélectionné, cela signifie que l'énergie interne résiduelle de l'ion produit de première génération est suffisamment importante pour produire la fragmentation ou bien, que nous avons affaire à des processus exothermiques[223], ce qui reste rare. C'est pour cela que généralement, il n'y a pas de décompositions consécutives observées lors d'une analyse MS/MS en mode résonant sur un piège à ions.

Activation des ions en mode non-résonant : Dans le cas du mode d'activation non-résonant, c'est-à-dire avec un faisceau d'ions traversant la zone de collision, il faut considérer l'énergie cinétique au centre de masse (E_{COM}). C'est cette énergie qui correspondra à l'Ec maximale qui sera convertie en énergie interne, ou en énergie vibrationelle dans l'état fondamental.

$$E_{COM} = E_{LAB} \times \left(\frac{masse\ de\ la\ cible}{masse\ de\ la\ cible + masse\ de\ l'ion}\right)$$

Où la cible est le gaz de collision non réactif

Cette conversion d'Ec en énergie interne sera d'autant plus efficace que la durée de vie du complexe de collision sera longue. Une collision est efficace lorsque la conversion de l'Ec en énergie interne est maximale. Cette situation correspond au cas où l'Ec est très faible (conditions de réaction ion-molécule). Récemment, un modèle simplifié a été proposé[336] : ou il est considéré qu'une collision « n » devient efficace lorsque l'énergie interne acquise pendant la collision conduit à une constante de vitesse de dissociation de l'ion parent suffisamment grande pour permettre alors sa décomposition avant qu'une collision consécutive « n+1 » ne se produise. Ce mode d'activation sera considéré comme un « chauffage rapide » car il est considéré qu'une ou deux collisions conduisent à une activation suffisante de l'ion pour que sa fragmentation soit quasi-spontanée. Ce type de «chauffage rapide» peut se produire à partir de faisceaux d'ions d'énergie cinétique allant de quelques eV à plus de 100 eV (au-delà les collisions deviennent trop inefficaces) dont la durée de résidence dans la zone de collision (multipole dit « RF only ») est considérée de l'ordre de 10 à 100 µsecondes[337].

Ce même mode d'excitation peut également se produire dans un piège à ions linéaire (mode HCD). Pour cela, un paquet d'ions sélectionné est injecté et piégé dans la cellule. L'activation se produit uniquement par la différence de potentiel entre les deux électrodes d'entrée et de sortie de la cellule. L'excitation dans ce piège n'est plus radiale (mode d'activation résonant) mais axiale comme pour un faisceau d'ions. Les décompositions consécutives deviennent alors possibles et une collision efficace peut se produire sur l'ion fragment de première génération, il n'est donc pas nécessaire d'augmenter la durée d'activation au delà de 10 µsecondes (valeur prise par défaut)[338]. L'ion parent et ses ions fragments sont relaxés en Ec et leurs activations collisionnelle se trouvent fortement réduites. Cette durée d'activation correspond souvent aux durées utilisées pour la relaxation des ions dans un piège entre chacune des manipulations des ions (injection, sélection, excitation, analyse etc…) lors d'analyses par séquence MS^n. Notons que dans ce mode non-resonant, la discrimination en bas rapport m/z due à l'existence du LMCO n'a plus lieu. Un même comportement est observé pour les ions lorsqu'ils sont activés par collision dans la cellule HCD ou dans un multipôle (souvent quadripolaire) « RF only », c'est notamment le cas d'un triple quadripole ou d'un Qq/TOF. Comme la fréquence de mouvement est inversement proportionelle à la masse, lors de la formation d'un ion fragment, dans un multipôle (ou dans un quadripôle, ou dans la cellule HCD) sa fréquence de mouvement dans le champ (quadripôlaire ou multipolaire) est plus grande que celle de son précurseur. Ainsi, le chemin parcouru par l'ion fragment dans la cellule de collision est plus grand que celui de son précurseur, donc sa probabilité collisionnelle est supérieure à celle de l'ion précurseur. Si son énergie au centre de masse est suffisante, une dissociation consécutive spontanée peut se

produire. Finalement, ce comportement est différent de celui dans un piège à ions où les décompositions compétitives sont plutôt mises en évidence.

5.2.1 Le triple quadripôle

Le triple quadripôle (QqQ) a été initié par Yost et Enke en 1979[230], il s'agit d'une association de 3 quadripôles montés en série : Q1, q2 (cellule de collision) et Q3 entre lesquels se trouvent des lentilles optiques (Figure 83). Cet analyseur est un outil de choix pour la mesure quantitative de composés d'origine biologique. Le triple quadrupôle peut être utilisé selon plusieurs modes d'acquisitions[20], pour mettre en évidence des réactions du type : spectre d'ions produits à partir d'un précurseur, spectre d'ions précurseurs à partir d'un ion fragment, ou spectre de perte de neutres constantes.

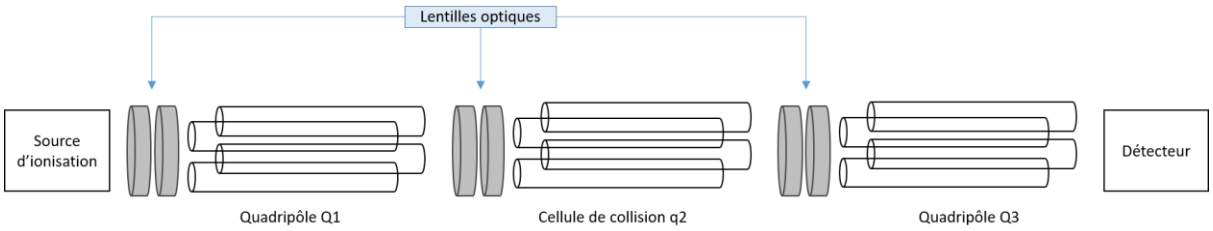


Figure 83 : Représentation schématique d'un analyseur en tandem triple quadripôle

Dans un spectromètre de masse triple quadripôle, les ions sont accélérés en sortie de source (tension souvent référencée à la terre T_0) et sont accompagnés jusqu'au détecteur par le simple accroissement des référentiels $T_{(1)}$ et $T_{(3)}$ (en valeurs absolues) de $U_{(1)}$ et $V_{RF(1)}$ de Q1 et de $U_{(3)}$ et $V_{RF(3)}$; de Q3 en considérant que les ions émergent aux bonnes tensions $U_{(i)}$ et $V_{RF(i)}$ à chaque correspondance.

Spectre de masse sur un QqQ : Afin de réaliser un spectre de masse sur un instrument de type triple-quadripôlaire il faut réaliser un balayage des tensions U et V_{RF} soit du filtre de masse Q1 soit de Q3. Pour cela les quadripôles fonctionnent en mode « RF only », pour seulement transmettre les ions d'une zone à une autre. Lors de l'utilisation en filtre d'ions Q1 (avec des valeurs fixes de $U_{(I)}$ et $V_{RF(I)}$), alors ce quadripôle permet de sélectionner un ou plusieurs ions spécifiquement, selon leurs valeurs de rapports masses-sur-charges. Un transfert des ions de la source qui correspondent aux ions précurseurs sélectionnés selon leurs rapports m/z est effectué dans Q1. Les quadripôles q2 et Q3 sont alors « transparents », c'est-à-dire que seule la tension V_{RF} est utilisée, ils servent uniquement à transmettre les ions : ce sont des filtres de masse. Cette étape permet d'obtenir un spectre de masse de tous

les ions précurseurs à partir de l'échantillon ionisé et émergeant de la source. De la même manière, lorsque l'on utilise Q3 plutôt que Q1 pour effectuer un spectre de masse alors c'est Q1 et q2 qui sont « transparents » servant uniquement à transmettre les composés ionisés vers Q3.

Spectre d'ions produits sur un QqQ : les ions sélectionnés M en Q1 ($U_{(1)M}$ et $V_{RF(1)M}$) fixés pour le rapport m/z de l'ion M choisi sont focalisés et transférés vers la cellule de collision, ou se produisent les collisions qui activent les ions afin de les dissocier. Elles sont produites par un gaz inerte, ainsi les fragmentations ont lieu par un processus de « dissociation induite par collision » (collision-induced dissociation, CID). L'énergie de collision (E_{Lab}) est définie par la différence de potentiel de référence du quadripole q2 de celui de la source (qui est nul - la terre). Les ions fragments ainsi formés sont transmis vers le filtre de masse Q3 qui sert uniquement à détecter les ions fragments produits.

Cette méthode de collisions activantes (mode non-résonant) repose sur la probabilité de collision entre les ions et les atomes de gaz. Cette collision, si elle est efficace, va induire une conversion partielle de l'énergie translationnelle en énergie vibrationnelle. Lorsque celle-ci dépasse le seuil énergétique de dissociation alors, l'ion précurseur se fragmente en ions produits. La fragmentation d'un ion peut se faire à plusieurs niveau d'énergie de collision : haute (de l'ordre du KeV) ou basse (de l'ordre de la centaine d'eV). En ce qui nous concerne et avec les instruments employés, seules les collisions sous faible énergies ont été utilisées (puisque nous n'avons pas utilisé de tandem TOF/TOF ou le mode d'activation EID).

Mode d'acquisition Selected/Multiple reaction monitoring sur un QqQ (Figure 84) : A partir du tandem QqQ, l'acquisition d'une seule transition correspondante aux rapports m/z d'un précurseur en Q1 et de son ion fragment associé, généralement l'ion le plus intense est possible. Pour cela, les tensions de balayage utilisées permettent le mode « selected reaction monitoring » (SRM). A l'inverse, pour la recherche de plusieurs transitions d'ions fragments produits pour un même précurseur, le mode MRM (multiple reaction monitoring) qui permet le passage d'une transition à une autre est utilisé. Pour cela les tensions U et V_{RF} de Q3 ne sont pas balayées mais sont « sautées » d'un couple de tensions (U/V_{RF}) spécifique, à un autre ce qui permet de « passer » d'une transition à l'autre pour un même précurseur. Dans ce mode d'acquisition, le dernier quadripôle (Q3) permet de séparer et transmettre en tant que filtre de masse les ions produits sélectionnés vers le détecteur. Ce quadripôle aura également sa valeur de tension de référence qui « flottera » de la même manière que celle de Q1 et q2. Le choix des rapports m/z des ions fragments détectés lors d'une analyse SRM ou MRM est réalisée en fonction de l'abondance des ions produits. C'est la transition la plus abondante qui sera sélectionnée, et demandée au spectromètre de masse de rechercher en mode SRM. Ce

mode d'acquisition est spécifique puisqu'il ne demande pas un balayage complet sur une gamme de tension. En effet, seules les tensions correspondanteds aux rapports m/z demandés par l'utilisateurs seront balayées. Ainsi, ils offrent l'avantage d'un temps de balayage par transition beaucoup plus long, permettant d'augmenter la sensibilité de ce type d'analyseurs, et donc d'abaisser la limite de détection de certains analytes. Ce mode d'acquisition a été utilisé durant mes travaux de thèse.

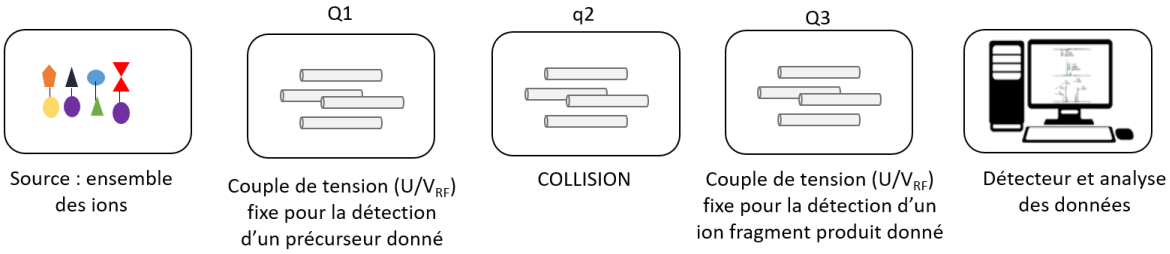


Figure 84 : Représentation schématique du mode d'acquisition SRM, sur un analyseur triple quadripôle

Mode d'acquisition « precursor ion scan » sur un QqQ (Figure 85) : Il s'agit d'un spectre d'ions précurseurs issus d'un ion fragment constant. Il permet de sélectionner dans Q3 un ion fragment (en fixant un couple de tensions U/V_{RF}), généré dans la cellule de collision. Tous les ions précurseurs possibles entrent dans Q1 par balayage continu des tensions U/V_{RF}, ils vont tous être expulsés consécutivement vers q2, où ils sont tous fragmentés. Ainsi quand un ion précurseur se fragmente en donnant l'ion sélectionné dans Q3, alors celui-ci sera détecté en sortie.

Le détecteur ne détecte donc que l'ion fragment sélectionné. La terre pourra « flotter », de la même manière que dans le mode d'acquisition MRM. Ce mode d'acquisition permet de détecter l'ensemble des ions fragments de structure commune ou isomères (voire même isobares) ayant un enchainement d'atome identique au sein de leurs structures. C'est les couples de tensions U/V_{RF}, qui sont mesurés à chaque fois qu'un signal est enregistré au détecteur. Ce sont ces couples qui permettent de remonter à la valeur des rapports m/z des ions précurseurs. Ce mode d'acquisition a été utilisé à de nombreuses reprises durant mes travaux de thèse.

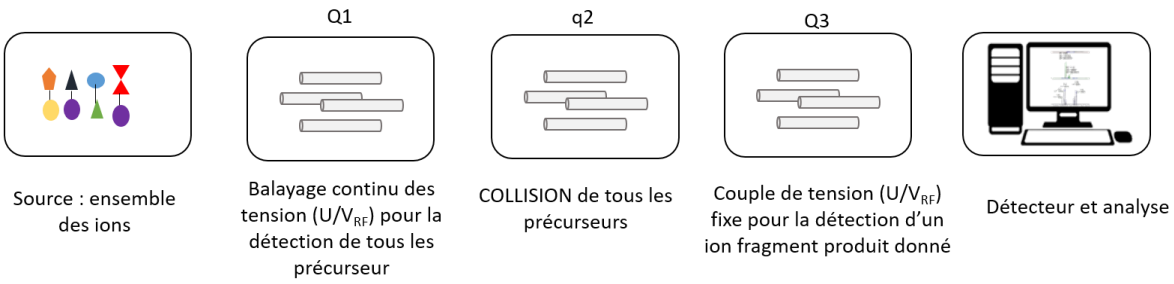


Figure 85 : Représentation schématique du mode de *balayage « Precursor ion scan »*, sur un spectromètre de masse triple quadripôle

Mode d'acquisition perte de neutre sur un QqQ : le mode d'acquisition « neutral loss » (NL), conduit au spectre de neutre constant, qui est basé sur la perte fixe d'un neutre associé à l'ion précurseur. Ce mode de balayage est une combinaison des deux modes détaillés précédemment. Sur Q1 et Q3, l'ensemble des couples de tensions U et V_{RF} sont balayés avec un décalage de la masse du neutre étudié. Ce mode de balayage est nommé « linked-scan » ou mode de balayage lié. Ce mode d'acquisition permet de détecter l'ensemble des ions parents ayant perdu une même neutre. Il peut être utilisé afin de rechercher des molécules ayant des chemins de fragmentations communs. Ce mode de balayage a été utilisé pendant mes travaux de thèse.

L'analyseur triple quadripolaire est l'analyseur le plus répandu en spectrométrie de masse pour la quantification des métabolites, car il est très sensible et qu'il possède une large gamme linéaire. Il est utilisé couramment en toxicologie et en pharmacologie. Il offre *via* ses différents modes de balayages tout un panel d'informations. Toutefois, certaines limites sont quand même bien présentes, puisqu'il s'agit d'un appareil à basse résolution. L'analyseur triple quadripôle ne permet pas d'effectuer des mesures de *m/z* précises, ce qui limite ses applications pour l'identification et la caractérisation de nouveaux marqueurs ou composés présents dans une matrice biologique complexe.

Le laboratoire est équipé de deux spectromètres de masse à basse résolution LC-QqQ : le premier est un spectromètre de masse Agilent K6460 (acheté en 2010), le second est spectromètre de masse Shimadzu 8060 (acheté en 2020), équipé d'une source Mikros ou le couplage avec la micro chromatographie liquide est possible, pour permettre un gain de sensibilité. Dans les instruments à faisceaux d'ions que nous avons utilisés (QqQ et Qq/TOF), l'activation collisionnelle d'ions sélectionnés a été faite dans des cellules de collision multipolaire « RF only » pour le mode non-résonant.

5.2.2 L'hybride quadripôle – temps-de-vol (QqTOF) : analyse en haute résolution des ions produits

Le spectromètre de masse hybride quadripôle temps-de-vol (Qq/TOF) est constitué d'un premier quadripôle (Q1, filtre de masse), d'une cellule de collision (q2, RF only), et d'un analyseur temps-de-vol[276]. La sélection des ions d'intérêts est effectuée dans le premier quadripôle pour être transférés dans la zone d'activation collisionnelle. Les ions activés se dissocient dans la cellule de collision q2. Notons que la zone qui suit la cellule de collision, permet d'obtenir une focalisation résultant en un angle plus étroit, permettant une diminution de la dispersion axiale à partir de là où le faisceau d'ions est injecté dans la zone où l'impulsion haute tension orthogonale dite « push-out pulse » se produit. Elle permet d'accélérer le paquet d'ions selon un axe orthogonal à l'injection du faisceau émergeant de la source ESI, pour « voler » vers le réflectron. Les ions après le premier trajet entrent dans le réflectron où ils sont réfléchis selon des vitesses : soit légèrement différentes pour les refocaliser en temps, afin d'arriver ensemble au détecteur, soit très différentes pour séparer en temps les ions précurseurs et les ions fragments formés en vol dans la première région libre de champs (zone avant le réflectron).

L'analyseur temps-de-vol ne permettant pas naturellement la sélection d'un ion parent, le seul mode d'acquisition possible pour une analyse ciblée (mode MS/MS) est la sélection de l'ion parent en Q1[276]. Un tandem hybride Qq/TOF possède une résolution avoisinante les 20 000, pour des *m/z* de 400. Il est couramment utilisé pour les analyses globales en lipidomique et pour la recherche de nouveaux composés en vue de leurs élucidations structurales.

Le laboratoire est équipé d'un spectromètre de masse Qq/TOF, Xevo G2-XS (Waters, Figure 86). Il est équipé d'une source ESI en Z spray, d'une zone de désolvatation « stepwave » qui permet de désolvater les agrégats chargés et d'éliminer les molécules non ionisées dans la source, les dernières traces de solvants ou de sels. L'instrument est ensuite équippé d'un quadripole (nécessaire pour le mode d'acquisition TOF-SRM ou MRM), d'une cellule de fragmentation de type XS et d'un détecteur de type QuanTOF. La cellule de collision retrouvée dans cet instrument, la XS collision cell, est un quadripôle segmenté, constituée d'une association de 37 petits quadripôles de 4mm, avec un espace d'un mm par segment, pour une longueur totale de 184 mm. Dans cette cellule segmentée un gradient de courant continu est appliqué, permettant d'attirer très rapidement les ions vers le TOF.

Ce spectromètre de masse permet l'acquisition selon plusieurs modes de balayage : le mode MS^E en est une spécificité. Il correspond au mode de balayage couramment appelé « data independant acquisition » (DIA). Dans cette configuration l'ensemble des ions de m/z

différents entrent dans la cellule de collision, et sont fragmentés sans distinction. Les ions fragments produits sont envoyés à l'analyseur QuanTOF. Un spectre moyennant les ions précurseurs et leurs ions fragments produits est obtenu. D'autres types d'acquisition sont possibles : un balayage des m/z des ions précurseurs ou le mode TOF-MRM, où un ion est sélectionné spéciquement dans Q1, envoyé dans q2 et ensuite dans le TOF. Ce mode d'acquisition permet un gain de sensibilité puisque seuls les ions sélectionnés sont analysés par le spectromètre de masse, évitant ainsi un balayage sur une large gamme de m/z.

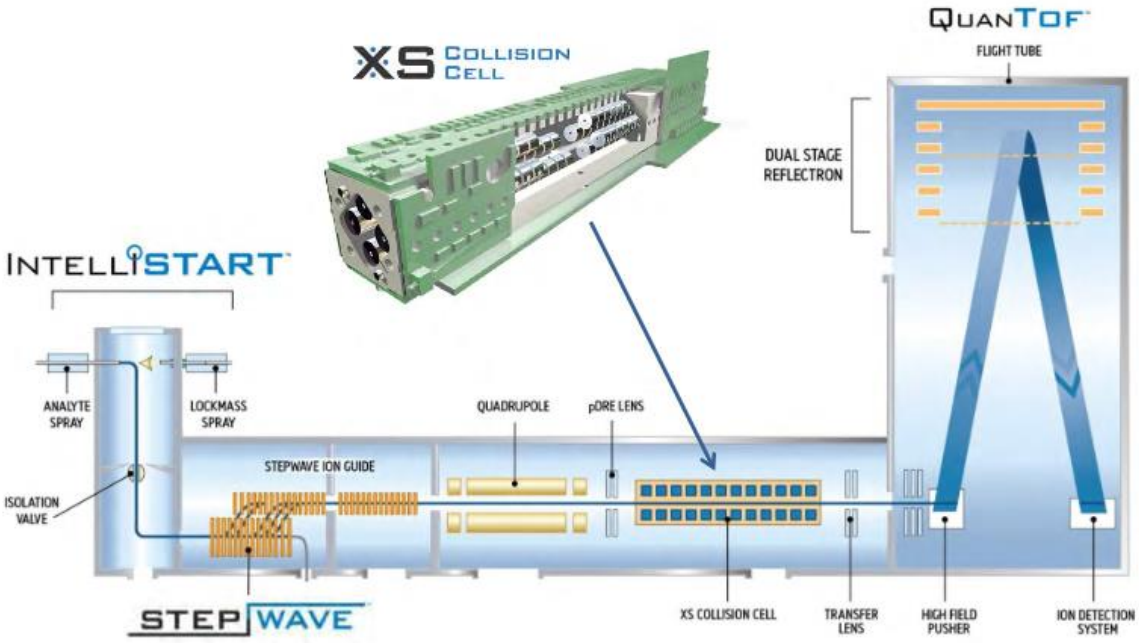


Figure 86 : Représentation schématique d'un Q-TOF Xevo G2-XS Waters et de sa cellule de collision segmentée XS Collision Cell[339]

5.2.3 Les analyseurs Orbitrap en tandem : le Q-Exactive Plus et Linear Trap Quadrupole Orbitrap VELOS

Aujourd'hui les spectromètres de masse constitués d'une source API et d'un analyseur très haute résolution tel que l'Orbitrap sont couplés à d'autres analyseurs pour en faire des tandems. Ce sont des instruments hybrides comme le Q-Exactive Plus et le Linear Trap Quadrupole (LTQ) Orbitrap Velos. Au cours de mes travaux de thèse, j'ai été amenée à utiliser ces deux spectromètres de masse. Le Q-Exactive Plus est composé après la source API, d'une lentille dite « S Lence » permettant une meilleure désolvatation des agrégats chargés et une meilleure focalisation du faisceau d'ions. Celui-ci sera injecté dans un guide d'ions courbé (Advanced Active Beam Guide AABG), afin de réduire le bruit de fond lié aux rayonnements des ions excités naturellement. L'AABG permet également d'éliminer des molécules neutres non ionisées dans la source. Par la suite, se trouve un quadripôle (HyperQuad Mass Filter) hautement sensible, avec en amont un préfiltre et en aval un post-filtre, permettant ainsi une meilleure focalisation des ions émergeant dans l'axe du quadripole. A la suite de l'HyperQuad Mass Filter se trouve la C-Trap qui permet de compacter les paquets d'ions comme décrit précédemment. Depuis la C-Trap deux actions sont possibles : soit d'envoyer vers l'analyseur Orbitrap un paquet d'ions émergeant de la source afin d'enregistrer un spectre de masse en très haute résolution, soit de transférer le paquet d'ions depuis la C-trap vers la cellule de collision HCD afin de produire les décompositions issues d'un ion sélectionné dans l'HyperQuad. Après dissociation les ions produits peuvent être recompactés dans la C-Trap afin d'être envoyés à l'analyseur Orbitrap. Dans les technologies VELOS et QExactive Plus la taille de l'analyseur Orbitrap est de 3 cm, contrairement aux technologies FUSION ou ELITE qui ont des analyseurs Orbitrap de 2 cm permettant une meilleure résolution en masse. L'ensemble de ces opérations de manipulations de paquets d'ions se déroulent en faisant « flotter » la terre de chacun des éléments du spectromètre de masse afin de diriger les ions dans un sens ou dans l'autre (Figure 87).

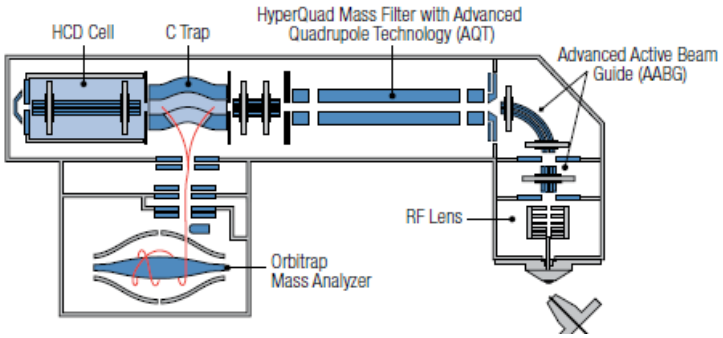


Figure 87 : Représentation schématique d'un spectromètre de masse Q-Exactive Plus Orbitrap

Le LTQ-Orbitrap VELOS (Figure 88) est équipé d'une source API, de S-Lens, d'un quadripôle parrallélépipédique à sections carrées (« Square quadrupole ») ayant le même rôle qu'un quadripôle de transfert mais plus facile à usiner ce qui amoindri son coût. A la suite de ce quadripôle se trouve un octopole permettant un transfert des ions vers le VELOS constitué de deux LTQ consécutifs. Le premier sous haute pression, constitue la zone d'excitation résonante et permet de produire la dissociation des ions. Le second LTQ sous plus basse pression permet notamment une meilleure analyse des ions lors de leur éjection radiale afin d'obtenir des spectres basse résolution, mais qui restent limités par le LMCO. A la sortie du LTQ se trouve un second octopole premettant de focaliser et transmettre les paquets d'ions vers la C-Trap. Les paquets d'ions peuvent suivre deux chemins à la sortie de la C-Trap, l'activation non-résonant par HCD ou l'analyse HR directe par l'Orbitrap.

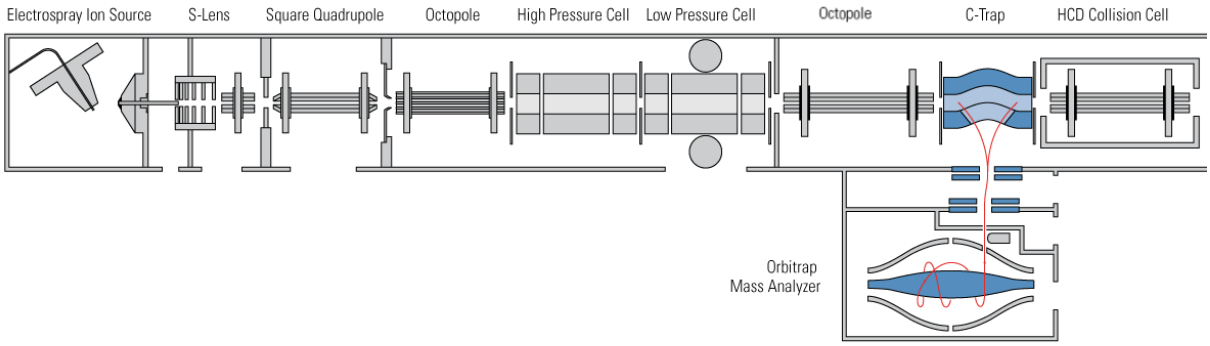


Figure 88 : Représentation schématique d'un spectromètre de masse VELOS couplé à un analyseur Orbitrap

Aujourd'hui le tandem VELOS-Orbitrap n'est pratiquement plus développé au profit du spectromètre de masse Fusion qui possède un quadripôle suivi d'une C-Trap, d'une cellule de collision multipolaire HCD, et d'un double piège VELOS. L'avantage de cette configuration est de permettre à partir des fragmentations en HCD de réaliser des analyses en basse résolution vers le piège, tout comme en haute résolution vers l'analyseur Orbitrap. La proximité de la cellule mulipolaire HCD (non résonant) et du VELOS (résonant) permet les opérations alternées d'isolement et d'activations d'un piège à l'autre.

5.2.4 Séquence de balayage pour les expériences MS^n à l'aide d'un piége à ions

Le principe de la MS/MS avec un piège à ions est le même qu'avec des tandems à faisceaux d'ions. La différence est que, dans ces derniers les ions traversent une cellule de collision, alors que dans les pièges à ions, ils sont stockés et manipulés (sélectionnés /activés/ éjectés) par des séquences temporelles et non plus spatiales. Ainsi pour une séquence MS^3, il s'agira d'ajouter une séquence MS à l'expérience en mode MS/MS.

Après un premier balayage permettant d'enregistrer le spectre de masse, un ion d'intérêt est choisi, puis l'ensemble des ions est réinjecté dans le piège à ions. L'augmentation de la tension V_{RF} permet d'éjecter l'ensemble des ions de rapports m/z plus petits que celui de l'ion d'intérêt. Lorsque ce dernier se retrouve à la limite de l'éjection quadripolaire (profondeur nulle du puits de pseudopotentiel) ou hexapolaire, une haute tension V_{RF} est appliquée en sens inverse afin de ramener l'ion à la valeur de q_z où se fera ensuite son activation (q_z = 0,25). Toutefois, il existe des ions de rapport m/z plus intenses (comme par exemple ses isotopes naturels). A la valeur de q_z < 0,25 (dont la valeur dépend de celle de l'ion mono-isotopique correspondant), une radiofréquence dipôlaire de forte amplitude est appliquée, pour permettre aux ions stockés non-souhaités, d'être éjectés du puits de pseudopotentiel. Cette étape constitue l'isolement de l'ion d'intérêt. Celui-ci, afin d'être fragmenté peut alors être activé, *via* l'application de sa fréquence de mouvements à la valeur de q_z=0,25. A ce moment-là, l'amplitude du mouvement de l'ion va augmenter, sans pour autant modifier sa fréquence, seule son énergie cinétique augmente. Cette étape est importante et constitue l'activation de l'ion d'intérêt. Une fois activé, cet ion se dissocie en ions fragments de rapports m/z plus faibles, dont la partie en dessous des valeurs de q_z = 0,908 peuvent rester stockés dans le piège à ions. Les autres ions fragments, ayant une valeur de q_z > 0,908 seront éjectés directement. Une augmentation de la valeur de V_{RF} permet alors d'éjecter du puits de pseudopotentiel l'ensemble des ions fragments ayant été stockés. C'est

l'ensemble de ces étapes qui produisent une séquence MS/MS. Pour effectuer une MS^3 ou une MS^n il suffit de répéter ces opérations 3 ou n fois en réinjectant préalablement des paquets d'ions après analyse des ions fragments de l'étape (n-1) pour choisir l'ion fragment qui deviendra un ion sélectionné. Entre chacune des manipulations ainsi effectuées lors d'une séquence MS^n, il est nécessaire de relaxer en énergie cinétique les ions parents survivants et les ions fragments produits pour permettre aux ions de se retrouver focalisés au centre du piège, c'est-à-dire au fond du puits de pseudo-potentiel. Cette étape est nécessaire pour ne pas risquer de perdre des ions qui seraient ainsi éjectés du puits de Dehmelt *via* des champs non linéaires et pour pouvoir appliquer efficacement la fréquence calculée par l'ordinateur puisque la fréquence des ions diminue lorsqu'ils s'écartent de ce centre et que l'excitation résonante devient plus ou moins efficace.

Chapitre IV : La spectrométrie de masse en lipidomique et les différentes types d'analyses possibles

Le lipidome est complexe et constitué d'un grand nombre de molécules aux propriétés physico-chimiques différentes, présentes à des quantités très différentes selon les échantillons. C'est pour cela que plusieurs types d'analyses coexistent en lipidomique : les analyses globales, les analyses ciblées et les analyses semi ciblées.

1. Les analyses globales

Les analyses de lipidomique dites « globales » ont pour but de détecter le plus grand nombres d'espèces moléculaires lipidiques possible dans un seul et même extrait[340]. Malheureusement, les protocoles utilisés pour la préparation des échantillons, souvent les protocoles de Bligh & Dyer[183] ou Folch[187], ne permettent pas d'extraire l'entièreté du lipidome, et donc certains lipides minoritaires ne seront pas détectés lors des analyses. C'est le cas par exemple des métabolites des acides gras polyinsaturés.

Il s'agit d'une analyse sans *a priori* sur la composition lipidique de l'échantillon. Ces analyses sont peu sensibles et ne permettent pas d'effectuer une quantification absolue de chaque composé. La capacité de chaque molécule à s'ioniser étant différente (même au sein d'une même sous famille), l'analyse quantitative est extrêmement compliquée lorsque les standards primaires de toutes les molécules recherchées ne sont pas disponibles et injectés en même temps. Seules les analyses en quantification relative (par rapport au standard interne) sont possibles lors des analyses globales. Dans certains cas, il est possible de réaliser des analyses semi-quantitatives, si lors de l'extraction des échantillons, un standard interne de chaque classe lipidique est ajouté.

Lors d'une analyse globale, le spectromètre de masse réalise un balayage de l'ensemble des composés sur une large gamme de masse allant de m/z 300 à m/z 12000. Les spectromètres de masse équipés d'un analyseur TOF ou Orbitrap sont les plus adaptés pour la mise en place d'une analyse globale, car ils disposent d'une large gamme dynamique. De manière générale, une haute résolution est nécessaire pour ce type d'approche, c'est pour cela que des méthodes d'analyses globales sont également réalisées sur des appareils équipés d'un détecteur Orbitrap, où la résolution et la sensibilité sont plus élevées[341]. Au laboratoire, une méthode de lipidomique globale par analyses Qq/TOF a été développée[180].

Ces études de lipidomique globales permettent d'obtenir une empreinte des lipides majoritaires d'un échantillon, qui sont en général comparées en aveugle à l'aide de méthodes statistiques poussées, permettant de mettre en avant les différences entre plusieurs échantillons. Les ions discriminants sont ensuite annotés, ce qui est très chronophage et compliqué à mettre en place. Cette approche est très développée en métabolomique, notamment *via* la mise en place de logiciels et d'outils informatiques permettant de faciliter le retraitement des données. L'application de ces workflows à la lipidomique reste très difficile, cette solution est peu utilisée à l'heure actuelle en lipidomique.

Il est en général préféré de confronter les différents profils lipidomiques obtenus à une banque de données (maison, ou disponible en ligne). Cette base de données est généralement théorique, car les standards primaires ne sont pas toujours disponibles. Les spectromètres de masses commercialisés aujourd'hui sont équipés d'outils permettant d'interroger des banques des données immédiatement à partir des données de sortie du spectromètre de masse, c'est le cas par exemple de UNIFI, le logiciel développé par Waters.

2. Les analyses semi-ciblées

Les analyses semi-ciblées en lipidomique, se focalisent sur certaines familles en utiisant une préparation d'échantillon et un mode de balayage spécifique sur les instruments de type triple quadripôle. Les analyses permettent de détecter des espèces lipidiques de certaines classes de molécules uniquement par spectrométrie de masse en tandem (MS/MS). Les analyses peuvent être effectuées selon deux modes de balayage : le mode MRM, ou le mode perte de neutre (NL) car certains ions fragments produits ou pertes de neutres sont spécifiques d'une classe de lipides. Par exemple, pour l'analyse des glycérophospholipides, le fragment majoritaire détecté correspond à la perte de la tête polaire (par exemple : sérine, choline, inositol). Il est ainsi possible de programmer des transitions MRM pour ces composés[3,342] et ainsi d'analyser spécifiquement ces familles

Cette méthode d'analyse, est généralement développée sur des spectromètres de masse de type triple quadripôle, puisque ce sont les seuls capables d'effectuer des balayages de pertes de neutres, et parfois sur des appareils de type Qq/TOF qui possèdent, tout comme les QqQ des modes de balayages MRM.

3. Les analyses ciblées

Les analyses ciblées en lipidomique, permettent de détecter spécifiquement certaines molécules peu abondantes et de les quantifier[343]. Ces études sont effectuées avec un *a priori* des espèces lipidiques attendues (puisque seule une classe de lipides est recherchée spécifiquement à l'aide de ces analyses). Dans un premier temps, les lipides d'intérêt sont extraits spécifiquement à l'aide d'une méthode d'extraction dédiée. Cette approche est basée sur des analyses MS/MS, généralement effectuées à l'aide de spectromètre de masse triple quadripôle, en utilisant le mode de balayage MRM. Lors de ces analyses, chaque ion précurseur et son ion fragment majoritaire sont recherchés spécifiquement, ces transitions sont spécifiques à chaque analyte. L'utilisation de ce mode de balayage permet de diminuer le bruit de fond et d'accroitre la sensibilité de détection des analytes. Ainsi, grâce à une analyse ciblée, il est possible de détecter spécifiquement certains métabolites, en atteignant des limites de détection basses de l'ordre de quelques picogrammes[179,191].

4. La quantification en lipidomique

La spectrométrie de masse n'est pas un bon outil pour la quantification, car chaque molécule analysée présente sa propre efficacité d'ionisation, qui peut être est très différente d'une espèce moléculaire à l'autre, même pour deux composés structuralement proches. Des gammes de calibration doivent donc être réalisées pour chaque composé que l'on souhaite quantifier. Ces gammes sont réalisées avec les standards primaires des analytes à analyser en utilisant des concentrations croissantes de métabolites, injectées en même temps, et, dans les mêmes conditions expérimentales. C'est la réalisation de cette gamme de calibration uniquement, qui permet d'effectuer une quantification absolue des analytes. L'ajout d'un standard interne est nécessaire afin de limiter les interférences dues au spectromètre de masse. En lipidomique, il est possible d'effectuer une quantification absolue, notamment pour les stérols, les acides gras, les acides biliaires[344], les oxylipines[179] ou les bases sphingoïdes.

Lorsque les standards primaires des molécules ne sont pas disponibles commercialement ou ne peuvent pas être synthétisés, il est possible d'effectuer une quantification relative. Celle-ci permet, de comparer l'abondance des lipides présents dans l'échantillon entre eux, en corrigeant leurs abondances respectives par celle du standard interne. C'est notamment le cas pour le dosage des lipides complexes ou particuliers comme les phospholipides et les sphingolipides.

5. Le niveau d'identification en lipidomique

En lipidomique, il existe 4 niveaux d'exigences permettant de classer les molécules identifiées et les molécules non identifiées. En réalité ce sont les règles de la métabolomique[345] qui sont appliquées à la lipidomique. Actuellement, des groupes de travail comme par exemple le Lipidomics Standards Initiative Consortium et la société internationale de lipidomique essaient de faire évoluer ces différents niveaux d'identification afin qu'ils soient spécifiques à la lipidomique[346,347].

Le premier niveau d'identification correspond aux composés totalement identifiés. Pour qu'un composé puisse être considéré comme identifié, il est nécessaire qu'il soit validé par *a minima* deux critères orthogonaux et indépendants d'identification et qu'il soit comparé à la molécule pure. Dans ce premier niveau d'identification, il est nécessaire de distinguer les isomères S et R pour un composé identifié. Pour les lipides il sera nécessaire d'avoir les informations sur l'isomérie *trans* et *cis*, ainsi que sur la position de la double liaison. Il faut donc par exemple, être capable de comparer son temps de rétention en LC ou GC ainsi que ses ions fragments obtenus en MS/MS à ceux d'un standard synthétisé, analysés tous deux dans les mêmes conditions expérimentales. En effet, la comparaison avec des données retrouvées dans la littérature ne permet pas de confirmer l'identification du composé[345]. Certains composés qui ont été identifiés pendant ma thèse sont identifiés avec ce premier niveau d'identification.

Le second niveau d'identification correspond aux composés putativement annotés. Il s'agit en général de composés pour lesquels certaines annotations sont très probables, mais pour laquelle le standard n'est pas disponible commercialement ou que sa synthèse est difficile. Il est possible d'annoter ces fragments à partir d'une banque de données, ou de mener l'étude « à la main » par les expérimentateurs. Une identification de niveau 2 correspond également à une identification à partir d'une banque de donnée externe au laboratoire. Par exemple, dans ce niveau d'identification se retrouvent les composés dont l'isomérie (R ou S) n'est pas connue, comme le 13(R)-HODE et le 13(S)-HODE[345]. Certains composés qui ont été identifiés pendant ma thèse sont identifiés avec ce second niveau d'identification.

Le troisième niveau d'identification correspond aux composés dont l'annotation est incomplète, ou lorsque plusieurs annotations sont proposées. Par exemple, en lipidomique c'est le cas pour les acides gras polyinsaturés lorsqu'ils sont analysés en LC-(ESI)-MS, ils ne se dissocient pas sous excitation collissionnelle en ions fragments produits. Nous obtenons alors un unique ion fragment correspondant à la chaine alkyle de l'AG. Avec cette méthode il est alors impossible de déterminer la position de la ou des doubles liaisons, ce qui rend

plusieurs annotations (structures) possibles pour un seul et même ion fragment (ou précurseur s'il ne s'agit que d'un AGI)[345]. Cependant, les informations dont nous disposons nous permettent de classer le composé dans une famille chimique. C'est le cas pour les composés identifiés dans ma thèse, possédant une insaturation sur la chaine carbonée : bien que l'annotation soit très probable et que la famille dont fait parti le composé est connu, les lipoaminoacides identifiés ont un niveau 3 d'identification.

Enfin, le dernier niveau d'identification correspond à un composé inconnu. Les informations à notre disposition ne nous permettent pas de conclure sur la nature, la famille ou la structure du composé.

Les règles proposées par la société internationale de Lipidomique sont différentes[347]. Elles permettent d'identifier les composés selon huit niveaux d'identification. Le niveau le plus bas d'identification est lorsque la composition élémentaire de la molécule est connue. D'autres niveaux sont plus spécifique à l'annotation en lipidomique, comme le niveau d'identification correspondant à l'annotation des phosphates (utilisée pour l'annotation des PIP_2, PIP_3 par exemple). Un autre niveau d'identification est quant à lui dédié à la position de la liaison *sn* (pour les glycérolipides et les TAG par exemple), un quatrième niveau d'identification est spécifique à la position de la double liaison sur la chaine d'acide gras (comme par exemple pour les AG bactériens). Un cinquième niveau d'identification est spécifique aux composés ayant une identification de l'ensemble des ions fragments d'intérêt, excepté pour la chaîne d'acide gras (par exemple pour un C12:1Aspartate, ou l'ensemble des ions fragments sont identifiés, à l'exception de la structure de la chaîne d'acide gras, ou la position de l'insaturation ainsi que la stéréochimie de la molécule reste inconnue)[347]. Un septième niveau permet de définir des composés dont la totalité des ions fragments sont identifiés, mais dont la stéréochimie de la molécule reste inconnue[347]. Enfin, le dernier niveau d'identification, comme en métabolomique correspond à l'annotation complète d'un composé avec son isomérie (par exemple 13R-HODE et 13S-HODE)[347].

Compte tenu de la structure des molécules analysées pendant ma thèse, nous avons utilisé l'identification conseillée par la métabolomique[345].

6. Les règles de chimie des réactions ions-molécules appliquées pendant ces travaux de thèse et nomenclature

Les lipoaminoacides se dissocient sous mode d'excitation non résonant et génèrent jusqu'à plus de 10 ions fragments. Identifier les pertes d'atomes entre l'ion précurseur et chaque ion fragment produit est essentiel pour avoir accès à sa composition élémentaire. Cependant, il est également nécessaire de connaitre la structure des ions fragments produits afin de permettre un niveau 1 d'identification. Elles sont obtenues *via* l'analyse des mécanismes de fragmentation qui sont inévitablement soumis à certaines règles de chimie organique, qui seront succintement décrites par la suite.

Règles de fragmentation utilisées pour décrire nos mécanismes : Afin de décrire les mécanismes proposés dans la suite du manuscrit concernant les différents composés identifiés, les règles utilisées se sont basées sur la position des charges. Ces règles sont incontournables pour l'interprétation de la fragmentation des ions. Plus précisément, elles se basent sur la position des différents sites déprotonés ou protonés. Ces règles sont imposées notamment lorsque la rupture d'une liaison se produit et affirment :

1) Qu'il n'est pas possible de maintenir une charge positive sur un atome de carbone en alpha d'un atome de carbone portant un hétéroatome. Si cela doit arriver, nécessairement un hydrure (H–) doit migrer sur cette position ou bien, une liaison doit se déplacer pour neutraliser cette charge. Par contre une charge négative peut être localisée sur ce site naturellement. En particulier si l'hétéroatome est doublement lié au carbone, il peut alors y avoir délocalisation de la charge qui se trouve dispersée et donc amenant a la stabilisation de l'anion.

2) Il n'est pas possible de maintenir une charge négative sur un atome de carbone portant un hétéroatome. Si cela arrive comme intermédiaire immédiatement un proton migre sur ce site ou une liaison covalente bascule pour neutraliser la charge. Cependant, une charge positive peut être portée par cet atome de carbone, alors elle sera stabilisée par délocalisation.

De plus, les agrégats chargés qui émergent des gouttelettes et qui survivent jusque dans la zone de désolvatation peuvent conduire à des anions où la charge peut être localisée sur différents sites en compétition. Bien qu'elles soient de mêmes compositions élémentaires, plusieurs structures (déprotomères[336]) peuvent être formées. Cela s'explique notamment par

la présence de différents sites (basiques ou acides), qui peuvent être en interaction avec le cluster chargé. Lors de la désolvatation la charge reste attachée sur le site basique ou acide initial. A partir de ce site, la charge peut migrer par prototropie afin de former d'autres protomères (ou déprotomères)[336,229,348]. Dans tous les cas, il y aura une distribution de positions simplement chargées : c'est-à-dire qu'il sera possible d'avoir plusieurs structures monochargées différentes pour la molécule déprotonée ou pour la molécule protonée. C'est à partir de ces positions que les fragmentations sont induites, et c'est pour cette raison que parfois, plusieurs mécanismes peuvent être proposé. Dans la mesure du possible, nous avons décidé de n'en proposer qu'un seul, correspondant à celui étant le plus probable.

Jamais des fragmentations non induites par la charge ne seront considérées car elles ne peuvent être produites qu'à partir de sites radicalaires. Elles sont possibles lors de collisions de très haute énergie ou lors d'un bombardement par électrons c'est-à-dire par electron induced dissociation (EID)[349], ou encore par interaction avec des atomes métastables[350], mais ces collisions n'ont jamais lieu dans nos conditions expérimentales. Certains des ions fragments produits par les lipoaminoacides font appels à des processus par étapes, qui peuvent s'interpréter par la formation d'ions dipôles.

Les ions dipôles : Il s'agit d'espèces décrites pour la première fois à la fin des années 1960 lorsque l'ionisation chimique a été introduite. Ce sont des intermédiaires, résultant de réactions ions-molécules, à partir desquels les transferts de protons peuvent se faire. Vers la fin des années 1970 le concept d'ions dipôles revient, lors de l'interprétation de certaines fragmentations. Pierre Longevialle travaillait sur des diamino-3,17-stéroïdes[351, 352]. Afin de permettre l'interprétation des spectres d'ionisation chimique de ces composés, il introduit alors la notion d'ion dipôles. Les ions fragments impliquent la perte de l'amine du cycle A et ensuite, celle de l'amine du cycle D bien que ces groupes soient très éloignés et que la conformation du squelette androstane est complètement figée dans un plan moyen empêchant tout transfert de proton entre les cycles A et D. A la même période le groupe de D. H. Williams[353] explore les ions dipôles dans le cas de plus petites molécules, afin d'expliquer certaines fragmentations inattendues. Par la suite, de tels intermédiaires ont été décrits à partir d'ions négatifs à nombre pair d'électrons (molécule déprotonée) formés à partir d'esters à longue chaine d'estradiol *via* une ionisation négative[352,354]. Ces résultats ont permis de mettre en évidence que le site de déprotonation se trouvait dans le cycle A, et que la fragmentation avait lieu dans le cycle D. Là encore, afin d'expliquer ce phénomène il a fallu introduire cette notion d'ions dipôles.

Ce concept est un modèle qui permet d'interpréter certaines réactions, et en général, des réactions où les charges éloignées du centre de fragmentation semblent être

spectatrices : ce qui est faux lors d'activation de basse énergie. En réalité cela signifie que l'espèce moléculaire en question s'isomérise en ion dipôle où des échanges de proton (H^+), d'hydrure (H^-), ou dans certains cas concernant les ions radicaux, des charges H•, peuvent avoir lieu entre les partenaires avant qu'ils se séparent pour produire les ions fragments compétitifs.

Lorsque les fragmentations impliquent la présence des ions dipôles dans les cas les plus simples, il est possible d'observer deux ions fragments dont la somme en m/z correspond à la masse de l'ion précurseur à plus ou moins une unité m/z. Lorsque la variation est de -1 m/z, un échange de proton a eu lieu entre les partenaires. Lorsque la variation est de +1 unité de masse, un échange d'hydrure a eu lieu entre les partenaires. Lorsque la somme correspond à la masse de l'ion moléculaire, c'est un électron qui a migré.

Les effets d'énergie sur l'orientation de leurs apparitions peuvent être une aide à leurs mises en évidence. En effet, lorsque l'ion dipôle est réalisé, la dissociation conduit à deux processus compétitifs dont l'un correspond à la rupture directe de l'ion dipôle, l'autre correspondant à un réarrangement *via* un transfert de proton, d'hydrure ou d'hydrogène radicalaire. Selon la théorie du quasi-équilibre (Cf. Annexe 1), le premier processus est favorisé à partir des ions de plus haute énergie interne alors que le second est favorisé à partir de ceux de plus faible énergie interne. Pour que ces intermédiaires puissent se former, ou conduire aux échanges de H^+, H^- il est nécessaire que les partenaires sous leurs formes neutres aient des basicités en phase gazeuse ou des acidités relativement proches les unes des autres. Dans ces conditions, leur durée de vie doit être suffisante afin de permettre de tels échanges.

Mise en place d'une nomenclature pour les LpAA : Afin de simplifier la lecture et d'harmoniser l'annotation des ions fragments produits par activation collisionnelles pour les lipoaminoacides, l'introduction d'une nomenclature a été nécessaire. Celle-ci se base sur celle utilisée pour l'annotation des ions fragments de peptides protonés ou déprotonés. Cette nomenclature avait été définie il y a une quarantaine d'années[355] pour les ions positifs pour des collisions de haute énergie (au-delà du keV) où la fragmentation des chaines latérales des résidus (acides aminés) étaient rompues.[356] Plus récemment, cette nomenclature avait été modifiée pour être cohérente avec la fragmentation des peptides déprotonés (pour les ions négatifs, Figure 89)[357]. Alors, ici nous avons proposé une nomenclature qui s'applique à nos conjugués tout comme pour les lipopeptides. Par exemple, la rupture de la liaison amide, lien entre les deux partenaires des LpAA, va conduire à deux ions compétitifs et

complémentaires : y^- (partie correspondant à l'acide aminé déprotoné) et $[Lp\text{b-2H}]^-$ (à l'ion fragment correspondant à la partie acide gras).

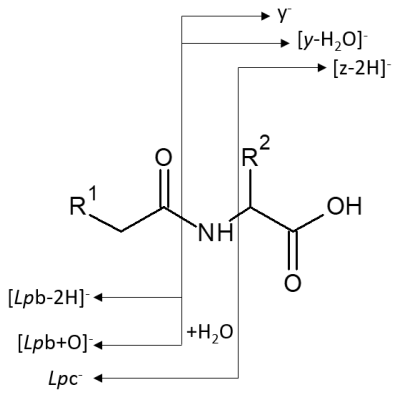


Figure 89 : Nomenclature mise en place spécifiquement pour les lipoaminoacides

PRESENTATION DES TRAVAUX DE THESE : résultats et discussion

Objectifs et travaux de thèse

Pour mieux comprendre les relations hôte-microbiote, la caractérisation des métabolites bactériens est essentielle pour déterminer les interactions fonctionelles. L'objectif principal de mes travaux de thèse, a été d'étendre l'identification, la caractérisation structurale et la quantification de nouveaux lipopeptides produits par les bactéries du microbiote intestinal. Une stratégie expérimentale de chimie analytique par LC-HRMS permettant l'identification et la caractérisation de ces lipopeptides a été mise en place. Pour cela, un workflow analytique spécifique aux lipopeptides a été développé. Il comprend plusieurs étapes complémentaires, dont une extraction sur phase solide et une analyse par spectrométrie de masse à l'aide d'un appareil Qq/TOF.

Les composés ont été analysés par des expériences d'activation collisionnelle des ions d'intérêt en modes MS^2 (avec un Qq/TOF) et MS^n (avec LTQ-Orbitrap VELOS). Les spectres d'ions produits de l'ion précurseur sélectionné soumis au processus CID (collision-induced dissociation)[214] ont été comparés avec ceux des standards purs, synthétisés spécifiquement, permettant d'identifier la majeure partie des lipopeptides (et lipoaminoacides) avec un niveau 1 d'identification selon les règles de la métabolomique[358].

Les mécanismes de fragmentation de chacune des familles (selon l'acide aminé présent quel que soit la taille et le niveau d'insaturation de la partie acide gras) des composés identifiés a été établi, afin de mettre en évidence les processus diagnostiques, indépendants de la partie lipidique, propres à chacune de ces familles. Pour une meilleure classification des fragmentations et leur généralisation, une nouvelle nomenclature spécifique aux lipopeptides a été proposée simplifiant ainsi l'annotation des spectres CID. Après leur identification et leur caractérisation, ces composés ont été dosés dans plusieurs matrices (bactéries, côlons de souris, fèces de souris, fèces humaines). Pour cela une méthode de quantification absolue a été mise en place basée sur l'analyse en basse résolution en utilisant un tandem LC-QqQ dont la robustesse a été déterminée.

Ce travail d'identification, partie la plus importante de notre étude, nous a amené à approfondir les dissociations majeures observées. Certaines fragmentations caractéristiques ont été détectées et interprétées. C'est le cas notamment de la régénération inattendue d'ions fragments permettant des attributions précises concernant la partie aminoacide. Pour les étudier en détails, les profils d'évolution des ions fragments, notamment ceux inattendus mais caractéristiques, en fonction de l'énergie de collision (energy-resolved mass spectrometry, ERMS) ont été réalisés. Le profil caractérisant les ions produits est Gaussien. Nous avons

observé pour la première fois que la largeur à mi-hauteur de ces Gaussiennes augmente d'autant plus que le rapport m/z de l'ion fragment est faible, le processus physique résultant des collisions est en cours d'étude et de modélisation. Pour certains ions fragments nous avons observés deux types de profils inattendus : (1) Gaussiens larges plus ou moins déformés, (2) composites (superposition non centrée de plusieurs formes probablement Gaussiennes). L'étude de ces profils particuliers inattendus montrent une certaine corrélation avec les processus impliqués dans la formation de ces ions produits. Une analyse plus poussée par mobilité ionique cyclique (cIMS) a été réalisée.

L'ensemble des résultats obtenus au cours de ces différentes études ont fait l'objet de trois articles dont deux sont soumis et un troisième qui devrait être soumis avant la soutenance de ma thèse.

Résultats

Dans cette partie, les résultats obtenus et en cours de valorisation sous forme de publications sont résumés et décrits les uns après les autres. Les publications associées sont insérées. Une discussion générale qui reprend l'ensemble des résultats est proposée en fin de chapitre.

Publication 1 : Mise en place d'un workflow analytique, spécifique à la recherche de nouveaux lipopeptides bactériens par chromatographie liquide couplée à la spectrométrie de masse à haute résolution (Lc-QqTOF)

Résumé :

L'amélioration des connaissances sur les métabolites produits par le microbiote est un point clé pour comprendre son rôle en santé humaine. Parmi eux, les lipoaminoacides (LpAA) contenant de l'asparagine et leurs dérivés sont des métabolites bactériens qui pourraient avoir un impact sur l'hôte. Dans cette étude, notre objectif était d'étendre la caractérisation de cette famille. Nous avons développé un « workflow » semi-ciblé afin d'identifier et de quantifier les nouveaux candidats. Tout d'abord, la préparation d'échantillon (étape d'extraction) et l'analyse des lipopeptides par chromatographie liquide (LC) couplée à spectrométrie de masse à haute résolution (HRMS) ont été optimisé. L'utilisation d'une banque de données théoriques « maison » a permis d'interroger les données brutes manuellement. Cette stratégie nous a permis de trouver vingt-trois nouveaux LpAA conjugués aux acides aminés Asn, Gln, Asp, Glu, His, Leu, Ile, Lys, Phe, Trp et Val. Ces métabolites ont ensuite été entièrement caractérisés par MS^2 et MS^n, et comparés aux standards synthétisés pour valider les annotations. Enfin, une méthode d'analyse quantitative a été développée par LC couplée à un instrument triple quadripôle, où la linéarité et limites de quantification et de détection ont été déterminées. Quatorze nouveaux composés ont ainsi été quantifiés dans des bactéries à Gram positif, *Lactobacilus animalis*, et douze dans *Escherichia coli* Nissle 1917.

Résultats détaillés :

Développement d'une méthode d'extraction et d'analyse par chromatographie couplée à la spectrométrie de masse à haute résolution spécifique aux lipopeptides :

Des travaux antérieurs à ceux de ma thèse ont permis l'identification et la caractérisation de deux lipoaminoacides (C14Asn et C12Asn), et de plusieurs lipopeptides (C12AsnGABA, C12AsnAla, C12AsnVal). Ces composés ont été extraits à partir de culots bactériens *via* une extraction sur phase solide en utilisant une plaque SPE polymérique HLB où les métabolites d'intérêts sont élués avec du méthanol. Dans le but d'étendre l'identification de nouveaux lipopeptides, cette méthode a été modifiée en utilisant un mélange de standards variés (lipopeptides, lipoaminoacides, acides gras hydroxylés, acyl-sérotonine, composés nitrés, anandamides) comme échantillon test. Plusieurs conditions de broyage et d'extraction décrites dans la littérature afin d'analyser les lipopeptides ont été comparés. Finalement, nous avons conservé les conditions de broyage utilisant un mélange tampon Tris (50 mM à pH 9) dans du méthanol selon des proportions 1:2.

Plusieurs protocoles d'extraction liquide-liquide (LLE) ont été testés, notamment le protocole de Bligh & Dyer[183] classique, modifié et adapté avec du méthanol acidifié (HCl ou CH_3COOH). Les rendements d'extraction de ces 3 LLE avoisinaient les 35%. Une LLE en utilisant de l'eau et de l'acétate d'éthyle a également été testée, bien que les résultats soient encourageants la méthode manquait de robustesse. L'utilisation d'une plaque SPE échangeuse d'anions (utilisée au laboratoire pour l'extraction des isoprostanes) n'a pas permis d'extraire nos composés d'intérêts. En revanche, les tests réalisés sur des plaques SPE polymériques HLB, après plusieurs essais de solvants éluant, ont permis d'atteindre un rendement d'extraction de 20 à 68% selon les métabolites. L'élution successive avec de l'acétonitrile, du méthanol et de l'acétate d'éthyle a conduit aux meilleurs rendements d'extraction avec une robustesse supérieure aux autres méthodes d'extractions testés. Nous avons donc décidé d'utiliser cette méthode d'extraction pour la suite de notre étude.

Afin de garantir une sensibilité maximale pour la détection des composés d'intérêt, une méthode analytique a été développée et optimisée avec un spectromètre de masse hybride Qq/TOF en utilisant le même mélange de standards. L'optimisation des paramètres de la source d'électro-nébulisation (où la polarité négative des ions a été choisie) du spectromètre de masse n'a eu que très peu d'effets au niveau du gain en sensibilité. Afin d'augmenter la sensibilité, la méthode analytique a été miniaturisée, *via* l'utilisation d'une colonne à plus faible diamètre interne : 1mm de diamètre interne au lieu de 2.1mm, permettant un gain de

sensibilité de l'ordre de 80%. Les analyses ont été effectuées selon plusieurs modes de balayages MS et MS^E soit en continuum soit en centroïde.

Standard interne :

Afin de suivre l'ensemble des étapes du workflow analytique, il est important d'utiliser un standard interne. Afin qu'il soit comparable aux lipopeptides, nous avions commandé à nos collaborateurs chimistes un C12AsnGABA où les hydrogènes de la chaine d'acide gras étaient totalement remplacés par des deutérium (d23). Une instabilité du composé se traduisant par la présence d'un massif isotopique caractéristique des pertes successives d'un, de deux, de trois et parfois même de quatre atomes de deutérium a été observée. Des variations étaient notables entre standard extrait et non extrait. Nous avons supposé qu'un échange entre les atomes d'hydrogène du méthanol et les deutériums du C12AsnGABAd23 était possible lors de l'extraction, nous empêchant alors de quantifier nos rendements d'extraction et de préparation.

Par la suite, nous avons alors synthétisé un second standard interne : le C16AsnGABA substitué par des atomes de ^{13}C sur la totalité de la chaine carbonée de l'acide gras, générant ainsi une espèce beaucoup plus stable. Par contre son rendement d'extraction était de 20%. Comme celui des autres standards (C14Asn, 12AsnGABA, C12AsnAla et C12AsnVal) était autour de 80%, nous avons supposé que la différence de longueur de chaine entre un acide laurique et un acide palmitique induisait suffisamment de différences de polarité pour modifier fortement l'élution sur la plaque SPE HLB. Ainsi, nous nous sommes aperçus que nous avions développé un workflow analytique davantage optimisé pour les lipopeptides de longueurs de chaine d'acide gras plus courtes que l'acide palmitique.

Mise en place d'une banque de données théoriques des lipopeptides :

Une base de données théorique comprenant 2300 lipopeptides potentiels a été constituée sur la base du premier lipopeptide identifié : le C12AsnGABA. Celle-ci a été imaginée en déclinant les différentes parties qui compose le lipopeptide : l'AG (allant de 8 à 20 atomes de carbone, mono insaturés, saturés, hydroxylés ou constitués d'un cyclopropane), l'acide aminé, et le neurotransmetteur (GABA, histamine, sérotonine et dopamine). L'ensemble de ces combinaisons a été envisagée, afin de produire une banque de données à interroger pour identifier et caractériser de nouveaux lipopeptides et LpAA. Cette banque de données maison composée de 2436 composés, contient la structure de la molécule, son code SMILES et les masses correspondantes aux molécules protonées et déprotonées (puisque seules ces deux formes ont été retrouvées pour les lipopeptides).

Acquisition des données en LC-HRMS² et retraitement :

Le retraitement des données a été effectué *via* le logiciel UNIFI, qui permet d'associer des ions précurseurs détectés (match entre un m/z de la données brute et la banque de données), à leurs ions fragments produits et d'en proposer une structure. Il est possible d'entrer une liste d'ions « fragments attendus », et une liste de pertes de neutre attendues qui seront automatiquement recherchées par le logiciel. De manière générale, le logiciel UNIFI présente un grand nombre de dysfonctionnement pour notre problématique, générant un nombre de faux positifs importants (allant parfois jusque 50.000 matchs par échantillon), à trier manuellement, ce qui s'avère être trop chronophage. De plus, les options d'ions fragments attendus et de perte de neutres, qui aurait pu nous être très utile, ne semblent pas fonctionner lorsque l'on recherche des composés minoritaires. A l'inverse le logiciel semble fonctionner dans les conditions où les composés sont concentrés, c'est-à-dire lorsqu'ils sont majoritaires.

Nous avons finalement utilisé le logiciel QuanLynx, qui fonctionne de la même manière : lorsqu'un m/z de la base de données est détecté, celui-ci est identifié. Cependant, lorsqu'un m/z est détecté, le logiciel n'associe pas les ions fragments correspondants. De plus, Quanlynx n'est pas capable de vérifier automatiquement si le rapport m/z identifié correspond à celui d'un ion monoisotopique ou à un isotopologue, cette vérification est réalisée manuellement, ce qui rend le retraitement contraignant et chronophage.

Identification d'un premier nouvel aminolipide : le C13Asparagine

Dans l'article, une grande partie est dédiée à la détection et à l'élucidation de la structure d'un nouvel LpAA, le C13Asn (Figure 90). Dans les conditions expérimentales initiales, c'est-à-dire avec un gradient chromatographique s'étalant sur 18 min, le précurseur $[(C13Asn\text{-}H)]^-$ était largement co-élué avec des composés isobares. Les analyses MS^2 dans ces conditions ne permettaient de distinguer quels ions fragments étaient issus du composé d'intérêt, de ceux des ions isobares. La modification des conditions expérimentales afin d'améliorer la séparation chromatographique du composé a été nécessaire. Le gradient chromatographique a été allongé et est passé de 18 min à 40 min. Cette modification a permis de séparer les molécules isobares, et ainsi d'identifier sans ambiguïté les différents ions produits issus de l'ion précurseur et non plus ceux produits par les interférents isobares issus d'une co-élution. En effet, les échantillons biologiques sont complexes, c'est pour cela qu'il est courant de retrouver au même temps de rétention des composés isobares à celui d'intérêt. Ainsi, les ions fragments caractéristiques du C13Asn sont décrits, et comparés avec ceux générés à partir d'un standard pur synthétisé spécifiquement. C'est de cette manière que nous avons pu confirmer l'identification de niveau 1 selon les règles de métabolomique[345].

L'identification et la caractérisation de ce premier composé nous a servi de base pour l'identification et la caractérisation d'autres composés, qui ont tous été identifiés par LC-HRMS en comparaison avec un standard primaire synthétisé spécifiquement.

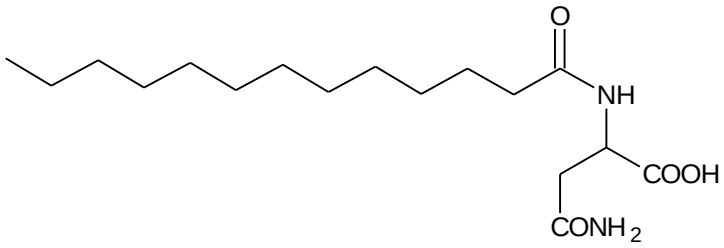


Figure 90 : Structure du C13Asn

Concernant les mécanismes de fragmentation des anions $[C13Asn\text{-}H]^-$ dérivant de C13Asn, ils sont décrits dans la publication et suivent tous les règles de fragmentation décrites préalablement. Il s'agit des processus de pertes directes d'H_2O (perte majeure, m/z 309), et de 3 pertes très faibles (moins de 3%) correspondant aux pertes de CO_2 (m/z 283), et des pertes consécutives de (H_2O+CO_2) (m/z 265), et de (H_2O+CO_2+HCN) (m/z 238).

La perte mineure de CO_2 est directe mais selon un mécanisme par réarrangement basé sur un transfert de proton de l'amide vers l'atome de carbone d'où CO_2 est éliminé. La présence d'une charge négative résultante sur un atome de carbone, d'où CO_2 est parti, portant un hétéroatome est pratiquement impossible compte tenu des règles de la chimie organique (Schema 1a). Par contre, la perte majeure d'eau (Schéma 1b) va se produire à partir de $[C13Asn\text{-}H]^-$ où l'amide est déprotonée et où la charge se situe sur l'atome d'oxygène. Par une attaque nucléophile sur l'acide carboxylique libre pour générer par cyclisation un intermédiaire tétrahédrique qui pour se dissocier, évolue selon un processus par étape *via* une isomérisation en ion dipôle. Au sein de celui-ci, l'ion OH^- libéré arrache un proton, après exploration de la surface de son partenaire neutre, à un site acide en phase gazeuse. Bien que différents sites portent des protons mobilisables ici, c'est le groupe imine qui sera favorablement déprotoné. Il en résulte une ouverture du cycle à 5-chainons régénérant un groupe carboxylate avec un groupement nitrile (ion m/z 309). Ce dernier peut perdre consécutivement le CO_2. Bien qu'*a priori* triviale, cette perte serait très défavorisée si un proton mobile de la position énolisable (position acide) de l'amide ne venait pas neutraliser la charge. Cette assistance conduit à la formation d'un énolate (ion m/z 265) qui ensuite, peut s'isomériser en ion dipôle en se cyclisant selon une substitution nucléophile intramoléculaire (formation d'alkyl-2-oxazoline) libérant ainsi l'anion CN^-, qui restant au sein l'ion dipôle. CN^- arrache alors un proton au groupe NH de l'hétérocycle pour être éliminé sous forme de HCN et forme un alkylidène-2-oxazoline déprotoné (ion m/z 238) (Schéma 1c). D'autres ions

abondants sont présents, dont le second plus abondant est l'amide de l'acide gras déprotoné (m/z 212). Les autres ions fragments abondants correspondent aux ions caractéristiques de l'acide aminé lui-même (m/z 131, 114).

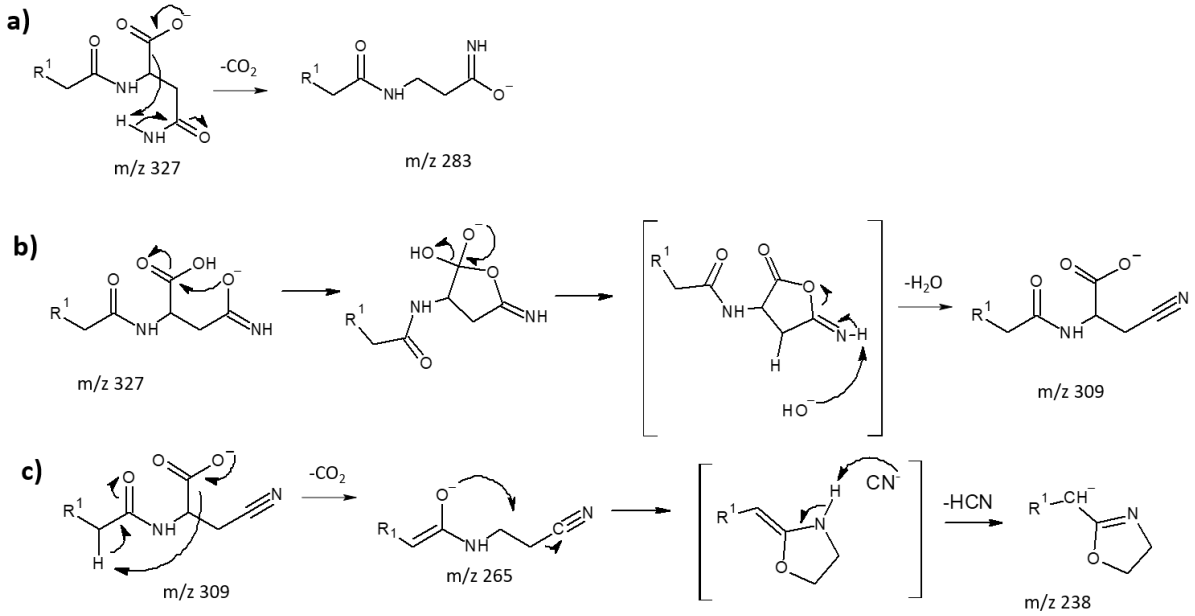


Schéma 1 : Mécanismes possibles de fragmentation par pertes directes compétitives (a) très mineure de CO_2 et (b) majeure d'H_2O qui est suivie de pertes consécutives (c) mineures de (H_2O+CO_2) et (H_2O+CO_2+HCN)

Extension à d'autres lipoaminoacides par LR-MS

Une fois le premier composé caractérisé, et son ion fragment le plus intense identifié, les transitions MRM théoriques d'autres LpAsn ont été prédites, permettant d'identifier en LC-QqQ les composés putatifs (puisque la résolution est faible et ne conduit qu'aux masses nominales) : C12:1Asn, C14:1Asn, C15Asn, C16Asn, C16:1Asn, C17Asn, et C17:1Asn. Il a alors été décidé d'étendre cette méthode, en déclinant les MRM théoriques pour les autres LpAA, pour tous les acides aminés et longueurs de chaînes potentielles confondus. Cette méthode a permis d'identifier en basse résolution 48 LpAA putatifs: C12:1Leu/Ile, C13Leu/Ile, C14Leu/Ile, C14:1Leu/Ile, C15Leu/Ile, C15:1Leu/Ile, C16Leu/Ile, C16:1Leu/Ile, C17Leu/Ile, C17:1Leu/Ile, C18Leu/Ile, C18:1Leu/Ile, C12Gln, C12:1Gln, C13Val, C14Val, C14:1Val, C15Val, C16Val, C16:1Val, C17:1Val, C18Val, C12Glu, C12:1Glu, C16Glu, C15:1Glu, C16Glu, C17:1Glu, C12Lys, C12:1Lys, C14Lys, C15Lys, C16Lys, C16:1Lys, C17Lys, C17:1Lys, C12:1Phe, C13Phe, C14Phe, C14:1Phe, C15Phe, C15:1Phe, C16Phe, C16:1Phe, C17Phe, C17:1Phe, C18Phe et C18:1Phe.

Afin de valider les structures de ces composés, il est nécessaire de réaliser des analyses en LC-HRMS², et de comparer les résultats à ceux des standards purs qui ont été

synthétisés à cet effet. Vu le grand nombre de composés putatifs identifiés, nous avons décidé de nous focaliser uniquement sur la famille entière des LpAsn, puis sur un composé de chaque famille (une seule longueur de chaine d'AG pour un amino-acide). La synthèse de leur standard spécifique, nous a permis d'atteindre le niveau 1 d'indentification pour les 22 composés suivants : C12:1Asn, C13Asn, C14:1Asn, C15Asn, C16Asn, C16:1Asn, C17Asn, C17:1Asn, C12Arg, C12Asp, C12Gln, C12Glu, C14Glu, C16Glu, C12His, C12Leu, C12Ile, C12Lys, C14Lys, C14Phe, C12Trp, C12Val (Figure 91). Le niveau d'identification des autres composés sera discuté dans la dernière partie de cette thèse.

Mise en place d'une méthode de quantification de 22 LpAA :

Après l'identification et la caractérisation de ces vingt-deux composés, une méthode de quantification absolue par LC-QqQ a été développée sur le Shimadzu TQ-8060. Les limites de quantification et de détection ont été calculées pour chacun des composés, et atteignent l'ordre du pg/µL. Une gamme de standards primaires composée de 10 points, allant de 0.004 à 1250 pg/µL est injectée en même temps que les échantillons afin de garantir les mêmes conditions d'analyses et de permettre une quantification absolue où la précision a été déterminée comme allant de 85 à 115%. Les coefficients de variations intra et inter jour ont été calculés et sont inférieurs à 5%, permettant ainsi de garantir la robustesse de la méthode d'analyse.

Quantification des nouveaux LpAA dans deux souches bactériennes :

Cette méthode de quantification a été utilisée afin de doser les LpAA dans des souches bactériennes à gram négative (*EcN*), et à gram positive (*Lactobacillus animalis*). Dans la souche EcN WT, les lipoaminoacides suivants C12Asn, C13Asn, C14Asn, C14:1Asn, C15Asn, C16Asn, C16:1Asn, C17:1Asn, C12Glu, et le C12Asp sont dosés. Dans la souche *Lactobacillus animalis*, il a été quantifié les lipoaminoacides C12Lys, C14Lys, C12His, C12Gln, C12Asp, C14:1Asn, C16:1Asn, C15Asn, C17Asn, C12Trp, C12Val, C12Leu, C16Asn, et C14Phe.

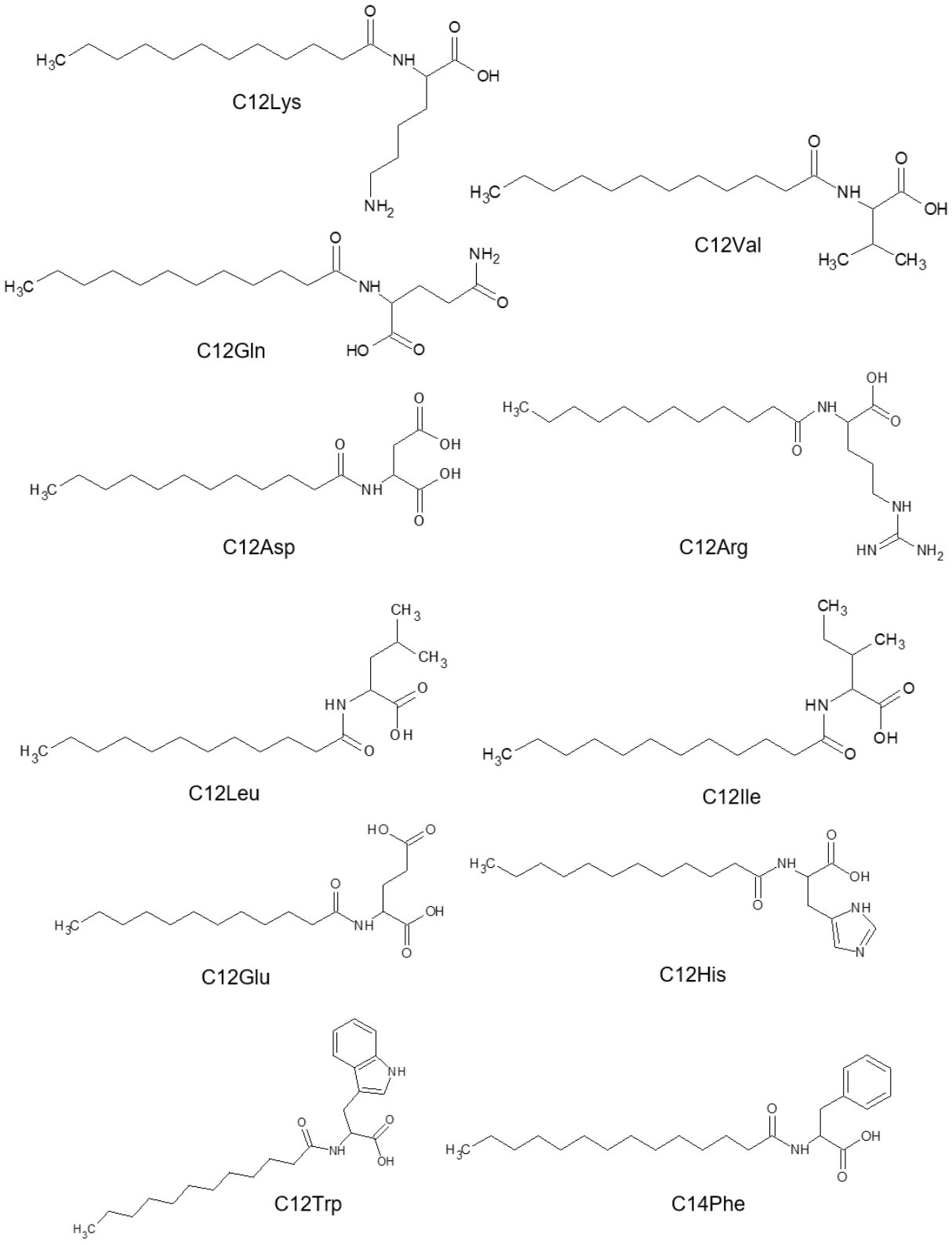


Figure 91 : Structures de certains LpAA identifiés à l'aide du Workflow lipidomique développé

EcN contient un ilot de gènes appelé pks codant pour des enzymes et des protéines accessoires permettant la synthèse de LpAsn et de la colibactine, une génotoxine. Le premier gène de ce cluster est ClbA qui code pour une PPtase (Phosphopantethéinyl transférase) qui permet l'activation des enzymes de synthèse. ClbN code pour une enzyme responsable de l'ajout de l'asparagine sur l'acide tétradécanoïque (acide myristique) menant à la formation du C14Asn. Ce composé entre dans la composition de la précolibactine. Le clivage du C14Asn par une peptidase codée par ClbP permet de libérer du C14Asn et la forme active et mature de la colibactine. Afin de déterminer l'implication des enzymes codées par ClbA, ClbN et ClbP dans la synthèse des aminoacides, le dosage des LpAA a été effectué dans des souches bactériennes sauvages ou délétées pour ces gènes. Les résultats ont été représentés sous forme d'une heat-map où 4 clusters se dégagent.

Le premier met en évidence les composés dont les concentrations sont diminuées chez les bactéries mutées versus sauvages et absentes de *Lactobacillus animalis* (C14Asn, C12Asn, C14:1Asn, C16Asn, C16:1Asn, C17:1Asn et C12Glu). Ces composés sont donc dépendants des enzymes codées par ClbA, ClbN et ClbP pour leurs synthèses. Le second cluster met en évidence les LpAA qui nécessitent les enzymes codées par ClbN et ClbP mais pas ClbA pour leurs synthèses, ces composés sont également absent des *Lactobacillus animalis* (C15Asn et C13Asn). Le troisième cluster permet de déterminer les LpAA retrouvés dans les EcN sauvages et délétées pour ClbA, mais absents des mutants ClbN, ClbP et des *Lactobacillus animalis* (C17Asn, C12Lys et C14Lys). Enfin, le dernier cluster est composé des LpAA retrouvés uniquement chez *Lactobacillus animalis* (C12His, C14Phe, C12Gln, C12Trp, C12Leu/Ile, C12Val et C12Asp).

Ces données permettent de mettre en évidence que la synthèse des LpAA n'est pas restreinte aux bactéries gram négatives, puisque *Lactobacillus animalis* en est également capable. De plus, ces résultats montrent que la délétion d'un gène modifie la concentration ainsi que la nature des LpAA produits par une bactérie.

Conclusion :

Un workflow analytique spécifique aux lipopeptides comportant une extraction optimisée ainsi qu'une analyse spécifique par LC-HRMS, incluant une base de données comportant 2436 composés potentiels a été développé. Ce workflow a permis de mettre en évidence la présence de quarante-neuf nouveaux lipopeptides putatifs dans certaines souches bactériennes. La caractérisation structurale de vingt-deux LpAA a été réalisée à l'aide de standards synthétisés, pour lesquels la fragmentation des composés a été étudiée en détails. Les LpAA régénèrent l'ion produit majoritaire correspondant à la perte de H_2O. Ensuite, une méthode de quantification absolue a été mise en place par LC-QqQ (Figure 92). L'analyse de bactéries ayant des mutations sur certaines protéines clés a permis de mettre en évidence leurs rôles dans la biosynthèse de ces LpAA. A ce jour, cette méthode analytique représente une opportunité d'étudier la production des LpAA, produits par le microbiote.

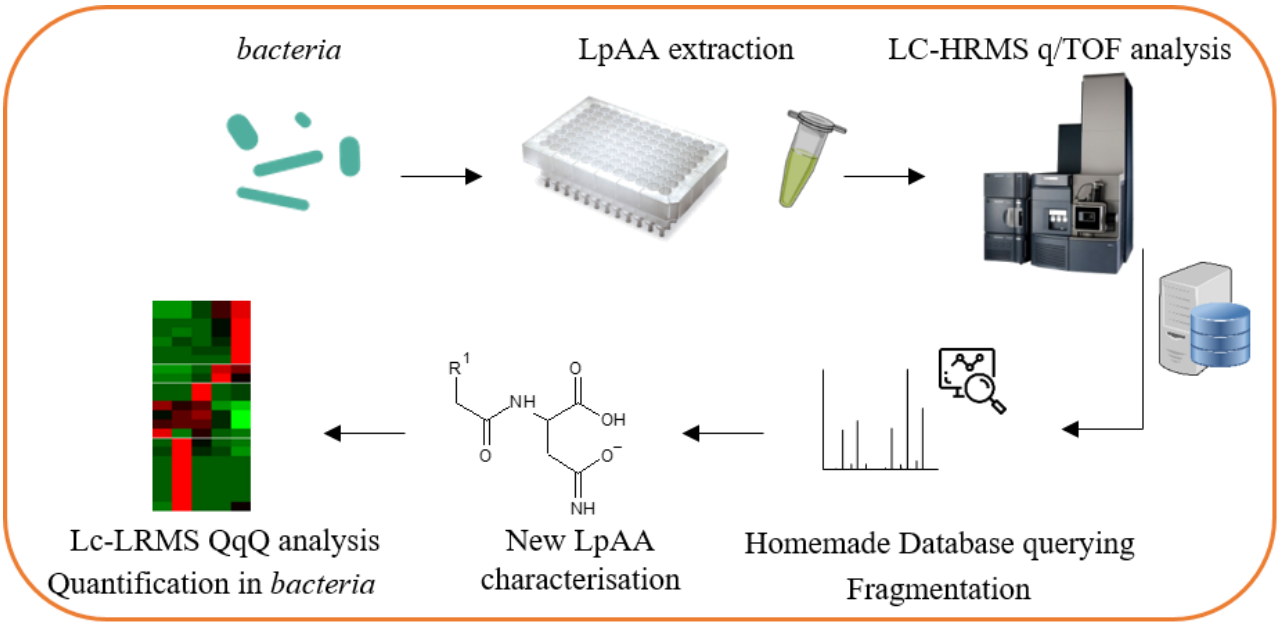


Figure 92 : Schéma récapitulatif du workflow analytique développé pour la recherche de nouveaux lipides bactériens, produits par le microbiote intestinal

Analytica Chimica Acta xxx (xxxx) xxx

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Discovery and quantification of lipoamino acids in bacteria

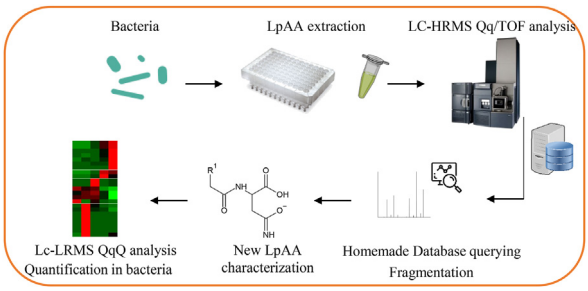
Amandine Hueber [a, b, c], Camille Petitfils [c], Pauline Le Faouder [a, b], Geoffrey Langevin [d], Alexandre Guy [d], Jean-Marie Galano [d], Thierry Durand [d], Jean-François Martin [a, e], Jean-Claude Tabet [a, f, g], Nicolas Cenac [c, 1], Justine Bertrand-Michel [a, b, 1, *]

[a] MetaboHUB-MetaToul, National Infrastructure of Metabolomics and Fluxomics, Toulouse, France
[b] I2MC, Université de Toulouse, Inserm, Université Toulouse 3 Paul Sabatier, Toulouse, France
[c] IRSD, Université de Toulouse, INSERM, INRA, INPENVT, Université de Toulouse 3 Paul Sabatier, Toulouse, France
[d] Institut des Biomolécules Max Mousseron IBMM, UMR 5247 CNRS, Université de Montpellier-ENSCM, Montpellier, France
[e] Toxalim (Research Centre in Food Toxicology), INRAE UMR 1331, ENVT, INP-Purpan, Paul Sabatier University (UPS), Toulouse, France
[f] Université Paris-Saclay, CEA, INRAE, Département Médicaments et Technologies pour la Santé (DMTS), MetaboHUB, F-91191, Gif sur Yvette, France
[g] Sorbonne Université, Faculté des Sciences et de l'Ingénierie, Institut Parisien de Chimie Moléculaire (IPCM), F-75005, Paris, France

HIGHLIGHTS

- Lipopeptides and Lipoaminoacids are a group of bacterial metabolites which could impact the host physiology.
- A lipidomic workflow by SPE, LC-HRMS has been developed to expand their identification.
- 22 new LpAA has been discovered and quantified in bacteria.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 August 2021
Received in revised form
4 November 2021
Accepted 21 November 2021
Available online xxx

Keywords:
Lipoaminoacids
Mass spectrometry
Suspect screening
Lipidomic
Microbiota

ABSTRACT

Improving knowledge about metabolites produced by the microbiota is a key point to understand its role in human health and disease. Among them, lipoamino acid (LpAA) containing asparagine and their derivatives are bacterial metabolites which could have an impact on the host. In this study, our aim was to extend the characterization of this family. We developed a semi-targeted workflow to identify and quantify new candidates. First, the sample preparation and analytical conditions using liquid chromatography (LC) coupled to high resolution mass spectrometry (HRMS) were optimized. Using a theoretical homemade database, HRMS raw data were manually queried. This strategy allowed us to find 25 new LpAA conjugated to Asn, Gln, Asp, Glu, His, Leu, Ile, Lys, Phe, Trp and Val amino acids. These metabolites were then fully characterized by MS^2, and compared to the pure synthesized standards to validate annotation. Finally, a quantitative method was developed by LC coupled to a triple quadrupole instrument, and linearity and limit of quantification were determined. 14 new LpAA were quantified in gram positive bacteria, *Lactobacilus animalis*, and 12 LpAA in *Escherichia coli* strain Nissle 1917.

© 2021 Published by Elsevier B.V.

Abbreviations: RT, retention time.

* Corresponding author. MetaboHUB-MetaToul, National Infrastructure of Metabolomics and Fluxomics, Toulouse, France.
E-mail address: justine.bertrand-michel@inserm.fr (J. Bertrand-Michel).

[1] Nicolas Cénac and Justine Bertrand-Michel jointly supervised this work.

https://doi.org/10.1016/j.aca.2021.339316
0003-2670/© 2021 Published by Elsevier B.V.

Please cite this article as: A. Hueber, C. Petitfils, P. Le Faouder *et al.*, Discovery and quantification of lipoamino acids in bacteria, Analytica Chimica Acta, https://doi.org/10.1016/j.aca.2021.339316

Abbreviations

ACN	acetonitrile
AcOEt	ethyl acetate
CE	collision energy
DCM	dichloromethane
ESI	electrospray
FA	formic acid FWMH: full width at half maximum
HBTU	N,N,N′,N'-tetramethyl-O-(1H-benzotriazol-1-yl) uronium hexafluorophosphate
Hept	heptane
HLB	hydrophobic lipophilic balanced
HOBT	1-hydroxybenzotriazole
IS	internal standard
LC-HRMS	liquid chromatography high resolution mass spectrometry
LC-LRM	liquid chromatography low resolution mass spectrometry
LOD	limit of detection
LLE	liquid liquid extraction
LLOQ	lower limit of quantification
LOD	limit of detection
LOQ	limit of quantification
MeOH	methanol
MS	mass spectrometry
NMM	N-methylmorpholine
PIS	product ion scan
Qq/TOF	Quadrupole coupled to time of flight
QqQ	triple quadrupole
RF	radiofrequency
RT	retention time
SPE	solid phase extraction
SRM	selected reaction monitoring
TFA	Trifluoroacid
TIPS	triisopropyl silane
TOF	time of flight
TRIS	2-Amino-2-hydroxymethyl-propane-1,3-diol

1. Introduction

A functional microbiota is formed by the repertoire of expressed genes and, even more importantly, by proteins and metabolites. For human microbiota, a lot of studies investigated the compositional level but few studies have investigated active genes, proteins, or metabolites [1]. However, studies of molecules produced by the microbiota may be more relevant than composition when it comes to understand the role played by microbiota in host function. In fact, the taxonomic composition of the human microbiome varies tremendously across individuals, while its functional capacity is highly conserved in a same individual [2]. This conserved functional capacity is defined as the functional redundancy of the microbiota [2]. Although this redundancy is important to maintain the stability and resilience of the human microbiome, it introduces the idea that, until a certain point, species are substitutable in a given microbiota in terms of function [3]. The general aim of our consortium is to identify new metabolites produced by the gut microbiota to determine their effect on the host. In several pathologies related to intestinal microbiota dysbiosis, no increase in bacterial translocation is observed in patients indicating that, if the bacteria have an impact on the host, it is through molecules, secreted and/or transformed by the bacteria, able to cross the intestinal barrier. Amongst bacterial metabolites, several lipids including lipopeptides [4], long chain [5] and short chain fatty acid [6] or secondary bile acids [7], have been described for their capacity to cross the epithelial barrier. As in the brain, lipoamino acids (LpAA) composed by long chain unsaturated fatty acids conjugated to amino acid (AA) [8] have been identified and that we characterized lipopeptides and LpAA containing asparagine produced by *Escherichia coli* Nisle 1917 (EcN) [4], we decided to extend the discovery of LpAA from bacteria to other AA than asparagine.

In this study, a specific workflow is developed to identify and characterize new LpAAs in different bacterial strains using Liquid Chromatography High Resolution Mass Spectrometry (LC-HRMS). Advances in high-resolution mass spectrometry and tandem mass spectrometry, such as linear trap quadrupole (LTQ)/Orbitrap or hybrid Quadrupole Time-Of-Flight (Qq/TOF) technologies, offer an opportunity to rapidly screen biological and environmental specimens for a large array of chemicals using untargeted method [9–11]. To apply these global approaches, metabolites need to be present in a sufficient quantity to be characterized or at least detected, which is not our case since the concentration of our lipids of interest is quite low (about 50–70 pg.mg^{-1} of protein). Therefore, it is essential to develop a semi-targeted or suspect screening method dedicated to LpAAs. The typical approach applied in this case is a large scale suspect screening to generate semi-quantitative data in order to prioritize targeted developments. “Suspects” are defined as known compounds (“known/unknowns”) in terms of chemical name and structure which are expected (“suspected”) to be present in a sample. These approaches are often used in environmental and toxicological studies [12–15]. Even if data are acquired by HRMS, suspect screening differs from untargeted analysis because the analytes are compared to a user-defined chemical database or existing chemical inventories with manual or software-matching algorithms using accurate mass and isotope patterns. In this semi-targeted approach sample preparation plays a key role, and it needs to be adapted to the metabolites of interest [16]. In lipidomic, the two main methods which could be combined are: a liquid-liquid extraction [17] and a solid phase extraction after protein precipitation in a mixture of water and organic solvent [8,18,19].

To characterize bacterial LpAA, the different steps of the workflow developed in this project were to: 1) optimize an adapted and miniaturized chromatographic separation, 2) develop a specific sample preparation for LpAA, 3) construct a theoritical LpAA homemade database, 4) efficiently query HRMS data to find candidates and 5) characterize and confirm the new LpAA. Finally, we assessed our quantitative methods of LpAA in gram negative and positive bacteria and decipher the synthesis of this new LpAA in EcN.

2. Material and methods

2.1. Chemical and reagents

Heptane, ethyl acetate (AcOEt), methanol (MeOH), Hank's balanced salt solution (HBSS), Trizma® hydrochloride, dichloromethane (DCM), N,N,N′,N'-Tetramethyl-O-(1H-benzotriazol-1-yl) uronium hexafluorophosphate, 1-Hydroxybenzotriazole (HOBT), N-Methylmorpholine (NMM), TriFluoroAcid (TFA), triisopropyl

silane (TIPS), piperidine and C12Arg were purchased from Sigma Aldrich (Saint Quentin Fallavier, France). Acetonitrile (ACN, HRMS grade), and formic acid (HRMS grade) were purchased from Thermo Scientific. Water used in this study was purified on a milliQ system (Millipore). The 96 well-plate for solid extraction (SPE) (Oasis HLB 30 mg) were purchased from Waters (Saint-Quentin en Yvelines, France).

2.2. Standards

2.2.1. Synthesis of lipoaminolipids and lipopeptides

LpAA were synthesized following the standard procedure described in literature, by either liquid phase synthesis (LPS) or solid phase synthesis (SPS) with 2-chlorotrityl resin. In short, for LPS, the appropriate terbutyl aminoester and fatty acid were coupled with HBTU/HOBt/NMM in DCM, followed by hydrolysis of ester (and deprotection of an eventual trityl primary amide) with DCM/TFA. Purification of LpAA was carried out by Reversed Flash Chromatography (C4 or C18 column) with H_2O/MeOH or H_2O/ACN + 0.1% TFA. For SPS, appropriate Fmoc-amino acid was immobilized to 2-chlorotrityl resin with NMM in DCM. Fluorenylmethoxycarbonyl (Fmoc) protection was removed with piperidine/DCM (2/8), and resulting amine was coupled to appropriate fatty acid with HBTU/HOBt/NMM. After intensive washing, resin was cleaved with TFA/TIPS/Water (95/2.5/2.5) to obtain desired LpAA with variable yield and high purity, mostly without further purifications. Detailed synthesis of LpAA is developed in supplementary data. Internal standard (IS) used in this experiment was (S)-4-(4-amino-4-oxo-2-palmitamidobutanamido)butanoic acid substituted with 16 atoms of ^{13}C on the fatty acid chain (for detailed standard synthesis see Supplementary Information 1).

2.2.2. Standard preparation to determine linearity and reproducibility

Standard solutions were prepared in MeOH at the following concentrations 0.004, 0.19, 0.07, 0.30, 1.22, 4.88, 19.53, 78.13, 312.5 and 1250 pg.μL^{-1} for all primary standards. The concentration of the IS *C16AsnGABA was 600 pg μL^{-1}. Calibration curves were calculated by the ratio between the analyte area and the internal standard area.

2.3. Bacterial samples

2.3.1. EcN bacteria strains and culture conditions

Escherichia coli strain Nissle 1917 (EcN) (serotype O6:K5:H1), the isogenic mutants EcNΔclbA, EcNΔclbN and EcNΔclbP were used in this study. Gene inactivation was engineered using the lambda red recombinase method [20] and deletions were confirmed by using flanking primers. EcNwt, EcNΔclbA, EcNΔclbN, and EcNΔclbP were grown on LB agar plates supplemented with kanamycin (50 $\mu g\ mL^{-1}$) when required. After overnight incubation at 37 °C, single colonies of each strain were seeded in 3 mL of LB with antibiotics (when needed) and incubated overnight at 37 °C in agitation (250 rpm). Bacteria were then inoculated at OD600 = 0.05 in 10 mL of minimal medium A (K_2HPO_4 10.5 $g.L^{-1}$; KH_2PO_4 4.5 $g.L^{-1}$; $(NH_4)_2SO_4$ 1 $g.L^{-1}$; sodium citrate 0.5 $g.L^{-1}$; $MgSO_4$ 0.2 $g.L^{-1}$, thiamine 0,1 $g.L^{-1}$, and glucose 3%) supplemented with kanamycin (when required) and cultures were grown for 12 h at 37 °C under shaking conditions (250 rpm). The last step was repeated one more time by inoculating the bacteria at OD600 = 0.02 for 24 h before sampling them. The bacteria were stored at -80 °C before the LpAA and lipopeptide extraction.

2.3.2. Anaerobic bacteria strain and culture conditions

Lactobacillus animalis (DSMZ strain n°20602) was grown on Lactobacilli MRS broth (Difco, Fisher scientific SAS, Illkirch, France) − agar plates supplemented with cysteine chloride monohydrate 0.5 $g.L^{-1}$. After 24 h of incubation at 37 °C, 3 colonies were seeded in 30 mL of MRS broth supplemented with cysteine chloride monohydrate 0.5 $g.L^{-1}$ (Merck − Sigma Aldrich, St. Quentin Fallavier, France) and incubated overnight at 37 °C. All the culture steps have been performed in a hypoxic environment ensured by using a Whitley H35 hypoxystation (don Whitley scientific, Bingley, United Kingdom) and all the incubations were done in an anaeropack jar (Fisher scientific SAS, Illkirch, France) with an anaeroGen sachet (Fisher scientific SAS, Illkirch, France) to ensure an anaerobic environment.

2.4. Sample preparation

2.4.1. Extraction of LpAA and lipopeptides

For the extraction, 500 µL of Tris buffer (pH = 9), 1 mL of MeOH and 10 µL of IS *C16AsnGABA were added to the bacterial pellets and crushed with a Fast Prep instrument (MP Biomedicals), using two cycles (6,5 m s^{-1}, 30 s). Ten µL of suspension were withdrawn for protein quantification with Biorad assays, then, 6.6 mL of water were added to the homogenate. Samples were centrifuged at 1016×*g* for 15 min (4 °C) and the supernatants submitted to SPE using HLB plates (Oasis HLB, 30 mg, Waters). Plates were conditioned with 750 µL ethyl acetate, 750 µL MeOH and 750 µL H_2O:MeOH (90:10; v/v). Samples were loaded at a flow rate of one drop per 2 s and SPE plates washed using 1 mL H_2O:MeOH (90:10; v/v) followed by 1 mL heptane. Lipopeptides were eluted using 1 mL AcN, 1 mL MeOH and 1 mL AcOEt. Eluent were carefully removed from the plate, and transferred to a Pyrex tube, dried under nitrogen gas and reconstituted in 10 µL MeOH for LC-HRMS analysis. Extract was stored at −20 °C before LC-MS analysis.

2.4.2. Validation of the LpAA specific sample preparation

The extraction yield and the matrix effect were determined to validate the preparation of bacteria. Briefly, three sets of bacteria (*ECN*) were prepared according to section 2.4.1 in triplicate: 1) bacteria were extracted without IS; 2) bacteria were extracted without IS and IS were spiked in the final extract at three different concentrations (1000 pg μL^{-1}, 500 pg.μL^{-1} and 250 pg.μL^{-1}) and 3) bacteria were spiked with the three different concentrations of IS and then extracted. The extraction yield was determined as the difference between peak areas of IS in pre-spiked (condition 3) and post-spiked samples (condition 2). Condition 1 was used to ensure the absence of IS in bacteria. The matrix effect was determined as the difference between peak areas of IS added to the extracted samples (condition 2) and pure IS. Two independent experiments have been performed.

2.5. High resolution mass spectrometry analysis

2.5.1. LC-MS/MS conditions

The acquity UPLC-I-Class system (Waters) was equipped with CSH C18 column Waters (1 × 100 mm; 2,7 µm), as analytical column, and maintained at 40 °C. The mobile phases consisted of H_2O:HCO_2H (99.9:0.1; v/v) (A) and ACN:HCO_2H (99.9:0.1, v/v) (B). Flow rate was 0.15 mL min^{-1}. The multi-step gradient was as follow: 25% B at 0 min, 100% B at 16 min, 100% B at 18 min, 25% B at 18,1 until 20 min (Condition A). The autosampler (Acquity UPLC-I-Class FTN) was set at 4 °C, and the injection volume was 1 µL per analysis. UPLC was coupled on-line to a Xevo G2-XS time of flight (Qq/TOF) (Waters, Saint-Quentin en Yvelines, France) equipped with electrospray ionization (ESI). The MS, MS^E and MS/MS analysis were achieved in positive and negative ionization mode, in 2 separate runs with a full-width at half maximum (FWHM)

resolution of 16 000 (*m/z* 400). For the MS analysis, source temperature was 120 °C, cone voltage was set at 40 kV, nebulizer gas flow rate was 400 L h^{-1} and desolvation temperature was set at 380 °C. The analysis was performed in continuum mode for both the MS and MS [2] experiments. Mass spectra were recorded in MS and MS^E modes from *m/z* 50 to *m/z* 750. For the MS [2] experiments, nitrogen was the target gas in the XS collision cell (as segmented RF only quadrupole), ion precursors were isolated at ± 5 m/z for the required sensitivity. MS [2] experiments were carried out under low energy collisional activation conditions. The collision energy (E_{LAB}) was optimized using synthesized standards to maintain the precursor ion relative abundance at 30%. The best condition was an E_{LAB} of 20 eV for LpAA and lipopeptides. For the full characterization of C13Asn, the same flow rate and solvent A and B were used, starting with 48% of A at 0 min, 60% at 40 min, and 2 min to return on initial conditions (Condition B) and the same mass spectrometer parameters were applied.

2.5.2. Data processing

Unifi software (Version 1.9.4.053): data in continuum mode were loaded to the software. For the automatic identification by Unifi, the homemade database was added to the Unifi library and then to the analysis method as expected precursor. The analysis parameters were set as follow: *find 3D peaks*: at low energy peak intensity threshold was 2500 counts and *m/z* range was set between 30 m/z and 600 *m/z*, at higher energy peak intensity threshold was 80 counts, and *m/z* range was set between 150 m/z and 600 m/z; noise level was set at high level: 100000 counts; *3D peak detection*: without any restriction of retention time; *3D isotope cluster* was set at 0.3. In addition to the protonated MH^+ or deprotonated $[M-H]^-$ molecules, the adduct $[M+Na]^+$ or $[M+HCOO]^-$ ions was chosen for each precursor molecules under positive and negative ionization modes, respectively. The allowed score was set at 8. The software has been set to automatically search expected fragments: deprotonated amino-acid (for the 20 AA), AA ammoniac loss $[AA-H)-NH_3]^-$ (noted as $(z-2H)^-$ with nomenclature showed Fig. 1), deprotonated gamma-aminobutyric acid $[GABA-H]^-$, and fatty amid from C8 alkyl chain to C19 alkyl chain (i.e $[C_{12}H_{24}NO-H]^-$ noted as Lpc^-). The software has been set to automatically search neutral loss corresponding to the GABA loss: 103,0633. The mass error accepted for the precursor ion was set at 20 ppm, and at 2 mDa for the expected fragment ions.

QuanLynx software (application from MassLynx software, Version 4.2): reprocessing method were created to query *m/z*, with 5 ppm error accepted for the deprotonated molecule from the theoritical homemade database, without any RT restriction.

2.6. Low resolution mass spectrometry analysis

2.6.1. LC-MS/MS conditions

High-performance liquid chromatography system was a Shimadzu Mikros LC system, equipped with a thermostatted autosampler SIL-30AC, a rack changer II, a Nexera Mikros binary pump, a degasser on-line DGU-20A3R and a LTO-Mikros column oven. As for HRMS analysis, the analytical column was a CSH C18 column Waters (1 × 100 mm; 2,7 µm) and was maintained at 40 °C. The mobile phases consisted of water:formic acid (99.9:0.1; v/v) (A) and acetonitrile:HCO_2H (99.9:0.1, v/v) (B). Flow rate was 0.1 mL.min^{-1}. The multi-step gradient starts with 30% B at 0 min, 80% B at 13 min, 100% B at 13.5 min, 100% B at 16 min, and at 16.5 min 30% B until 18 min. The chromatography system was coupled online to a triple quadrupole mass spectrometer Shimadzu 8060 equipped with a ESI source in negative mode and was optimized as follow: *nebulizer*: 2 L.min^{-1}; *desolvation line*: 250 °C; *heating-block*: 450 °C, *heating gaz flow*: 10 L.min^{-1}, and no drying gas was used. Argon gas (purity, >99.9995%) was used for collision-induced dissociation (CID).

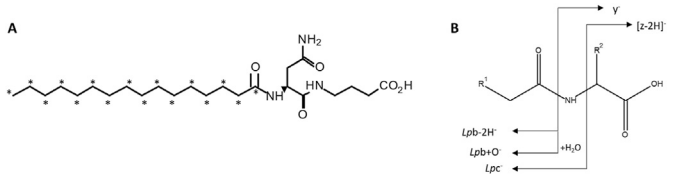


Fig. 1. (**A**) Structure of the labelled $^{13}C16:0$-AsnGABA internal standard (* = ^{13}C); (**B**) nomenclature of product ions from dissociation of the deprotonated lipo-amino-acids (Hueber et al., Amino Acid, Submitted).

Identification tools: the detector was programmed to profile putative compounds using theoretical SRM transition deducted from LpAA fragmentation with variation of fatty acid length. For instance, for C12Asn (previously characterized [4], the MRM transitions were $\{[M-H]^-; [(M-H_2O)-H]]^-\}$ and $\{[M-H]^-; Lpc^-\}$ corresponding to {313; 295} and {313; 198}. So to identify the other compounds of the series, we determined the theoretical MRM transitions for various alkyl chain length, e.g. the theoretical MRM of the C13Asn were {327; 309} and {327; 212}. This method was applied to all the amino acids identified in the bacterial LpAA.

Quantification tools: for each metabolite the MRM transition was optimized on the pure standard to get the best selectivity and the best sensitivity with a qualitative and a quantitative transition and was programmed to monitor a 2 min-window (expected chromatographic retention time ± 1 min). The dwelltime were optimized for each compound and were set between 30 and 100 msec.

2.6.2. Data processing

Peak detection, integration and quantitative analysis were performed by Labsolution Post run and Insight software (Version 5.99 SP2, Shimadzu, Kyoto, Japan). Concentration of the analytes in samples were calculated by calibration curves (Section 2.2.2). LOQ (Limit of Quantification) was determined using 3 calibration ranges in triplicate. LLOQ (Lower Limit of Quantification) were determined for accuracy (coefficient of variation) and precision (CV) in the range of 80–120% and LOQ for the accuracy and precision in the range of 85–115%. Linear dynamic range was determined as the range from LOQ to the highest calibration point reached. LOD (Limit of Detection) measured in biological samples was defined as a level with a signal-to-noise ratio (S/N) of 3.

3. Results and discussion

3.1. Analytical optimization to improve LpAA detection

3.1.1. Chromatographic separation and detection

A mixture of pure metabolites was used to develop the semi-targeted analysis. We tried to choose a broad spectrum of structures with different biophysical properties such as: C12AsnGABA, C12Asn, C14Asn, C16 and C18 hydroxy fatty acids, C20:4 or C18:1 acyl serotonin and 20:4 or 18:1 acyl anandamide at a concentration of 1 ng.μL^{-1}. As an internal standard, we need a molecule close to the one studied but absent of our matrix then a labelled derivative of CxAsnGABA was chosen. First, we synthetized a C12AsnGABA with the C12:0 fully deuterated. Due to a poor quality of labelled precursor, we finally synthetized a fully $^{13}C16:0$AsnGABA labelled on the fatty acid (Fig. 1). This internal standard was added to the mixture of standards to develop the analytical method.

The first step of the workflow was to optimize the chromatographic separation. The classical method used to separate free oxylipins and hydroxylated fatty acids under low resolution LC/ESI-MS/MS conditions [21,22] was adapted for the high-resolution Qq/

TOF instrument. To improve the sensitivity, the C18 2.1 mm diameter column was replaced by a 1 mm diameter columns with a reduction of flow rate from 0.4 mL.min^{-1} to 0.15 mL.min^{-1}. The total ionic current obtained for a EcN extract was increased by 87% (data not shown).

The mixture of standards was used to adjust the MS parameters in order to obtain the best sensitivity. Both the desolvation voltage and source temperature modifications did not improved the sensitivity. The gas flow was optimized between 500 and 800 L.h^{-1} to an optimal value at 800 L.h^{-1} in negative ionization mode and at 600 L.h^{-1} in positive ionization mode. The radiofrequency (RF) only quadrupole offset voltage was set with an optimum value at 130 V in both ionization modes. These optimizations allowed to achieve an average detection improvement of 15% for both the LpAA and lipopeptides compounds.

3.1.2. Sample preparation

Two main protocols are typically used to extract fatty acid metabolites: liquid-liquid extraction (LLE) and solid phase extraction (SPE). Bligh and Dyer LLE using a mixture of MeOH, $CHCl_3$ and H_2O is used in limited number of studies [23], others used heptane and ethyl acetate, instead of MeOH and $CHCl_3$ to extract endocannobinoid with efficient yield [24]. Nevertheless, the use of SPE after protein precipitation in a mixture of water and solvent (e.g. methanol, acetone or acetonitrile) is classically used in lipidomic for this family of lipids. Different types of SPE sorbent are used: polymeric reverse-phase sorbent plate for free oxylipins [21], C18 cartridge combined with ACN elution for surfactine, a bacterial lipopeptide [25], C18 sorbent combined with hexane and isopropanol elution for anandamide and N-Acyl-ethanolamine [26], or normal phase SPE to extract endocannabinoid from human tissues [27]. LLE and SPE could be combined to extract isoprostane [28] or acyl aminoacid with a series of complex SPE enrichment step [8].

Due to the ubiquity of lipids, no blank matrix is available to measure extraction yield, then a mixture of C12Asn, C14Asn, C14AsnGABA and IS at a concentration of 1 ng. μL^{-1} in solvent was used to assay the extraction. First, classic Bligh and Dyer extraction was tested. Even with addition of an acid (0.1% of acetic acid or 0.1% of HCl) to extract acidic metabolite after protonation, the extraction yield was very low and variable from one metabolite to another (data not shown). Then, two different strategies were tested: one using HLB SPE purification and another one using LLE extraction (Table S1). For SPE, standards and IS were homogenized in MeOH:H_2O (2:1; v/v) or in MeOH:TRIS pH 9 (2:1; v/v). After a centrifugation to precipitate the proteins, a dilution to 15% of methanol was performed. The metabolites were eluted by ACN, MeOH and AcOEt in a first method, by MeOH and CH_2Cl_2 in a second, or by ACN, MeOH and CH_2Cl_2 in a last one. The best yield of recovery (between 29% and 68%) for the 3 metabolites tested and for the IS were obtained with the first method (Table S1). For LLE, the best results were obtained with ethyl acetate in presence of water. However, to obtain clear extract ready to inject in case of complex matrix such as bacteria or tissue, the protocol has been completed by a washing step with heptane to remove neutral lipid [29]. This latter step induced the loss of C14Asn and IS and a lack of reproducibility. So, SPE combined with ACN, MeOH and AcOEt elution was chosen.

3.1.3. Validation of the sample preparation

To validate the sample preparation, extractions of EcN pellets (n = 3; 2 independent experiments) were performed with different concentrations of IS (250, 500 and 1000 pg μL^{-1}). In parallel, EcN pellets were extracted without internal standard which was added at the three concentrations at the end of the preparation. These different protocols compared to the non-extracted ISTD allowed the calculation of preparation yield and of the matrix effect. The preparation yield was similar between the different concentrations, with a recovery of 29.6% ± 3.3% for 1000 pg.μL^{-1}, 26.3% ± 7.2% for 500 pg.μL^{-1} and 43.6% ± 14.2% for 250 pg.μL^{-1}. The matrix effect was also similar between the different concentrations: with a recovery of 35.7% ± 1.3% for 1000 pg.μL^{-1}, 30.8% ± 3.2% for 500 pg.μL^{-1} and 26.2% ± 8% for 250 pg.μL^{-1}.

3.2. Workflow for the discovery of new LpAA

3.2.1. Theoretical lipopeptides and LpAA library inspired by C14Asn and C12AsnGABA

To query the LpAA of interest in lipid extracts, a homemade database has been constructed based on an amide linkage between a fatty acid and an amino acid or/and neurotransmitter (serotonin, dopamine, GABA or histamine). The length of fatty acid, as usually described in bacteria [30], was between 8 and 19 carbon atoms, saturated or with one unsaturation (or one cyclopropane which present the same mass default), with or without one hydroxyl group. 28 putative fatty acids were generated and one of the 20 natural amino-acids or an ornithine has been added (588 molecules) or one neurotransmitter such as GABA, serotonin, dopamine or histamine (84 molecules). Finally, each fatty acid linked to one amino acid were conjugated to one neurotransmitter (serotonin, dopamine, GABA or histamine) to propose 1764 molecules. These different associations generated at the end 2436 possible structures. For each putative metabolite, structural drawing, elemental composition of the $[M\text{-}H]^-$ and $[M+H]^+$ molecular species with the corresponding accurate *m/z* values have been generated and added to Unifi or QuanLynx Waters tools (Database not shown).

3.2.2. Acquisition of the data to provide putative candidate by spectral analogy

The lipid extracts were injected on the LC-Qq/TOF mass spectrometer to acquire data in MS^E continuum and centroid recording, in both positive and negative ionization modes. The continuum scan is required for the sample reprocessing with Unifi software, but not for Quanlynx software. The analysis on continuum mode generated heavy data (8 Giga octet per analyzed sample), compared to the centroid one (1 Giga octet per analyzed sample), which is not suitable to analyze large series of sample under large throughput conditions. First, we used the Unifi software to increase probability to find new LpAA or lipopeptides using the research of the neurotransmitter's common neutral losses or the expected fragment ions by analogy to the fragmentations describes for the C12Asn, C14Asn and C12AsnGABA [31] assuming that the same fragmentation pathways can occur under low energy collision conditions for the LpAA homologues. Nevertheless, after several tests on bacterial samples, this application was unable to find candidates. Candidates were only identified in a concentrated standards mixture (1 ng μL^{-1}). Even when a match occurred in the database, Unifi proposed a structure for the MS/MS associated fragment ions with an erroneous interpretation. In our hands, identification of new compounds using this software function worked only for metabolites present at a higher concentration than our metabolites of interest. If the intensity threshold was enhanced at the maximum to identify a maximum of compounds, more than 50 000 matches, impossible to check manually, were generated. In contrast, if this threshold was not enhanced at the maximum, some LpAA were not identified. In addition, when several compounds characterized by different molecular masses co-eluate, their respective molecular ionic species were activated together and their respective product ions were recorded together. Consequently, for each precursor ion, the software attributes the same product ions without distinction of their origin. Sometimes, for one fragment ion of interest, the

software indicated more than 50 precursors ions that should be checked manually. Finally, we concluded that the Unifi software was unsuitable for the identification of LpAA in complex mixtures.

Finally, we used QuanLynx (an application from MassLynx software, Waters usually applied to quantify metabolites) to automatically query our candidates in the homemade database where the precursor ions as deprotonated $[M-H]^-$ molecules were added regardless of retention time. For each automatic match, which could be the natural isotopic peak or an isobaric isotopomer, we had to check manually the presence of the natural isotopic cluster for each putative candidate in order to confirm that the match was the deprotonated molecule. Even if it was time consuming, the use of QuanLynx seems to be the best option for our purpose. This workflow has been used to identify LpAA in lipid extracted from *EcN* and *L. animalis*. We obtained 162 potential matches from the 2436 available compounds in the database, and after the isotopologue verification, 25 of them were real matches that needed to be fully characterized.

3.2.3. Characterization of the structure of a new LpAA: C13-Asn

The putative identification of C13-Asn by MassLynx was first done by attribution of its monoisotopic molecular mass (Mw = 328.2317 Da) from the mixture separation in μLC-HRMS and molecular deprotonation of compounds under negative ESI conditions. MS/MS analyses suggest the proposed LpAA structure *via* its gas phase fragmentations (through simple cleavages or rearrangements) of the deprotonated [LpAA-H] molecule.

3.2.3.1. Improvement of the chromatographic separation of C13-Asn. Using the chromatographic separation "conditions **A**", the analyzed selected ions were a broad peak cluster (ranging from *m/z* 326 to *m/z* 330, Fig. 2(a), **(b)**). Even if the peaks between *m/z* 328 and *m/z* 330 were the most intense signals persisting at higher E_{lab} values, it was possible to detect peaks at the *m/z* 327 nominal values as expected for the [C13Asn-H] ion (window of Fig. 2(a)). In addition to the peak at *m/z* 327.2274 (i.e., the [C13Asn-H] ion of interest), few isobaric ions, less intense than the neighboring satellite peaks ranging from *m/z* 328 and *m/z* 330, appeared between *m/z* 326.9867 and *m/z* 327.9867. Consequently, the product ion spectrum exhibited a lot of unexpected peaks impossible to interpret on the base of the [LpAA-H] ion structure. Indeed, recorded at E_{lab} = 5 eV, the product ion spectrum displays only few fragment ions at *m/z* 311.2354 (15.9920 Da loss), *m/z* 295.2689 (31.9587 Da loss) and *m/z* 190.8965 (136.3309 Da loss) which were inconsistent with dissociations of the putative [C13Asn-H] ion at *m/z* 327.2274. These losses cannot be attributed to fragmentations of the expected [C13Asn-H] ion (*m/z* 327.2274) but rather, either to its isobaric ions (at *m/z* 327.2682, *m/z* 327.1608, and *m/z* 326.9867) or to the other selected ions at the *m/z* 328-m/z 330 scale. At higher collision energy (Fig. 2(b)), the previous unknown signals were maintained except *m/z* 190 which was shifted at *m/z* 183.0229. Furthermore, additional peaks at *m/z* 255.2102 and *m/z* 269.2231 appear in addition to the expected peaks at *m/z* 309, *m/z* 265, *m/z* 238, *m/z* 212, *m/z* 131, *m/z* 114 and *m/z* 96 (Table S2). The results obtained at higher collision energy confirmed that fragmentations take place from either the isobaric ion mixture or/and higher mass precursor ions together, according to different relative stabilities under low energy collision conditions.

By improvement of chromatographic conditions (conditions B), peaks at the *m/z* 327 nominal values (at 5 eV Fig. 2(c)) were almost exclusively constituted by the [C13Asn-H] ion (*m/z* 327.2274). The CID spectrum exhibited very weak intense peaks at *m/z* 311 and *m/z* 285 which were not expected. This suggested that the precursor ions related to the isobaric compounds which co-elutes with C13Asn, were more fragile toward the CID processes. The peaks at *m/z* 311 and *m/z* 285 disappear at E_{lab} = 20 eV (Fig. 2(d)) and then, the product ion spectrum was characterized by the exclusive presence of the expected fragment ions of [C13Asn-H] with less than 5 ppm of error (Table S2). However, a main signal arises at *m/z* 327.1645 in addition to the peak at *m/z* 327.2274 ion (window in Fig. 2(d)). The intensity of the peak at *m/z* 327.2274 was strongly decreased at the high collision energy compared to the low collision energy (window in Fig. 2(c)) while the intensity of its satellite peak at *m/z* 327.1645 was very low at low energy. This suggests that the *m/z* 327.2274 ion was significantly less stable than its isobaric ion at *m/z* 327.1645 with an unknown structure.

Interestingly, at low collision energy (5 eV, Fig. 2(e)), the CID spectrum of *m/z* 327 of the synthetic C13Asn did not display fragmentation which was not the case for the CID spectrum of the natural product under same collision conditions (Fig. 2(a)). This confirmed the fragility of unknown isobaric ions relative to [C13Asn-H] and its satellite peaks of the *m/z* 328 - *m/z* 330 zone. Finally, both the product ion spectra (Fig. 2(d) and (f)) of *m/z* 327 provided from biological mixture and from synthetic sample were stackable as two identical fingerprints. The structure proposed to the *m/z* 327.2274 ion was confirmed and the different product ions of its CID spectrum were completely rationalized on the basis of the [C13Asn-H] structure.

3.2.3.2. Main displayed fragmentations in product ion spectra of the deprotonated C13-Asn. General interpretations of even-electron molecular ionic species are based on the following considerations: (i) deprotonation can occur competitively at carboxylic acid, amide and enolizable ester sites resulting in a mixture of various [LpAA-H] deprotomers [32], (ii) the negative (or positive) charge exclusively promotes fragmentations under low collision energy conditions and (iii) dissociations may yield simple cleavages as well as rearrangements. The latter can occur directly or through stepwise processes (e.g. *via* isomerization into ion-dipole intermediates prior to complete dissociation) [33]. Due to the role of the charge, different competitive bond cleavages can take place according to the charge location.

The CID spectrum rationalization can allow the selection of the most relevant collision energy to reach the best sensitivity for the elucidation of the LpAA molecule structure and the best specificity for its quantification. To illustrate this purpose, we detailed the dissociations of the [C13Asn-H] ion (*m/z* 327) activated under the E_{Lab} = 20 eV conditions.

Favored loss of water from [C13Asn-H]. This abundant loss yielding *m/z* 309 was particularly specific and was observed independently to the length of fatty acid moiety (its relative abundance is larger than 50%). Indeed, its signal appeared as the base peak (at E_{Lab} = 20 eV, Fig. 2(f)) in the CID spectrum. It reached an optimum of 56.8% intensity relative to the total ionic current (TIC) at E_{Lab} = 20 eV as shown in the energy resolved mass spectrometry (ERMS) breakdown curves (Fig. S1). Such water release did not exceed an optimum of 10% of TIC (Table S2) for the most abundant one as demonstrated from ERMS of the other studied deprotonated [LpAA-H] species (ERMS not reported). Thus, the water loss must be assisted by the functional group of the AA side chain since this loss is totally hinder from LpAA with non-polar AA moiety. Consequently, the amide group must be deprotonated to drive a nucleophilic attack at the carboxyl site yielding *via* cyclization, a tetrahedral intermediate (Scheme 1(a)). This one isomerized by C–OH bond cleavage into an OH ion/[(C13Asn-H)-OH] neutral complex from which OH removed one proton to the neutral and the resulting complex dissociated by water release and formation of [(C13Asn-H)-H_2O] (*m/z* 309). This carboxylate derivate carries a nitrile leaving group capable of being subsequently released.

Consecutive losses of (H_2O + CO_2) and (H_2O + CO_2+HCN) yielding

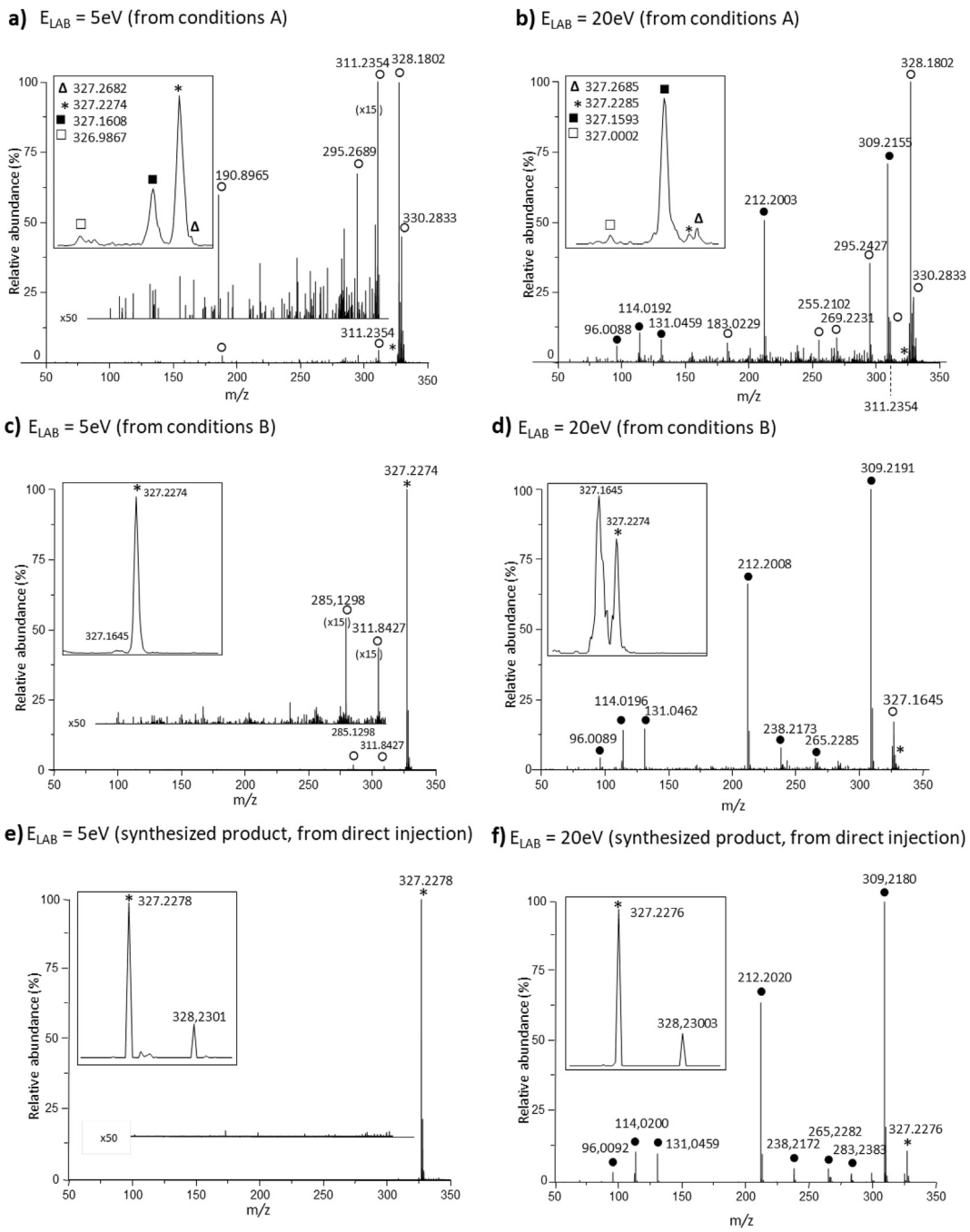
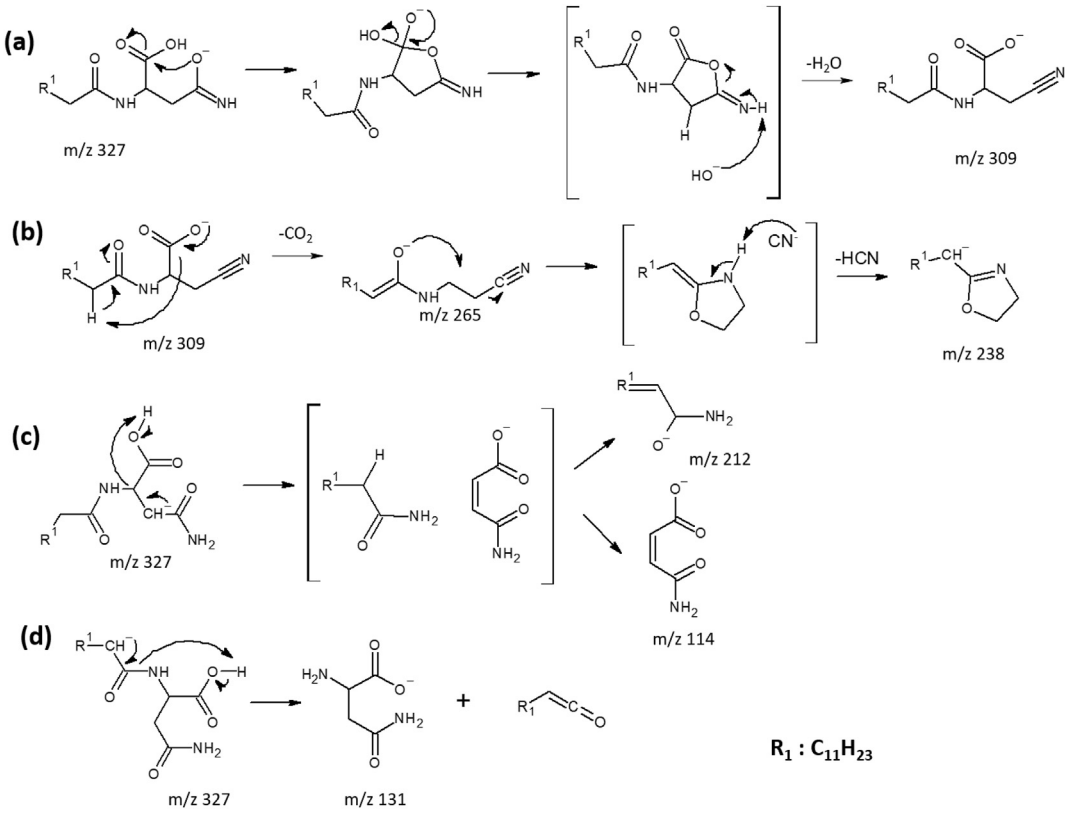


Fig. 2. Product ion spectra of *m/z* 327 generated from μLC-HRMS in negative ESI mode under the A separative conditions (a)–(b) and under the B conditions (c)–(d) from biological sample, and by direct injection from synthetic C13Asn sample (e)–(f) under low energy collision conditions: (a)-(c)-(e) E_{Lab} = 5 eV and (b)-(d)-(f) at E_{Lab} = 20 eV. Normalized *m/z* molecular zone in window of each product ion spectrum and in (a) and (b) the peak intensities of the *m/z* 100-m/z 310 scale were amplified by a factor of 50). In product ion spectra (b) and (d), the used annotations of peaks (with intensity >5% of base peak) are (i) the asterisk (*) for expected precursor ion, (ii) the empty circle (○) for unexpected peaks and (iii) the full circle (●) for expected peaks attributed to [C13Asn-H] (according to Table S2 in Supplementary Information).

m/z 265 and m/z 238. The CO_2 loss consecutive to that of H_2O was assisted by a 1–4 proton transfer from the amide enolizable α position to the carbon atom which initially carried the CO_2 lost to obtain the [(C13Asn-H)-H_2O-CO_2] enolate (*m/z* 265, Scheme 1(b)). Consecutively, the alkoxy site can assist the loss of the leaving CN group through internal nucleophilic substitution. By this way, the CN ion/[(C13Asn-H)-(H_2O + CO_2+CN] neutral complex was formed, and internal proton transfer from enamine site to CN allowed the elimination of HCN and led to formation of *m/z* 238 as a deprotonated 2-alkyl-oxazoline form (Scheme 1(b)).

Competitive formation of complementary m/z 212 and m/z 114 ion pair. From the deprotomer molecular species (originated from the water loss, with the charge at the enolizable position of the amide group of Asn), the moiety side chain migrated concomitantly with proton from the neighbored carboxylic acid group. This step led to the RCONH-CH bond cleavage yielding ion/dipole complex



Scheme 1. Proposed formation of major product anions from the LpAA C13Asn (*m/z* 327) product anion.

constituted by deprotonated maleic acid monoamide and fatty acid amide. This complex can directly dissociated to release the fatty acid amide leading to formation of *m/z* 114 (ion $[z\text{-}2H]^-$) or after an inter-partner proton transfer, the isomerized complex can decomposed into fatty acid amidate (ion *Lpc*$^-$, *m/z* 212) (Scheme 1(c)).

Formation of deprotonated asparagine (m/z 131). The CO—NH bond cleavage was promoted by the neighbored negative charge at the enolizable position of the amide linkage combined with proton migration from carboxylic group to nitrogen site. By this pathway, the fatty acid ketene was directly released (Scheme 1(d)) and was not enough acid to form an ion/dipole by molecular species isomerization prior to dissociation.

After C13Asn full fragment pattern elucidation, the same identification has been expanded to other LpAA found by suspect screening experiments performed on a LC-LRMS/MS using triple quadrupole instrument more sensitive than the Qq/TOF tandem. From previous MS [2] characterization, it could be possible to predict theoretical MS/MS according to the alkyl chain length and to the deprotonated amino acid as product anions. These MRM transitions were programmed on μLC-QqQ method to analyze bacterial extracts. Due to the improved sensitivity, this approach permitted to find 54 putative new LpAA: C14:1Asn, C15Asn, C16Asn, C16:1Asn, C17Asn, C17:1Asn, C12:1Leu/Ile, C13Leu/Ile, C14Leu/Ile, C14:1Leu/Ile, C15Leu/Ile, C15:1Leu/Ile, C16Leu/Ile, C16:1Leu/Ile, C17Leu/Ile, C17:1Leu/Ile, C18Leu/Ile, C18:1Leu/Ile, C12Gln, C12:1Gln, C13Val, C14Val, C14:1Val, C15Val, C16Val, C16:1Val, C17:1Val, C18Val, C12Glu, C12:1Glu, C16Glu, C15:1Glu, C16Glu, C17:1Glu, C12Lys, C12:1Lys, C14Lys, C15Lys, C16Lys, C16:1Lys, C17Lys, C17:1Lys, C12:1Phe, C13Phe, C14Phe, C14:1Phe, C15Phe, C15:1Phe, C16Phe, C16:1Phe, C17Phe, C17:1Phe, C18Phe and C18:1Phe. These compounds were identified in LC-LRMS, without a level 1 of identification which correspond to metabolites fully characterized and identified. Two orthogonal analysis methods of an authentic chemical standard analyzed in the same laboratory is mandatory for this level of identification. In addition, the standard should be compared to experimentally identified metabolite in the same laboratory with the same analytical method [34]. These 54 compounds were identified with a level 3 of identification corresponding to metabolites with incomplete or putative annotation. However, this method didn't allow the identification of LpAA with serine, ornithine, methionine, alanine, cysteine and threonine.

We decided to characterize the most abundant LpAA and at least one per family. We focused on the whole LpAsn family (C12:1Asn, C14:1Asn, C15Asn, C16:1Asn, C17Asn, C17:1Asn), and on C12Lys, C14Lys, C12Glu, C14Glu, C16Glu, C12Asp, C12Gln, C12Val, C12Trp, C12Arg, C12His, C12Leu, C12Ile, and C14Phe. These compounds were synthetized (except C12Arginine which were purchased from Sigma Aldrich) to be fully characterized by MS [2], with an identification level 1 or level 2, to permit construction of the quantitative method (Table S2).

3.2.3.3. Full characterization of 22 new bacterial LpAA. Following the principle of elucidation detailed for the C13Asn, 22 LpAA were fully characterized with their fragmentation pattern (Table S2). These 22 molecules were composed by a fatty acid (from C12 to C17) conjugated to one of 12 following amino acid: asparagine, arginine, aspartate, glutamine, glutamate, histidine, leucine/isoleucine, lysine, phenylalanine, tryptophan and valine. The dissociations presented in section 3.2.3.2 (Scheme 1) were similar for all the LpAsn (except the product ions containing the fatty acid moiety) to those displayed in the product ion spectra of the deprotonated C12Asn, C12:1Asn, C14Asn, C14:1Asn, C15Asn, C16Asn, C17Asn and C17:1Asn molecules.

The consecutive losses of (H_2O + CO_2) and (H_2O + CO_2+HCN) presented in section 3.2.3.2 (Scheme 1 (b)) yielding *m/z* 265 and *m/z* 238 for the C13Asn, were specific to the LpAsn family (Lp being C12:1, C12, C13, C14:1, C14, C15, C16:1, C16, C17:1 and C17) and to LpGln and LpHis even if the abundance of these ions were low (Table S2). The loss of water of the $[C12Gln-H]^-$ (*m/z* 327) was strongly reduced to 4% compared to the 56.8% observed for [C13Asn-H] (*m/z* 327). Such a difference has been already described for the dissociation of the free [Asp-H] and [Glu-H] anion. Concerning the observed (H_2O + CO_2+HCN) loss sequence of the [C12His-H] ion (*m/z* 336), the previous proposed mechanisms cannot be involved. The water loss was possible, due to the large gas phase acidity of the imidazole group of the histidine moiety and induced a bicyclic system from which, the consecutive losses of CO_2 and (CO_2+HCN) were produced by dissociation pathways as reported in Scheme S1.

The formation of the fatty acid amidate (ion Lpc^-, Fig. 1(b)) and its complementary [z-2H] ion specifically occurred from LpAA with amide group on the side chain of amino-acid moiety (Asn and Gln) *via* a similar mechanism (Scheme 1(c)). Note that in competition with this ion pair, formation of deprotonated amino acid (i.e., y^-) took place favorably from [LpGln-H]. When this amide group was exchanged by a carboxyl group (i.e., Asp and Glu), only the fatty acid amidate, Lpc^-, was detected with an intensity of 38% of TIC for [LpAsp-H] and 4% at maximum for [LpGlu-H]. However, the abundance reduction of Lpc^-, [z-2H] and y^- (or their disappearance) was accompanied by a strong emergence of fatty acid carboxylate (i.e., [*Lp*b + O]) and deprotonated amino-acid anhydride (i.e., [*y*-H_2O]) as a complementary product ion pair. This regeneration of deprotonated fatty acid was specific to the [LpAsp-H] and [LpGlu-H] precursor ions. For LpAA conjugates with basic AA moieties (i.e., Lys, Arg and His), with aromatic side chain (Phe and Trp) and with non-polar amino acids (Val, Leu and Ile), these above competitive ion pairs were replaced by deprotonated amino-acid (y) with abundance higher than 30% (except for Arg). The y abundance was re-enforced (>50% of TIC) mainly from non-polar side chain AA.

To conclude, the 22 LpAA discovered in this study were characterized by MS [2], but to reach the level 1 of identification, it was necessary to compare the candidates to the pure synthesized compounds which allow to verify the fragmentation pattern and the retention time.

Table 1
Comparison of discovered aminolipids measured in bacterial sample and pure synthetized compounds: retention time (RT) comparison with RSD% calculation, *m/z* with error ppm.

		RT LC-HRMS (min)	% RSD	RT LC-LRMS (min)	% RSD	Elementary composition	*m/z* calculated $[M-H]^-$	*m/z* experimental (HRMS)	Erreur (ppm)
C12-ASN	Synthesized	3.72	0,56	5.89	0,36	$C_{16}H_{30}N_2O_4$	313.2132	313.2119	4.4
	Sample	3.75		5.86				313.2120	4.1
C12:1-ASN	Synthesized	5.70	0,12	4.81	0,14	$C_{16}H_{28}N_2O_4$	311.1976	311.1972	1.4
	Sample	5.69		4.80				311.1980	1.2
C13-ASN	Synthesized	5.42	1,16	7.04	0,40	$C_{17}H_{32}N_2O_4$	327.2289	327.2274	4.7
	Sample	5.51		7.00				327.2274	4.7
C14-ASN	Synthesized	6.60	0,63	8.18	0	$C_{18}H_{34}N_2O_4$	341.2445	341.2443	0.8
	Sample	6.66		8.18				341.2446	0.1
C14:1-ASN	Synthesized	7.68	0,27	6.79	0	$C_{18}H_{32}N_2O_4$	339.2289	339.2278	3.3
	Sample	7.65		6.79				339.2296	2.0
C15-ASN	Synthesized	8.69	0,24	9.34	0,07	$C_{19}H_{36}N_2O_4$	355.2602	355.2599	0.9
	Sample	8.66		9.33				355.2605	0.8
C16-ASN	Synthesized	7.27	0,88	10.51	0	$C_{20}H_{38}N_2O_4$	369.2758	369.2752	1.8
	Sample	7.18		10.51				369.2769	2.8
C16:1-ASN	Synthesized	9.66	0,14	8.91	0	$C_{20}H_{36}N_2O_4$	367.2602	367.2600	2.1
	Sample	9.64		8.91				367.2602	0.1
C17-ASN	Synthesized	8.14	0,77	11.67	0,06	$C_{21}H_{40}N_2O_4$	383.2915	383.2914	0.3
	Sample	8.23		11.68				383.2914	0.3
C17:1-ASN	Synthesized	8.18	0,25	9.99	0	$C_{21}H_{38}N_2O_4$	381.2758	381.2758	0.1
	Sample	8.15		9.99				381.2752	1.8
C12-ARG	Synthetized	2.30	0,61	3.09	0,45	$C_{18}H_{36}N_4O_3$	355.2714	355.2714	0.1
	Sample	2.32		3.07				355.2714	0.1
C12-ASP	Synthesized	5.48	0,25	6.83	0	$C_{16}H_{29}NO_5$	314.1972	314.1974	0.1
	Sample	5.46		6.83				314.1941	10.2
C12-GLN	Synthetized	4.63	0,30	5.86	0,36	$C_{17}H_{32}N_2O_4$	327.2289	327.2288	0.4
	Sample	4.61		5.83				327.2277	3.8
C12-GLU	Synthesized	5.42	0,39	6.84	0,10	$C_{17}H_{30}NO_5$	328.2129	328.2122	2.2
	Sample	5.45		6.83				328.2136	2.0
C14-GLU	Synthesized	7.36	0,28	9.10	0	$C_{19}H_{34}NO_5$	356.2422	356.2427	4.3
	Sample	7.39		9.10				356.2420	6.3
C16-GLU	Synthesized	9.33	0,07	11.37	0; 06	$C_{21}H_{38}NO_5$	384.2755	384.2756	0.2
	Sample	9.32		11.38				384.2755	0.1
C12−HIS	Synthesized	2.19	1,59	2.96	0,47	$C_{18}H_{31}N_3O_3$	336.2292	336.2290	0.8
	Sample	2.24		2.94				336.2303	3.1
C12-LYS	Synthesized	2.12	0,99	2.89	0,49	$C_{18}H_{36}N_2O_3$	327.2653	327.2646	0.3
	Sample	2.15		2.87				327.2653	0.1
C14-LYS	Synthesized	3.57	0,98	4.57	0,76	$C_{20}H_{40}N_2O_3$	355.2966	355.2961	1.4
	Sample	3.62		4.62				355.2959	2.0
C12-LEU/ LEU	Synthesized	8.73	0,24	8.68	0,16	$C_{18}H_{35}NO_3$	312.2544	312.2540	1.3
	Sample	8.76		8.66				312.2542	0.7
C14-PHE	Synthesized	10.84	0,32	13.00	0,05	$C_{23}H_{37}NO_3$	374.2700	374.2699	0.4
	Sample	10.79		12.99				374.2700	0.1
C12-TRP	Synthesized	8.26	0,51	10.04	0	$C_{23}H_{34}N_2O_3$	385.2496	385.2496	0.1
	Sample	8.20		10.04				385.2497	0.1
C12-VAL	Synthetized	7.95	0	9.81	0	$C_{17}H_{33}NO_3$	298.2387	298.2379	2.9
	Sample	7.95		9.81				298.2387	0.1

3.2.4. Validation for the discovered structures

So pure synthesized standards were analyzed in the same conditions than bacterial extracts, and HRMS analyses permitted to confirm *m/z* ratios and retention time (Table 1). Herein, even with the use of standards, the double bond or hydroxyl positions on fatty acid moiety could not be determined. The level of annotation for each LpAA characterized is indicated in Table 1.

Even if LpAA composed by Phe and Val associated to C20:4, C22:6 or C18:1 fatty acid have been described in brain [8], from the 2200 theoretical aminolipids database, we identified 10 new LpAA in EcN (*Escherichia coli* strain Nissle 1917): C13Asn, C14:1Asn, C15Asn, C16Asn, C16:1Asn, C17:1Asn, C12Glu, C12Lys, C12His and C12Asp; and 15 LpAA in *Lactobacillus animalis*: C12Arg, C14Glu, C12Lys, C14Lys, C12His, C12Gln, C12Asp, C14:1Asn, C16:1Asn, C15Asn, C12Trp, C12Val, C12Leu/Ile, C16Asn and C14Phe.

3.3. Quantification of new aminolipid in bacteria

3.3.1. Development of a sensitive method by LC-MS/MS

To study these new aminolipid in bacteria or in complex matrixes such as intestines or stools it is essential to develop a robust sensitive and quantitative method. As we gave priority to sensitivity in front of resolution, we developed this method on a LR-MS triple quadrupole detection Shimadzu 8060 system coupled to a μLC. In order to achieve the necessary selectivity and sensitivity of the method, the source parameter and the mass detection were optimized. For the source parameters nebulizer was assayed from 1 to 5 L min^{-1}; desolvation line from 150 to 250 °C; interface temperature from 100 to 300 °C; heating-block from 300 to 450 °C, heating gas flow from 0 to 10 L min^{-1}, and finally dry gas at 0 L min^{-1} in order to obtain the best compromise. Analyses were performed in Selected Reaction Monitoring (SRM) detection mode using argon as a collision gas. Optimization for each compounds were done automatically with LabSolution software. The quantification transition chosen for each compound was the more abundant, even if it was not a specific transition. The 2 transitions used for qualitative analysis were chosen among the most abundant with exclusion of the quantitative transition. The values of Q1 pre bias, energy collision and Q3 pre bias were automatically optimized by the software. For example for the C13Asn metabolite, the quantitative transition was {*m/z* 327}/{*m/z* 309} corresponding to the $[M - H]^-/[(M - H_2O)-H^-]$ and the two qualitative transitions were {*m/z* 327}/{*m/z* 212} corresponding to the $[M - H]^-/[(RCONH_2)-H^-]$ and {*m/z* 327}/{*m/z* 131} corresponding to the $[M - H]^-/[(\text{Amino acid})-H^-]$. The same transition (qualitative and quantitative) scheme was effective for all the LpAA in this study, and are report in Table 2 (the second qualitative transition is not shown). Collision energy were optimized to obtain the best detection for each standard (Table 2).

3.3.2. Sensitivity, linearity and reproducibility of the method

To study the sensitivity and the linearity of the quantitative method, calibration lines with 10 calibration points from 0.004 to 1250 pg μl^{-1} (with 600 pg μL^{-1} of IS) were analyzed 3 times in triplicates. The limit of detection (LOD) used in the analytical method of biological samples was defined as the level with a signal-to-noise ratio (S/N) of 3. The limit of quantification (LOQ) was determined by measuring the level with a signal-to-noise ratio (S/N) greater than 10 and with an accuracy and precision within the range of 80–120%. The accuracy and the precision of all 10 calibration points above LOQ were in the range of 85–115%. The Linear dynamic range was determined as the range from LOQ to the highest calibration point corresponding to the conditions described above (Table 2). Repeatability and precision were respectively calculated using relative standard deviation (% RSD) and accuracy for all range of concentrations of pure standards in triplicates, they were under 2% (data not shown). The same experiment was done to assay the inter-day variations, calibration curves were analyzed twice, with 15 days between analyses, and RSD were under 5%. The method allowed the quantification of LpAA with a great sensitivity with a LOQ between 0.39 and 1.56 pg injected, and a large linearity range from LOD until 1250 pg injected.

3.3.3. LpAA quantification in EcN and L. animalis

Our quantitative method has been used to study LpAA synthesis by a gram negative bacteria EcN and a gram positive Lactobacillus

Table 2
SRM characterization with quantitative and qualitative transition, energy applied to generate the transition (EC et Q1/Q3PreBias), Linear regression equation and correlation R2.

Compound	Linear regression	R^2	LOD ($pg.\mu L^{-1}$)	LOQ ($pg.\mu L^{-1}$)	Quantitative MRM transition	Q1 Pre Bias (V)	EC (eV)	Q3 Pre Bias (V)	Qualitative MRM transition
C12:1Asn	y = 49290x - 173866	0.9990	0.39	1.56	312 → 293	30	15	13	311 → 183
C13Asn	y = 63564x - 222566	0.9984	0.19	1.56	327 → 309	34	18	15	327 → 212
C14:1Asn	y = 132014x - 110079	0.9977	0.19	0.78	339 → 321	14	17	22	339 → 224
C15Asn	y = 25510x - 65143	0.9996	0.39	1.56	355 → 240	25	24	16	355 → 337
C16Asn	y = 64907x - 151246	0.9990	0.19	0.39	369 → 351	15	20	16	369 → 254
C16:1Asn	y = 53376x - 199389	0.9995	0.78	1.56	367 → 349	28	20	26	367 → 252
C17Asn	y = 137315x - 20177	0.9978	0.39	0.78	383 → 365	15	20	18	383 → 96
C17:1Asn	y = 40608x - 117983	0.9989	0.19	0.78	381 → 363	14	25	10	381 → 266
C12Arg	y = 15767x - 105656	0.9988	0.04	0.78	355 → 131	20	26	20	355 → 131
C12Asp	y = 44937x - 317123	0.9950	0.19	0.78	314 → 198	24	23	21	314 → 199
C12Gln	y = 22434x - 40015	0.9985	0.19	0.39	327 → 198	24	28	23	327 → 127
C12Glu	y = 27518x - 86993	0.9991	0.19	1.56	328 → 198	10	20	20	328 → 128
C14Glu	y = 49107x - 214246	0.9983	0.19	0.78	356 → 226	26	22	24	356 → 128
C16Glu	y = 1262x - 19015	0.9943	0.19	0.78	384 → 255	30	23	8	384 → 128
C12His	y = 30227x - 77911	0.9987	0.39	0.78	336 → 110	25	27	10	336 → 154
C12Leu	y = 181029x - 418227	0.9990	0.04	1.56	313 → 130	13	29	26	312 → 268
C12Ile	y = 236326x - 425633	0.9985	0.04	1.56	313 → 130	12	29	26	312 → 268
C12Lys	y = 45468x - 202933	0.9976	0.19	1.56	327 → 145	13	25	27	327 → 283
C14Lys	y = 70865x - 401950	0.9988	0.19	0.78	355 → 145	14	25	27	355 → 311
C14Phe	y = 81754x - 29953	0.9981	0.19	0.78	374 → 164	15	24	29	374 → 226
C12Pro	y = 77314x - 52780	0.9975	0.19	1.56	296 → 114	21	21	11	296 → 198
C12Trp	y = 46967x - 173680	0.9987	0.19	1.56	385 → 203	14	21	20	285 → 256
C12Val	y = 180951x - 433184	0.9988	0.04	0.39	298 → 116	22	22	12	298 → 254

animalis. In wild-type EcN, we quantified the C12Asn, C13Asn, C14Asn, C14:1Asn, C15Asn, C16Asn, C16:1Asn, C17:1Asn, C12Glu, C12Asp, C12AsnBABA and C12AsnGABA, and in Lactobacillus animalis the C12Lys, C14Lys, C12His, C12Gln, C12Asp, C14:1Asn, C16:1Asn, C15Asn, C17Asn, C12Trp, C12Val, C12Leu, C16Asn and C14Phe (Fig. 3, Table S3).

EcN is known to harbour a genomic island, named pks, which carries a cluster of genes that enables the synthesis of LpAA and of a genotoxin called colibactin [35–38]. This toxin is produced by a complex biosynthetic machinery involving the sequential action of proteins ClbA to ClbS [39]. ClbA, a phosphopantetheinyl transferase (PPTase) essential for the activation of the NRPS and PKS enzymes [39], is mandatory for the biosynthesis of colibactin [35] but is also involved in the biosynthesis of other bioactive metabolites such as the siderophores enterobactin, salmochelin and yersiniabactin [40]. Following activation by ClbA, the initiating NRPS ClbN uses Asn as a substrate to generate an N-myristoyl-D-Asn (C14Asn). The NRPS-PKS assembly line continues the synthesis of precolibactin compound(s) using malonyl-coA and different amino acids [36–38] as substrates. The precolibactin is then cleaved by peptidase ClbP to liberate colibactin and N-myristoyl-D-Asn [36–38].

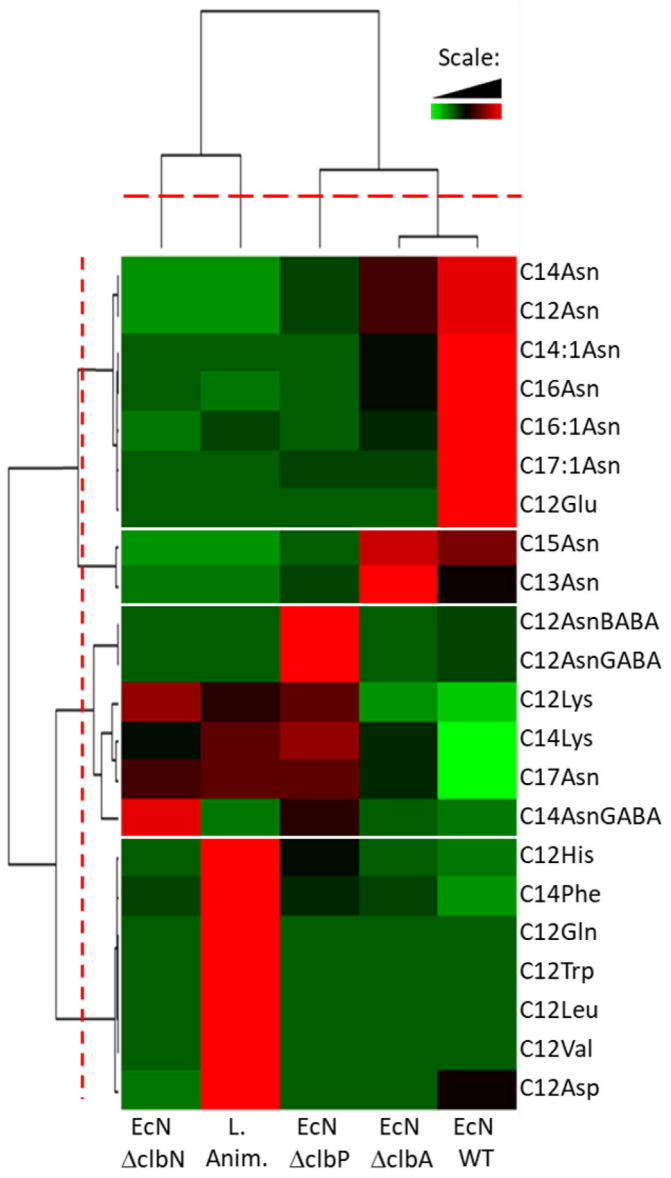


Fig. 3. Heat map of LpAA and Lipoeptides quantified by LC-MS/MS. Data are shown in a matrix format: Each row represents a single lipid, and each column represents a bacterial strain: Escherichia coli strain Nissle 1917 wild type (EcN WT) or delated for specific gene of the pks island (clbN, clbA or clbP) and Lactobacillus animalis (L. Anim.). Each color patch represents the normalized quantity of lipid (row) in a bacterial strain (column), with a continuum of quantity from bright green (lowest) to bright red (highest). The pattern and length of the branches in the left dendrogram reflect the relatedness of the lipids and in the top dendogram, the realatedness of the bacterial strains. The dashed red lines are the dendrogram distance used to cluster lipids or bacteria. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In order to determine the implication of ClbA, ClbN and ClbP in the synthesis of the different identified lipids, LpAA were quantified in EcN wild type, in EcN deleted for ClbA, ClbN or ClbP, and in L. animalis. Hierarchical clustering of LpAA amounts quantified in bacterial pellets (pg/mg of proteins) was used to reveal the main differences between them (Fig. 3). LpAA formed 4 different clusters. The first cluster was constituted by LpAA with a concentration decreased in EcN ΔclbA, ΔclbP and ΔclbN compared to EcN WT and not expressed in L. animalis. The synthesis of these LpAA was dependent on clbA, clbN and clbP as previously described for the C14asn. The second cluster discriminated LpAA which need clbN and clbP but not clbA to be produce and not present in L. animalis. The third cluster was composed by LpAA expressed in EcN deleted for ClbN and ClbP, and in L. animalis. Finally, the fourth cluster was composed by LpAA expressed mainly in L. animalis. All together these data indicated that these gram negative and gram positive bacteria produced LpAA. In addition, a single deletion of genes modified not only the concentration but also the nature of LpAA demonstrating the importance to quantify a large panel of LpAA when studying bacterial metabolites.

4. Conclusion

In conclusion, we developed a workflow which potentially allows the identification of the 2436 compounds of our database using mass spectrometry. The structural determination of the compounds by mass spectrometry and its validation with pure synthetized standards are two important steps for the characterization of LpAA and lipopeptides as their low concentration do not allowed their analysis either by NMR or/and by X-ray, which would require more than several tens of μg and several mg, respectively. This method allowed the characterization and quantification of 22 new LpAA in bacteria. Our quantitative method could be applied to more complex matrix such as patient feces or biopsies in order to determine the role of these bacterial molecules in diseases related to microbiota dysbiosis. This method represents a unique opportunity to connect an innovative and quantitative method to clinical applications to investigate LpAA synthesis by patient microbiota.

Funding resources

This work was supported by the French National Institutes of Health (Inserm) and Region Occitanie. NC is a recipient of the grant from ANR (ANR-18-CE14-0039-01) and NC & JBM are recipient of the grant from ANR (ANR-20-CE14-0011-01).

CRediT authorship contribution statement

Amandine Hueber: Conceptualization, Data curation, Visualization, Investigation, Writing – original draft, Writing – review & editing. **Camille Petitfils:** Resources. **Pauline Le Faouder:** Resources, Conceptualization. **Geoffrey Langevin:** Organic Synthesis. **Alexandre Guy:** Organic Synthesis. **Jean-Marie Galano:** Organic Synthesis. **Thierry Durand:** Organic Synthesis. **Jean-François Martin:** Visualization, Investigation. **Jean-Claude Tabet:** Writing –

original draft, Writing – review & editing. **Nicolas Cenac:** Writing – original draft, Writing – review & editing, Supervision, Conceptualization, Project administration, Funding acquisition. **Justine Bertrand-Michel:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Conceptualization, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Mass spectrometry analysis were done on MetaToul (Toulouse metabolomics & fluxomics facilities, www.metatoul.fr) which is part of the French National Infrastructure for Metabolomics and Fluxomics MetaboHUB-ANR-11-INBS-0010. We would like to thank Hanna Kulyk-Barbier for helpful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2021.339316.

References

[1] P. Lepage, M.C. Leclerc, M. Joossens, S. Mondot, H.M. Blottière, J. Raes, D. Ehrlich, J. Doré, A metagenomic Insight into our gut's microbiome, Gut 62 (1) (2013) 146–158, https://doi.org/10.1136/gutjnl-2011-301805.

[2] S.D. Allison, J.B.H. Martiny, Resistance, resilience, and redundancy in microbial communities, Proc. Natl. Acad. Sci. Unit. States Am. 105 (Supplement 1) (2008) 11512–11519, https://doi.org/10.1073/pnas.0801925105.

[3] A.E. Pérez-Cobas, A. Artacho, H. Knecht, M.L. Ferrús, A. Friedrichs, S.J. Ott, A. Moya, A. Latorre, M.J. Gosalbes, Differential effects of antibiotic therapy on the structure and function of human gut microbiota, PLoS One 8 (11) (2013), e80201, https://doi.org/10.1371/journal.pone.0080201.

[4] T. Pérez-Berezo, J. Pujo, P. Martin, P. Le Faouder, J.-M. Galano, A. Guy, C. Knauf, J.C. Tabet, S. Tronnet, F. Barreau, M. Heuillet, G. Dietrich, J. Bertrand-Michel, T. Durand, E. Oswald, N. Cenac, Identification of an analgesic lipopeptide produced by the probiotic Escherichia coli strain Nissle 1917, Nat. Commun. 8 (1) (2017) 1314, https://doi.org/10.1038/s41467-017-01403-9.

[5] J. Pujo, C. Petitfils, P. Le Faouder, V. Eeckhaut, G. Payros, S. Maurel, T. Perez-Berezo, M. Van Hul, F. Barreau, C. Blanpied, S. Chavanas, F. Van Immerseel, J. Bertrand-Michel, E. Oswald, C. Knauf, G. Dietrich, P.D. Cani, N. Cenac, Bacteria-derived long chain fatty acid exhibits anti-inflammatory properties in colitis, Gut 70 (6) (2021) 1088–1097, https://doi.org/10.1136/gutjnl-2020-321173.

[6] E. Pouteau, I. Meirim, S. Métairon, L.-B. Fay, Acetate, propionate and butyrate in plasma: determination of the concentration and isotopic enrichment by gas chromatography/mass spectrometry with positive chemical ionization: SCFA concentration and isotopic enrichment in plasma, J. Mass Spectrom. 36 (7) (2001) 798–805, https://doi.org/10.1002/jms.181.

[7] R. Berkecz, F. Tömösi, T. Körmöczi, V. Szegedi, J. Horváth, T. Janáky, Comprehensive phospholipid and sphingomyelin profiling of different brain regions in mouse model of anxiety disorder using online two-dimensional (HILIC/RP)-LC/MS method, J. Pharmaceut. Biomed. Anal. 149 (2018) 308–317, https://doi.org/10.1016/j.jpba.2017.10.043.

[8] B. Tan, D.K. O'Dell, Y.W. Yu, M.F. Monn, H.V. Hughes, S. Burstein, J.M. Walker, Identification of endogenous acyl amino acids based on a targeted lipidomics approach, J. Lipid Res. 51 (1) (2010) 112–119, https://doi.org/10.1194/jlr.M900198-JLR200.

[9] S.S. Andra, C. Austin, D. Patel, G. Dolios, M. Awawda, M. Arora, Trends in the application of high-resolution mass spectrometry for human biomonitoring: an analytical primer to studying the environmental chemical space of the human exposome, Environ. Int. 100 (2017) 32–61, https://doi.org/10.1016/j.envint.2016.11.026.

[10] E.L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H.P. Singer, J. Hollender, Identifying small molecules via high resolution mass spectrometry: communicating confidence, Environ. Sci. Technol. 48 (4) (2014) 2097–2098, https://doi.org/10.1021/es5002105.

[11] E.L. Schymanski, H.P. Singer, J. Slobodnik, I.M. Ipolyi, P. Oswald, M. Krauss, T. Schulze, P. Haglund, T. Letzel, S. Grosse, N.S. Thomaidis, A. Bletsou, C. Zwiener, M. Ibáñez, T. Portolés, R. de Boer, M.J. Reid, M. Onghena, U. Kunkel, W. Schulz, A. Guillon, N. Noyon, G. Leroy, P. Bados, S. Bogialli, D. Stipaničev, P. Rostkowski, J. Hollender, Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis, Anal. Bioanal. Chem. 407 (21) (2015) 6237–6255, https://doi.org/10.1007/s00216-015-8681-7.

[12] M.G.E. Guardian, P. He, A. Bermudez, S. Duan, S.S. Kaushal, E. Rosenfeldt, D.S. Aga, Optimized suspect screening approach for a comprehensive assessment of the impact of best management practices in reducing micropollutants transport in the potomac river watershed, Water Res. X 11 (2021) 100088, https://doi.org/10.1016/j.wroa.2021.100088.

[13] A. Wang, R.R. Gerona, J.M. Schwartz, T. Lin, M. Sirota, R. Morello-Frosch, T.J. Woodruff, A suspect screening method for characterizing multiple chemical exposures among a demographically diverse population of pregnant women in San Francisco, Environ. Health Perspect. 126 (7) (2018), 077009, https://doi.org/10.1289/EHP2920.

[14] J.M. Colby, K.L. Thoren, K.L. Lynch, Suspect screening using LC–QqTOF is a useful tool for detecting drugs in biological samples, J. Anal. Toxicol. 42 (4) (2018) 207–213, https://doi.org/10.1093/jat/bkx107.

[15] M. Pourchet, L. Debrauwer, J. Klanova, E.J. Price, A. Covaci, N. Caballero-Casero, H. Oberacher, M. Lamoree, A. Damont, F. Fenaille, J. Vlaanderen, J. Meijer, M. Krauss, D. Sarigiannis, R. Barouki, B. Le Bizec, J.-P. Antignac, Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues, Environ. Int. 139 (2020) 105545, https://doi.org/10.1016/j.envint.2020.105545.

[16] I. Liakh, T. Sledzinski, L. Kaska, P. Mozolewska, A. Mika, Sample preparation methods for lipidomics approaches used in studies of obesity, Molecules 25 (22) (2020) 5307, https://doi.org/10.3390/molecules25225307.

[17] P. Manirakiza, A. Covaci, P. Schepens, Comparative study on total lipid determination using soxhlet, roese-gottlieb, Bligh & dyer, and modified Bligh & dyer extraction methods, J. Food Compos. Anal. 14 (1) (2001) 93–100, https://doi.org/10.1006/jfca.2000.0972.

[18] C.C. Teo, W.P.K. Chong, E. Tan, N.B. Basri, Z.J. Low, Y.S. Ho, Advances in sample preparation and analytical techniques for lipidomics study of clinical samples, TrAC Trends Anal. Chem. (Reference Ed.) 66 (2015) 1–18, https://doi.org/10.1016/j.trac.2014.10.010.

[19] A. Dupuy, P. Le Faouder, C. Vigor, C. Oger, J.-M. Galano, C. Dray, J.C.-Y. Lee, P. Valet, C. Gladine, T. Durand, J. Bertrand-Michel, Simultaneous quantitative profiling of 20 isoprostanoids from omega-3 and omega-6 polyunsaturated fatty acids by LC–MS/MS in various biological samples, Anal. Chim. Acta 921 (2016) 46–58, https://doi.org/10.1016/j.aca.2016.03.024.

[20] K.A. Datsenko, B.L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products, Proc. Natl. Acad. Sci. Unit. States Am. 97 (12) (2000) 6640–6645, https://doi.org/10.1073/pnas.120163297.

[21] P. Le Faouder, V. Baillif, I. Spreadbury, J.-P. Motta, P. Rousset, G. Chêne, C. Guigné, F. Tercé, S. Vanner, N. Vergnolle, J. Bertrand-Michel, M. Dubourdeau, N. Cenac, LC–MS/MS method for rapid and concomitant quantification of pro-inflammatory and pro-resolving polyunsaturated fatty acid metabolites, J. Chromatogr. B 932 (2013) 123–133, https://doi.org/10.1016/j.jchromb.2013.06.014.

[22] J. Pujo, C. Petitfils, P. Le Faouder, V. Eeckhaut, G. Payros, S. Maurel, T. Perez-Berezo, M. Van Hul, F. Barreau, C. Blanpied, S. Chavanas, F. Van Immerseel, J. Bertrand-Michel, E. Oswald, C. Knauf, G. Dietrich, P.D. Cani, N. Cenac, Bacteria-derived long chain fatty acid exhibits anti-inflammatory properties in colitis, Gut 70 (6) (2021) 1088–1097, https://doi.org/10.1136/gutjnl-2020-321173.

[23] E.G. Bligh, W.J. Dyer, A rapid method of total lipid extraction and purification, Can. J. Biochem. Physiol. 37 (8) (1959) 911–917, https://doi.org/10.1139/o59-099.

[24] A.A. Zoerner, S. Batkai, M.-T. Suchy, F.-M. Gutzki, S. Engeli, J. Jordan, D. Tsikas, Simultaneous UPLC–MS/MS quantification of the endocannabinoids 2-arachidonoyl glycerol (2AG), 1-arachidonoyl glycerol (1AG), and anandamide in human plasma: minimization of matrix-effects, 2AG/1AG isomerization and degradation by toluene solvent extraction, J. Chromatogr. B 883–884 (2012) 161–171, https://doi.org/10.1016/j.jchromb.2011.06.025.

[25] M. Alajlani, A. Shiekh, S. Hasnain, A. Brantner, Purification of bioactive lipopeptides produced by Bacillus subtilis strain BIA, Chromatographia 79 (21–22) (2016) 1527–1532, https://doi.org/10.1007/s10337-016-3164-3.

[26] C. Garst, M. Fulmer, D. Thewke, S. Brown, Optimized extraction of 2-arachidonyl glycerol and anandamide from aortic tissue and plasma for quantification by LC-MS/MS: endocannabinoid extraction from tissue and plasma, Eur. J. Lipid Sci. Technol. 118 (5) (2016) 814–820, https://doi.org/10.1002/ejlt.201500115.

[27] V. Mutemberezi, J. Masquelier, O. Guillemot-Legris, G.G. Muccioli, Development and validation of an HPLC-MS method for the simultaneous quantification of key oxysterols, endocannabinoids, and ceramides: variations in metabolic syndrome, Anal. Bioanal. Chem. 408 (3) (2016) 733–745, https://doi.org/10.1007/s00216-015-9150-z.

[28] A. Dupuy, P. Le Faouder, C. Vigor, C. Oger, J.-M. Galano, C. Dray, J.C.-Y. Lee, P. Valet, C. Gladine, T. Durand, J. Bertrand-Michel, Simultaneous quantitative profiling of 20 isoprostanoids from omega-3 and omega-6 polyunsaturated fatty acids by LC–MS/MS in various biological samples, Anal. Chim. Acta 921 (2016) 46–58, https://doi.org/10.1016/j.aca.2016.03.024.

[29] J. Dalli, R.A. Colas, M.E. Walker, C.N. Serhan, Lipid mediator metabolomics via LC-MS/MS profiling and analysis, in: M. Giera (Ed.), Clinical Metabolomics, Methods in Molecular Biology, vol. 1730, Springer New York, New York, NY,

2018, pp. 59–72, https://doi.org/10.1007/978-1-4939-7592-1_4.
[30] H. Goldfine, Bacterial membranes and lipid packing theory, J. Lipid Res. 25 (13) (1984) 1501–1507, https://doi.org/10.1016/S0022-2275(20)34423-0.
[31] J. Pujo, C. Petitfils, P. Le Faouder, V. Eeckhaut, G. Payros, S. Maurel, T. Perez-Berezo, M. Van Hul, F. Barreau, C. Blanpied, S. Chavanas, F. Van Immerseel, J. Bertrand-Michel, E. Oswald, C. Knauf, G. Dietrich, P.D. Cani, N. Cenac, Bacteria-derived long chain fatty acid exhibits anti-inflammatory properties in colitis, Gut 70 (6) (2021) 1088–1097, https://doi.org/10.1136/gutjnl-2020-321173.
[32] E. Darii, Y. Gimbert, S. Alves, A. Damont, A. Perret, A.S. Woods, F. Fenaille, J.-C. Tabet, First direct evidence of interpartner hydride/deuteride exchanges for stored sodiated arginine/fructose-6-phosphate complex anions within salt-solvated structures, J. Am. Soc. Mass Spectrom. 32 (6) (2021) 1424–1440, https://doi.org/10.1021/jasms.1c00040.
[33] R.B. Cole (Ed.), Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications, second ed., Wiley, Hoboken, N.J, 2010.
[34] R.M. Salek, C. Steinbeck, M.R. Viant, R. Goodacre, W.B. Dunn, The role of reporting standards for metabolite annotation and identification in metabolomic studies, GigaScience 2 (1) (2013) 13, https://doi.org/10.1186/2047-217X-2-13.
[35] J.-P. Nougayrede, Escherichia coli induces DNA double-strand breaks in eukaryotic cells, Science 313 (5788) (2006) 848–851, https://doi.org/10.1126/science.1127059.
[36] X. Bian, J. Fu, A. Plaza, J. Herrmann, D. Pistorius, A.F. Stewart, Y. Zhang, R. Müller, In vivo evidence for a prodrug activation mechanism during colibactin maturation, Chembiochem 14 (10) (2013) 1194–1197, https://doi.org/10.1002/cbic.201300208.
[37] C.A. Brotherton, E.P. Balskus, A prodrug resistance mechanism is involved in colibactin biosynthesis and cytotoxicity, J. Am. Chem. Soc. 135 (9) (2013) 3359–3362, https://doi.org/10.1021/ja312154m.
[38] M.I. Vizcaino, P. Engel, E. Trautman, J.M. Crawford, Comparative metabolomics and structural characterizations illuminate colibactin pathway-dependent small molecules, J. Am. Chem. Soc. 136 (26) (2014) 9244–9247, https://doi.org/10.1021/ja503450q.
[39] F. Taieb, C. Petit, J.-P. Nougayrède, E. Oswald, The enterobacterial genotoxins: cytolethal distending toxin and colibactin, EcoSal Plus 7 (1) (2016), https://doi.org/10.1128/ecosalplus.ESP-0008-2016 ecosalplus.ESP-0008-2016.
[40] P. Martin, I. Marcq, G. Magistro, M. Penary, C. Garcie, D. Payros, M. Boury, M. Olier, J.-P. Nougayrède, M. Audebert, C. Chalut, S. Schubert, E. Oswald, Interplay between siderophores and colibactin genotoxin biosynthetic pathways in Escherichia coli, PLoS Pathog. 9 (7) (2013), e1003437, https://doi.org/10.1371/journal.ppat.1003437.

Table 1: Comparison of discovered aminolipids measured in bacterial sample and pure compounds: RT comparison, m/z with error ppm and %RSD.

		RT LC-HRMS (min)	%RSD	RT LC-LRMS (min)	%RSD	Elementary composition	m/z calculated $[M-H]^-$	m/z experimental (HRMS)	Erreur (ppm)
C12-ASN	Synthesized	3.72	0,56	5.89	0,36	$C_{16}H_{30}N_2O_4$	313.2132	313.2119	4.4
	Sample	3.75		5.86				313.2120	4.1
C12:1-ASN	Synthesized	5.70	0,12	4.81	0,14	$C_{16}H_{28}N_2O_4$	311.1976	311.1972	1.4
	Sample	5.69		4.80				311.1980	1.2
C13-ASN	Synthesized	5.42	1,16	7.04	0,40	$C_{17}H_{32}N_2O_4$	327.2289	327.2274	4.7
	Sample	5.51		7.00				327.2274	4.7
C14-ASN	Synthesized	6.60	0,63	8.18	0	$C_{18}H_{34}N_2O_4$	341.2445	341.2443	0.8
	Sample	6.66		8.18				341.2446	0.1
C14:1-ASN	Synthesized	7.68	0,27	6.79	0	$C_{18}H_{32}N_2O_4$	339.2289	339.2278	3.3
	Sample	7.65		6.79				339.2296	2.0
C15-ASN	Synthesized	8.69	0,24	9.34	0,07	$C_{19}H_{36}N_2O_4$	355.2602	355.2599	0.9
	Sample	8.66		9.33				355.2605	0.8
C16-ASN	Synthesized	7.27	0,88	10.51	0	$C_{20}H_{38}N_2O_4$	369.2758	369.2752	1.8

	Sample	7.18		10.51				369.2769	2.8
C16:1-ASN	Synthesized	9.66	0,14	8.91	0	$C_{20}H_{36}N_2O_4$	367.2602	367.2600	2.1
	Sample	9.64		8.91				367.2602	0.1
C17-ASN	Synthesized	8.14	0,77	11.67	0,06	$C_{21}H_{40}N_2O_4$	383.2915	383.2914	0.3
	Sample	8.23		11.68				383.2914	0.3
C17:1-ASN	Synthesized	8.18	0,25	9.99	0	$C_{21}H_{38}N_2O_4$	381.2758	381.2758	0.1
	Sample	8.15		9.99				381.2752	1.8
C12-ARG	Synthetized	2.30	0,61	3.09	0,45	$C_{18}H_{36}N_4O_3$	355.2714	355.2714	0.1
	Sample	2.32		3.07				355.2714	0.1
C12-ASP	Synthesized	5.48	0,25	6.83	0	$C_{16}H_{29}NO_5$	314.1972	314.1974	0.1
	Sample	5.46		6.83				314.1941	10.2
C12-GLN	Synthetized	4.63	0,30	5.86	0,36	$C_{17}H_{32}N_2O_4$	327.2289	327.2288	0.4
	Sample	4.61		5.83				327.2277	3.8
C12-GLU	Synthesized	5.42	0,39	6.84	0,10	$C_{17}H_{30}NO_5$	328.2129	328.2122	2.2
	Sample	5.45		6.83				328.2136	2.0
C14-GLU	Synthesized	7.36	0,28	9.10	0	$C_{19}H_{34}NO_5$	356.2422	356.2427	4.3

	Sample	7.39		9.10				356.2420	6.3
C16-GLU	Synthesized	9.33	0,07	11.37	0;06	$C_{21}H_{38}NO_5$	384.2755	384.2756	0.2
	Sample	9.32		11.38				384.2755	0.1
C12-HIS	Synthesized	2.19	1,59	2.96	0,47	$C_{18}H_{31}N_3O_3$	336.2292	336.2290	0.8
	Sample	2.24		2.94				336.2303	3.1
C12-LYS	Synthesized	2.12	0,99	2.89	0,49	$C_{18}H_{36}N_2O_3$	327.2653	327.2646	0.3
	Sample	2.15		2.87				327.2653	0.1
C14-LYS	Synthesized	3.57	0,98	4.57	0,76	$C_{20}H_{40}N_2O_3$	355.2966	355.2961	1.4
	Sample	3.62		4.62				355.2959	2.0
C12-LEU/LEU	Synthesized	8.73	0,24	8.68	0,16	$C_{18}H_{35}NO_3$	312.2544	312.2540	1.3
	Sample	8.76		8.66				312.2542	0.7
C14-PHE	Synthesized	10.84	0,32	13.00	0,05	$C_{23}H_{37}NO_3$	374.2700	374.2699	0.4
	Sample	10.79		12.99				374.2700	0.1
C12-TRP	Synthesized	8.26	0,51	10.04	0	$C_{23}H_{34}N_2O_3$	385.2496	385.2496	0.1
	Sample	8.20		10.04				385.2497	0.1
C12-VAL	Synthetized	7.95	0	9.81	0	$C_{17}H_{33}NO_3$	298.2387	298.2379	2.9

	Sample	7.95		9.81				298.2387	0.1

Table 2: SRM characterisation with quantitative and qualitative transition, energy applied to generate the transition (EC et Q1/Q3PreBias), Linear regression equation and correlation R2.

Compound	Linear regression	R^2	LOD ($pg.\mu L^{-1}$)	LOQ ($pg.\mu L^{-1}$)	Quantitative MRM transition	Q1 Pre Bias (V)	EC (eV)	Q3 Pre Bias (V)	Qualitative MRM transition
C12:1Asn	y = 49290x - 173866	0.9990	0.39	1.56	312 → 293	30	15	13	311 → 183
C13Asn	y = 63564x - 222566	0.9984	0.19	1.56	327 → 309	34	18	15	327 → 212
C14:1Asn	y = 132014x - 110079	0.9977	0.19	0.78	339 → 321	14	17	22	339 → 224
C15Asn	y = 25510x - 65143	0.9996	0.39	1.56	355 → 240	25	24	16	355 → 337
C16Asn	y = 64907x - 151246	0.9990	0.19	0.39	369 → 351	15	20	16	369 → 254
C16:1Asn	y = 53376x - 199389	0.9995	0.78	1.56	367 → 349	28	20	26	367 → 252
C17Asn	y = 137315x - 20177	0.9978	0.39	0.78	383 → 365	15	20	18	383 → 96
C17:1Asn	y =40608x - 117983	0.9989	0.19	0.78	381 → 363	14	25	10	381 → 266
C12Arg	y =15767x - 105656	0.9988	0.04	0.78	355 → 131	20	26	20	355 → 131

C12Asp	y = 44937x - 317123	0.9950	0.19	0.78	314 → 198	24	23	21	314 → 199
C12Gln	y = 22434x - 40015	0.9985	0.19	0.39	327 → 198	24	28	23	327 → 127
C12Glu	y = 27518x - 86993	0.9991	0.19	1.56	328 → 198	10	20	20	328 → 128
C14Glu	y = 49107x - 214246	0.9983	0.19	0.78	356 → 226	26	22	24	356 → 128
C16Glu	y =1262x - 19015	0.9943	0.19	0.78	384 → 255	30	23	8	384 → 128
C12His	y = 30227x - 77911	0.9987	0.39	0.78	336 → 110	25	27	10	336 → 154
C12Leu	y = 181029x - 418227	0.9990	0.04	1.56	313 → 130	13	29	26	312 → 268
C12Ile	y = 236326x - 425633	0.9985	0.04	1.56	313 → 130	12	29	26	312 → 268
C12Lys	y = 45468x - 202933	0.9976	0.19	1.56	327 → 145	13	25	27	327 → 283
C14Lys	y =70865x - 401950	0.9988	0.19	0.78	355 → 145	14	25	27	355 → 311
C14Phe	y = 81754x - 29953	0.9981	0.19	0.78	374 → 164	15	24	29	374 → 226
C12Pro	y = 77314x - 52780	0.9975	0.19	1.56	296 → 114	21	21	11	296 → 198
C12Trp	y = 46967x - 173680	0.9987	0.19	1.56	385 → 203	14	21	20	285 → 256
C12Val	y = 180951x - 433184	0.9988	0.04	0.39	298 → 116	22	22	12	298 → 254

Supplementary Table S1:

Yield of extraction (%) for C12-Asn, C14-Asn, C14:1-Asn, C14-AsnGABA, IS in different sample preparation conditions.

	C12-Asn	**C14-Asn**	**C14 1-Asn**	**C14-AsnGABA**	**IS**
Crush : Tris/MeOH (2/1) **Solubilisation :** MeOH/H_2O (85/15) **SPE HLB** **Elution :** 1 mL AcN, MeOH, AcOEt	56 ±3	29±2	68±4	56±3	43±2
Crush: MeOH/H_2O (85/15) **SPE HLB** **Elution:** 1 mL MeOH $CHCl_2$	63±6	12±8	33±3	31±3	1±1
Crush: MeOH/H_2O (85/15) **SPE HLB** **Elution:** 1 mL AcN MeOH $CHCl_2$	47±14	6±6	68±17	31±2	0.5±0.5
Crush: Tris/MeOH (2/1) **LLE** **Washing**: heptane **Elution**: AcOEt/H_2O	76±5	14±5	24±6	33±5	20±0

Supplementary Table S2: Accurate experimental m/z ratios from the main product ions[a],[b],[c] for deprotonated LpAA molecules display in low collision energy product ion spectra under (E_{LAB} = 20eV and high resolution conditions, using Q/TOF MS) and corresponding elemental composition of product ions and associated neutral losses. Experimental m/z deviation (giving in ppm) for all product ions, by comparison with calculated m/z ratios for the proposed elemental composition.

Precursor [LpAA-H]$^-$ ions selected from ESI source	Product ions					Formal neutral losses
	m/z values (experimental)	Relative abundances (%)	m/z values (calculated)	Error (ppm)	Elemental compositions	
C12:1Asn MW : 312.2049 [M-H]$^-$: 311.1972 (1.4 ppm ; 4.4%) $C_{16}H_{27}N_2O_4$	293.1870	56.4	293.1870	0.2	$C_{16}H_{25}N_2O_3$	H_2O
	267.2074	[b]	267.2078	1.5	$C_{15}H_{27}N_2O_2$	CO_2
	249.1972	[b]	249.1972	0.1	$C_{15}H_{25}N_2O$	(H_2O + CO_2)
	222.1860	3.9	222.1863	1.5	$C_{14}H_{24}NO$	(CO_2 + H_2O + HCN)
	196.1709	15.7	196.1706	1.1	$C_{12}H_{22}NO$[d]	$C_4H_5NO_3$
	131.0460	8.4	131.0456	1.6	$C_4H_7N_2O_3$ [f]	$C_{14}H_{20}O$
	114.0195	8.0	114.0191	1.5	$C_4H_4NO_3$	$C_{14}H_{23}NO$
	96.0090	3.1	96.0086	1.1	$C_4H_2NO_2$	$C_{14}H_{25}NO_2$
C12Asn[h] MW : 314.2205 [M-H]$^-$: 313.2119 (4.4 ppm ; 3.4%) $C_{16}H_{29}N_2O_4$	295.2023	44.2	295.2027	1.4	$C_{16}H_{27}N_2O_3$	H_2O
	269.2219	[a]	269.2234	5.7	$C_{15}H_{29}N_2O_2$	CO_2
	251.2113	[a]	251.2128	6.3	$C_{15}H_{27}N_2O$	(H_2O + CO_2)
	224.2006	[a]	224.2019	6.2	$C_{14}H_{26}NO$	(CO_2 + H_2O + HCN)
	198.1855	34.6	198.1863	2.5	$C_{12}H_{24}NO$[d]	$C_4H_5NO_3$
	131.0452	5.0	131.0456	7.5	$C_4H_7N_2O_3$[f]	$C_{12}H_{22}O$
	114.0188	5.1	114.0191	7.6	$C_4H_4NO_3$	$C_{12}H_{25}NO$
	96.0093	[a]	96.0086	2.1	$C_4H_2NO_2$	$C_{12}H_{27}NO_2$
C13Asn MW : 328.2362 [M-H]$^-$: 327.2274 (4.7 ppm ; 4.3%) $C_{17}H_{31}N_2O_4$	309.2191	56.8	309.2184	2.4	$C_{17}H_{29}N_2O_3$	H_2O
	283.2376	[a]	283.2376	0.1	$C_{16}H_{31}N_2O_2$	CO_2
	265.2285	[a]	265.2285	0.1	$C_{16}H_{29}N_2O$	(H_2O + CO_2)
	238.2173	[b]	238.2176	1.4	$C_{15}H_{28}NO$	(CO_2 + H_2O + HCN)
	212.2008	23.5	212.2019	5.6	$C_{13}H_{26}NO$[d]	$C_4H_5NO_3$
	131.0454	3.8	131.0456	6.2	$C_4H_7N_2O_3$[f]	$C_{13}H_{24}O$
	114.0196	4.1	114.0191	0.6	$C_4H_4NO_3$	$C_{13}H_{27}NO$

	113.0345	[a]	113.0356	2.2	$C_4H_5N_2O_2$	$C_{13}H_{26}O_2$
	96.0089	[a]	96.0086	2.1	$C_4H_2NO_2$	$C_{13}H_{29}NO_2$
C14Asn[h] MW : 342.2518 $[M-H]^-$: 341.2443 (0.8 ppm ; 7.7%) $C_{18}H_{33}N_2O_4$	323.2343	50.5	323.2340	0.6	$C_{18}H_{31}N_2O_3$	H_2O
	297.2536	[a]	297.2547	3.9	$C_{17}H_{33}N_2O_2$	CO_2
	279.2440	[a]	279.2441	0.7	$C_{17}H_{31}N_2O$	$(H_2O + CO_2)$
	252.2325	[a]	252.2332	3.1	$C_{16}H_{30}NO$	$(CO_2 + H_2O + HCN)$
	226.2174	27.6	226.2176	1.1	$C_{14}H_{28}NO$[d]	$C_4H_5NO_3$
	131.0453	4.0	131.0456	7.0	$C_4H_7N_2O_3$[f]	$C_{14}H_{26}O$
	114.0191	3.9	114.0191	5.0	$C_4H_4NO_3$	$C_{14}H_{29}NO$
	96.0086	[a]	96.0086	8.2	$C_4H_2NO_2$	$C_{14}H_{31}NO_2$
C14:1ASN MW : 340.236 $[M-H]^-$: 339.2278 (3.3 ppm ; 8.4%) $C_{18}H_{31}N_2O_4$	321.2175	60.5	321.2183	2.7	$C_{18}H_{29}N_2O_3$	H_2O
	295.2376	[b]	295.2391	5.0	$C_{17}H_{31}N_2O_2$	CO_2
	277.2279	[a]	277.2285	2.2	$C_{17}H_{30}N_2O$	$(H_2O + CO_2)$
	250.2159	[b]	250.2176	6.9	$C_{16}H_{28}NO$	$(CO_2 + H_2O + HCN)$
	224.2020	19.2	224.2019	0.1	$C_{14}H_{28}NO$[d]	$C_4H_5NO_3$
	131.0462	6.2	131.0456	0.1	$C_4H_7N_2O_3$[f]	$C_{14}H_{24}O$
	114.0196	5.7	114.0191	0.6	$C_4H_4NO_3$	$C_{14}H_{27}NO$
	96.0080	[b]	96.0086	2.5	$C_4H_2NO_2$	$C_{14}H_{29}NO_2$
C15Asn MW : 356.2675 $[M-H]^-$: 355.2599 (0.9 ppm ; 13%) $C_{19}H_{35}N_2O_4$	337.2496	53.2	337.2496	0.2	$C_{19}H_{33}N_2O_3$	H_2O
	311.2697	[b]	311.2704	2.3	$C_{18}H_{35}N_2O_2$	CO_2
	293.2593	[b]	293.2598	1.8	$C_{18}H_{33}N_2O$	$(H_2O + CO_2)$
	266.2482	[b]	266.2489	2.8	$C_{17}H_{32}NO$	$(CO_2 + H_2O + HCN)$
	240.2333	24.4	240.2338	0.1	$C_{15}H_{30}NO$[d]	$C_4H_5NO_3$
	131.0454	4.7	131.0456	6.2	$C_4H_7N_2O_3$[f]	$C_{15}H_{28}NO$
	114.0196	4.6	114.0191	0.6	$C_4H_4NO_3$	$C_{15}H_{31}NO$
	96.0088	[b]	96.0086	3.1	$C_4H_2NO_2$	$C_{15}H_{33}NO_2$
C16Asn MW : 370.2831 $[M-H]^-$: 369.2752 (1.8 ppm ; 18.3%) $C_{20}H_{37}N_2O_3$	351.2648	52.8	351.2653	1.4	$C_{20}H_{37}N_2O_3$	H_2O
	325.2851	[b]	325.2860	2.9	$C_{19}H_{37}N_2O_2$	CO_2
	307.2751	[b]	307.2754	1.3	$C_{19}H_{35}N_2O$	$(H_2O + CO_2)$
	280.2639	[b]	280.2645	2.5	$C_{18}H_{34}NO$	$(CO_2 + H_2O + HCN)$
	254.2485	20.5	254.2489	1.7	$C_{16}H_{32}NO$[d]	$C_4H_5NO_3$

	131.0455	4.2	131.0456	6.2	$C_4H_7N_2O_3$[f]	$C_{16}H_{30}O$
	114.0191	4.1	114.0191	0.2	$C_4H_4NO_3$	$C_{16}H_{33}NO$
	96.0089	[b]	96.0086	1.1	$C_4H_2NO_2$	$C_{16}H_{35}NO_2$
C16:1Asn MW : 368.2675 $[M-H]^-$: 367.2600 (2.1 ppm ; 29.8%) $C_{20}H_{35}N_2O_3$	349.2498	53.2	349.2496	0.4	$C_{20}H_{33}N_2O_3$	H_2O
	323.2701	5.3	323.2704	0.9	$C_{19}H_{35}N_2O_2$	CO_2
	305.2592	[b]	305.2598	2.1	$C_{19}H_{33}N_2O$	$(H_2O + CO_2)$
	278.2489	[b]	278.2489	0.1	$C_{18}H_{32}NO$	$(CO_2 + H_2O + HCN)$
	252.2327	6.0	252.2328	2.3	$C_{16}H_{30}NO$[d]	$C_4H_5NO_3$
	131.0457	3.6	131.0456	3.9	$C_4H_7N_2O_3$[f]	$C_{16}H_{28}O$
	114.0194	[b]	114.0191	2.3	$C_4H_4NO_3$	$C_{16}H_{31}NO$
	96.0089	[b]	96.0086	2.1	$C_4H_2NO_2$	$C_{16}H_{33}NO_2$
C17:1Asn MW : 382.2831 $[M-H]^-$: 381.2758 (0.2 ppm ; 20.5%) $C_{21}H_{37}N_2O_4$	363.2645	54.3	363.2653	2.2	$C_{21}H_{35}N_2O_3$	H_2O
	337.2848	3.2	337.2860	3.7	$C_{20}H_{37}N_2O_2$	CO_2
	307.2747	[b]	307.2754	2.6	$C_{19}H_{35}N_2O$	$(H_2O + CO_2)$
	292.2628	[b]	292.2645	6.1	$C_{19}H_{34}NO$	$(CO_2 + H_2O + HCN)$
	266.2480	15.0	266.2489	3.5	$C_{17}H_{32}NO$[d]	$C_4H_5NO_3$
	131.0451	3.8	131.0456	8.5	$C_4H_7N_2O_3$[f]	$C_{17}H_{30}NO$
	114.0185	3.2	114.0191	10.2	$C_4H_4NO_3$	$C_{17}H_{33}NO$
	96.0087	[b]	96.0086	0.3	$C_4H_2NO_2$	$C_{17}H_{35}NO_2$
C17Asn MW : 384.2988 $[M-H]^-$: 383.2914 (0.3 ppm ; 23.5%) $C_{21}H_{39}N_2O_4$	365.2808	50.6	365.2809	0.5	$C_{21}H_{37}N_2O_3$	H_2O
	339.3013	3.6	339.3017	1.2	$C_{20}H_{39}N_2O_2$	CO_2
	321.2908	[b]	321.2911	1.0	$C_{20}H_{37}N_2O$	$(H_2O + CO_2)$
	294.2798	[b]	294.2802	1.5	$C_{19}H_{36}NO$	$(CO_2 + H_2O + HCN)$
	268.2646	15.2	268.2645	0.1	$C_{17}H_{34}NO$[d]	$C_4H_5NO_3$
	131.0460	3.7	131.0456	1.6	$C_4H_7N_2O_3$[f]	$C_{17}H_{32}NO$
	114.0197	3.4	114.0191	0.3	$C_4H_4NO_3$	$C_{17}H_{35}NO$
	96.0092	[b]	96.0086	1.0	$C_4H_2NO_2$	$C_{19}H_{37}NO_2$
C12Arg MW : 356.2787 $[M-H]^-$: 355.2715 (0.1 ppm ; 21.2%)	338.2441	(a)	338.2449	2.4	$C_{18}H_{32}N_3O_3$	NH_3
	313.2483	34.5	313.2496	0.8	$C_{17}H_{33}N_2O_3$	(CH_2N_2)
	311.2816	[b]	311.2816	0.1	$C_{17}H_{35}N_4O$	CO_2
	269.2568	14.2	269.2598	7.2	$C_{16}H_{33}N_2O$	$C_2H_2N_2O_2$

$C_{18}H_{35}N_4O_3$	252.2330	[b]	252.2332	1.1	$C_{16}H_{30}NO$	$C_2H_5N_3O_2$
	198.1845	[b]	198.1863	2.5	$C_{12}H_{24}NO$[d]	$C_6H_{11}N_3O_2$
	173.1022	16.7	173.1043	6.4	$C_6H_{13}N_4O_2$[f]	$C_{12}H_{22}O$
	156.0769	[b]	156.0778	6.0	$C_6H_{10}N_3O_2$	$C_{12}H_{25}NO$
	131.0806	13.5	131.0826	6.9	$C_5H_{10}N_2O_2$	$C_{13}H_{25}N_2O$
C12Asp MW : 315.2045 $[M-H]^-$: 314.1971 (0.1 ppm ; 5.0%) $C_{16}H_{28}NO_5$	296.1842	[b]	296.1867	8.7	$C_{16}H_{26}NO_4$	H_2O
	270.2060	16.5	270.2074	5.4	$C_{15}H_{28}NO_3$	CO_2
	252.1946	[b]	252.1969	9.2	$C_{15}H_{26}NO_2$	$(H_2O + CO_2)$
	199.1706	27.4	199.1703	1.2	$C_{12}H_{23}O_2$	$C_4H_5NO_3$
	198.1853	37.8	198.1863	5.2	$C_{12}H_{24}NO$[d]	$C_4H_4O_4$
	132.0298	[b]	132.0302	3.2	$C_4H_6NO_4$[f]	$C_{12}H_{22}O$
	115.0031	4.5	115.0036	5.1	$C_4H_3O_4$	$C_{12}H_{25}NO$
	88.0401	8.7	88.0404	3.4	$C_3H_6NO_2$	$C_{13}H_{22}O_3$
C12Gln MW : 328.2362 $[M-H]^-$: 327.2288 (0.4 ppm ; 21.7%) $C_{17}H_{31}N_2O_4$	309.2188	4.0	309.2183	1.4	$C_{17}H_{29}N_2O_3$	H_2O
	283.2388	4.0	283.2391	1.1	$C_{16}H_{31}N_2O_2$	CO_2
	265.2283	[b]	265.2285	0.3	$C_{16}H_{29}N_2O$	$(H_2O + CO_2)$
	238.2173	[b]	238.2176	1.4	$C_{15}H_{28}NO$	$(CO_2 + H_2O + HCN)$
	198.1859	21.6	198.1863	2.2	$C_{12}H_{24}NO$[d]	$C_5H_7NO_3$
	145.0615	28.5	145.0618	2.5	$C_5H_9N_2O_3$[f]	$C_{12}H_{22}O$
	127.0513	20.2	127.0513	0.1	$C_5H_7N_2O_2$	$C_{12}H_{24}O_2$
C12Glu MW : 329.2202 $[M-H]^-$: 328.2122 (2.2 ppm ; 6.0%) $C_{17}H_{30}NO_5$	310.2023	[b]	310.2023	0.3	$C_{17}H_{28}NO_4$	H_2O
	284.2231	17.3	284.2231	0.2	$C_{16}H_{30}NO_3$	CO_2
	266.2122	[b]	266.2125	1.3	$C_{16}H_{28}NO_2$	$(H_2O + CO_2)$
	240.2333	12.8	240.2332	0.1	$C_{15}H_{30}NO$	(C_2O_4)
	199.1702	31.2	199.1703	0.8	$C_{12}H_{23}O_2$[e]	$C_5H_7NO_3$
	198.1854	3.5	198.1863	4.7	$C_{12}H_{24}NO$[d]	$C_5H_6O_4$
	146.0459	3.8	146.0458	0.1	$C_5H_8NO_4$[f]	$C_{12}H_{22}O$
	128.0359	18.3	128.0353	4.5	$C_5H_6NO_3$	$C_{12}H_{24}NO_2$
	102.0554	7.1	102.0555	0.1	$C_4H_8NO_2$	$C_{13}H_{22}O_3$
C14Glu MW : 357.2515	338.2332	3.2	338.2336	1.4	$C_{19}H_{32}NO_4$	H_2O
	312.2539	22.0	312.2544	1.7	$C_{18}H_{34}NO_3$	CO_2

$[M-H]^-$: 356.2427 (4.3 ppm ; 13.0%) $C_{19}H_{34}NO_5$	294.2434	[b]	294.2438	1.5	$C_{18}H_{32}NO_2$	$(H_2O + CO_2)$
	268.2647	8.0	268.2645	0.4	$C_{17}H_{34}NO$	$(CO_2 + CO_2)$
	227.2016	33.1	227.2016	0.2	$C_{14}H_{26}O_2$[e]	$C_5H_8NO_3$
	226.2162	[b]	226.2176	6.3	$C_{14}H_{28}NO$[d]	$C_5H_6NO_4$
	146.0455	[a]	146.0458	2.6	$C_5H_8NO_4$[f]	$C_{14}H_{26}O$
	128.0350	14.3	128.0353	2.5	$C_5H_6NO_3$	$C_{14}H_{28}O_2$
	102.0557	3.4	102.0555	0.1	$C_4H_8NO_2$	$C_{15}H_{26}O_3$
C16Glu MW : 385.2828 $[M-H]^-$: 384.2756 (0.2 ppm ; 25.6%) $C_{21}H_{38}NO_5$	366.2664	3.7	366.2649	3.8	$C_{21}H_{36}NO_4$	H_2O
	340.2848	19.9	340.2857	2.7	$C_{20}H_{38}NO_3$	CO_2
	322.2749	[b]	322.2751	0.8	$C_{20}H_{36}NO_2$	$(H_2O + CO_2)$
	296.2957	3.7	296.2958	0.6	$C_{19}H_{38}NO$	$(CO_2 + CO_2)$
	255.2321	33.9	255.2329	3.3	$C_{16}H_{31}O_2$[e]	$C_5H_7O_3N$
	254.2488	[b]	254.2489	1.0	$C_{16}H_{32}NO$[d]	$C_5H_4O_4$
	146.0455	13.3	146.0458	2.6	$C_5H_8NO_4$[f]	$C_{16}H_{30}O$
	128.0350	[b]	128.0353	2.4	$C_5H_6NO_3$	$C_{16}H_{32}O_2$
	102.0568	[b]	102.0555	7.3	$C_4H_8NO_2$	$C_{17}H_{30}O_3$
C12His MW : 337.2365 $[M-H]^-$: 336.2310 (5.5 ppm ; 27.2%) $C_{18}H_{30}N_3O_3$	318.2190	[b]	318.2187	0.9	$C_{18}H_{28}N_3O_2$	H_2O
	292.2409	11.8	292.2394	5.0	$C_{17}H_{30}N_3O$	CO_2
	274.2302	7.7	274.2288	5.5	$C_{17}H_{28}N_3$	$(H_2O + CO_2)$
	247.2185	[b]	247.2179	2.1	$C_{16}H_{27}N_2$	$(CO_2 + H_2O + HCN)$
	199.1712	5.7	199.1703	4.2	$C_{12}H_{23}O_2$	$C_6H_7N_3O$
	198.1862	[b]	198.1863	0.1	$C_{12}H_{24}NO$[d]	$C_6H_6N_2O_2$
	154.0624	32.3	154.0621	1.3	$C_6H_8N_3O_2$[f]	$C_{12}H_{22}O$
	137.0356	[b]	137.0356	1.0	$C_6H_5N_2O_2$	$C_{12}H_{25}NO$
	110.0721	15.1	110.0723	2.5	$C_5H_8N_3$	$C_{13}H_{22}O_3$
C12Leu MW : 313.2616 $[M-H]^-$: 312.2531 (4.2 ppm ; 19.1%) $C_{18}H_{34}NO_3$	294.2417	[b]	294.2438	7.3	$C_{18}H_{32}NO_2$	H_2O
	268.2628	13.8	268.2645	1.3	$C_{17}H_{34}NO$	CO_2
	198.1847	[a]	198.1863	1.3	$C_{12}H_{24}NO$[d]	$C_6H_{10}O_2$
	130.0873	65.4	130.0873	6.5	$C_6H_{12}NO_2$[f]	$C_{12}H_{22}O$
	294.2443	[b]	294.2438	1.5	$C_{18}H_{32}NO_2$	H_2O

C12Ile MW : 313.2616 $[M-H]^-$: 312.2540 (1.3 ppm ; 5.1%) $C_{18}H_{34}NO_3$	268.2642	23.2	268.2645	1.4	$C_{17}H_{34}NO$	CO_2
	198.1862	3.4	198.1863	0.3	$C_{12}H_{24}NO^{(d)}$	$C_6H_{10}O_2$
	130.0880	68.1	130.0873	5.0	$C_6H_{12}NO_2^{(f)}$	$C_{12}H_{22}O$
C12Lys MW : 328.2725 $[M-H]^-$: 327.2645 (2.5 ppm ; 18.0%) $C_{18}H_{35}N_2O_3$	310.XXXX	(c)	310.2387		$C_{18}H_{32}NO_3$	NH_3
	309.2552	(b)	309.2547	1.4	$C_{18}H_{33}N_2O_2$	H_2O
	283.2747	22.5	283.2754	2.8	$C_{17}H_{35}N_2O$	CO_2
	198.1836	(a)	198.1863	1.3	$C_{12}H_{24}NO^{(d)}$	$C_6H_{11}NO_2$
	145.0984	57.0	145.0982	1.0	$C_6H_{13}N_2O_2^{(f)}$	$C_{12}H_{22}O$
C14Lys MW : 356.3038 $[M-H]^-$: 355.2959 (2.0 ppm ; 38.6%) $C_{20}H_{39}N_2O_3$	338.XXXX	(c)	338.2700		$C_{20}H_{36}NO_3$	NH_3
	337.2856	(b)	337.2860	1.3	$C_{20}H_{37}N_2O_2$	H_2O
	311.3059	16.7	311.3068	2.8	$C_{19}H_{39}N_2O$	CO_2
	226.2175	(a)	226.2176	0.2	$C_{14}H_{28}NO^{(d)}$	$C_6H_{11}NO_2$
	145.0980	44.7	145.0982	1.7	$C_6H_{13}N_2O_2^{(f)}$	$C_{14}H_{26}O$
C14Phe MW : 375.277 $[M-H]^-$: 374.2697 (1.9 ppm ; 41.6%) $C_{23}H_{36}NO_3$	356.2577	(b)	356.2595	5.0	$C_{23}H_{34}NO_2$	H_2O
	330.2798	(b)	330.2802	1.3	$C_{22}H_{36}NO$	CO_2
	238.2171	4.3	238.2173	2.3	$C_{15}H_{28}NO$	$C_8H_8O_2$
	226.2164	12.3	226.2176	0.6	$C_{14}H_{28}NO^{(d)}$	$C_9H_8O_2$
	164.0718	35.8	164.0717	0.3	$C_9H_{10}NO_2^{(f)}$	$C_{14}H_{26}O$
	147.0451	5.9	147.0451	0.1	$C_9H_7O_2$	$C_{14}H_{29}NO$
	91.0554	(b)	91.0553	1.2	C_7H_7	$C_{16}H_{29}$
C12Trp MW : 386.2569 $[M-H]^-$: 385.2493 (1.9 ppm ; 30.6%) $C_{23}H_{33}N_2O_3$	367.2382	(b)	367.2391	2.5	$C_{23}H_{31}N_2O_2$	H_2O
	341.2587	7.9	341.2598	0.2	$C_{22}H_{33}N_2O$	CO_2
	323.2477	(b)	323.2492	4.9	$C_{22}H_{31}N_2$	$(H_2O + CO_2)$
	256.1908	19.4	256.1918	0.3	$C_{14}H_{26}NO_3$	C_9H_7N
	238.1812	6.7	238.1812	4.4	$C_{14}H_{24}NO_2$	C_9H_9NO
	212.2008	4.8	212.2019	0.4	$C_{13}H_{26}NO^{(d)}$	$C_{10}H_7NO_2$
	203.0816	30.6	203.0826	0.4	$C_{11}H_{11}N_2O_2^{(f)}$	$C_{12}H_{22}O$
	159.0915	(b)	159.0927	0.2	$C_{10}H_{11}N_2$	$C_{13}H_{22}O_3$
	116.0494	(b)	116.0505	0.1	C_8H_6N	$C_{15}H_{27}N$

	74.0242	[b]	74.0247	0.5	$C_2H_4NO_2$	$C_{21}H_{29}NO$
C12Val MW : 299.2460 $[M-H]^-$: 298.2379 (2.9 ppm ; 23.4%) $C_{23}H_{33}N_2O_3$	280.2281	[b]	280.2282	0.4	$C_{17}H_{30}NO_2$	H_2O
	254.2482	25.7	254.2489	1.4	$C_{16}H_{32}NO$	CO_2
	236.2375	[b]	236.2383	3.7	$C_{16}H_{30}N$	$(H_2O + CO_2)$
	198.1856	[b]	198.1863	3.7	$C_{12}H_{24}NO$[d]	$C_5H_8O_2$
	116.0712	51.0	116.0717	4.3	$C_5H_{10}NO_2$[f]	$C_{12}H_{22}O$

[a] Detected product ion, with relative abundance lower than 3% at E_{LAB} = 20eV but may be higher than this abundance limit at different E_{LAB} values.

[b] Detected product ion with relative abundance lower than 3% of relative abundance.

[c] Possible product ion never detected.

[d] Deprotonated fatty acid amines with the side chain size related to the LpAA considered

[e] Deprotonated fatty acid with the side chain size related to the LpAA considered

[f] Deprotonated amino acid related to the LpAA considered

[g] Fragment ions (without anticipating their origin) of the amino acid corresponding of LpAA

[h] Deprotonated LpAA already described from Pujo et al, nature com. but recorded in other conditions (HCD mode using quadrupole/orbitrap FT/MS instrument)

Supplementary Table 3: Quantification of LpAA in EcN and L. animalis. Data are represented as the mean ± SEM in pg/mg of proteins. ND: not detectable, NQ: Not Quantifiable

	EcN WT	EcN DclbA	EcN DclbN	EcN Dclbp	L. animalis
C12Arg	ND	ND	ND	ND	NQ
C16Glu	ND	ND	ND	ND	NQ
C16PheGABA	ND	ND	ND	ND	ND
C12Lys	ND	0.03 ± 0.01	0.22 ± 0.12	0.19 ± 0.09	0.15 ± 0.09
C12His	ND	0.12 ± 0.05	0.14 ± 0.06	0.76 ± 0.29	3.083 ± 1.23
C14Lys	ND	0.04 ± 0.02	0.08 ± 0.03	0.07 ± 0.04	0.07 ± 0.03
C12:1Asn	ND	ND	ND	ND	ND
C12AsnAABA	ND	ND	ND	ND	ND
C12AsnBABA	0.13 ± 0.01	0.019 ± 0.02	ND	20.4 ± 7.93	ND
C12AsnGABA	0.12 ± 0.005	0.003 ± 0.001	ND	1.16 ± 0.47	ND
C12Asn	4.55 ± 1.88	2.74 ± 1.43	ND	0.81 ± 0.29	ND
C12Gln	ND	ND	ND	ND	0.61 ± 0.04
C16GluGABA	ND	ND	ND	ND	ND
C12Asp	0.07 ± 0.01	0.004 ± 0.002	0.001 ± 0.0009	0.004 ± 0.003	0.23 ± 0.11
C14:1Asn	3.22 ± 0.44	0.75 ± 0.09	ND	0.11 ± 0.04	0.03 ± 0.01
C13Asn	0.26 ± 0.03	0.67 ± 0.12	ND	0.14 ± 0.05	0.002 ± 0.001
C13AsnDopamine	ND	ND	ND	ND	ND
C12Glu	0.03 ± 0.007	ND	ND	ND	ND
C14AsnGABA	ND	0.005 ± 0.002	0.06 ± 0.04	0.03 ± 0.01	ND
C14:1AlaGABA	ND	ND	ND	ND	ND
C14Asn	69.42 ± 4.23	41.49 ± 0.96	0.006 ± 0.003	12.00 ± 4.33	ND
C12ValGABA	ND	ND	ND	ND	ND
C16:1Asn	1.89 ± 0.20	0.39 ± 0.09	ND	0.07 ± 0.02	0.23 ± 0.11
C12MetDopamine	ND	ND	ND	ND	ND
C12LeuGABA	ND	ND	ND	ND	ND
C15Asn	47.09 ± 3.92	59.66 ± 5.76	ND	11.64 ± 4.49	1.19 ± 0.61

C12Pro	ND	ND	ND	ND	NQ
C14Glu	ND	ND	ND	ND	NQ
C17:1Asn	0.29 ± 0.04	0.04 ± 0.01	ND	0.04 ± 0.01	ND
C12Trp	ND	ND	ND	ND	0.05 ± 0.02
C12Val	ND	ND	ND	ND	0.04 ± 0.02
C14GABA	ND	ND	ND	ND	ND
C12Leu	ND	ND	ND	ND	0.24 ± 0.09
C16Asn	0.58 ± 0.05	0.17 ± 0.02	0.05 ± 0.02	0.06 ± 0.02	0.02 ± 0.01
C14IleGABA	ND	ND	ND	ND	ND
C17Asn	ND	0.011 ± 0.004	0.017 ± 0.007	0.019 ± 0.007	0.019 ± 0.009
C14Phe	ND	0.006 ± 0.002	0.007 ± 0.003	0.009 ± 0.003	0.03 ± 0.01
C16LeuGABA	ND	ND	ND	ND	ND
C12AlaGABA	ND	ND	ND	ND	ND

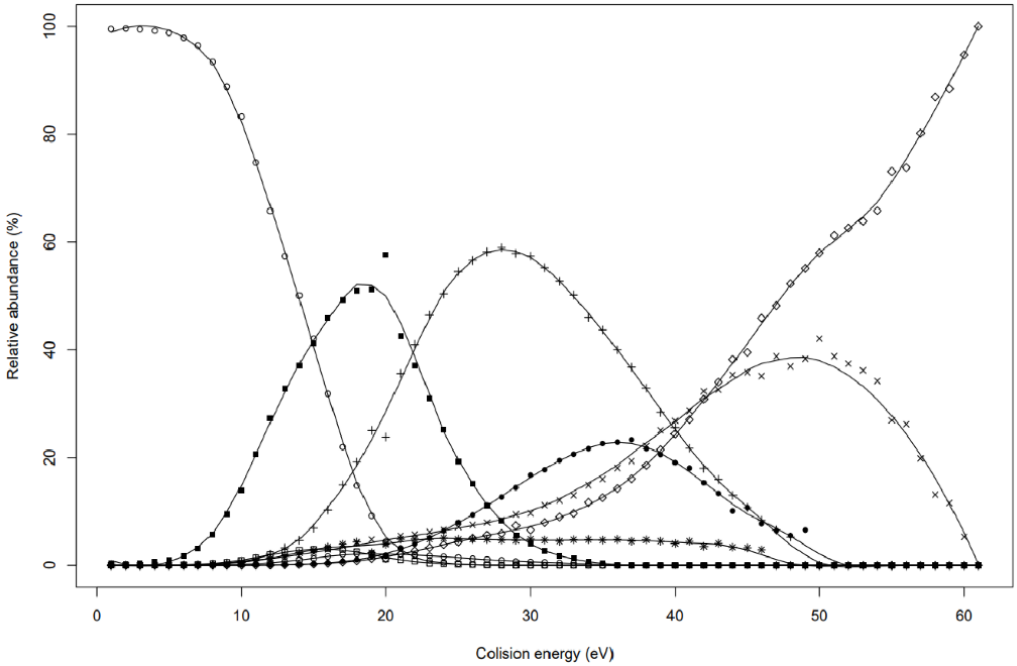
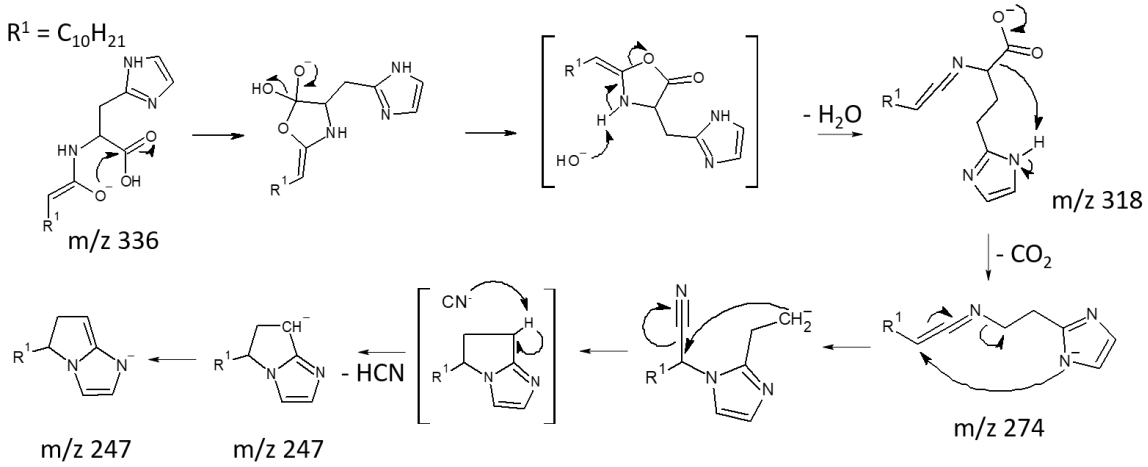


Figure S1. Product ion spectra of m/z 327 generated from µLC-HRMS in negative ESI mode under the A separative conditions (a)-(b) and under the B conditions (c)-(d) from biological sample, and by direct injection from synthetic C13Asn sample (e)-(f) under low energy collision conditions: (a)-(c)-(e) E_{Lab} = 5 eV and (b)-(d)-(f) at E_{Lab} = 20 eV. Normalized m/z molecular zone in window of each product ion spectrum and in (a) and (b) the peak intensities of the m/z 100-m/z 310 scale were amplified by a factor of 15 (ou 50). In product ion spectra (b) and (d), the used annotations of peaks (with intensity > 5% of base peak) are (i) the asterisk (*) for expected precursor ion, (ii) the empty circle (○) for unexpected peaks and (iii) the full circle (●) for expected peaks.

Scheme S1: Proposed formation of major product anions from the LpAA C12His (m/z 336) product anions



Les informations concernant la synthèse organique des composés sont en annexe 2.

Publication 2 : Nouveaux lipoaminoacides produits par *Escherichia coli :* origine de l'ion fragment carboxylate d'acide gras, régénéré à partir de la dissociation de l'anion lipo-glutamate

Résumé

L'amélioration des connaissances sur les lipoaminoacides est un point clé afin de comprendre le rôle du microbiote intestinal dans la santé humaine à travers ses diverses fonctions. L'élucidation structurale de ces composés est souvent réalisée en ESI à l'aide de la spectrométrie de masse en tandem, généralement après fragmentation par dissociation induite par collision. Généralement, les acides aminés sont analysés *via* le mode d'ionisation positif, cependant nous avons préféré le mode d'ionisation négatif, qui procure l'avantage de ne produire que peu d'ions fragments, mais suffisamment pour caractériser les structures des composés. De cette manière, lors de notre précédente étude, nous avons identifié les LpGlu. Dans cette présente étude, nous nous sommes intéressés aux mécanismes de fragmentation, permettant ainsi d'obtenir des informations sur les structures des ions fragments du LpAA N-palmytol acyl, lié à l'acide glutamate (noté C16Glu). Le comportement singulier du $[C16Glu-H]^-$ sous CID mène à la production d'une dizaine d'ions fragments. Certains d'entre eux sont décrits par des fragmentations attendues, alors qu'une paire abondante d'ions complémentaires et inattendus est observée correspondant aux ions fragments $[y-H_2O]^-$ et $[Lpb+O]^-$ selon notre nomenclature. Spécifique au glutamate, ces ions produits sont formés par un processus de dissociation par étapes, comportant une isomérisation moléculaire par cyclisation formant un cycle à 7 chaînons, suivie de la formation d'un ion dipôle. Ce complexe ion dipôle peut se dissocier soit directement par séparation des partenaires, soit après le transfert de proton d'un partenaire vers l'autre. Selon le mécanisme direct de dissociation, la régénération inattendue de l'acide gras déprotoné a lieu, de la même manière qu'elle est observée lors de la fragmentation de peptides déprotonés quand il existe un ou plusieurs résidus Glu (ou Asp) dans sa séquence. Les mécanismes proposés ont été confirmés à l'aide des calculs quantiques (DFT), expliquant le comportement inattendu de cet anion $[C16Glu-H]^-$. La longueur de la chaine lipidique ou la présence d'insaturation(s) sur cette chaine, semble ne pas avoir une influence déterminante sur l'importance de ces ions fragments, ce qui n'est pas le cas de l'acide aminé conjugué.

Résultats détaillés :

Formation d'ions fragments inattendus à partir du [C16Glu-H]– :

Lors de l'étape d'ionisation/désorption dans la source électrospray, seule la forme déprotonée $[C16Glu-H]^-$ est détectée. Malgré la présence de deux acides carboxyliques, la forme déprotonée n'est jamais accompagnée de la forme sodée $[C16Glu+Na]^-$). Cette forme annotée $[(C16Glu-H+Na)-H]^-$ peut être distinguée de la forme *cation solvatée* qui est annotée par l'écriture générique $[C16Glu+Na]^-$)[359,260]. Ces deux formes se distinguent à l'aide des spectres d'ions fragments (spectres CID), la première forme conduit aux ruptures de liaisons covalentes alors que la seconde, sous collision régénère le cation Na^+[360]. Ainsi contrairement aux composés constitués d'au moins deux groupes carboxylates qui conduisent à de telles molécules sodées, le C16Glu n'est jamais cationisé même si des sels de sodium sont ajoutés ou que la concentration de l'échantillon est fortement augmentée. Le sel déprotoné n'a donc pas pu être utilisé pour démontrer quel site était déprotoné pour conduire à certains ions fragments. L'alternative a été d'utiliser la méthylation régiosélective du groupe carboxylate de la chaine latérale du résidu Glutamate.

Le spectre d'ions produits de l'ion sélectionné $[C16Glu-H]^-$ permet de mettre en évidence la formation d'une dizaine d'ions fragments. Par comparaison avec les autres aminolipides décrits comme les $[LpAsn-H]^-$, il était attendu de retrouver les ions fragments $[Lpc]^-$ (m/z 254), $[z-2H]^-$ (m/z 129) et l'ion $[(C16Glu-H_2O)-H]^-$ (Figure 93).

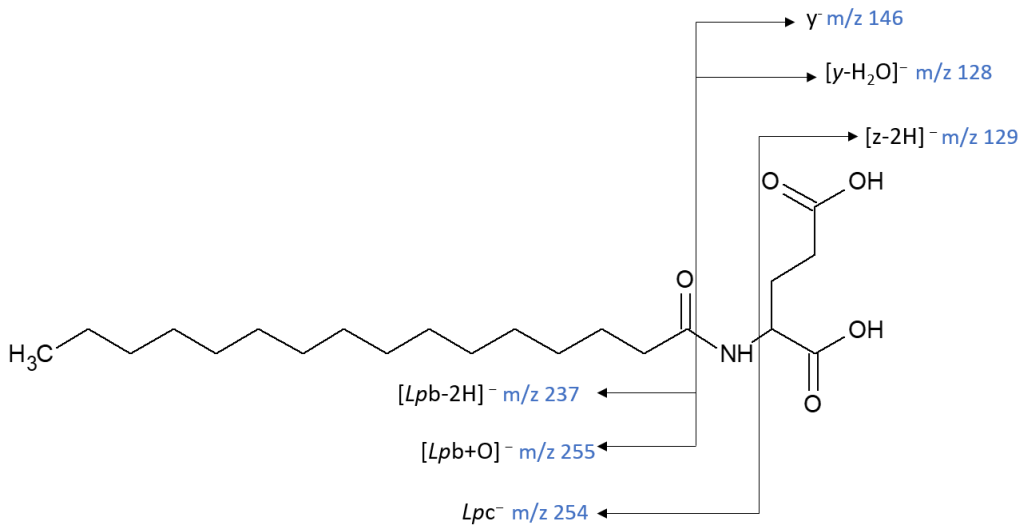


Figure 93 : Annotation des ions fragments produits à partir du C16Glutamate

De manière surprenante, l'abondance relative de ces ions à 20 eV est fortement réduite comparée aux $[LpAsn\text{-}H]^-$, puisqu'avec l'acide glutamique conjugué, ils sont inférieurs à 3,7%. L'ion m/z 129 que nous observons correspond en réalité à l'isotopologue d'un ion m/z 128, dont la composition élémentaire est de $^{12}C_4{}^{13}C_1H_6NO_3$ (1.7 ppm d'erreur), l'ion m/z 128 sera noté comme étant $[y\text{-}H_2O]^-$. De la même manière, nous détectons un ion fragment inattendu de rapport m/z 255, qui pourrait correspondre à l'ion $[Lpc+1]^-$ (de composition élémentaire : $C_{15}{}^{13}CH_{32}NO$) mais qui correspond à l'ion isobare de composition $C_{16}H_{31}O_2$ et sera annoté comme $[Lpb+O]^-$. L'abondance relative de ces deux ions fragments sont respectivement de 34% ($[Lpb+O]^-$, m/z 255) et de <3% ($[Lpc]^-$, m/z 254).

Mécanisme de fragmentations des ions m/z 255 et m/z 128 à partir du [C16Glu-H]– :

Les ions m/z 255 ($Lpb+O]^-$) et m/z 128 ($[y\text{-}H_2O]^-$) sont spécifiques au résidu glutamate. Ils sont formés selon un processus de dissociation par étapes dû à une isomerisation de l'espèce moléculaire $[C16Glu\text{-}H]^-$ en ion dipôle avant dissociation (Schéma 2). La première étape consiste en une attaque nucléophile régiosélective du carboxylate de la chaine latérale de l'acide aminé sur la liaison amide (liaison C-N). Cette première étape mène à la formation d'un intermédiaire hétérocyclique tétravalent à 7-chainons qui devient défavorisé dès que le groupement carboxyle de la chaine latérale est méthylé. La seconde étape est l'ouverture du cycle, induit par la charge de l'alcoolate, qui se fait par rupture de la liaison C-N. Cette rupture est concomitante avec un transfert de proton de l'acide carboxylique voisin sur le groupe NH libéré (ce groupe ne pouvant pas porter la charge négative car les amines en phase gazeuse font parties des fonctions chimiques les moins acides).

Ainsi un anhydride alicyclique est formé avec un groupe carboxylate sur l'extrémité de la chaine. Ce dernier peut à nouveau faire une attaque nucléophile sur le groupe C=O de l'anhydride le plus proche, pour générer à nouveau un intermédiaire tétraédrique cyclique à 6-chainons. Le groupe alcoolate de cet intermédiaire induit la libération de l'acide gras ainsi régénéré sous forme de carboxylate et d'un anhydride aminé cyclique à 6-chainons, qui reste en interaction non covalente au sein d'un complexe ion-dipôle. Ce complexe peut évoluer de deux manières, conduisant soit à sa dissociation directe donnant l'ion m/z 255 $[Lpb+O]^-$, soit par transfert de proton inter-partenaires, conduisant à la déprotonation de l'anhydride aminé cyclique et la neutralisation du carboxylate de l'acide gras régénéré. Cette étape conduit à l'ion m/z 128 $[y\text{-}H_2O]^-$.

Afin d'élucider le mécanisme mis en jeu lors de la fragmentation du C16Glu, la synthèse d'un ester C16OMeGlu (Figure 94).

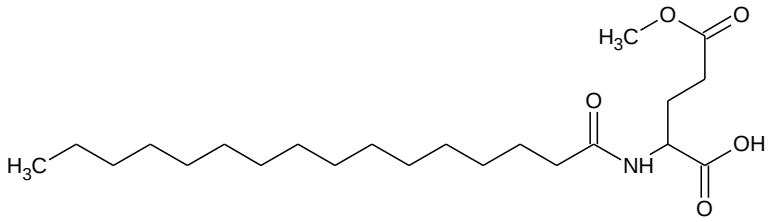


Figure 94 : Structure du C16OMeGlu

La méthylation entraine, outre l'empêchement des pertes consécutives de CO_2, la quasi disparition de l'ion fragment m/z 255 montrant la régiosélectivité de la première étape du mécanisme (Schema 2). L'ensemble de ces résultats ont été confirmés par les calculs quantiques qui ne seront pas développés ici car ils ont été réalisés en collaboration avec un groupe extérieur : M. Green, K. Richardson et J. Ujma de la société Waters (Wilmslow, UK).

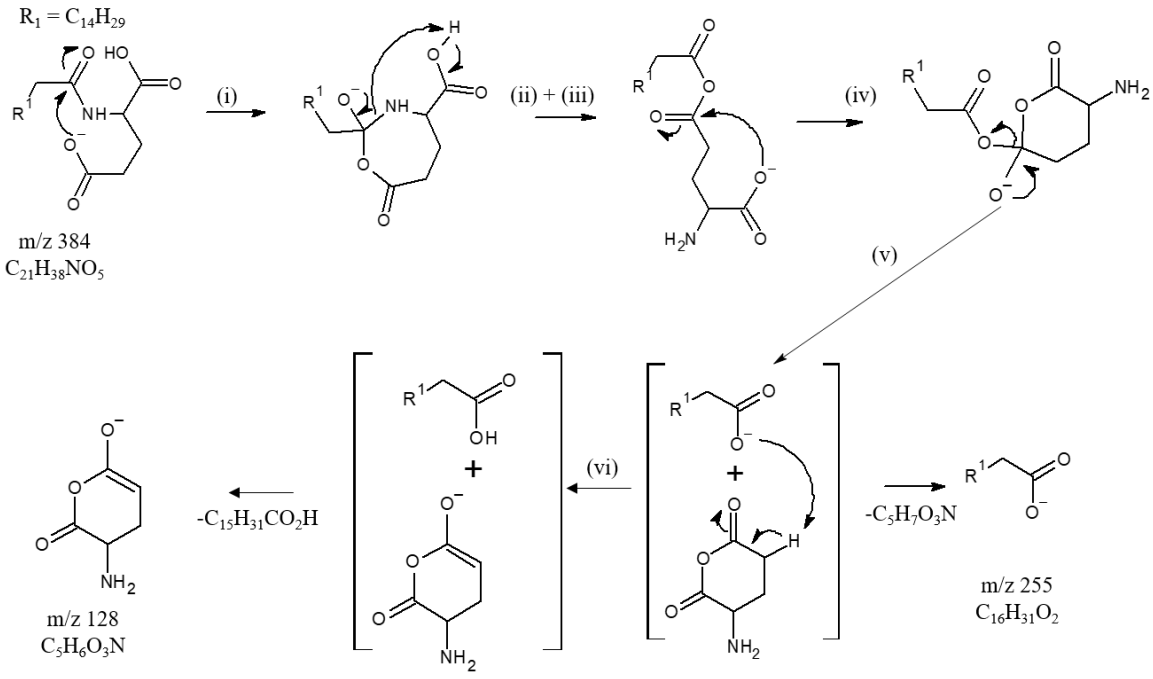


Schéma 2 : Mécanisme régiosélectif proposé pour la formation des ions complémentaires ions $[Lpb+O]^-$ et $[y\text{-}H_2O]^-$

Conclusion :

Dans cette étude, la LC-ESI-HRMS/MS réalisée en ionisation négative pour l'analyse des lipoaminoacides présente certains avantages en terme de sensibilité et de spécificité des ions fragments obtenus. Dans les conditions de dissociation induite par collision à basse énergie, l'anion $[C16Glu-H]^-$ se dissocie principalement en une paire d'ions produits, complémentaires, permettant de qualifier sans ambiguïté les deux composants du lipoaminoacide. La mise en place d'une annotation spécifique aux lipopeptides permet de définir cette paire d'ions complémentaires inattendue, correspondant aux ions fragments $[Lpb+O]^-$ qui représente l'amide déprotoné d'acide gras et $[y-H_2O]^-$ qui représente l'acide aminé déshydraté et déprotoné. De plus, la paire d'ions correspondantes aux ions fragments $[Lpc]^-$ / $[z-2H]^-$ attendue disparait de façon surprenante dans le cas de la dissociation de $[C16Glu-H]^-$. La formation des ions $[Lpb+O]^-$ et $[y-H_2O]^-$ complémentaires implique un processus par étapes, dont la première est une cyclisation par attaque nucléophile régiosélective sur la liaison amide exclusivement par le carboxylate de la chaine latérale . Il en résulte un intermédiaire cyclique tétravalent, qui évolue par isomérisation en un ion dipôle à partir duquel en compétition, les ions fragments $[Lpb+O]^-$ et $[y-H_2O]^-$ sont régénérés. La régiosélectivité ainsi que la pertinence des mécanismes de fragmentations proposés ont été confirmés expérimentalement et à l'aide de calculs quantiques.

Identification of bacterial lipo-amino acids: origin of regenerated fatty acid carboxylate from dissociation of lipo-glutamate anion

Amandine Hueber[a,b,c], Yves Gimbert[d,e], Geoffrey Langevin[f], Jean-Marie Galano[f], Alexandre Guy[f], Thierry Durand[f], Nicolas Cenac[b], Justine Bertrand-Michel[a,c*], Jean-Claude Tabet[a,d, g]

[a]. MetaboHUB-MetaToul, National Infrastructure of Metabolomics and Fluxomics, Toulouse, F-31077, France

[b]. IRSD, Université de Toulouse, INSERM, INRA, INPENVT, Université de Toulouse 3 Paul Sabatier, F-31024 Toulouse, France

[c]. I2MC, Université de Toulouse, Inserm, Université Toulouse 3 Paul Sabatier, F-31432 Toulouse, France

[d]. Sorbonne Université, CNRS, Institut Parisien de Chimie Moléculaire (UMR 8232), 4 place Jussieu, F-75005 Paris, France

[e]. Département de Chimie Moléculaire (UMR 5250), CNRS, Université Grenoble Alpes, 38610 Gières, France

[f]. Institut des Biomolécules Max Mousseron, UMR 5247 CNRS, Université de Montpellier-ENSCM, F-34093 Montpellier, France

[g]. Université Paris-Saclay, CEA, INRAE, Département Médicaments et Technologies pour la Santé, F-91191 Gif-sur-Yvette, France

*Corresponding author: Justine Bertrand-Michel, justine.bertrand-michel@inserm.fr

Keywords: Lipidomic, Lipoamino acid, Tandem Mass Spectrometry, Ion-dipole

Abstract: **Background:** The identification of bacterial metabolites produced by the microbiota is a key point to understand its role in human health. Among them, lipo-amino acids (LpAA), which are able to cross the epithelial barrier and to act on the host, are poorly identified. **Methods**: Structural elucidation of few of them was performed by high resolution tandem mass spectrometry based on electrospray combined with selective ion dissociations reach by collision-induced dissociation (CID). The negative ions were used for their advantages of yielding only few fragment ions sufficient to specify each part of LpAA with sensitivity. In order to find specific processes that help structural assignment, the negative ion dissociations have been scrutinized for a LpAA: the N-palmitoyl acyl group linked to glutamic acid (C16Glu). **Results:** the singular behavior of $[C16Glu-H]^-$ towards CID showed tenth product ions, eight were described by expected fragment ions. In contrast, instead of the expected product ions due to CONH-CH bond cleavage, an abundant complementary dehydrated glutamic acid and fatty acid anion pair were observed. Specific to glutamic moiety, they were formed by a stepwise dissociation *via* molecular isomerization through ion-dipole formation prior to dissociation. This complex dissociated

by partner splitting either directly or after inter-partner proton transfer. By this pathway, surprising regeneration of deprotonated fatty acid takes place. **Conclusion**: such regeneration is comparable to that occurred from dissociation to peptides containing acid amino-acid. Modeling allow to confirm the proposed mechanisms explaining the unexpected behavior of this glutamate conjugate.

Introduction

Among intestinal microbiota metabolites, lipopeptides (Pérez-Berezo et al. 2017) and lipo-amino acids (LpAA) (Vizcaino et al. 2014) have been described for their capacity to regulate host homeostasis. The major technical barrier for the identification of bacterial LpAA or lipopeptides is the difficulty or even the impossibility to isolate a sufficient sample amount to make the analysis either by NMR or/and by X-ray, which would require more than several tens of µg and several mg, respectively. In a previous study, identification of LpAA and lipopeptides produced by *Escherichia coli* Nissle 1917, such as the N-lauroyl acyl group linked to asparagine (C12Asn), has been done in negative ESI mode yielding abundant $[C12Asn\text{-}H]^-$ anions (Pérez-Berezo et al. 2017).

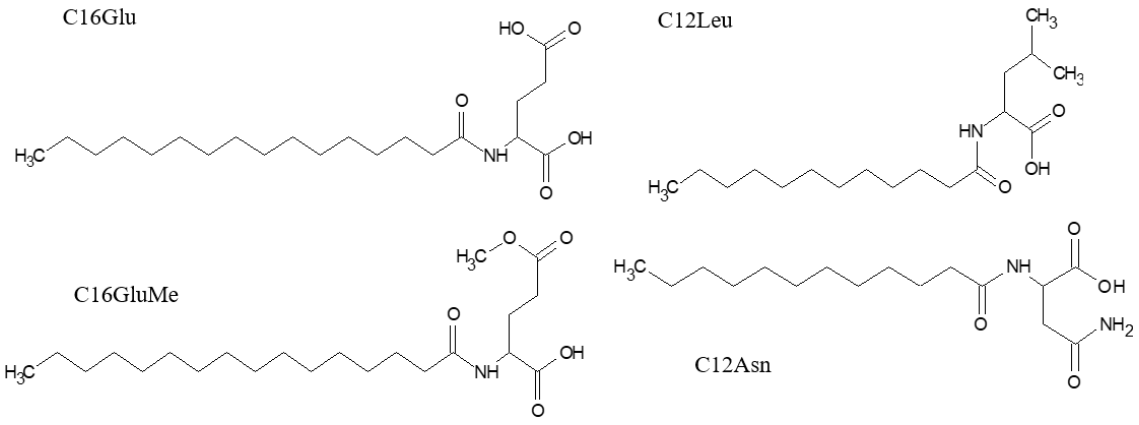
Analysis of these metabolites by hyphenated methods as liquid chromatography-tandem high-resolution mass spectrometry LC-HRMS/MS is especially versatile for complex mixture analysis. This method is characterized by a large selectivity, specificity and sensitivity when it is combined to desorption/ionization as electrospray (ESI). Indeed, this mode is suitable for ionization of a broad variety of more or less polar metabolites. Furthermore, it may be a powerful method for structural elucidation of *do novo* compounds (Thomas et al. 2019) especially, by using ion collisional activation (i.e. collision-induced dissociation, CID). Firstly, by using low resolution instrumentation based on linear ion trap quadrupole, an encouraging study of ornithine and glutamine lipids (from *Rhodobacter sphaeroides*) proposed mechanisms (Zhang, Ferguson-Miller, et Reid 2009) or a possible rationalization of ion dissociations under low energy collision conditions in resonant excitation mode.

In this mode, only the parent ion is excited by its own frequency. Therefore, since the produced ions are not excited, their consecutive dissociations are inhibited unlike competitive dissociations (Bennaceur et al. 2013). However, they occur when the dissociative process is exothermic and/or the residual internal energy carried by the produced ions is sufficient to drive consecutive dissociations (Darii et al. 2021) (Tabet et al. 2021). A previous study (Boukerche et al. 2016) investigated the detected fragment ions of protonated LpAA ($[LpAA+H]^+$) resonantly excited under low-energy collisional activation conditions. They interpreted formation of these product ions using exclusively mechanisms based on charge-promoted dissociations. In certain cases, as a consequence, molecular isomerization into ion-dipole as intermediate of stepwise dissociation pathways is considered (Afonso, Cole, et Tabet 2010). Such an isomerization explains competitive dissociations mainly based on internal proton transfer into ion/neutral complex responsible of the complementary product ions. This approach is in fact an alternative to the one previously proposed (Zhang et al. 2009) for some fragmentations where the charge remained spectator as it happened during collision process with energies in the keV range

(Wysocki et al. 1988) (Gross et al. 1992). Interestingly, to test the hypothesis that many lipo-amino acids originate in brain, the identification of new LpAAs was conducted in LC/HRMS/MS by Tan et al (Tan et al. 2010) and explained the observed fragmentations by charge-driven dissociations after CID of selected ions.

With the future aim to identify more metabolites from bacteria of the microbiota and based on these above results, the LpAA characterization will be extended to different amino acids conjugated to various fatty acyl groups using LC-ESI/HRMS/MS. For this, a broader study will be conducted to build a database of product ion spectra (or fragment ion spectra (Murray et al. 2013)) of various LpAA ions with a specific and clear fragmentations of $[LpAA\text{-}H]^-$ to detect, identify and elucidate the *do novo* molecule structure of LpAA (Thomas et al. 2019). The negative ESI mode was chosen to provide abundant deprotonated LpAA useful to deliver, after CID, fragment ions which are characteristic of both the amino-acid and fatty acyl moieties.

For our purpose, dissociations of deprotonated LpAA standards were investigated under low collision energy conditions (i.e., non-resonant mode) by considering previous studies. In our previous study (Pérez-Berezo et al. 2017), the major dissociative processes of the $[C12Asn\text{-}H]^-$ from N-lauroyl-asparagine occurred, in addition to the water loss (corresponding to the base peak), through competitive covalent bond cleavages around the amide linkage with competitive formations of complementary ion pairs. This takes place through stepwise dissociation *via* ion/neutral molecular isomerization resulting in the $[(C12Asn\text{-}H)\text{-}(Asn\text{-}NH_3)]^-$ and $[(Asn\text{-}H)\text{-}NH_3]^-$ product ion pair through competitive proton transfer allowing to define the two parts of the LpAA. In this study, behavior of the system constituted by both the glutamic residue and the C16 fatty acyl moieties (**Scheme 1**) was deeply scrutinized. It will be compared to a polar, asparagine (C12Asn) and non-polar amino acid, leucine (C12Leu) (**Scheme 1**) to reach a pertinent interpretation of the $[C16Glu\text{-}H]^-$ fragmentations based on elemental composition of precursor ion and product ions using an hybrid Qq/TOF tandem mass spectrometer.



Scheme 1. Structure of the studied lipoamino acids: N-palmitoyl acyl group linked to glutamic acid (C16Glu) and monoester derivate (C16GluMe), N-lauroyl acyl group linked to leucine (C12Leu) and to asparagine (C12Asn).

Our study will focus on the production of the observed product ions and especially those that are unexpected i.e., product ions generated from LpAA skeleton rearrangement and in particular the carboxylate of the fatty acid. This ion has also been observed and described by Tan et al. (Tan et al. 2010). Their work can be considered as a landmark since their proposed interpretation for regeneration of the fatty acid carboxylate from $[LpGlu\text{-}H]^-$. Indeed, without involving the second carboxylic acid group, the proposed tetrahedral intermediate of dissociative $[LpGlu\text{-}H]^-$ evolved directly towards the fatty acid carboxylate product ion. However, dissociation of $[LpGABA\text{-}H]^-$ (containing only one carboxylic acid group) did not lead to regeneration of the carboxylate product ion (Tan et al. 2010). This suggests that the second carboxylic acid in $[LpGlu\text{-}H]^-$ must play important role in the regeneration of the unexpected fatty acid carboxylate product ion. In addition, they suggested that the $[(Glu\text{-}H)\text{-}H_2O]^-$ ion was generated from $[Glu\text{-}H]^-$ by water loss. However, product ion spectrum of $[C16Glu\text{-}H]^-$ (m/z 384) under resonant excitation conditions displayed, in addition to the H_2O and CO_2 losses (ions at m/z 366 and m/z 340), only the pair of the deprotonated fatty acid and dehydrated glutamate product ions (Figure S1) which evidences that the latter ions is directly formed from the precursor ion rather than from dehydration of glutamate anion as intermediate.

To rationalize formation of this particular species, the product ion spectrum of the deprotonated $[C16GluMe\text{-}H]^-$ mono ester (Scheme 1) was investigated and compared to $[C16Glu\text{-}H]^-$. For a better description of the origin of the unexpected regeneration of the deprotonated C16 fatty acid, the respective product ion spectra of the C16Glu and C16GluMe anions were investigated under different collision energy conditions. Finally, this study provided an analytical approach to determine without ambiguity the structure of lipoamino acids and lipopeptides within the framework of mass spectrometry analysis essential for bacterial LpAA identification. For this purpose, it is necessary to better account for the dissociation processes of these compounds by means of well-defined systematic mechanisms. The described mechanisms must explain formation of both the expected and unexpected product ions with similar mechanisms based on the chemical reactivity.

Material and method

Chemical synthesis of LpAA

Synthesis of lipoaminoacids such as N-lauroyl-asparagine (C12Asn), N-lauroyl-leucine (C12Leu), N-palmitoyl-glutamic acid (C16Glu) and N-palmitoyl-glutamic acid 5-methyl ester (C16GluMe) were done by solid phase chemistry in a SPS reactor using a 2-chlorotrityl resin (max loading 1.6 mmol/g, 250 mg, 400 µmol, 2.0 eq.). Very briefly, after the fixation of the appropriate Fmoc-amino acid (200 µmol, 1.0 eq.) in presence of NMM (800 µmol, 4.0 eq.) in DCM for 3h, the resin was capped by

DCM/MeOH/NMM (17/2/1) for 5 min and the Fmoc group removed with Piperidine/DMF (2/8) for 10 min. After washing, the fatty acid is added in presence of HBTU (750 µmol, 5.0 eq.), HOBt (15 µmol, 0.1 eq.) and NMM (7.5 mmol, 10.0 eq.) in NMP (5mL) for 3 hours to overnight monitoring with a Kaiser's test. After washing, the resin is cleaved with TFA/TIPS/Water (95/2.5/2.5) for 10 min. Obtained filtrate were concentrated and triturated in ACN to obtained desired LpAA without further purifications: C16Glu (53 mg, Yield = 69%); C12Leu (112 mg, Yield = 67 %); C12-Asn (63mg, Yield = 100%) C16GluMe (30mg, Yield = 58 %). LpAA were dissolved in methanol with a concentration of 10ng/µL.

Mass spectrometry analysis:

Direct infusion mass-spectrometry analysis: were done in negative mode ESI(-) using a Xevo G2-XS time of flight (Qq/TOF) mass spectrometer (Waters, Milford, MA) equipped with electrospray ionization (ESI). The source parameters were set as follows: source temperature = 120°C, capillary voltage = -2.6 kV, desolvation gas flow rate = 300 L/h, cone gas flow = 40 L/h, desolvation temperature = 380°C, collision gas nitrogen pressure = 10^{-1} mbar, voltage QqTOF = 9.00 kV, pressure in flying tube = $2.7x10^{-7}$ mbar, average number of spectra = 35 (1 spectra every 14 msec for 500 ms),

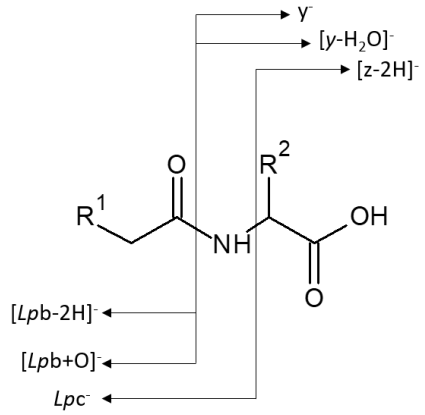
MS/MS experiments: product ion spectra of the precursor $[C16Glu-H]^-$ ion were performed in segmented RF only quadrupole (Waters , Milford, MA) under collision-induced-dissociation conditions (i.e., CID, using non-resonant excitation mode) within a full width at half maximum (FWHM) resolution of 16 000 resolution at *m/z* 400 for product ion analysis. Precursor ions were selected within ± 5 *m/z*. The product ion spectra (Murray et al. 2013) (called also CID spectra) were recorded at two collision energies (noted as E_{Lab}): 17 eV and 30 eV. These CID spectra were inspected manually using MassLynx software (Waters) to confirm annotations (*vide infra*). Only the signal higher than 5% of base peak were annotated in CID spectra and in the supplementary **Table S1**, the reported signals were those with relative abundances ≥ 0.4% of the total ionic current TIC at one E_{lab} values. The ion abundances were relative to TIC which included precursor and product ions and were given in percent of TIC. Relative and absolute ion abundances were reported in the supplementary **Table S1**.

Modeling

DFT calculations were performed using Gaussian16 package, employing the B3LYP functional with the 6-31+G(d,p) basis set. Zero Point Energies (ZPE) are added as computed at the optimizing level. The nature of all stationary points was confirmed by analyzing the harmonic vibrational frequencies. The energies reported in this work are E energies in kcal/mol, corrected from ZPE. For reasons of computational time and under the reasonable assumption that the length of the aliphatic chain will not significantly influence the nature of the pathways studied, the structures were modeled with an ethyl group (for computational details see **Supplementary Information S1**).

Nomenclature and notation of product anions

The notations used to describe and discuss mass spectrometry results were based on IUPAC recommendations (Murray et al. 2013) to provide more accurate and consistent text throughout the manuscript. In addition, it was proposed a simple nomenclature which could be applied to the LpAA anion dissociation (Tan et al. 2010). In this nomenclature only some fragmentations are described and concern only the loss of small neutrals (H_2O, NH_3, CO_2….) which can be also consecutive losses (H_2O+CO_2) from the amino-acid (AA) residue. Such an annotation suggests their formations occur through consecutive fragmentations from this AA residue generated by dissociation of LpAA anions. This is not true since some of these fragment ions (the abundant ones) are generated from direct cleavages of the deprotonated lipoamino-acid and not from the amino-acid as intermediate. This nomenclature cannot be applied for deprotonated lipopeptide dissociations since, N-terminus (linked to fatty acyl moiety) or C-terminus (related to peptide moiety) product ions from peptide bond cleavages cannot be attributed. In order to meet this requirement , the nomenclature on deprotonated peptide dissociations (Chu et al. 2015) based on fragmentations of protonated peptide ions (with (i+j) residues in sequence) yielding was adapted, e.g. fragment ions as $[b_i\text{-}2H]^-$, c_i^- for the N-terminus and y_j^-, $[z_j\text{-}2H]^-$ for the C-terminus. This was based on the nomenclature of product ions generated under keV collision energy conditions introduced by Roepstorff et Fohlman (Roepstorff et Fohlman 1984) and Biemann (Biemann 1988). This nomenclature as reported in **Scheme 2** allows to describe both the lipo-amino acid and lipo-peptide fragmentations.



Scheme 2. Annotation of the observed product ions from dissociation of deprotonated lipoamino acids

Indeed, annotation distinction between (i) fragment ions constituted by the fatty acyl moiety (related to N-terminus of peptide sequence) was annotated by the *Lp* prefix linked to peptide fragment (e.g. $[Lpb_i\text{-}2H]^-$, Lpc_i^- …) and (ii) fragment ions related to the C-terminus peptide sequence contains exclusively a peptide fragment sequence (e.g. y_j^-, $[z_j\text{-}2H]^-$ …). When Glu or Asp are included in the peptide

sequence, then a complementary ion pair such as $[Lp\mathrm{b}_i+O]^-$ and $[\mathrm{y}_j\text{-}H_2O]^-$ appeared. From the lipoamino-acids with AA=Glu, the particular product ions specific to (i) the fatty acyl moiety such as fatty acid amidate and fatty acid carboxylate formally annotated as $lp\mathrm{c}^-$ and as $[lp\mathrm{b}+O]^-$, respectively and (ii) the amino acid residue such as deprotonated AA (i.e., $[\text{Glu-H}]^-$), and its fragment $[(\text{Glu-H})\text{-}NH_3]^-$ ion annotated as y^- and $[\text{z-2H}]^-$ were detected. The consecutive fragment ions due to the small neutral loss, e.g. CO_2 or H_2O, from the first generation product y^- and $[\text{z-2H}]^-$ ions, will be annotated as $[\text{y-}CO_2]^-$ and $[(\text{z-2H})\text{-}H_2O]^-$, respectively.

Result and discussion

The mass spectrum of the C16Glu compound (Mw 385 u) displayed essentially the $[\text{C16Glu-H}]^-$ anions (base peak at m/z 384, **Figure S1**). Under the source conditions used, adduct ions $[\text{C16Glu+anion}]^-$ (anions as inorganic or organic acids) as well as deprotonated $[(\text{C16Glu+Na-H})\text{-H}]^-$ salt were not observed even if the concentration of C16Glu was increased. This behavior characterized the other LpAA studied. The absence of adduct ions was due to the presence of one or several sufficiently acidic sites able to protonate the reactive anions from the used solvents and allowed to reach a large sensibility of the $[\text{C16Glu-H}]^-$ detection. In MS/MS mode, this is one of the advantages for the investigation of the dissociations of the $[\text{C16Glu-H}]^-$ anion in order to perform structural analysis of this family of conjugates. Such a detailed studied were only performed for protonated LpAA (Boukerche et al. 2016) which as expected was driven different dissociation pathways. The $[\text{C16Glu-H}]^-$ anion (m/z 384) submitted to collisional activation (CID with E_{lab}=17 eV, and E_{lab}=30 eV, **Figure 1(a),(b)**) was characterized by a product ion spectrum (at E_{lab}=17 eV) displaying tenth peaks representative to product ions. The peaks with intensities lower than 5% of the base peak at E_{lab}=17 eV were annotated in the CID spectrum only when their intensities became higher than 5% of base peak at E_{lab}=30 eV, and conversely for the intensities of base peaks lower than 5% at E_{lab}=30 eV and higher than 5% at E_{lab}=17 eV.

In our previous study (Pérez-Berezo et al. 2017), two intense peaks at m/z 198 (ion $lp\mathrm{c}^-$) and m/z 114 (ion $[\text{z-2H}]^-$) emerged in addition to the base peak at m/z 295 (related to the H_2O release) from CID spectrum of $[\text{C12Asn-H}]^-$ (m/z 313, **Figure S2**). The complementary $lp\mathrm{c}^-/[\text{z-2H}]^-$ ion pair were interpreted to be formed *via* ion/dipole intermediate which allows the production of these complementary fragment ions. By analogy with C12Asn, the product ion spectrum of the $[\text{C16Glu-H}]^-$ anion (m/z 384) (**Figure 1(a),(b)**) should display the fragment ions at m/z 254.2483 (ion $lp\mathrm{c}^-$ as $C_{16}H_{32}NO$ with error of 2.5 ppm), and m/z 129,0193 (ion [z-2H]- as $C_5H_5O_4$, with error of 151.6 ppm. This product fragment ion m/z 129 may match, with the first ^{13}C isotopologue of the m/z 128 ion (ion $[(\text{y-}H_2O)\text{-1}]$- as $^{12}C_4{}^{13}C_1H_6NO_3$ with error of 1.7 ppm). Surprisingly, the abundance of this ion $lp\mathrm{c}^-$ and $[\text{z-2H}]^-$ pair was of very low. A pair of product ions corresponding to $[lp\mathrm{c}+1]^-$ (at m/z 255) and $[(\text{z-2H})\text{-1}]^-$ (at m/z 128) appeared. Less abundant fragment ions corresponding to either small size neutral losses or cleavage of bonds close the amide linkage were displayed in the CID spectra (**Figure 1(a)(b)).**

The latter corresponded to product ions with charge retention at the amino acid moiety or at the fatty acyl part as confirmed by high resolution measurements (**Table S1** and interpretations were reported in **Schemes S1 and S2**).

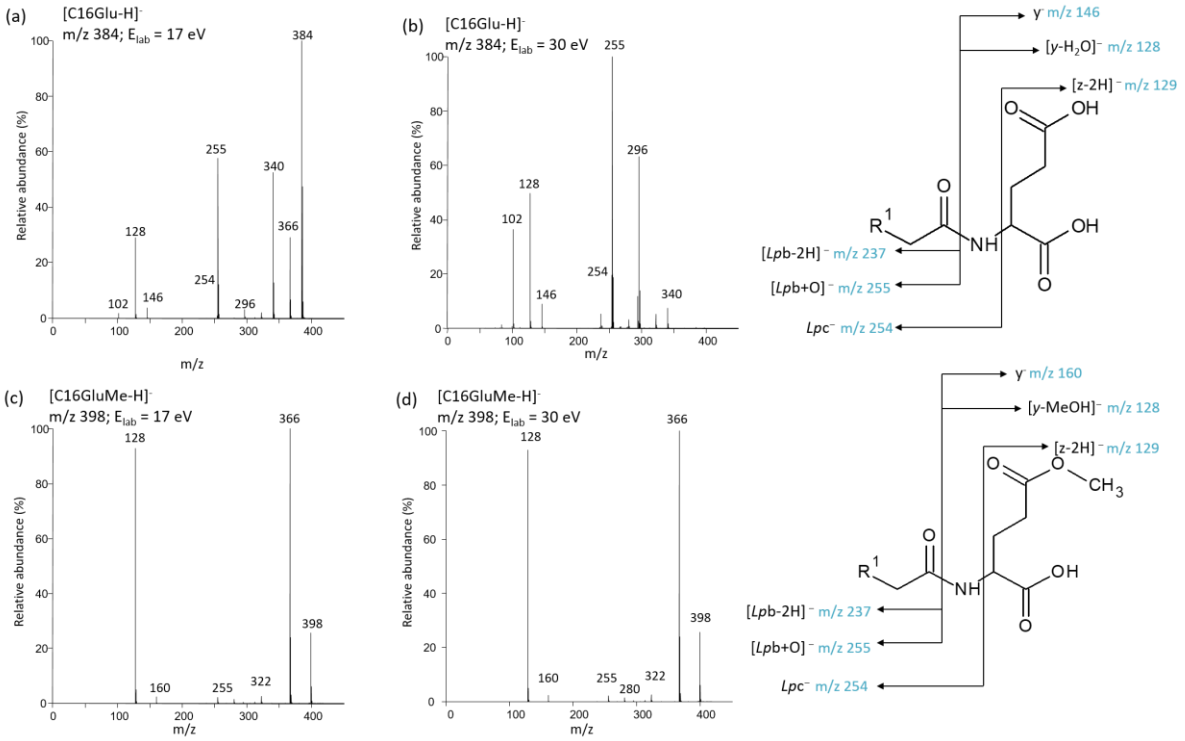


Figure 1. Product ion spectra (i.e., MS^2 fragmentations) of (a)-(b) $[C16Glu\text{-}H]^-$ (m/z 384 with R_1 = $CH_3(CH_2)_{13}$-) at E_{Lab} = 17 eV and E_{Lab} = 30 eV, and (c)-(d) $[C16GluMe\text{-}H]^-$ (m/z 398) at E_{Lab} = 17 eV and E_{Lab} = 30 eV, respectively (accurate m/z values and corresponding elemental composition reported in Table S1).

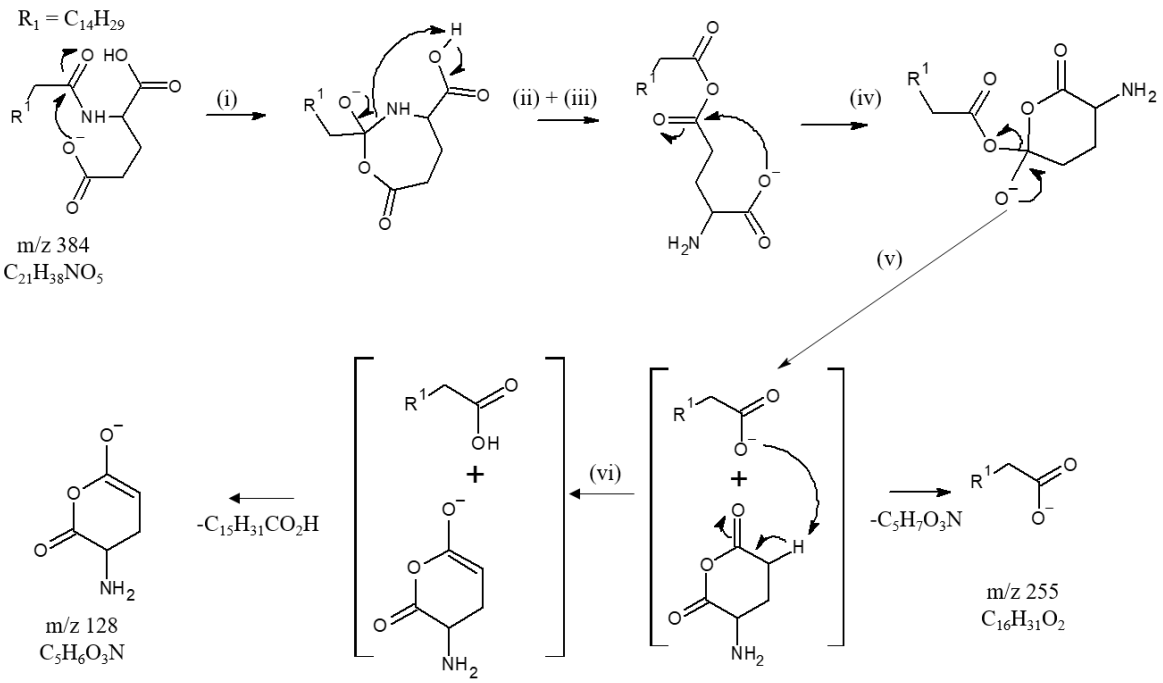
These signals were detected at:

(i) m/z 366 and m/z 340 are abundant species accompanied by weaker ion at m/z 322 (less than 5% of TIC at this energy at E_{lab}=17 eV and increases to 1.8 at 30 eV) and m/z 296 (~ 1% at 17 eV and 21% at 30 eV) respectively due to the expected H_2O, CO_2, (H_2O + CO_2) and $2CO_2$ small size neutral losses (proposed fragmentation mechanisms in **Scheme S1**);

(ii) m/z 146 (i.e., y^-) and m/z 102 i.e., $[y\text{-}CO_2]^-$, < 1% at 17 eV and 13% at 30 eV) resulting from the loss of fatty acyl moiety (proposed fragmentation mechanisms reported **Scheme S2**). The detected m/z 129 ion which was the ^{13}C natural isotopologue of the abundant m/z 128 ion (i.e., $[y\text{-}H_2O]^-$) and not the $[y\text{-}NH_3]^-$ isobaric ion, is not reported in **Table S1**.

Since the formation mechanisms of product ions described above have already been studied (Pérez-Berezo et al. 2017), we focused on the origin and formation mechanism of the abundant unexpected m/z 255 and m/z 128 product ions displayed in the CID spectra recorded at 17 eV and 30 eV. These complementary ions are considered together rather than separately as proposed by Tan et al. (Tan et al. 2010).The elemental compositions attributed to the ions at m/z 255.2321 and at m/z 128.0363 were

respectively $C_{16}H_{31}O_2$ (error of 3.3 ppm) and $C_5H_6NO_3$ (error of 2.4 ppm) (**Table S1**). Thus, they may be respectively annotated as the complementary $[lp\text{b}+O]^-$ and $[\text{y-H}_2O]^-$ product ions rather than $[lp\text{c}+1]^-$ and $[(\text{z-2H})-1]^-$. Indeed, formation of the $[\text{y-H}_2O]^-$ and $[lp\text{b}+O]^-$ ion pair can be rationalized by a stepwise dissociation process from $[\text{C16Glu-H}]^-$ ion through molecular isomerization into ion/dipole complex intermediate followed by competitive splitting processes (**Scheme 2**) (Afonso et al. 2010). The first step is a regioselective nucleophilic attack of amino acid side chain carboxylate site to amide linkage resulting in cyclic tetravalent intermediate. The negative alkoxide site migration induces the C-N bond cleavage and ring opening concomitant with proton transfer to nitrogen atom and resulting in an anhydride linkage. The end group carboxylate interacts with the closer anhydride C=O group yielding a 6-membered tetravalent system which isomerizes in ion/dipole consisting in α amino anhydride neutral and long chain carboxylate. This complex can evolve either to a direct partner splitting to yield m/z 255 (i.e., the fatty acid carboxylate as $[lp\text{b}+O]^-$) or after internal proton transfer from the amino anhydride neutral to the C16carboxylate partner, release fatty acid neutral to drive to the deprotonated 6-membered amino-anhydride (m/z 128 annotated as $[\text{y-H}_2O]^-$) (Scheme 2).



Scheme 2. Proposed regioselective stepwise formation of complementary $[\text{y-H}_2O]^-$ (m/z 128) and $[lp\text{b}+O]^-$ (m/z 255) product anions through ion-dipole complex molecular isomerization.

The regioselectivity of the first step was confirmed by the CID spectrum of $[\text{C16GluMe-H}]^-$, m/z 398 for which the peak at m/z 255 was strongly reduced in this product ion spectrum (**Figure 1(c)(d)**). Indeed, the peak at m/z 255 which was the base peak in the CID spectra of $[\text{C16Glu-H}]^-$ did not reach 1% of base peak at E_{Lab}=17 eV and 9.3% at E_{Lab}=30 eV for $[\text{C16GluMe-H}]^-$. This confirmed that the nucleophilic reactivity of the carboxylic acid in α position of amide linkage is significantly hindered compared to the carboxylic group of the side chain. In addition, the CID spectra of [C12Leu-

H]$^-$ (m/z 312), a non-polar LpAA, (**Scheme S2 and Table S1**) did not display peak at m/z 199 ion corresponding to the [*lp*b+O]$^-$ ion confirming the very weak reactivity of the amino α carboxyl group. On the other hand, this explained also why the dissociation of the [LpGABA-H]$^-$ ion did not yield the *lp*c$^-$ product ion (Tan et al. 2010).

In order to confirm the validity of the proposed mechanism above, modeling was performed. A first investigation was carried out starting from the structure **A** corresponding to the LpGlu (pathway in black, **Figure 2**). For reasons of computational time and under the reasonable assumption that the length of the aliphatic chain will not significantly influence the nature of the pathways studied, the structures were modeled with an ethyl group.

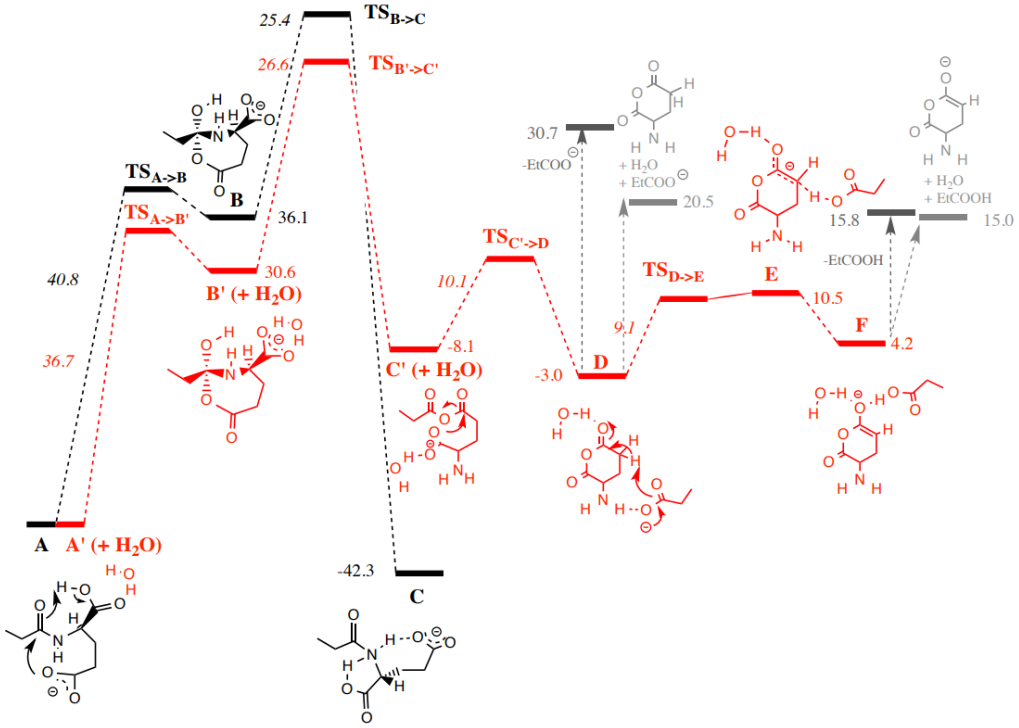
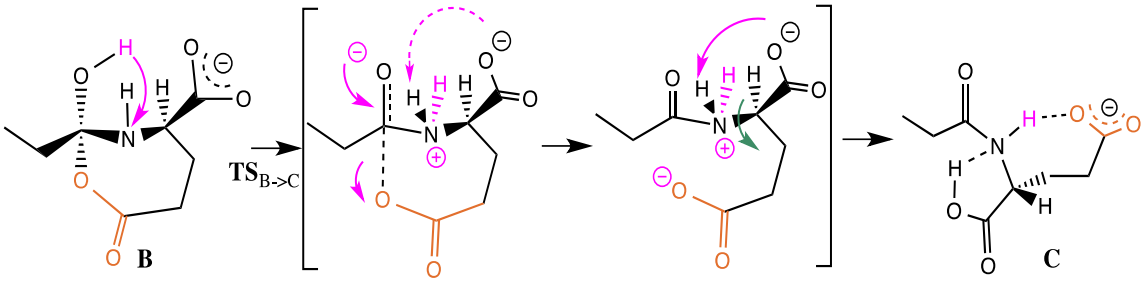


Figure 2. Energetic profiles (B3LYP/6-31+G(d,p)), ZPE at the optimizing level): (a) the *black pathway*, an evolutionary dead-end yielding the A carboxylate isomerization into the deprotomer **C** and (b) the *red pathway*: competitive fragmentations of the hydrated (**A** +H$_2$O) carboxylate towards the product anions as (i) propionic carboxylate and (ii) the deprotonated α amino glutaric anhydride (energy in kcal/mol)

A transition state (**TS$_{A->B}$**) leading to the cyclization into a tetrahedral intermediate **B** situated 40.8 kcal/mol above **A** and corresponding to an endothermic transformation by 36.1 kcal/mol could be located. This cyclization is facilitated by an assistance of the proton of the COOH group of the amino acid which stabilizes the produced alkoxide by its neutralization. In a second step, we explored the possibility of the C-N bond cleavage by proton migration from the OH group of tetrahedral intermediate to the neighbored nitrogen atom. This migration results in the creation of an ammonium group which should promote the anticipated cleavage to open the ring and to form an anhydride group (i.e., the **C'**

intermediate). A transition state ($TS_{B->C}$) for such a proton transfer was found (25.4 kcal/mol above **B**). However, the opened ring product **C** to which it connects is not the result of the C-N bond cleavage wished. Indeed, this proton transfer leads to the opening of the anhydride, immediately followed by the migration of the proton initially on -NH- groups which is captured by the $-CO_2^-$ group (**Scheme 3**).



Scheme 3. Formation of the **C** anion (deprotomer of the initial **A** anion) connected by $TS_{B->C}$ as a dead end of this pathway.

In order to strongly hinder protonation of the carboxylate group by an ammonium proton for favoring C-N cleavage and ring-opening, a water molecule was introduced (**A'** + H_2O) in calculations to check its effect on the re-orientation of reaction toward formation of the (**C'**+H_2O) intermediate rather than that the **C** deprotomer. In a similar way, a first cyclization (*via* $TS_{A'->B'}$ + H_2O) leads to structure (**B'** + H_2O) with a lower energy cost than in the absence of water (36.7 versus 40.8 kcal/mol). From the 7-membered ring tetrahedral (**B'** + H_2O) intermediate, the proton transfers to the nitrogen atom triggers, this time, the rupture of the C-N bond to yield, through rearrangement process, the anhydride (**C'** + H_2O) intermediate. This cleavage corresponds to an exothermic transformation by 8.1 kcal/mol with respect to the solvated (**B'** + H_2O) form. At the end, to reach deprotonated fatty acid (ion $[lpb+O]^-$) from the (**C'** + H_2O) intermediate, the carboxylic group must drive a nucleophilic attack at the adequate C-O bond of the anhydride group yielding ion/neutral complex (**D** intermediate stabilized by hydrogen bond) consisting in deprotonated fatty acid (herein propionic carboxylate, $EtCOO^-$) and the α amino glutaric anhydride neutral. This cyclization leading to **D** and the formation of the carboxylate ($[lpb+O]^-$) occurs *via* $TS_{C'->D}$ with a barrier of 10.1 kcal/mol in an exothermic process of 3.0 kcal/mol. In the **D** intermediate (**Figure 3**), the $EtCOO^-$ group is favorably positioned to attract a mobile proton from the enolizable CH_2 site of the carbonyl group. This deprotonation requires a cost of 9.1 kcal/mol to form an intermediate **E** located very slightly (1.4 kcal/mol) above the $TS_{D->E}$. This is probably due to a persistent interaction between the OH group of the carboxylic acid and the CH that prevents the full formation of the enolate form. This is evidenced by the analysis of the value of the dihedral angle OC-CH in this structure which is 20.2 degrees while it would expect a value very close to 0° for an enolate form. The system gains energy by migration of the acid towards the oxygen atom of the enolate by developing a hydrogen bond type interaction. In structure **F**, the enolate is this time completely formed, the previously discussed dihedral angle being now close to 0 degrees (-1.0). Evolutions of the C-C and C-O distances between **E** and **F** are also logical.

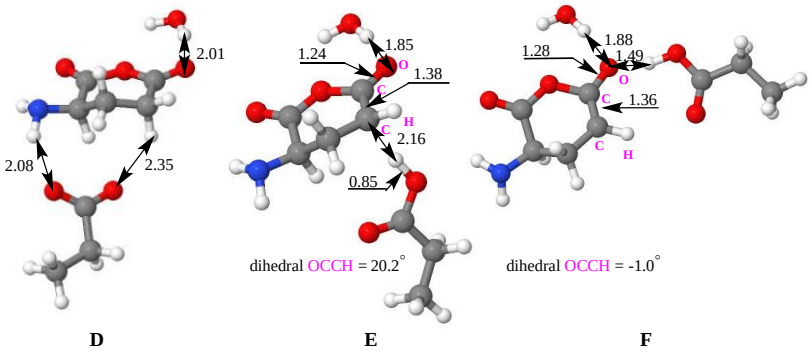


Figure 3: Structures **D**, **E** and **F**. Distances in Å.

Finally, the competitive dissociation of both the **D** and **F** intermediates are only desolvation steps (i.e., cleavage of hydrogen bond) characterized by level lower than the highest transition state (i.e**.,** $\mathbf{TS_{B'->C'}}$) of these stepwise dissociations of the mono-hydrated model system of lipo-glutamic acid.

Conclusion

In this study, LC-ESI-HRMS/MS performed in negative ion mode for the LpAA analysis has some advantages in terms of sensibility and specificity. Under low energy collision-induced-dissociation (CID) conditions, $[LpAA\text{-}H]^-$ anion dissociate mainly into complementary product ion pair allowing to unambiguously qualify the two conjugated components of LpAA except for non-polar AA residues which present only one characteristic product ion. Using the peptide dissociation annotation, the complementary $[lpc]^-$ and $[z\text{-}2H]^-$ ion pair respectively corresponding to the fatty acid amidate and deaminated amino-acid carboxylate are abundant in the product ion spectra of $[LpAA\text{-}H]^-$. However, from deprotonated glutamate conjugates such as $[C16Glu\text{-}H]^-$, the $[lpc]^-$ and $[z\text{-}2H]^-$ anion pair disappears in favor to another diagnostic ion pair. Indeed, the complementary $[lpb+O]^-$ and $[y\text{-}H_2O]^-$ product anions appears due to regeneration of fatty acid carboxylate and formation of dehydrated glutamate *via* a nucleophilic attack of amide group by one of the two carboxylic groups. This annotation was the one used for deprotonated peptide dissociation. Interestingly, dissociation of deprotonated C16GluMe mono methyl-ester located at the glutamic side chain leads to reduction of the $[lpb+O]^-$ ion abundance which evidences the strong regioselectivity of the δ carboxylic acid on the nucleophilic attack on amide linkage resulting in 7-membered ring with a tetrahedral reactive site. From this intermediate evolves toward a stepwise process *via* isomerization into ion-dipole intermediate from which through internal proton transfer and complex splitting yields in competition, the $[lpb+O]^-$ and $[y\text{-}H_2O]^-$ product anions. The specific behavior of the glutamate residue allows it to be distinguished from other amino acid residues independently to the fatty acid amide chain length. Thanks to the modeling, the pertinence of the above proposed stepwise process mechanism leading to the complementary product ions especially the regeneration of the fatty acid carboxylate is confirmed. All these results allow to improve the identification of LpAA using mass spectrometry which is essential for the study of this low concentrated bacterial metabolites.

Abbreviations : ACN: acetonitrile; CID : collision induced dissociation;, DCM : Dichloromethane; DFT: density functional theory ; EcN : Escherichia coli ; ESI : electrospeay ; GABA : γ- aminobutyric acid; HOBt: 1-hydroxybenzotriazole hydrate; Fmoc: 9-fluorenylmethoxycarbonyl; LC-HRMS : liquid-chromatography high resolution mass spectrometry ; LpAA : Lipoaminoacids ; MeOH: Methanol; NMM: N-methylmorpholine ; NMP: N-methyl-2-pyrrolidone; RF : radio-frequence ;TFA: trifluoroacetic acid TIPS: Triisopropylsilane ; ZPE : zero point energy.

Acknowledgements: Mass spectrometry analysis were done on MetaToul (Toulouse Metabolomics & Fluxomics facilities, www.metatoul.fr) which is part of the French National Infrastructure for Metabolomics and Fluxomics MetaboHUB-ANR-11-INBS-0010. All the computations presented in this article were performed using the Ceciccluster platform of the PCECIC infrastructure, which is supported by ICMG (FR 2607). We would like to thank Hanna Kulik-Barbier, Jean-François Martin and Pauline Le Faouder for helpful discussions.

Declarations:

Conflict of interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.

Research involving Human Participants and/or Animals & informed consent: No biological neither human material were used in this study

Bibliography

Afonso, Carlos, Richard B. Cole, et Jean Claude Tabet. 2010. « Even-Electron Dissociation in ESI and MALDI/MS : Fundamental, Instrumentation, Practicalities and Applications ». P. 631-84 in *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications, 2nd Edition*. Wiley.

Bennaceur, Chafia, Carlos Afonso, Sandra Alves, Anne Bossée, et Jean-Claude Tabet. 2013. « Instrumental Dependent Dissociations of N-Propyl/Isopropyl Phosphonate Isomers: Evaluation of Resonant and Non-Resonant Vibrational Activations ». *Journal of The American Society for Mass Spectrometry* 24(8):1260-70. doi: 10.1007/s13361-013-0656-3.

Biemann, Klaus. 1988. « Contributions of Mass Spectrometry to Peptide and Protein Structure ». *Biological Mass Spectrometry* 16(1-12):99-111. doi: 10.1002/bms.1200160119.

Boukerche, Toufik Taalibi, Sandra Alves, Pauline Le Faouder, Anna Warnet, Justine Bertrand-Michel, Mohamed Bouchekara, Mohammed Belbachir, et Jean-Claude Tabet. 2016. « Atypical Cleavage of Protonated N-Fatty Acyl Amino Acids Derived from Aspartic Acid Evidenced by Sequential MS3 Experiments ». *Amino Acids* 48(12):2717-29. doi: 10.1007/s00726-016-2286-0.

Chu, Ivan K., Chi-Kit Siu, Justin Kai-Chi Lau, Wai Kit Tang, Xiaoyan Mu, Cheuk Kuen Lai, Xinhua Guo, Xian Wang, Ning Li, Yu Xia, Xianglei Kong, Han Bin Oh, Victor Ryzhov, František Tureček, Alan C. Hopkinson, et K. W. Michael Siu. 2015. « Proposed Nomenclature for Peptide Ion Fragmentation ». *International Journal of Mass Spectrometry* 390:24-27. doi: 10.1016/j.ijms.2015.07.021.

Darii, Ekaterina, Yves Gimbert, Sandra Alves, Annelaure Damont, Alain Perret, Amina S. Woods, François Fenaille, et Jean-Claude Tabet. 2021. « First Direct Evidence of Interpartner Hydride/Deuteride Exchanges for Stored Sodiated Arginine/Fructose-6-Phosphate Complex Anions within Salt-Solvated Structures ». *Journal of the American Society for Mass Spectrometry* 32(6):1424-40. doi: 10.1021/jasms.1c00040.

Gross, Michael L. 1992. « Charge-Remote Fragmentations: Method, Mechanism and Applications ». *International Journal of Mass Spectrometry and Ion Processes* 118-119:137-65. doi: 10.1016/0168-1176(92)85060-D.

Murray, Kermit K., Robert K. Boyd, Marcos N. Eberlin, G. John Langley, Liang Li, et Yasuhide Naito. 2013. « Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013) ». *Pure and Applied Chemistry* 85(7):1515-1609. doi: 10.1351/PAC-REC-06-04-06.

Pérez-Berezo, Teresa, Julien Pujo, Patricia Martin, Pauline Le Faouder, Jean-Marie Galano, Alexandre Guy, Claude Knauf, Jean Claude Tabet, Sophie Tronnet, Frederick Barreau, Maud Heuillet, Gilles Dietrich, Justine Bertrand-Michel, Thierry Durand, Eric Oswald, et Nicolas Cenac. 2017. « Identification of an Analgesic Lipopeptide Produced by the Probiotic Escherichia Coli Strain Nissle 1917 ». *Nature Communications* 8(1):1314. doi: 10.1038/s41467-017-01403-9.

Roepstorff, P., et J. Fohlman. 1984. « Letter to the Editors ». *Biological Mass Spectrometry* 11(11):601-601. doi: 10.1002/bms.1200111109.

Tabet, Jean Claude, Yves Gimbert, Annelaure Damont, David Touboul, et Amina S. Woods. in press. « Combining chemical knowledge and quantum calculation for interpreting low-energy product ion spectra of metabolite adduct ions: Sodiated diterpene diester species as a case study ». *J. Amer. Soc. Mass Spectrom*.

Tabet, Jean-Claude, Yves Gimbert, Annelaure Damont, David Touboul, François Fenaille, et Amina S. Woods. 2021. « Combining Chemical Knowledge and Quantum Calculation for Interpreting Low-Energy Product Ion Spectra of Metabolite Adduct Ions: Sodiated Diterpene Diester Species as a Case Study ». *Journal of the American Society for Mass Spectrometry* 32(10):2499-2504. doi: 10.1021/jasms.1c00154.

Tan, Bo, David K. O'Dell, Y. William Yu, M. Francesca Monn, H. Velocity Hughes, Sumner Burstein, et J. Michael Walker. 2010. « Identification of Endogenous Acyl Amino Acids Based on a Targeted Lipidomics Approach ». *Journal of Lipid Research* 51(1):112-19. doi: 10.1194/jlr.M900198-JLR200.

Thomas, Marion, Lucille Stuani, Ekaterina Darii, Christophe Lechaplais, Emilie Pateau, Jean-Claude Tabet, Marcel Salanoubat, Pierre-Loïc Saaidi, et Alain Perret. 2019. « De Novo Structure Determination of 3-((3-Aminopropyl)Amino)-4-Hydroxybenzoic Acid, a Novel and Abundant Metabolite in Acinetobacter Baylyi ADP1 ». *Metabolomics* 15(3):45. doi: 10.1007/s11306-019-1508-3.

Vizcaino, Maria I., Philipp Engel, Eric Trautman, et Jason M. Crawford. 2014. « Comparative Metabolomics and Structural Characterizations Illuminate Colibactin Pathway-Dependent Small Molecules ». *Journal of the American Chemical Society* 136(26):9244-47. doi: 10.1021/ja503450q.

Wysocki, Vicki H., Mark E. Bier, et R. Graham Cooks. 1988. « Internal Energy Requirements for Remote Site Fragmentation ». *Organic Mass Spectrometry* 23(9):627-33. doi: 10.1002/oms.1210230902.

Zhang, Xi, Shelagh M. Ferguson-Miller, et Gavin E. Reid. 2009. « Characterization of Ornithine and Glutamine Lipids Extracted from Cell Membranes of Rhodobacter Sphaeroides ». *Journal of the American Society for Mass Spectrometry* 20(2):198-212. doi: 10.1016/j.jasms.2008.08.017.

Supplementary Informations

New lipoamino acid from *E.Coli* Bacteria: origin of regenerated fatty acid carboxylate from dissociation of lipo-glutamate anion

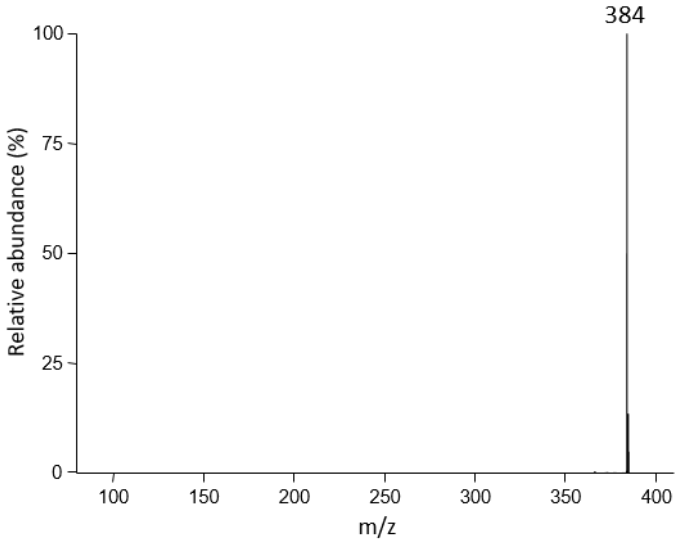


Figure S1. Negative ESI mass spectrum of the C16Glu standard (Mw=385 u)

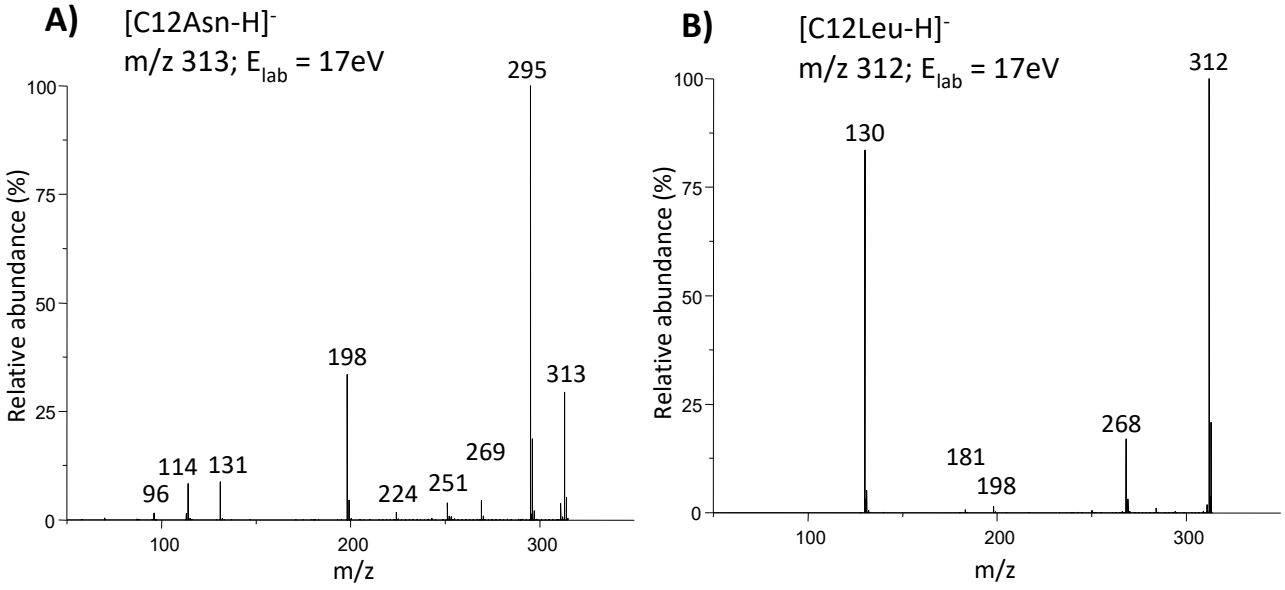
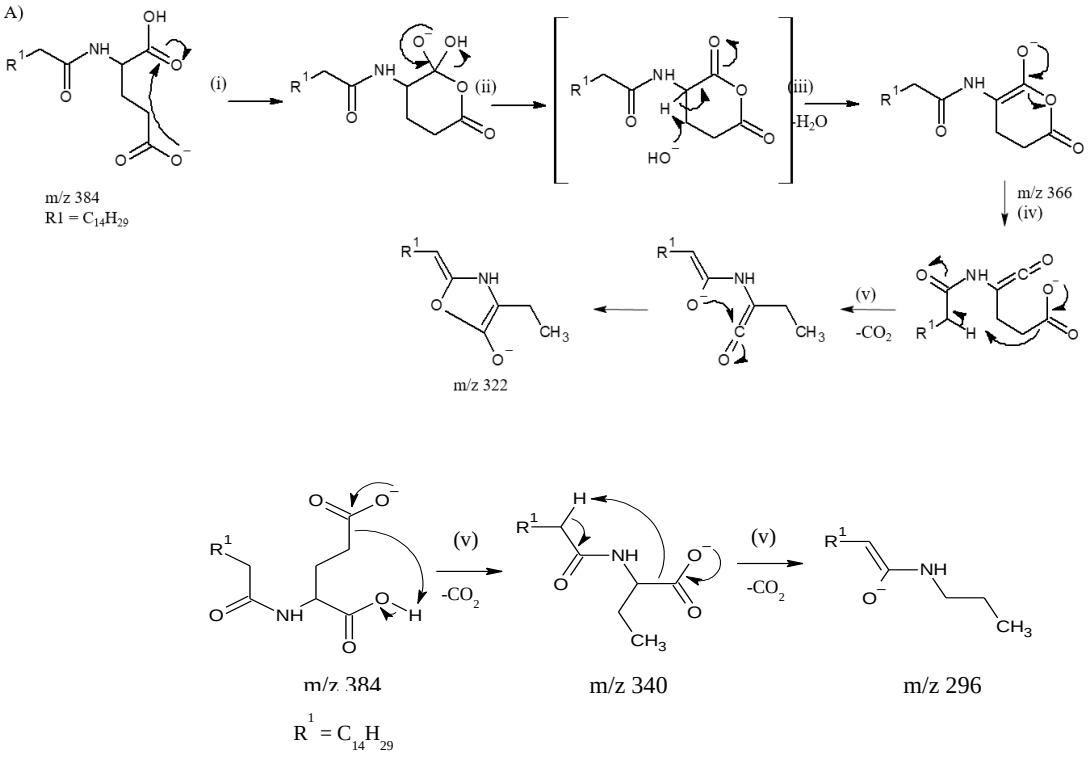
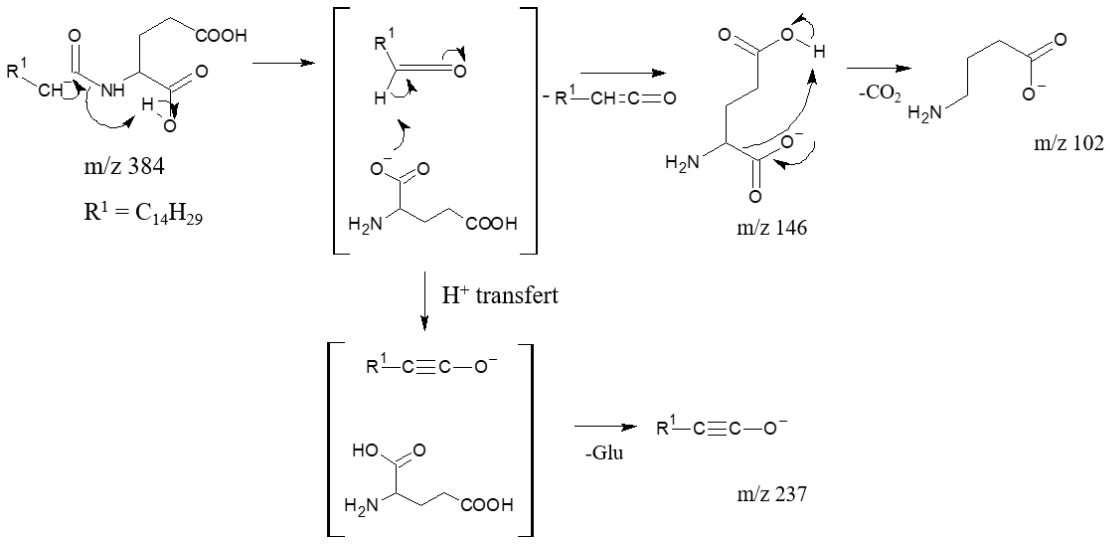


Figure S2. Product ion spectra of A) $[C12Asn-H]^-$ (m/z 313), at E_{Lab} = 17 eV (a), and B) $[C12Leu-H]^-$ (m/z 312) at E_{Lab} = 17 eV (b)



Scheme S1. Possible interpretations of the small size losses from competitive and consecutive dissociations of the $[C16Glu-H]^-$ anion: (a) stepwise processes (i) nucleophilic attack yielding cyclisation and (ii) ion-dipole intermediate resulting in (iii) the H_2O release (ion m/z 366) followed consecutively (iv)+(v) by the CO_2 loss (ion m/z 322); and (b) the direct loss of CO_2 (v) (m/z 340), competitive to the stepwise water loss (a), followed consecutively by a second loss of CO_2 (v) (m/z 296).



Scheme S2 : Possible interpretation of the complementary product m/z 146 and m/z 237 ion pair *via* molecular species isomerization into ion-dipole intermediate yielding either by direct partner splitting, product m/z 146 ion (able to release consecutively CO_2 into m/z 102) or by internal proton transfer and isomerized intermediate splitting in the ynolate m/z 237 ion.

Les informations concernant les calculs quantiques sont en annexe 3.

Publication 3 : Courbes d'évolution des profils de mobilité ionique des ions après collision en fonction de l'énergie d'excitation, un nouveau concept : allures de courbes composites permettant la distinction d'ions produits isomères

Résumé :

Nous avons démontré l'importance de l'analyse des lipoaminoacides en mode d'ionisation négatif, pour l'interprétation des mécanismes de fragmentation, permettant d'obtenir des informations sur les structures des ions fragments. Par exemple l'acide glutamique conjugué à l'acide laurique (C12Glu) après déprotonation donne $[C12Glu\text{-}H]^-$, qui sous CID mène à la production d'une dizaine d'ions fragments attendus et inattendus, comme les ions fragments complémentaires $[y\text{-}H_2O]^-$ et $[Lpb+O]^-$. Afin de mieux comprendre le comportement de ce couple d'ions fragments, les courbes d'évolution d'abondances relatives à l'ensemble des ions produits et de l'ion précurseur en fonction de l'énergie de collision ont été analysées. L'ensemble de ces courbes d'évolution sur un même graphique est appelé couramment « energy-resolved mass spectrometry » (ERMS). Les courbes qui constituent l'ensemble ERMS de l'anion $[C12Glu\text{-}H]^-$ présentent des profils Gaussiens pour la majorité d'entre-elles, comme obtenu à partir d'un même traitement en utilisant les abondances en valeurs absolues. Toutefois, si cette tendance est générale avec les valeurs absolues, ce n'est plus le cas de certains profils notamment des ions m/z 128 et m/z 199 (Figure 95), quand le retraitement des données est fait à partir de leurs abondances relatives comparé au courant ionique total. Ces profils étaient Gaussiens déformés et élargis, ou composites (avec deux composantes ou plus). Ce comportement jamais décrit, semble être indépendant de l'instrument à condition que la sensibilité soit suffisante, que ce soit un tandem à faisceau d'ions ou à piégeage d'ions. Nous avons émis l'hypothèse que la forme élargie pour l'ion m/z 199 était dû à une structure commune bien que formées selon des chemins différents. La forme composite du profil de l'ion m/z 128, est la conséquence de la formation de deux structures isomères produites selon un chemin direct ou par décompositions consécutives dont les seuils de dissociation diffèrent. Pour tester ces hypothèses, les calculs quantiques ont été effectués pour la production des ions m/z 199 et m/z 128. Concernant l'ion m/z 128, les calculs démontrent bien la possibilité de deux structures isomères. Des expériences de spectrométrie de mobilité ionique (IMS ion mobility spectrometry) ont été réalisées et nous avons créé un nouveau concept : la spectrométrie de mobilité ionique résolue en énergie (ERIMS, energy resolved ion mobility spectrometry). Les courbes d'évolution de mobilité ionique en mode cyclique en fonction de l'énergie de collision ont été réalisées. Après 15 cycles, les temps de dérive (mesures de sections efficaces) nous ont permis de mettre en

évidence, pour l'ion fragment m/z 128, deux structures isomères, dont l'intensité des signaux variaient selon l'énergie de collision. Pour l'ion fragment m/z 199, un seul temps de dérive le caractérise bien qu'il soit produit selon des chemins différents. Ainsi la structure de l'ion m/z 199 reste indépendante du processus de dissociation.

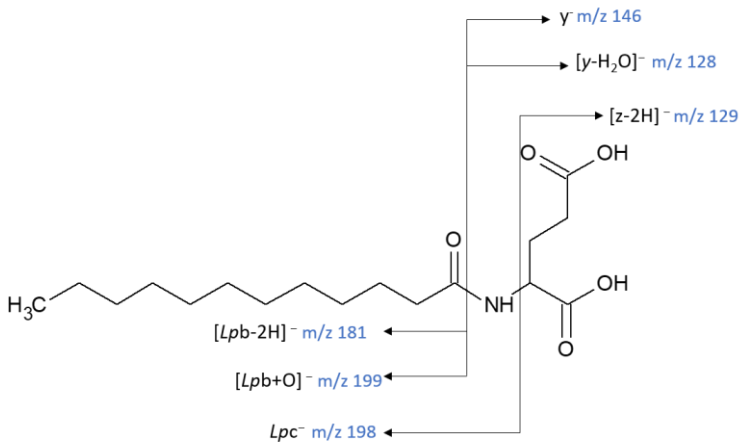


Figure 95 : Annotation des ions fragments produits à partir du C12Glutamate

Résultats détaillés :

Courbes ERMS:

Les courbes ERMS qui représentent l'évolution de l'abondance relative de chaque ion fragment par rapport à l'abondance totale des ions en fonction de l'énergie de collision du $[C12Glu-H]^-$ (m/z 312) et du $[C16Glu-H]^-$ (m/z 384) ont été réalisées à l'aide d'un spectromètre de masse Qq/TOF. Celles-ci ont été réalisées pour l'ensemble des ions fragments caractérisés précédemment. Les courbes ERMS pour les ions fragments correspondant aux pertes d'H_2O, de CO_2, (H_2O + CO_2), (CO_2 + CO_2), celles des fragments acides gras amides et de l'aminoacide déprotoné présentent des profils Gaussiens, et cela à partir des précurseurs déprotonés du C12Glu et du C16Glu. L'ion m/z 102, correspondant à l'acide aminé ayant perdu le CO_2, est caractérisé par une évolution continument croissante, tendant vers un maximum (100%). Par contre, les profils caractérisant la paire d'ions $[y\text{-}H_2O]^-$ et $[Lpb+O]^-$ (Figure 96) sont particuliers : ils sont dissymétriques en «Gaussiennes élargies » ou « composites ». Ces expériences sont parfaitement reproductibles sur deux Qq/TOF Waters Xevo G2-XS et sur un Synapt G2-Si.

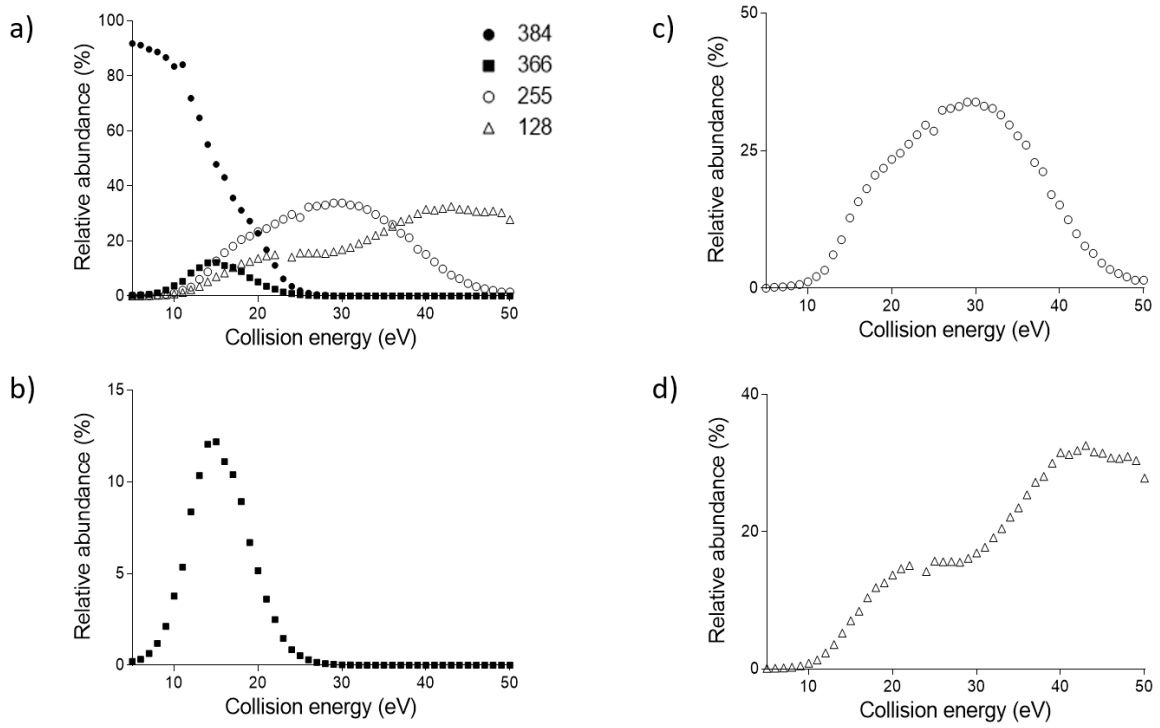


Figure 96 : Courbes ERMS issus de l'anion précurseur [C12Glu-H]- en a), et des ions fragments m/z 310 en b), m/z 199 en c) et m/z 128 en d).

Recherche des précurseurs de l'ion m/z 128 :

Notre première hypothèse pour expliquer la courbe composite est qu'il existe plusieurs chemins de fragmentations différents, l'un impliquant un mécanisme à basse énergie, et un second à plus haute énergie. Nous avons émis l'hypothèse que le premier était issu d'un mécanisme à plusieurs étapes, impliquant des réarrangements, alors que le second était issu d'un processus direct de fragmentation. Afin de valider cette hypothèse, des expériences de recherche d'ions précurseurs ont été réalisées à l'aide d'un triple quadripôle (TQ 8060 Shimadzu). Deux ions forment l'ion fragment m/z 128 : la molécule déprotonée m/z 384 et, l'ion m/z 366 correspondant à la perte de H_2O. Les mêmes résultats sont retrouvés pour l'anion $[C12Glu\text{-}H]^-$ m/z 328, où les ions m/z 328 et m/z 310 sont les précurseurs de l'ion m/z 128.

Largeur à mi-hauteur des courbes ERMS :

La largeur à mi-hauteur (en eV) de chacune des courbes ERMS a été mesurée. Les largeurs mesurées pour les profils Gaussiens et reportées sur un graphique, semblent présenter une tendance linéaire, alors que les ions m/z 102, m/z 128 et m/z 199 (correspondant respectivement aux ions $[(Glu\text{-}H)\text{-}CO_2]^-$, $[y\text{-}H_2O]^-$ et $[Lpb+O]^-$) s'éloignent de cette tendance. Cette droite obtenue dans ce domaine de rapports m/z est définie par la corrélation $y = 0.06428x + 26.97$, avec $R^2 = 0.9411$.

Courbes ER-IMS :

Pour lever l'ambiguïté sur la présence d'une ou de plusieurs structures pour les ions fragments m/z 199 et m/z 128, des expériences de mobilité ionique ont été réalisées en sélectionnant les ions selon différentes énergies de collision. Dans un premier temps en utilisant un Synapt G2-XS, où les ions m/z 199 et m/z 128 montraient chacun une seule valeur de section efficace, quel que soit l'énergie de collision appliquée.

Dans un second temps, les mêmes expériences réalisées à l'aide d'un instrument de mobilité ionique cyclique (Waters). Après 15 cycles dans la cellule de mobilité, deux valeurs de sections efficaces (DT = 205 et 212 msec) ont pu être attribuées pour l'ion m/z 128 (Figure 97). L'évolution de l'énergie de collision permet de voir l'évolution des deux composés pour le m/z 128.

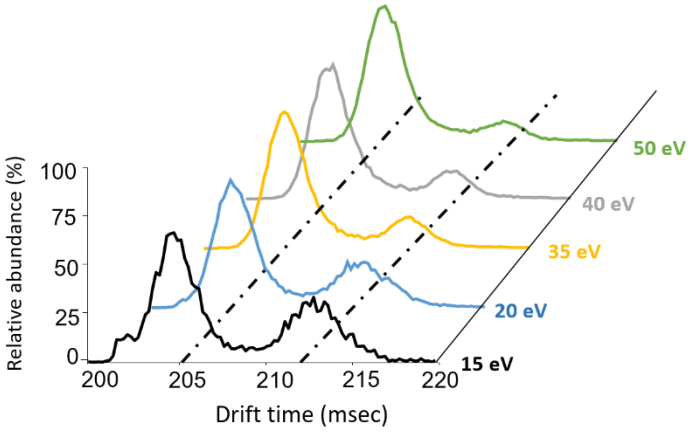


Figure 97 : Profils d'évolution de l'abondance relative de l'ion fragment m/z 128 issu du précurseur $[C12Glu-H]^-$ en mobilité ionique en fonction de l'énergie de collision

L'ion m/z 199 présente toujours une seule valeur de collision efficace (DT = 351 msec), indépendamment de l'énergie de collision appliquée (de 0 à 50 eV, Figure 98).

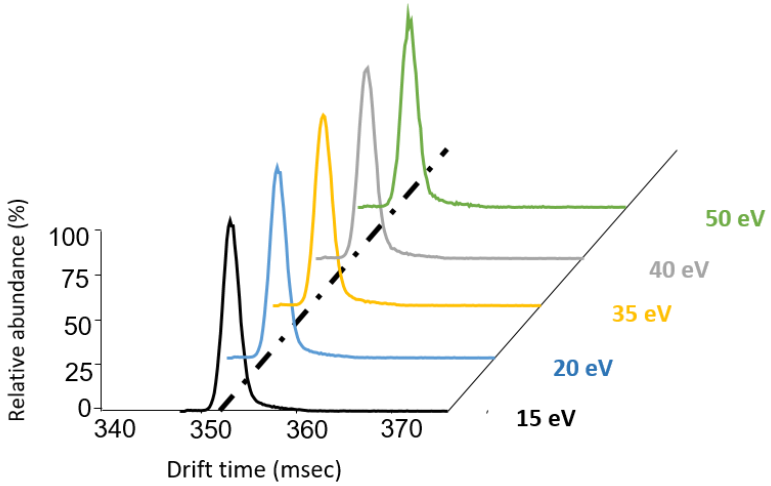


Figure 98 : Profil d'évolution de l'abondance relative de l'ion fragment m/z 199 issu du précurseur $[C12Glu-H]^-$ en mobilité ionique en fonction de l'énergie de collision

Afin de d'assurer que nous n'étions pas en présence d'un mélange d'ions précurseurs isomères (m/z 328) et que la conformation du précurseur est unique selon l'énergie de collision une expérience par mobilité ionique cyclique avec 90 cycles dans la cellule de mobilité a été réalisée, et une seule valeur de section efficace a été détectée pour l'ion précurseur.

Conclusion :

Dans cette étude, les courbes d'évolution des profils en fonction de l'énergie ont démontrées pour la première fois des profils « Gaussiens élargies » ou « composites » pour un couple d'ions complémentaires, respectivement les ions $[Lp\text{b+O}]^-$ m/z 199 , et $[\text{y-H}_2\text{O}]^-$ m/z 128. Par mobilité ionique cyclique, il a été possible de mettre en évidence la présence de deux isomères aux sections efficaces différentes pour l'ion m/z 128, permettant de mettre en avant la présence de deux isomères produits selon deux chemins de fragmentation différents pour lesquels nous avons proposé deux mécanismes de fragmentation. Les mêmes expériences ont également été réalisées sur l'ion m/z 199, cependant il n'a pas été possible de confirmer la présence de deux structures, permettant une meilleure compréhension du profil Gaussien élargi obtenue.

Abstract: Tandem mass spectrometry based on collisional activation (in non-resonant mode) is commonly used to determine molecular structures. Energy resolved mass spectrometry (ERMS) is useful to elucidate the mechanism(s) of unexpected fragment ion formations such as the regeneration of fatty acid carboxylate and its complementary ion (dehydrated glutamic acid anion) from deprotonated N-acyl glutamic acid [1]. The absolute abundance evolution of these ions according to the collision energy from the ERMS breakdown of this precursor ion usually shows Gaussian-like profiles. This trend remains generally true for the product ion abundances relative to the total ionic current (TIC) and their width at half maximum of these Gaussian profiles increases as the product ion m/z ratios decrease. In some case the breakdown profiles of studied unexpected ions become broad or composite. Being never described, they are used to demonstrate that these complementary fragment ions are generated by two (or more) distinct pathways having different appearance thresholds. Since it is not clear whether these pathways lead to product ions with common or isomeric structures, this approach combined with cyclic ion mobility spectrometry (cIMS) removes this ambiguity. This allows to show that each one of the double pathways leads to a single structure for the deprotonated fatty acid and isomeric forms for the dehydrated glutamate. This coupling, available for instruments having a collision cell in front to the ion mobility drift tube, introduce a new concept: the energy resolved ion mobility spectrometry (ER-IMS), a powerful tool for structural elucidation.

Introduction: In the 1970's, during the early days of tandem mass spectrometry based on double sector instrument (electric (Eo) and magnetic (B) sectors with both the fixed and scanned fields), the product m_i^+ ion resulting from spontaneous dissociation of precursor $M^{+\bullet}$ ion (accelerated by a fixed high voltage Vo) along its flight to the detector and especially in the field-free-regions (FFR) of instrument was detected as metastable ion. This ion characterized by its kinetic energy, $E_{kin,i}$ ($E_{kin,i} = zeVo(m_i/M)$) was displayed in the Ion Kinetic Energy (IKE) spectrum [2],[3] using E/B double sector Nier-Johnson geometry or in the Mass-Analyzed IKE (MIKE) spectrum by E scanning (after precursor ion selection) using reversed B/E geometry [4]. The product ion is detected as a signal of which the width and its shape [5] depend on the internal energy excess required above the dissociation transition state, TS (Transition state plus the kinetic shift related to the instrument) to reach the energy threshold, $E_{Threshold}$, combined with the reversed energy barrier (related to its mechanism) [5],[6]. The signal width results from the conversion of some of the excess internal energy (essentially in terms of vibrational energy carried out by the product ion) into translational energy called kinetic energy release (KER). The different shapes and KER values [4],[5] characterizing the signals can be summarized for the most common, by the following characteristics. First for product ion formed from simple cleavage (negligible reversed energy barrier), the corresponding signal can be relatively narrow with a Gaussian shape contrasting with those generated from characterizing rearrangement processes for which the significant reversed energy barrier results in a large increase in signal width [8]. Consequently, the signal appears strongly broader with a “flat-topped” (or “dished”) shape [8],[4],[8]. Finally, if several channels give common product ions (isomeric or not), their corresponding KER components are superimposed and a symmetrical composite signal results. The shape of these signals has been widely used to interpret

fragmentations and to specify TS, especially for stepwise processes [7]. High energy collision activation (in the keV range) is also used and is called as MIKE-collision-induced dissociation (MIKE-CID) [10]. Since the advent of quadrupole instruments at the end of the 70's, energy-resolved mass spectrometry (ERMS) breakdown curves [11] of a selected precursor ion, collisional activated either in resonant or non-resonant excitation mode under low energy conditions (modes called as CID), have been developed to gradually replaced MIKE spectra and KER measurements while sector instruments were replaced by smaller benchtop instruments. However, they do not provide the same information of MIKE signals but, are complementary. The ERMS profile consists of a set of breakdown curves, each curve being related to evolution of a product ion m_i abundance *versus* collision energy. ERMS profiles [9] allow (i) to distinguish isomeric precursor ions [9],[10], (ii) to contribute to a better understanding of dissociations even those taking place along complex pathways [14] and (iii) to determine fragmentation $E_{Threshold}$ of selected ions [15], [16]. However, to our knowledge, a given product ion has never been demonstrated to be generated by different mechanisms from only the direct analysis of a ERMS profile of a selected precursor ion.

In this study, the behavior of a deprotonated lipoamino acid, glutamic acid conjugated to a C16 fatty acid (C16Glu), towards low energy collisions gives us the opportunity to explore ERMS profiles in which a composite evolution has been highlighted for certain fragment ions. Recently, dissociations resulting from non-resonant collisional activation of the $[C16Glu\text{-}H]^-$ anion were studied using a Qq/TOF instrument [17]. Under low energy CID conditions, among the ten product ions observed from the $[C16Glu\text{-}H]^-$ anion (m/z 384), the unexpected fatty acid m/z 255 carboxylate [18], $[(C16Glu\text{-}H)\text{-}(Glu\text{-}H_2O)]^-$ annotated as $[Lp\text{b}+O]^-$ (Scheme S1) [17] mainly appeared. In addition, its complementary anion species as dehydrated glutamic m/z 128 carboxylate ($[(Glu\text{-}H)\text{-}H_2O]^-$ annotated as $[y\text{-}H_2O]^-$) was detected within less abundance (Figure S1a,b). The $[Lp\text{b}+O]^-$ ions are similar to the $[\text{b}+O]^-$ product ion generated from dissociation peptide containing glutamic residue [19], [20].

Experimental section:

Chemicals and Reagents: Formic acid and analytical solvents (acetonitrile and methanol) were purchased from Fischer Scientific (Thermo Fisher Scientific, Illkirch, France). Ultrapure water was generated from a Milli-Q system (Millipore, Saint-Quentin, Yvelines, France). N-palmitoyl acyl and N-lauroyl acyl group linked to glutamic acid were synthetized as described in [17].

Tandem Mass Spectrometry analysis: The compounds were analyzed by mass spectrometry in electrospray source in negative mode by direct infusion using a Hamilton syringe, at 5µL/min flow rate, at the concentration of 10ng/µL. Analysis were performed on a Xevo^{TM} G2-XS mass spectrometer (Waters, Wimslow, UK), source conditions were: gas temp was set at 120°C; nebulizer gas flow was set at 400 L/h; cone gas flow was set at 40 L/h; desolvation temp was set at 380°C; resolution R=16 000 at m/z 400. Analysis was performed on a SYNAPT^{TM} G2-XS mass spectrometer (Waters, Wimslow, UK); source temperature was set at 100°C; nebulizer pressure was set at 6.5 Bar; cone gas flow was set at 50 L/h, desolvation temp was set at 200°C, Desolvation gas flow 600L/hr). CID was performed in the trap cell of the TriWave region. Precursor ion scan analysis were performed on a triple quadrupole tandem 8060 (Shimadzu), experimental

source conditions were: gas temp, 325°C; gas flow, 10 L/min; sheath gas temp, 400°C; sheath gas flow, 12 L/min; capillary, 3500 V; voltage nozzle, 2000V; fragmentor, 135 V. MS/MS experiments were performed on a linear trap quadrupole (LTQ) Orbitrap Velos mass spectrometer (ThermoFisher Scientific, Bremen, Germany), experimental conditions: source voltage was set at 2.7 kV, capillary temp was set at 300°C, sheath gas flow was set at 9 L/min. Product ion spectra (from m/z 50 to 400) were acquired in Orbitrap FT/MS with a resolution R=30,000 at m/z 400. ERMS experiment conditions: selected m/z 384 and m/z 328 precursor ions were resonantly excited during 10 ms at $q_{z,exc}$=0.20 for normalized collision energy (NCE) was ramped from 0% to 50%. For acquisition, Xcalibur 2.0 software (ThermoFisher Scientific, Bremen, Germany) was used).

Cyclic Ion mobility spectrometry

Ion mobility analysis of product ions was performed using SELECT SERIES™ Cyclic™ IMS (cIMS) Qq/TOF instrument using electrospray ionization in negative mode. Samples were introduced by direct infusion using a Hamilton syringe, at 10µL/min flow rate, at the concentration of 10 ng/µL. Ion source condition were: ion source temperature was set to 100°C; nebulizer gas flow was set at 400 L/h; cone gas flow was set at 50 L/h, desolvation temperature was set at 200°C, This instrument has a Q-cIM-TOF geometry, as described elsewhere [21]. The precursor m/z ion was selected using the quadrupole mass filter prior to the IMS separator and subjected to CID in the trap ion guide (non-resonant lab frame collision energy from 0 eV to 50 eV). Product ions where then injected into the cIM device operating at 1.75 mbar of N_2, with a travelling wave height of 10 V and wave velocity of 375 m/s. Ions may be subjected to different numbers of passes of the cIMS separation device with more passes resulting in higher IMS resolution. The cIMS cell is 98 cm in length and a single pass of the device results in a resolution (CCS/ΔCCS) of ~ 65. 15 passes around the cIM yield's a resolving a power of ~250.

Collision Cross Section calculation

Preliminary structures were obtained in Avogadro [22] using the MMFF94 force field. Structures were then fully optimized in Gaussian 16 [23] using the B3LYP functional and the 6-31++G(d,p) basis set. Partial charges were assigned using the Merz-Singh-Kollman scheme, constrained to reproduce the dipole moment. CCS values were calculated using IMoS v1.10 [24] for both of them using the "Trajectory Method" [25]. IMoS software utilizes a variant of the trajectory method that has been adapted for use with N_2 gas [26], [27] and further tuned against Travelling Wave Ion Mobility data [28]. Two calculation methods were used: the first - "Trajectory method (no ion-induced quadrupole)" - models the dipole induced on the N_2 molecules by the ion, while the second - "Trajectory method (ion-induced quadrupole)" - additionally models the induced quadrupole.

Results and discussion:

Various origins of the [*Lp*b+O]$^-$ and [*y*-H_2O]$^-$ product ions from [C16Glu-H]$^-$ and [C12Glu-H]$^-$: the precursor ion spectra of these complementary m/z 255 ([*Lp*b+O]$^-$) and m/z 128 ([*y*-H_2O]$^-$) ions were performed using triple quadrupole tandem under non-resonant conditions (Figure S1). These spectra shows

that $[C16Glu\text{-}H]^-$ is the unique precursor ion of m/z 255 (Figure S2a) and that the complementary m/z 128 ion is generated competitively from the $[(C16Glu\text{-}H)\text{-}H_2O]^-$ (m/z 366) and $[C16Glu\text{-}H]^-$ (m/z 384) ions (Figure S2b). Nevertheless, detection of m/z 384 as precursor did not necessarily mean that it was a direct precursor ion of m/z 128. Indeed, if very early in the collision cell, m/z 384 rapidly decomposes in the m/z 366 ion by a loss of water, which itself promptly dissociates into m/z 128 by efficient additional collision. Then m/z 384 ion is detected as a precursor ion instead of m/z 366 which no longer appears in the precursor ion spectrum of m/z 128. Using resonant excitation in ion trap cell, it is possible to evidence the exclusive occurrence of direct competitive dissociations when they are endothermic. In contrast, the consecutive pathways can be detected exclusively if they are exothermic [29], [30]. Resonant excitation on the stored m/z 384 precursor ion (Figure S3a, b) resulted mainly in the endothermic formation of the m/z 366, m/z 255, and m/z 128 ions, strictly provided from the competitive dissociations of the m/z 384 ion. Abundance of the m/z 340 due to the CO_2 loss is too close of the noise to be observed in the product ion spectra. By increasing excitation energy, the relative abundances of the product m/z 128 and m/z 255 ions similarly increased relatively to the selected m/z 384 ion abundance (Figure S3). Finally, if the m/z 128 ion was directly produced from m/z 384 in both the resonant and non-resonant modes, the contribution of the m/z 366 product ion in the m/z 128 formation should be considered exclusively in the later excitation mode. Thus, the double origin of the m/z 128 ion (Figure S2) may explain why the relative abundances of both m/z 255 ($[Lp\text{b+O}]^-$) and m/z 128 ($[y\text{-}H_2O]^-$ product ions were not significantly affected by the collision energy increased from E_{Lab}=17 eV to E_{Lab}=30 eV (Figure S1). Indeed, if the m/z 255 and m/z 128 product ions were competitively formed only from a stepwise dissociation of $[C16Glu\text{-}H]^-$ (m/z 384) isomerized into ion-dipole complex ([1] Figure S4a), then the direct complex splitting resulting in the $[Lp\text{b+O}]^-$ ion must be preferred at the highest energies, in contrast to $[y\text{-}H_2O]^-$ generated by rearrangement through internal proton transfer between partners of the ion-dipole intermediate (Figure S4a) [31]. This mechanism has been corroborated by modeling density functional theory (DFT) level calculations that detailed these competitive stepwise pathways shown as endothermic [1]. Thus, the $[y\text{-}H_2O]^-$ ion would decrease relative to that of $[Lpb+O]^-$ with increasing collision energy, which is not experimentally shown, meaning that another dissociation pathway must contribute to the dehydrated glutamate $[y\text{-}H_2O]^-$ formation.

A similar behavior was observed for the dissociations of $[C12Glu\text{-}H]^-$ (m/z 328), homologue of the deprotonated C16Glu, since the complementary product m/z 199 and m/z 128 ion pair was abundantly displayed in CID spectra recorded at 14 eV and 27 eV (Figure S1c,d). To compare the CID spectra of homologues, the E_{Lab} values were chosen to roughly correspond to the energy at the center of mass ($E_{com}=E_{Lab}xm_A/(m_{Ar}+384)$ with E_{Lab}=17 eV and 30 eV) of the activated $[C16Glu\text{-}H]^-$ ion. It was expected that the dehydrated glutamate m/z 128 formed by internal proton transfer through prototropy [28],[29] between the ion-dipole intermediate partners (by analogy with mechanisms proposed for dissociations of $[C16Glu\text{-}H]^-$, Figure S4a), should disappear by increasing E_{Lab} in contrast to the fatty acid m/z 199 carboxylate produced by direct splitting of the same ion-dipole intermediate. By increasing E_{Lab} from 14 eV to 27 eV, their relative abundances were not strongly modified as observed for the $[C16Glu\text{-}H]^-$ ion (Figure S1a,b). On the other hand, the m/z 199 ion (as well as m/z 255 from $[C16Glu\text{-}H]^-$) provided from

regeneration of fatty acid carboxylate ([*Lp*b+O]$^-$) associated to its complementary dehydrated amino-acid product anion ([*y*-H_2O]$^-$) should be from two or more dissociative pathways.

A 7-membered tetrahedral intermediate (Figure S4a) is formed at the lower energy from the side chain carboxylate group to competitively yield m/z 128 (form **a**) and m/z 199 (or m/z 255). This selectivity was evidenced by dissociation of its methyl ester derivative at the side chain ([C16GluMe-H]$^-$,[1]) which exhibits a strong hindrance to the m/z 255 formation (less than 1% at E_{lab}=17 eV which increases to 9% at 30 eV) compared to its abundance from [C16Glu-H]$^-$ as the greatest of the product ions at these collision energy. This can mean that the α carboxylate group may play a weak role in the m/z 255 formation through a possible 5-membered tetrahedral intermediate (Figure S4b) at E_{lab}=30 eV. Alternatively, m/z 128 (within the same form **a**, Figure S4b) due to internal proton transfer must be hindered and not considered under these collision energy conditions. A same orientation characterizes the fatty acid carboxylate (m/z 199, [*Lp*b+O]$^-$) regeneration and the dehydrated glutamate (m/z 128, [*y*-H_2O]$^-$) formation from the [C12Glu-H]$^-$ (m/z 328) precursor ion (Figure S4a,b).

The m/z 310 (or m/z 366) ion ([(C12Glu-H)-H_2O]$^-$) yielding consecutively the m/z 128 ion ([*y*-H_2O]$^-$) must result by water loss from the molecular deprotomer with an amidate structure (Figure S4c) rather than the carboxylate deprotomer considered for the above competitive m/z 199 (or m/z 255) and m/z 128 product ions (Figure S4a). To dissociate the nitrogen atom of amidate group may yield nucleophilic attack at the nearest carboxylic acid group allowing the 5-membered cyclisation and the tetrahedral intermediate formation (Figure S4c). From this intermediate, water release occurs after formation of the OH^-/neutral complex to result in m/z 366 ion carrying the negative charge at the amide enolizable position. Through fatty acid ketene loss, the m/z 366 ion possibly dissociated to give rise to formation of the dehydrated m/z 128 glutamate (with γ lactam ring, Figure S4c) either by proton transfer (deprotomer **b**, expected favored at lower energy) or by simple cleavage (deprotomer **c**, favored at higher energy) (Figure S4c).

Finally, the regeneration of the fatty acid carboxylate ([*Lp*b+O]$^-$) associated to its complementary dehydrated amino-acid ([*y*-H_2O]$^-$) characterized the conjugated acidic amino-acids. However, for the conjugated aspartate, this process occurred in competition with the fatty acid amidate formation. Finally, [*Lp*b+O]$^-$) and [*y*-H_2O]$^-$ were not formed from other conjugated basic and non-polar amino-acid anions which mainly formed the complementary fatty acid amidate and deaminated glutamate anions [1], [17]. The pertinence of the proposed mechanisms to explain the different formations of this product ion pair will be evidenced below.

Different pathways yielding the [Lpb+O]$^-$) and [y-H_2O]$^-$ ion confirmed from the ERMS breakdown of [C12Glu-H]$^-$: In order to obtain more information about the competitive dissociations responsible to the [Lpb+O]$^-$) and [y-H_2O]$^-$ ion origins, ERMS breakdown of the [C12Glu-H]$^-$ precursor ion was performed under non-resonant excitation conditions, first using absolute ion abundances (Figure 1a). The shapes of the breakdown curves of product ions are similar to more or less symmetrical Gaussian-like profile (Figure 1b, c and d). Furthermore, it shows that beyond E_{lab}=35 eV, the larger absolute abundance of product ion does not exceed 5% of initial abundance of the precursor ion. The loss of product ion signals at higher collision

energies is attributed to a reduced transmission resulting from their dispersion occurring at higher kinetic energies.

The abundance relative to the total ionic current (TIC) of the ions contributing to the product ion spectrum was also used to highlight their relative evolution with E_{lab} allowing a better comparison of the shape and width of their respective profile (Figure 1e). Besides the most abundant ion at m/z 102 ($[y\text{-}CO_2]^-$, Figure 1e), only m/z 199 ($[Lpb+O]^-$) and m/z 128 ($[y\text{-}H_2O]^-$) have a relative abundance greater than 15% of TIC of the surviving product ions corresponding to about 6% of the initial TIC measured at E_{Lab}=5 eV.

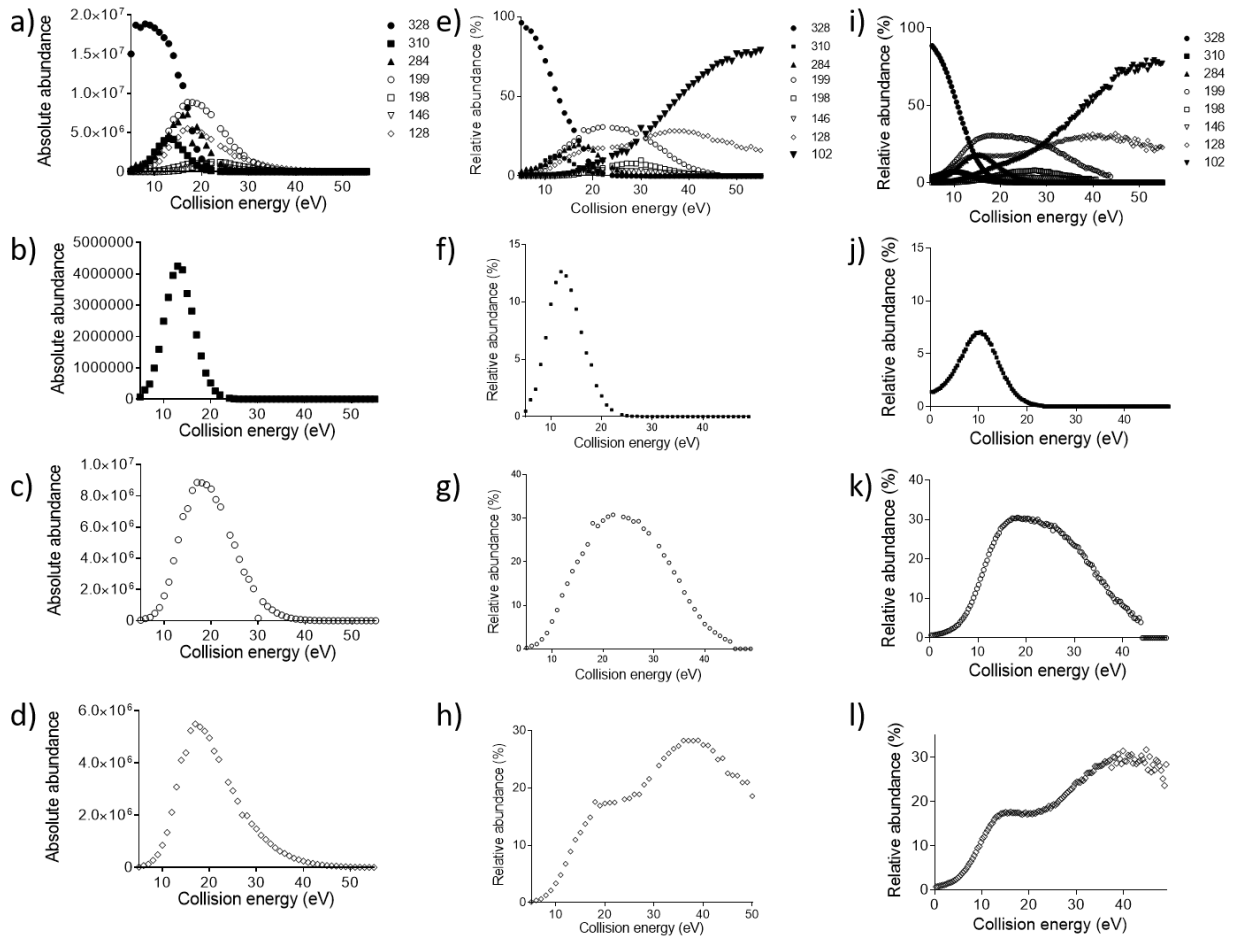


Figure 1. ERMS (a), (e), (i) of the precursor $[C12Glu\text{-}H]^-$ ion (m/z 328), and breakdown profiles of product ions: (b),(f),(j) m/z 310 (water loss), with a Gaussian-like shape, (c),(g),(k) m/z 199 ($[Lpb+O]^-$), with broader Gaussian-like shape and (d),(h),(l) m/z 128 ($[y\text{-}H_2O]^-$), with a multi-component shape. In the ERMS and product ion breakdown profiles, the ion abundances reported in the left column [(a), (b), (c), (d)] are given in absolute values (using the Xevo G2-XS mass spectrometer); abundances are relative to TIC in the middle column [(e), (f), (g), (h)] (same experiment of left column), and the right one [(i), (j), (k), (l)] (experiments using the Synapt G2-XS instrument),

The breakdown curves of each product ions are too Gaussian-like (e.g. m/z 310, Figure 1f) except for certain as m/z 199, and m/z 128 whom the profile shapes are broad Gaussian (more or less symmetric) (Figure 1g), and composite signal (Figure 1h), respectively. Note that these profiles are similar from one instrument (Figure 1g,h) to another (Figure 1k,l) despite the low TIC beyond 35 eV. This composite signal can be

considered as two partially overlapping Gaussian-like components which can be deconvoluted into two components. Each of the product ion profiles are characterized by their respective full width at half maximum (FWHM) and apex (collision energy, $E_{Lab,apex}$ and relative intensity, Table 1). Except a few measurements which deviated to the general trend, the statistical precision of the product ion breakdown profiles is sufficient to give reproducible measurements. However, the abundance of m/z 181 and m/z 84 was insufficient to accurately obtain the breakdown profiles above E_{Lab}=45 eV, and therefore they were not included in these results.

Table 1. Evolution of the characteristics[a] (profile shape, FWHM, collision energy at the apex, $E_{Lab,apex,}$ and its normalized and absolute intensities) of product ion breakdown profiles from ERMS of the $[C12Glu-H]^-$ ion (m/z 328)[b]

Product ion (m/z)	Elemental composition	Neutral loss	Profile shape	Apex intensity:		FWHM (eV)	$E_{Lab,apex}$ (eV)
				Absolute (arbitrary unit)	Normalized (%)		
310,2023	$C_{17}H_{28}NO_4$	H_2O	Gaussian	3956770	12.6	7	12
284,2231	$C_{16}H_{30}NO_3$	CO_2	Gaussian	7353592	19.9	11	17
266,2140	$C_{16}H_{28}NO_2$	$(H_2O + CO_2)$	Gaussian	794143	3.6	10	22
240,2350	$C_{15}H_{30}NO$	$(2CO_2)$	Gaussian	2128050	20.0	13	27
199,1718	$C_{12}H_{23}O_2$	$C_5H_7NO_3$	Broad[c]	61545688	30.7	24	23
198,1873	$C_{12}H_{24}NO$	$C_5H_6O_4$	Gaussian	842783	7.9	14	27
146,0458	$C_5H_8NO_4$	$C_{12}H_{22}O$	Non-Gaussian	652903	3.7	19	22
128,0353	$C_5H_6NO_3$	$C_{12}H_{24}NO_2$	multi-component[d]	302868	28.3[d]	40[d]	39[d]
102,0565	$C_4H_8NO_2$	$C_{13}H_{22}O_3$	[e]	39283	78.0	20[e]	55[e]

[a] When the intensity at half height measured on the negative slope of Gaussian signal is less than 5% of TIC for a corresponding E_{Lab} higher than 45 eV then, the accuracy of the estimated FWHM is poor, as for m/z 181 and m/z 84 then, not considered

[b] Sigmoid inflexion point at 12.5 eV for the survivor $[C12Glu-H]^-$ ion

[c] Broad Gaussian shape

[d] Two or more components may contribute to this broad signal, the intensity and apex of which are artificially measured from the greatest height of the broad signal. However, FWHM was measured over the whole signal.

[e] Monotonically increasing intensity starting from 10 eV to 55 eV reaching the height-half of the increase at 35 eV. The FWHM was 20 eV formally measured between the height-half of profile at 35 eV and 55 eV

The plot as a function of collision energy of the evolution of the FWHM value of each of the profiles related to the product ions (Figure 2) indicated that the lower m/z value was associated with the higher FWHM value. In particular, the Gaussian shape profile FWHM (empty circle ○, Figure 2), as a function of m/z, decreased as a quasi-linear manner with negative slope in this m/z ratio range. FWHM of the breakdown profile related to m/z 146 being weakly accurate, it is omitted.

As the two other values were far from this linear trend (solid circle ●, Figure 2), the corresponding product ions (m/z 199 and m/z 128) were formed by two or more mechanisms. Each of the components of the m/z 128 ion profile (deprotonated glutamic anhydride, $[y\text{-}H_2O]^-$) could be associated to a particular dissociation (either simple cleavage *vs.* rearrangement or direct mechanism *vs.* consecutive decompositions). Thus, this product ion could be characterized either by a common structure or by isomeric forms depending on the mechanism. The shape of the breakdown curve of the very abundant product ion at m/z 199 (regenerated fatty acid carboxylate, $[Lpb+O]^-$ likely with one or two structures resulting through different mechanisms) was broader (24 eV) than the other Gaussian profiles (Figure 2).

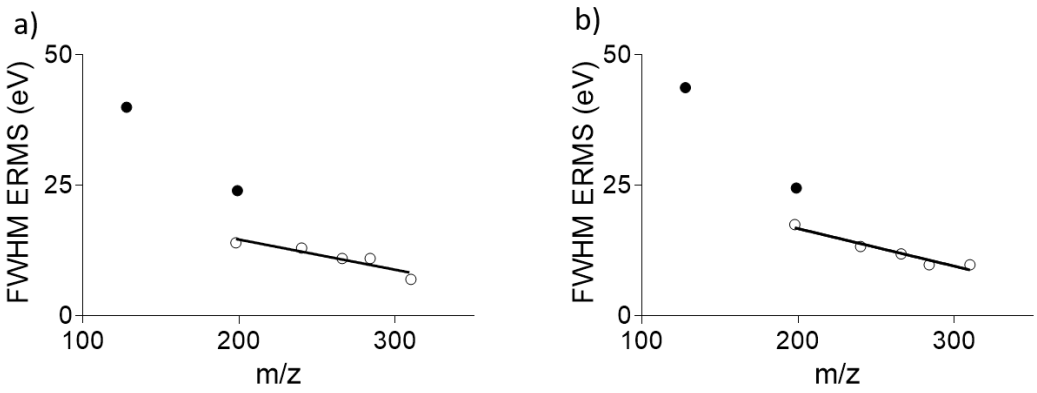


Figure 2. Evolution of the FWHM of the breakdown curves of the ERMS profiles *versus* the m/z ratio of corresponding product ion of $[C12Glu\text{-}H]^-$, m/z 328 (empty circle for Gaussian shape ○, solid circle for broad or composite signal ●): a) measured with the Xevo G2-XS (Qq/TOF) of the lipidomic platform, the linear dependence as y= -0.0579x + 26.2, with $R^2 = 0.85$; and b) measured on a Synapt G2-XS (IMS instrument), the linear dependence as y= -0.0720x + 31.2 with $R^2 = 0.93$)

The consistence of the $[C12Glu\text{-}H]^-$ results (Gaussian-like shape, ○) provided from a Qq/TOF Xevo G2-XS and a Synapt G2-XS illustrated in Figure 2 indicates similar trend toward collisions. This suggests that the ERMS are independent to the collision cell device (RF only segmented quadrupole and the stacked ring ion guide) working in non-resonant collision mode. In addition, the small variations of (i) source conditions (desolvation voltage) can affect internal energy carried out by the precursor ion [34] and eventually its tertiary structure (hydrogen bond *vs.* salt bridge) [35] and (ii) lifetime of survivor ions in collision cell influences the kinetic shift. Consequently, the competitive dissociation rate constants can be modified and thus, the relative contribution of pathway for each one of the m/z 199 and m/z 128 ions resulting in the shape symmetry change. Indeed, the more or less symmetric breakdown profile related to the broad Gaussian shape of m/z 199 can be rationalized by considering the different pathways occurring under variable contributions without presuming the structures of m/z 199 (isomers or common). This may account for the small differences in the breakdown profile of m/z 128 provided from the two used instruments.

Interestingly, ERMS of the homologous $[C16Glu\text{-}H]^-$ ion (m/z 384) provided from two platforms are very similar (Figure S5(a),(e)) and are comparable to that of $[C12Glu\text{-}H]^-$(Figure 1(e)). Especially, breakdown profiles of m/z 366 (Figure S5(f)), and m/z 255 (Figure S5(g)) are respectively comparable to those of the

homologous m/z 310 (Figure 1f) and m/z 199 (Figure 1g) product ions of $[C12Glu\text{-}H]^-$ (m/z 328), as well as the common m/z 128 ion from $[C16Glu\text{-}H]^-$ and $[C12Glu\text{-}H]^-$ (Figure S5(h) and Figure 1(h)). The set of these results indicated that the breakdown signal shape is independent to the fatty acid moiety size.

Close energy thresholds for the competitive $[^{0,2}A_3+O]^-$ and $[y\text{-}H_2O]^-$ ion formations directly from $[C12Glu\text{-}H]^-$: it is expected that the product ion spectra, will be markedly different when recorded under resonant excitation conditions. Consequently, ERMS breakdown curves are also expected to be different from those generated by non-resonant conditions (known as "fast heating" processes yielding prompt dissociations after one efficient collision) [36]. Indeed, under resonant excitation conditions, the ERMS breakdowns were completely modified (Figure S6) with the product ion breakdown profile shapes as sigmoidal. Such a difference was expected due to a very large number of very low energy collisions (competitive activation/relaxation called as "slow heating") [36] occur to reach ion dissociation threshold. As a consequence, only few fragment ions at m/z 310, m/z 284, m/z 199 and m/z 128, with positive sigmoid shape, were formed directly from $[C12Glu\text{-}H]^-$. When 10% of precursor ion population survived from activation, the maxima of the product ion sigmoid curves were reached at a common collision energy, NCE=27.5%. This means that as soon as all the precursor ion population is consumed (or resonantly ejected from the Dehmelt pseudo-potential well) [12], the product ions are relaxed and stored in constant relative abundances. Interestingly, the estimated $E_{threshold}$ (normalized collision energy [37], NCE, given in %), of each product ion can be roughly determined by the log (product ion relative abundance) *vs.* the NCE values (not reported) corresponding to 15% and 19.5% respectively, for m/z 310 (H_2O loss) and m/z 284 (CO_2loss) while $E_{threshold}$ was roughly similar (i.e., 17.3%) for the m/z 199 and m/z 128 ions. This common values for $E_{threshold}$ of formation of these two ions suggested that they are generated from a common intermediate from the precursor m/z 328 ion isomerization prior to dissociation.

The FWHM evolution of the Gaussian profiles could be fortuitous unless they could be directly related the KER associated to each fragmentation. This is unlikely since the water loss yielding m/z 310 is a stepwise process requiring proton transfer [20], whereas the CO_2 loss (m/z 284 formation) involves rather a direct cleavage. The first process must be accompanied by a larger back energy barrier (a non-negligible KER) than for the second process with a most likely lower KER, yet the profile associated with the m/z 310 ion is narrower than that observed for the m/z 284 product ion. Thus, the FWHM evolution must be attributed to another origin than KER associated to the fragment ions. A study underway to explain this trend is based on a simplistic model: the lower the fragment ion mass compared to its associated neutral during the dissociation of a given precursor ion, the more the fragment ion is deviated from its initial trajectory during its collision. In other words our results could be related to those obtained on the evolution of the fragment ion abundances in MIKE spectra according to the detection angle [40]. The discrimination can be applied for energy resolved of isomeric m/z 128 structures originate either directly from $[C12Glu\text{-}H]^-$ or through the $[(C12Glu\text{-}H)\text{-}H_2O]^-$ intermediate ion. To answer this question, ion mobility spectrometry analysis of these ions formed at different collision energies is studied.

ERMS/IMS coupling (ER-IMS) to evidence the various competitive pathways resulting in the [*Lp*b+O]$^{-}$ and [y-H_2O]$^{-}$ ion : the ion mobility spectrum [41] for the m/z 328 product ions through non-resonant excitation (E_{lab}=35 eV) after a single pass (98 cm of ion cloud flight, within resolution as (CCS/ΔCCS) ~ 65) of the cIMS device (Figure 3a) displays the main expected fragmentations. Note it does not exhibit several signals (different cross-sections) for the m/z 128 and m/z 199 ions. Moreover, this ion mobility spectrum does not display the water loss (resulting in m/z 310) at this energy, which is consistent with the disappearance of this ion as early as E_{lab}=25 eV (Figures 1f and 1g). Interestingly, the arrival time distribution (ATD) after 90 passes of the cIMS for the m/z 328 precursor ion at E_{Lab}=5eV (Figure 3b) presents a single peak suggesting that it has a single structure. Indeed, the precursor ion surviving to collision for roughly 90% is characterized by one single peak meaning only one structure (nor a isomer or conformer mixture) that would be injected into the source. It also means that this ion has not yet isomerized, at E_{lab}=5 eV, into a 7-membered cyclic tetrahedral intermediate (or an ion-dipole form) that would be characterized by another ATD. Figure 3c, d shows the ATD recorded for the m/z 199 and m/z 128 ions respectively after 15 passes of the cIMS device for E_{lab} as 15 eV, 20 eV, 35 eV, 40 eV, 50 eV (each cIMS ATD profiles are superposed) with normalized ion abundances.

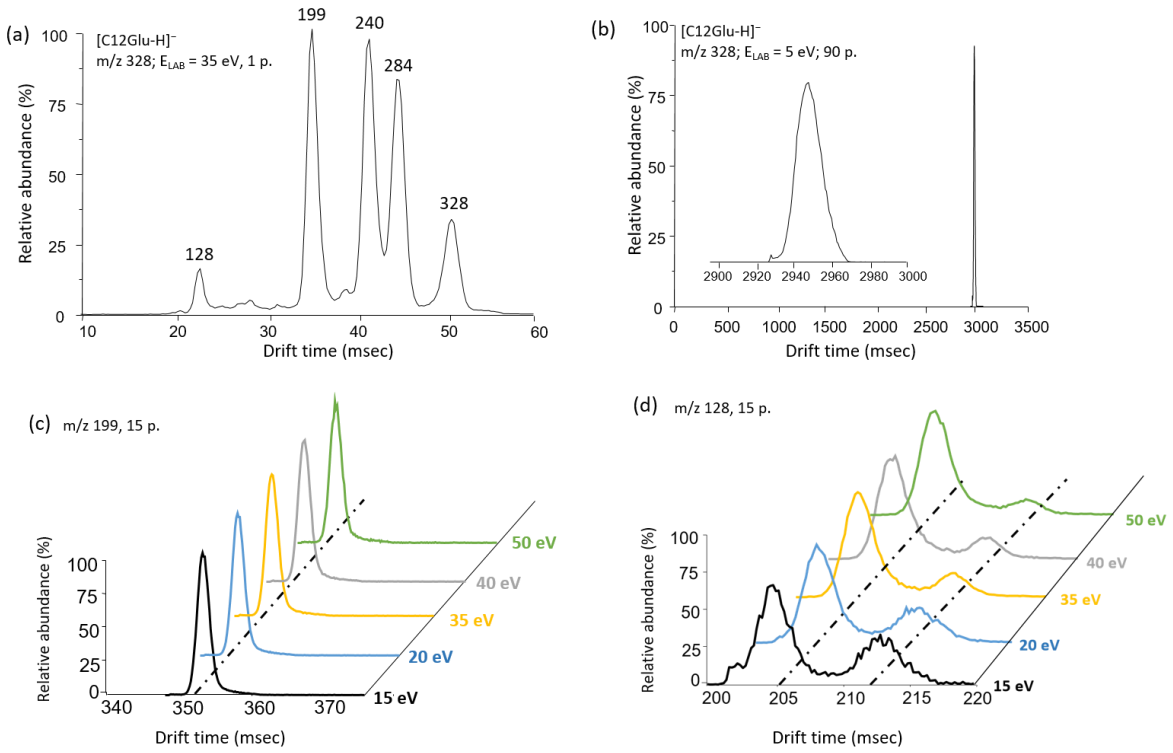


Figure 3. Extracted ions from the product ion mobility spectrum of the [C12Glu-H]$^{-}$ ion (m/z 328): (a) ion mobility spectrum of various product ions at E_{lab}=35 eV after 1 pass (p.), (b) the precursor m/z 328 ion after 90 passes, (c) [*lp*b+O]$^{-}$ (m/z 199) and (d) [y-H_2O]$^{-}$ (m/z 128), after 15 passes at different E_{lab} values as 15 eV, 20 eV, 35 eV, 40 eV, 50 eV (each ions profiles are superposed).

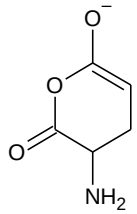
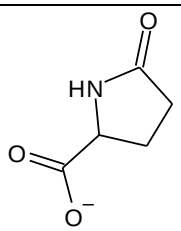
The ATD for the fatty acid carboxylate [*lp*b+O]$^{-}$, m/z 199 (after fifteen passes) is independent of collision energy, only the signal abundance increases from E_{lab}=15 eV to reach optimum at E_{lab}=35 eV. Furthermore, signals centered on the common ATD are characterized by the superimposable shapes for all collision energies (Figure 3c). This supports the hypothesis that a common structure characterizes the m/z 199 ion

(the fatty acid carboxylate) although generated from the energy-dependent dissociation pathways as above proposed in Figure S4a,b.

By contrast, Figure 3d shows that the ATD for the m/z 128 ion, $[y\text{-}H_2O]^-$ (dehydrated glutamate) has two structures revealed after 15 passes of the cIMS. The relative abundances of the two ATD peaks evolve with the collision energy: ER-IMS presents a composite profile for $[y\text{-}H_2O]^-$ (Figure 3d). As the collision energy increases, the abundance of the higher drift time component decreases relative to the abundance of the lower drift time component. This suggests that the former is the $[y\text{-}H_2O]^-$ ion directly produced from $[C12Glu\text{-}H]^-$ from splitting of the ion-dipole intermediate after inter-partner proton transfer, a low energy process yielding the deprotonated α amino glutaric anhydride (Figure S4a). The component at lower drift time could be attributed the isomeric $[y\text{-}H_2O]^-$ ion generated from consecutive dissociation of $[C12Glu\text{-}H]^-$ *via* the intermediate m/z 310 ion, an high energy process resulting in formation of the 2-pyrrolidone-5-carboxylate (Figure S4c). However, it is not possible to rule out the occurrence of various competitive pathways through the direct dissociation yielding a common $[y\text{-}H_2O]^-$ structure.

To probe these interpretations, calculation of the respective cross-section for the proposed isomeric m/z 128 structures compared to those provided in the IMS experiments were performed using variants of the trajectory method incorporating ion-induced dipole and quadrupole contributions to the ion-gas potential (Table 2).

Table 2. CCSs calculated in helium for the isomeric $C_5H_6NO_3^-$ ions (m/z 128) using different trajectory methods

Isomeric m/z 128 structures		Trajectory method CCS calculations	
		Ion dipole-induced potential model	ion-induced quadrupole potential model
Deprotonated α amino glutaric anhydride (form **a**)		128.6 Å^2	132.9 Å^2
2-pyrrolidone-5-carboxylate (deprotomers **b** and **c**)		127.2 Å^2	125.5 Å^2

The isomeric ion which disappears by E_{Lab} increase, corresponds to the one characterized by the larger CCS value, independently of the potential model used for its calculation. Its structure is attributed to deprotonated α-amino glutaric anhydride formed by a stepwise rearrangement through an ion-dipole intermediate evolving to proton transfer prior to ion-dipole splitting (Figure 4Sa). Such a dissociation pathway which in

competition with fatty acid m/z 199 carboxylate, is favored from precursor ion carrying the lower internal energy. Furthermore, the second pathway of the m/z 199 formation as proposed in Figure S4b appears only at higher collision energy and thus, the direct dissociation of the ion dipole intermediate is energetically favored and its complementary m/z 128 ion cannot be produced in these energy conditions. Finally, a third pathway is based on consecutive dissociation after a water loss (ion m/z 310) which yields m/z 128 (Figure S4c). This process requires higher collision energy due to its consecutive character and can only appear in such conditions. The structure for this m/z 128 ion corresponds to the 2-pyrrolidone-5-carboxylate (Figure S4c) which is characterized by the lowest CCS value. This ion appears also at lower energy due to the prompt loss of water leading to formation of the m/z 310 and thus, its consecutive m/z 128 fragment ion can emerge early at low energy (by rearrangement through proton transfer process) and late in high energy (by direct cleavage) (Figure S4c), and thus it can be detected in the studied collision energy range.

Conclusion

The combination of the energy resolved mass spectrometry, ERMS, and ion mobility spectrometry, IMS (associated to a new acronym from energy-resolved ion mobility spectrometry ERIMS) was applied to the analysis of negative ions of lipoaminoacids. In this study we investigate the dissociation of the deprotonated glutamic acid conjugated to lauric acid (C12Glu, Mw 328 Da), and demonstrate the importance to scrutinize with details the ERMS breakdown based only the product ion abundances relative to the total ionic current. Such a treatment allows a better appreciation of the complexity of dissociations thanks to the reproducibility of shape of profile which is characteristic to each product ion. The shape of the signal profiles representative to each product ion has never been scrutinized and described in detail to obtain information on their mechanisms of formation. In this way, a reproducible breakdown curve attributed to a product ion of the generated through one dissociation pathway can be characterized by a Gaussian-like profile shape whose FWHM increases as the m/z ratio of the associated product ion decreases. This was interpreted as resulting from dispersion of product ion, the smaller the m/z ratio of the product ion, the greater the dispersion of the ions, which leads to a broadening of the signal. In addition, the position of the apex of this signal in the ERMS could be attributed to the fragmentation mechanism (TS) leading to the considered fragment ion. In addition, broad Gaussian-like and composite shapes can be respectively observed for profiles related to the unexpected fatty acid carboxylate (m/z 199) and its complementary dehydrated glutamate (m/z 128), occurring in competition. Based on the proposed fragmentation mechanisms it is possible to rationalize their formation and proposed possible ion structures with some ambiguities due to various possible isomeric structure for m/z 128. IMS analysis can "lift the veil" on assumptions made about the proposed structures of the m/z 199 and m/z 128 ions produced depending on collision energy. On the other hand, the significance of the experimental results on the shape of these particular profiles is obtained thanks to the ER-IMS coupling consisting in the concomitant combination of ERMS and IMS analyses possible for instruments having a collision cell in front of the drift tube. First, the m/z 199 ion presents a single structure since the IMS signal is unique and constant whatever the collision energy, indicating that two (or more) dissociation pathways lead to this single structure ion according to relatively close appearance

thresholds as indicated by the more or less symmetrical broad Gaussian shape. Alternatively, its complementary m/z 128 ion presents a different behavior since two IMS signals appears to be associated to this product ion meaning that two isomeric 6- and 5 membered structures contribute in the m/z 128 ion, in accordance with the proposed mechanisms. The former is promoted only at low energy via an ionic-dipole intermediate through inter-partner proton transfer and the latter is consecutively generated by direct cleavage from the $[(C12Glu\text{-}H)\text{-}H_2O]^-$ product ion, along the collision energy range used which seems to correlate with the extracted signals of ion mobility analysis. Today, the IMS (and cIMS) analysis of product ions from collisional spectra at one energy was used without the use of the collision energy variation [21], [42], [43] to clarify some structural ambiguities. The ERMS/IMS(cIMS) coupling ERIMS analysis should be very useful to study the fragmentation mechanisms of selected precursor ions, especially for the distinction of isomeric and diastereomeric compounds that show very similar product ion spectra.

References

[1] Amandine Hueber, Yves Gimbert, Jean-Marie Galano, Alexandre Guy, Thierry Durand, Nicolas Cenac, Justine Bertrand-Michel, Jean-Claude Tabet, « Identification of bacterial lipo-amino acids: origin of regenerated fatty acid carboxylate from dissociation of lipo-glutamate anion », Aminoacids, 2021.

[2] J. H. Beynon, R. M. Caprioli, W. E. Baitinger, et J. W. Amy, « The ion kinetic energy spectrum and the mass spectrum of argon », *International Journal of Mass Spectrometry and Ion Physics*, vol. 3, nº 5, p. 313-321, déc. 1969, doi: 10.1016/0020-7381(69)80077-X.

[3] K. R. Jennings, « Metastable Transitions in the Mass Spectrum of Benzene », *The Journal of Chemical Physics*, vol. 43, nº 11, p. 4176-4177, déc. 1965, doi: 10.1063/1.1696663.

[4] J. H. Beynon, R. M. Caprioli, et T. Ast, « The effect of deuterium labeling on the width of a 'metastable peak' », *Org. Mass Spectrom.*, vol. 5, nº 2, p. 229-234, févr. 1971, doi: 10.1002/oms.1210050213.

[5] D. T. Terwilliger, J. F. Elder, J. H. Beynon, et R. G. Cooks, « The shapes of metastable peaks », *International Journal of Mass Spectrometry and Ion Physics*, vol. 16, nº 3, p. 225-242, mars 1975, doi: 10.1016/0020-7381(75)87022-7.

[6] E. G. Jones, J. H. Beynon, et R. G. Cooks, « Kinetic Energy Release in Unimolecular Ionic Reactions. Thermochemical Aspects », *The Journal of Chemical Physics*, vol. 57, nº 7, p. 2652-2658, oct. 1972, doi: 10.1063/1.1678646.

[7] J. Holmes, C. Aubry, P.M. Mayer, CRC;, *Assigning structure to ions in mass spectrometry*, vol. J. Chem. Phys. 41, 1964, 117. Press, Boca Raton, 2007.

[8] John. L. Holmes et Johan. K. Terlouw, « Metastable ion studies. IV—thermochemistry and ion structures among fragmenting [C3H6]+· ions », *Org. Mass Spectrom.*, vol. 10, nº 9, p. 787-796, sept. 1975, doi: 10.1002/oms.1210100915.

[9] R. G. Cooks et J. H. Beynon, « Metastable ions and ion kinetic energy spectrometry: The development of a new research area », *J. Chem. Educ.*, vol. 51, nº 7, p. 437, juill. 1974, doi: 10.1021/ed051p437.

[10] K. K. Murray, R. K. Boyd, M. N. Eberlin, G. J. Langley, L. Li, et Y. Naito, « Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013) », *Pure and Applied Chemistry*, vol. 85, nº 7, p. 1515-1609, juin 2013, doi: 10.1351/PAC-REC-06-04-06.

[11] Y. X. Jiang, K. V. Wood, R. G. Cooks., « Characterization of and energy deposition in [C2H4O2]+• ions using energyresolved mass spectrometry. », 1986.

[12] C. Bennaceur, C. Afonso, S. Alves, A. Bossée, et J.-C. Tabet, « Instrumental Dependent Dissociations of n-Propyl/Isopropyl Phosphonate Isomers: Evaluation of Resonant and Non-Resonant Vibrational Activations », *J. Am. Soc. Mass Spectrom.*, vol. 24, nº 8, p. 1260-1270, août 2013, doi: 10.1007/s13361-013-0656-3.

[13] A. Schwarzenberg, H. Dossmann, R. B. Cole, X. Machuron-Mandard, et J.-C. Tabet, « Differentiation of isomeric dinitrotoluenes and aminodinitrotoluenes using electrospray high resolution mass spectrometry: Nitroaromatic isomers differentiation by ESI-HRMS », *J. Mass Spectrom.*, vol. 49, nº 12, p. 1330-1337, déc. 2014, doi: 10.1002/jms.3471.

[14] E. Darii *et al.*, « First Direct Evidence of Interpartner Hydride/Deuteride Exchanges for Stored Sodiated Arginine/Fructose-6-phosphate Complex Anions within Salt-Solvated Structures », *J. Am. Soc. Mass Spectrom.*, vol. 32, nº 6, p. 1424-1440, juin 2021, doi: 10.1021/jasms.1c00040.

[15] P. B. Armentrout, K. M. Ervin, et M. T. Rodgers, « Statistical Rate Theory and Kinetic Energy-Resolved Ion Chemistry: Theory and Applications », *J. Phys. Chem. A*, vol. 112, nº 41, p. 10071-10085, oct. 2008, doi: 10.1021/jp805343h.

[16] C. Ruan, Z. Yang, N. Hallowita, et M. T. Rodgers, « Cation−π Interactions with a Model for the Side Chain of Tryptophan: Structures and Absolute Binding Energies of Alkali Metal Cation−Indole Complexes † », *J. Phys. Chem. A*, vol. 109, nº 50, p. 11539-11550, déc. 2005, doi: 10.1021/jp053830d.

[17] Amandine Hueber, Camille Petitfils, Pauline Le Faouder, Geoffrey Langevin, Alexandre Guy , Jean-Marie Galano, Thierry Durand, Jean-François Martin, Jean-Claude Tabet, Nicolas Cenac, Justine Bertrand-Michel, « Discovery and quantification of lipoamino acids in bacteria », *Analytica Chimica Acta*, 2021.

[18] B. Tan *et al.*, « Identification of endogenous acyl amino acids based on a targeted lipidomics approach », *Journal of Lipid Research*, vol. 51, nº 1, p. 112-119, janv. 2010, doi: 10.1194/jlr.M900198-JLR200.

[19] R. D. Hiserodt, S. M. Brown, D. F. H. Swijter, N. Hawkins, et C. J. Mussinan, « A study of b1+H2O and b1-ions in the product ion spectra of dipeptides containing N-terminal basic amino acid residues », *J Am Soc Mass Spectrom*, vol. 18, nº 8, p. 1414-1422, août 2007, doi: 10.1016/j.jasms.2007.04.018.

[20] A. G. Harrison et A. B. Young, « Fragmentation reactions of deprotonated peptides containing aspartic acid », *International Journal of Mass Spectrometry*, vol. 255-256, p. 111-122, sept. 2006, doi: 10.1016/j.ijms.2005.12.037.

[21] K. Giles *et al.*, « A Cyclic Ion Mobility-Mass Spectrometry System », *Anal. Chem.*, vol. 91, nº 13, p. 8564-8573, juill. 2019, doi: 10.1021/acs.analchem.9b01838.

[22] *Avogadro: an open-source molecular builder and visualization tool. Version 1.2.0. http://avogadro.cc/*.

[23] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 16, Revision B.01*. 2016.

[24] *imospedia available on : http://www.imospedia.com/*.

[25] M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, et M. F. Jarrold, « Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential », *J. Phys. Chem.*, vol. 100, nº 40, p. 16082-16086, janv. 1996, doi: 10.1021/jp961623v.

[26] H. I. Kim *et al.*, « Structural Characterization of Unsaturated Phosphatidylcholines Using Traveling Wave Ion Mobility Spectrometry », *Anal. Chem.*, vol. 81, nº 20, p. 8289-8297, oct. 2009, doi: 10.1021/ac900672a.

[27] H. Kim *et al.*, « Experimental and Theoretical Investigation into the Correlation between Mass and Ion Mobility for Choline and Other Ammonium Cations in N $_2$ », *Anal. Chem.*, vol. 80, nº 6, p. 1928-1936, mars 2008, doi: 10.1021/ac701888e.

[28] I. Campuzano *et al.*, « Structural Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross Sections », *Anal. Chem.*, vol. 84, nº 2, p. 1026-1033, janv. 2012, doi: 10.1021/ac202625t.

[29] J.-C. Tabet, Y. Gimbert, A. Damont, D. Touboul, F. Fenaille, et A. S. Woods, « Combining Chemical Knowledge and Quantum Calculation for Interpreting Low-Energy Product Ion Spectra of Metabolite Adduct Ions: Sodiated Diterpene Diester Species as a Case Study », *J. Am. Soc. Mass Spectrom.*, p. jasms.1c00154, sept. 2021, doi: 10.1021/jasms.1c00154.

[30] P. Barbier Saint Hilaire *et al.*, « Mechanistic study of competitive releases of H 2 O, NH 3 and CO 2 from deprotonated aspartic and glutamic acids: Role of conformation », *Journal of Chromatography B*, vol. 1047, p. 64-74, mars 2017, doi: 10.1016/j.jchromb.2016.08.036.
[31] R. B. Cole, Éd., *Electrospray and MALDI mass spectrometry: fundamentals, instrumentation, practicalities and biological applications*, 2. ed. Hoboken, NJ: Wiley, 2010.
[32] A. Maquestiau et R. Flammang, « Studies of prototropy by mass spectrometric methods », *Mass Spectrom. Rev.*, vol. 1, nº 3, p. 237-255, 1982, doi: 10.1002/mas.1280010303.
[33] E. D. Raczyńska, J.-F. Gal, et P.-C. Maria, « The guanylated bioamine agmatine – A theoretical investigation of its structure and exceptional high basicity in the gas phase », *Computational and Theoretical Chemistry*, vol. 1109, p. 10-18, juin 2017, doi: 10.1016/j.comptc.2017.03.032.
[34] J. Naban-Maillet *et al.*, « Internal energy distribution in electrospray ionization: Internal energy distribution in electrospray ionization », *J. Mass Spectrom.*, vol. 40, nº 1, p. 1-8, janv. 2005, doi: 10.1002/jms.773.
[35] E. Darii, S. Alves, Y. Gimbert, A. Perret, et J.-C. Tabet, « Meaning and consequence of the coexistence of competitive hydrogen bond/salt forms on the dissociation orientation of non-covalent complexes », *Journal of Chromatography B*, vol. 1047, p. 45-58, mars 2017, doi: 10.1016/j.jchromb.2016.09.038.
[36] McLuckey, S. A.; Goeringer, D. E, « Slow Heating Methods in Tandem Mass Spectrometry », 1997.
[37] Linda L. Lopez, Philip R. Tiller, Michael W. Senko and Jae C. Schwart, « Automated Strategies for obtaining standardized collisionally induced dissociation spectra on a benchtop ion trap mass spectrometer », 1998.
[38] A. Colorado et J. Brodbelt, « An empirical approach to estimation of critical energies by using a quadrupole ion trap », *J. Am. Soc. Mass Spectrom.*, vol. 7, nº 11, p. 1116-1125, nov. 1996, doi: 10.1016/S1044-0305(96)00077-3.
[39] P. Bayat, D. Lesage, et R. B. Cole, « TUTORIAL: ION ACTIVATION IN TANDEM MASS SPECTROMETRY USING ULTRA-HIGH RESOLUTION INSTRUMENTATION », *Mass Spec Rev*, vol. 39, nº 5-6, p. 680-702, sept. 2020, doi: 10.1002/mas.21623.
[40] D. M. Fedor et R. G. Cooks, « Angle resolved mass spectrometry with a reversed geometry spectrometer », *Anal. Chem.*, vol. 52, nº 4, p. 679-682, avr. 1980, doi: 10.1021/ac50054a021.
[41] V. Gabelica *et al.*, « Recommendations for reporting ion mobility Mass Spectrometry measurements », *Mass Spec Rev*, vol. 38, nº 3, p. 291-320, mai 2019, doi: 10.1002/mas.21585.
[42] A. L. Rister et E. D. Dodds, « Ion Mobility Spectrometry and Tandem Mass Spectrometry Analysis of Estradiol Glucuronide Isomers », *J. Am. Soc. Mass Spectrom.*, vol. 30, nº 10, p. 2037-2040, oct. 2019, doi: 10.1007/s13361-019-02272-w.
[43] S. Ollivier, M. Fanuel, H. Rogniaux, et D. Ropartz, « Molecular Networking of High-Resolution Tandem Ion Mobility Spectra: A Structurally Relevant Way of Organizing Data in Glycomics? », *Anal. Chem.*, vol. 93, nº 31, p. 10871-10878, août 2021, doi: 10.1021/acs.analchem.1c01244.

Acknowledgements: Valérie Steiner for her helpful discussion, Hanna Kulyk-Barbier for her helpful technic help in the using LTQ/Orbitrap and Qq/TOF waters instrument for experimental repetitions; Geoffrey Langevin of the Durand's group for synthesis of C12Glutamate and C16Glutamate conjugates. Part of mass spectrometry analysis were done on MetaToul (Toulouse metabolomics & fluxomics facilities, www.metatoul.fr) which is part of the French National Infrastructure for Metabolomics and Fluxomics MetaboHUB-ANR-11-INBS.

Abreviation: ATD: arrival time distribution; CCS: collision cross section; CID: collision induced dissociation; cIMS: cyclic ion mobility spectrometry; DFT: density functional theory; FFR: field free region; FWHM : full width at half maximum; IKE: ion kinetic energy; KER: kinetics energy release; LTQ: linear trap quadrupole; MIKE: mass-analyzed ion kinetic energy; NCE: normalized collision energy; TIC: total ion current ; TS : transition state.

DISCUSSION GENERALE

Dans cette dernière partie du manuscrit, j'aborderai en premier temps certains points concernant le workflow analytique spécifique aux lipopeptides. Dans un second temps, nous discuterons de la fragmentation des lipoaminoacides en général, puis de certains cas particuliers. Enfin, nous aborderons le cas du C12Glutamate, et des possibilités apportées par la mobilité ionique pour la distinction d'ions fragments isomères.

I) Workflow analytique

Le développement analytique a été réalisé en deux temps, à l'aide d'un mélange de standards de lipopeptides, de lipoaminoacides, d'acyl-serotonine, d'acides gras et de composés nitrés. La première étape a été de développer une méthode d'extraction spécifique à ces molécules, la seconde d'optimiser les conditions d'analyses sur le spectromètre de masse hybride Qq/TOF Xevo G2-XS.

Choix de la matrice biologique : Une fois ce workflow développé et optimisé, nous avons recherché les lipopeptides produits par le microbiote intestinal de souris. Dans ce type d'approche, deux stratégies sont envisagées : la recherche de composés dans les côlons de souris ou dans des bactéries du microbiote en culture. Notre première approche a été de les rechercher dans les côlons de souris. Plusieurs quantités de tissus ont été testées allant de 20 mg jusqu'à la totalité du tissu (environ 90 mg). L'utilisation de 20 mg de côlon, n'a pas permis l'identification de candidats potentiels. De plus, les lipopeptides et lipoaminoacides couramment quantifiés[93] sur un appareil triple quadripolaire Agilent 6460 (C12AsnGABA, C12Asn et C14Asn) n'étaient pas retrouvés *via* l'analyse par LC-HRMS sur le Qq/TOF. Nous avons supposé que l'analyse en LC-HRMS manquait de sensibilité pour détecter ces lipopeptides retrouvés à quelques picogrammes/mg de protéines dans les côlons. L'augmentation de la quantité de tissus conduisait à la perte totale du standard interne. Les plaques SPE utilisées pour le protocole d'extraction étaient préconisées pour 20 mg de tissu uniquement. Si nous avions voulu continuer avec une telle quantité de tissus, nous aurions dû ré-optimiser le protocole d'extraction, en utilisant une plaque SPE appropriée, (60 mg de dépôt de tissu). D'autres stratégies peuvent aussi être envisagées, comme plusieurs SPE

successives (interactions polaires puis apolaires). C'est le protocole expérimental utilisé par Tan[361], permettant l'identification de lipoaminoacides dans le cerveau.

Suite à ces premiers résultats, nous nous sommes focalisés sur des bactéries du microbiote. En effet, s'affranchir de l'effet matrice notable du côlon permet un gain de sensibilité important, pendant l'extraction et lors de l'ionisation en electrospray. Afin d'augmenter nos chances de découvrir de nouveaux composés, nous avons sélectionné plusieurs bactéries. Nous avons travaillé sur culot bactérien plutôt que sur surnageant bactérien, pour détecter ces composés minoritaires plus facilement.

Dans un premier temps, nous avons continué les travaux sur *Escherichia coli* Nissle 1917, puis nous avons sélectionné d'autres bactéries, majoritaires du microbiote intestinal et appartenant à des phylas différents, comme les *Bacteroidetes*, les *Bifidobacterium Longum*, ou encore les *Enterococcus faecalis*. En parallèle, nous nous sommes intéressés à la souche bactérienne *Lactobacillus animalis* car notre équipe a mis en évidence que l'abondance de cette bactérie était diminuée dans le microbiote de souris soumise à un modèle de syndrome de l'intestin irritable (SII) (données non publiées actuellement). De plus, l'abondance de cette bactérie était inversement corrélée à l'hypersensibilité viscérale mécanique des souris. Dans le but d'accroitre le nombre de composés conjugués au GABA, les *Lactobacillus animalis* ont été cultivées dans un milieu de culture enrichi en GABA.

Identification et caractérisation de nouveaux lipopeptides : L'étude de la bactérie probiotique *Escherichia coli* Nissle 1917, nous a permis d'identifier un premier LpAA jamais décrit, le C13Asn et de quantifier toute sa famille. L'analyse des autres souches bactériennes, notamment des *bacteroidetes* a démontré la présence de ce même composé. Ces deux bactéries partagent donc la production de certains LpAA bien qu'elles soient extrêmement éloignées d'un point de vu génétique. Ces quantifications de LpAA dans les bactéries renforcent la notion de redondance fonctionnelle du microbiote[110].

Concernant *Lactobacillus animalis*, nous avons émis l'hypothèse que cette bactérie est capable de produire une molécule lipidique ayant une activité analgésique car la diminution de son abondance dans notre modèle de SII était associée à l'augmentation de l'hypersensibilité viscérale. L'application du workflow n'a pas permis de mettre en évidence la présence de lipopeptide contenant du GABA dans cette bactérie. Comme *L. animalis* ne possède les enzymes nécessaires à la synthèse du GABA à partir de la glutamine contenue dans le milieu, nous avons cultivé cette bactérie en présence de GABA. Dans ces conditions, nous avons pu identifier du C16PheGABA, du C12AsnGABA, du C16GluGABA, du

C14AsnGABA, du C12IleGABA, du C14IleGABA et du C12AlaGABA. Ces résultats vont être utilisés dans un article en cours de rédaction auquel je suis associée.

Pour conclure, les premiers essais sur côlons murins n'ont pas permis l'identification de nouveaux lipides bactériens. La seconde stratégie focalisée sur les bactéries a permis de démontrer que plusieurs souches de bactéries, appartenant à des phylas différents sont capables de produire des lipopeptides et des LpAA.

Analyse des lipopeptides par spectrométrie de masse à haute résoltion : Le mode d'ionisation préférentiellement utilisé dans cette étude à été l'électro-nébulisation en mode négatif. Ce mode a été sélectionné car les lipopeptides et lipoaminoacides sont des composés contenant un groupement carboxylique facilement « déprotonable » en phase gazeuse, ainsi qu'un groupement amide relativement acide. Contrairement aux nombreux sites qui peuvent être protonés (chacun des hétéroatomes le sont) en mode positif, seuls trois ou quatre sites peuvent être déprotonés en compétition. L'ionisation en mode négatif permet donc de limiter les modes possibles de dissociation évitant la dispersion du courant ionique sur trop d'ions fragments (comme cela arrive lors de dissociations des molécules protonées). Toutefois, les composés de notre base de données constitués de sérotonine, s'ionisent uniquement en mode positif[362]. Cela est la conséquence de la présence des sites relativement basiques comme l'atome d'azote indolique, mais également du résidu NH_2 de la chaine latérale[363]. En parallèle les analyses des extraits bactériens ont été effectuées à l'aide deux modes d'ionisation. Le mode négatif a permis l'identification du C13Asn, et du C14AsnGABA. Le mode d'ionisation positif, cela a permis de valider les structures des composés identifiés en mode négatif, cependant aucun composé constitué de sérotonine n'a été détecté. Ainsi, seules les molécules déprotonées $[LpAA\text{-}H]^-$ ont été analysées en mode négatif, et leurs dissociations sous collision de basse énergie ont été interprétées à l'aide du mode négatif également.

Retraitement des données via des logiciels spécialisés : L'étape cruciale et la plus chronophage de ce workflow a été le retraitement des données. Plusieurs logiciels de retraitement ont été essayés et comparés afin d'optimiser nos chances de découvrir de nouveaux lipopeptides. Le logiciel UNIFI, qui est difficile d'utilisation et sur lequel énormément de tests ont été réalisés, a finalement été abandonné, car le « processing » des données n'était pas adapté à nos exigences, nos conditions expérimentales et à nos composés minoritaires. Le logiciel de quantification de Waters QuanLynx s'est finalement révélé le plus efficace pour interroger nos données.

Par exemple, pour l'analyse en négatif de l'extrait d'EcN 1917, 162 composés ont pu être identifiés comme lipopeptides potentiel de notre base de données. Après vérification des massifs isotopiques, 25 composés uniquement correspondaient aux molécules déprotonées. Les MS^2 en spectrométrie de masse à haute résolution ont été réalisées, et 5 composés ont finalement été validés et leurs élucidations structurales réalisées. Le même travail a été réalisé avec le mode d'ionisation positif, où aucun nouveau lipopeptide supplémentaire n'a été identifié.

Dans un second temps, le logiciel MZ-MINE2 a été utilisé, afin de comparer les groupes entre-eux. Ce logiciel permet après le choix de plusieurs paramètres pas toujours bien définis et difficiles à gérer, de générer des listes de rapports masse sur charge et de comparer leur présence dans différents échantillons. De façon très visuelle, le logiciel attribue un rond vert ou rouge pour chaque m/z de la liste s'il est présent ou absent de l'échantillon. Ainsi, nous avons pu comparer des côlons de souris contrôles et ceux de souris gavées par injection péritonéales de bactéries connues pour leur production accrue de GABA. La liste de rapports masse-sur-charge des composés détectés uniquement dans les souris gavées a été réalisée. Des analyses MS^2 ont été réalisées sur cette liste de composés, par LC-HRMS sur un QqTOF et un QExactive Plus, en supposant qu'il s'agissait potentiellement de lipopeptides ou d'aminolipides. Cependant, lors des analyses par LC-HRMS aucun composé n'a été caractérisé ou identifié comme étant un lipopeptide.

Quel niveau d'identification pour les composés de notre étude ? L'utilisation du mode d'acquisition SRM a permis d'étendre cette identification à vingt-deux nouveaux composés. L'analyse des ions fragments réalisée, soit à l'aide d'un mode d'excitation résonante, soit d'un mode d'excitation non résonante a été faite afin de mieux comprendre les mécanismes de fragmentations impliqués. En effet, le mode non résonant permet de mettre en évidence les processus consécutifs, alors que lors d'une excitation en mode résonant seuls les processus compétitifs sont mis en évidence. Sous excitation résonante, seul l'ion précurseur est excité par sa propre fréquence[338,364].

Grâce à ces deux modes d'ionisation, la majorité des molécules ont pu être caractérisées à l'aide d'un niveau d'identification 1 (c'est le cas pour seize molécules) selon les règles appliquées pour la caractérisation de nouvelles molécules en métabolomique[345]. Cependant, certaines structures d'ions fragments, et donc des ions précurseurs ne peuvent être totalement élucidées : c'est le cas pour les lipoaminoacides où l'acide gras conjugué est insaturé. En effet, il n'est pas possible de dissocier ces chaines pour localiser les insaturations en collision de basse énergie, comme cela est possible avec des collisions de haute énergie

(au-delà du KeV)[365]. Néanmoins, des collisions par électron (EID ou EIEIO)[366] ou par atome métastable[367] permettent ces processus à charge piégée, impliquant l'accès à des états électroniques excités. Ces modes d'activation sont soit limités d'accès soit non commerciaux, et n'ont donc pas pu être utilisés lors de ces travaux. Notons que l'utilisation de collision basse énergie sur les lipides cationisés par du Cu^{II} conduisent aussi à la localisation d'insaturation, mais la sensibilité de la méthode reste limitée[368]. Une autre méthode, par GC-MS permet l'identification de la position des doubles liaisons après une dérivatisation par du réactif 4,4-diméthyloxazoline[369].

En ce qui concerne mon travail, la structure des LpAA ayant une chaine d'acide gras insaturée n'a pas pu être définie : les composés C12:1Asn, C14:1Asn, C16:1Asn et C17:1Asn, sont donc identifiés avec un niveau d'identification 3 correspondant car la position de la double liaison n'est pas connue. Nous supposons néanmoins que celle-ci se situe en position w7 puisque que ce sont les AGI les plus abondants de nos profils d'acides gras totaux bactériens. Cependant, il ne s'agit ici que d'une hypothèse puisque nous avons également pu doser de faibles quantité l'acide gras C16:1w9, qui pourrait parfaitement être l'acide gras à l'origine de la formation des lipoaminoacides. Ces interrogations pourraient, potentiellement être levées en comparant les candidats aux différents isomères synthétisés en GC-MS avec un gradient et une dérivatisation adaptée.

Les lipoaminoacides conjugués à la leucine et à l'isoleucine n'ont pas pu être identifiés à un niveau 1. Les ions fragments générés en MS^2, à partir des deux standards synthétisés, sont identiques pour les deux composés conjugués. De plus, il n'a pas été possible d'obtenir une séparation chromatographique suffisante pour ces deux composés ($\Delta Tr < 0.5$ secondes) dans nos conditions expérimentales. Ainsi, le C12Leu et le C12Ile ont été caractérisés avec un niveau d'identification 2, correspondant à des annotations putatives, très probables. Pour les caractériser complètement il faudrait être capable de les séparer chromatographiquement en utilisant une colonne C18 plus longue, une colonne C30 ou des colonnes chirales.

Analyses par pertes de neutres : A l'aide de l'analyse de la fragmentation du C12AsnGABA et du C14AsnGABA, nous avons identifié une perte de neutre systématique correspondant au GABA (-103,0633). Les extraits bactériens ont été analysés en utilisant le mode de balayage perte de neutre sur le spectromètre de masse triple quadripôlaire Agilent K6460 du laboratoire. Cette stratégie consiste à fixer la perte de neutre constante à 103.0633 Da. Un grand nombre de pics chromatographiques ont été détectés avec ce mode de balayage, et chacun d'entre eux a été sélectionné et analysé en MS^2 en haute résolution (Qq/TOF). Cependant, les composés identifiés par le QqQ, comme comportant hypothétiquement (car basse résolution)

un GABA au sein de leurs structures, étaient des faux positifs. Lors des analyses MS^2 sur le Qq/TOF, aucun des m/z sélectionnés ne permettait d'identifier un lipopeptide, ni aucun des composés de notre banque de données. Le manque de résolution de notre appareil de type QqQ qui a plus de 10 ans et qui ne donne accès qu'aux masses nominales, n'est pas adapté pour ce type d'approche.

Recherche de composés par chemins communs de fragmentation : Les connaissances apportées par la caractérisation des C12AsnGABA, C14AsnGABA, C12Asn, C14Asn et C13Asn permettent de prédire certains ions fragments attendus à partir de la dissociation d'anions pour les molécules de notre banque de données théoriques. En effet, ces composés vont se fragmenter après activation collisionnelle pour produire des spectres CID. Plusieurs ions fragments sont générés : le lipoaminoacide déprotoné et déshydraté (perte H_2O), l'amide déprotoné de l'acide gras, ou encore l'aminoacide lui-même déprotoné. Ainsi, en suivant ce chemin de fragmentation, et en le déclinant à l'ensemble des composés lipoaminoacides et lipopeptides GABA de notre banque de données, nous avons imaginé des méthodes d'analyse par QqQ avec une mode de balayage « précurseur ion ». Pour cela, nous avons fixé les ions fragments attendus, déclinaison des pertes H_2O pour tous les composés, des acides gras aminés pour tous les composés, des acides aminés déprotonés et du neurotransmetteur déprotoné. Par exemple, pour les lipoaminoacides ou lipopeptides constitués de lysine, nous avons sélectionné l'ion fils correspondant au fragment $[Lys\text{-}H]^-$. Le spectromètre de masse, ainsi paramétré recherche les rapports masse-sur-charge des ions parents qui se dissocient en un ion fragment $[Lys\text{-}H]^-$. Lorsque l'un des pic semblait correspondre à la masse d'une molécule déprotonée contenant de la lysine, nous les avons analysés par LC-$HRMS^2$. Lors de chaque analyse de haute résolution sur le Qq/TOF, comme pour les pertes de neutre, nous nous sommes aperçus que les correspondances annoncées étaient surement liées à un manque de sensibilité du QqQ, puisqu'aucun des composés n'a pu être caractérisé en utilisant cette méthode. Cependant, par cette méthode, nous sommes parvenus à retrouver certains des composés ayant été identifiés préalablement à l'aide des SRM théoriques. Finalement, les analyses dites de « SRM théoriques » nous ont permis d'accroitre le nombre de lipoaminoacides identifiés.

Synthèse de lipoaminoacides et de lipopeptides : La caractérisation des candidats a pu être validée dans certains cas par spectrométrie de masse en tandem (Qq/TOF). Dans d'autres cas afin de lever certains doutes concernant les mécanismes de fragmentation, nous avions réalisé des expériences de MS^n en mode résonant sur le LTQ-VELOS-Orbitrap de la plateforme MétaToul réseaux métaboliques de Toulouse. L'ensemble de ces données nous

ont permis de proposer des mécanismes de fragmentations pour chacun des composés identifiés, et ainsi, de les caractériser avec différents niveaux d'identifications dont le niveau 1. Les composés néo-identifiés ont tous été comparés à des standards primaires identiques, synthétisés spécifiquement. La synthèse des lipoaminoacides et des lipopeptides a été réalisée à Montpellier, par le groupe de Thierry Durand. Celles-ci ont été fastidieuses, et ont, dans certains cas nécessité un long travail d'optimisation. C'est notamment le cas pour les lipopeptides GABA, les lipopeptides sérotonine, dopamine et histamine. Pour ces synthèses spécifiques, de nombreux essais de résines ont été effectués, avec des rendements parfois inférieurs à 10%. L'obtention de produits purs, essentielle à notre projet, a rajouté une contrainte supplémentaire. A ce jour, les chimistes travaillent toujours pour essayer de développer une méthode de synthèse efficace pour les dérivés de la dopamine et pour les dérivés histamines. L'absence de ces standards nous a empêché de valider la méthode d'extraction et d'analyse pour ces composés (dopamine, histamine). C'est peut-être l'une des raisons pour lesquelles aucun lipopeptide contenant ces neurotransmetteurs n'a été détecté à l'aide de notre stratégie expérimentale. Celui-ci mériterait peut-être une optimisation spécifique pour ces composés aux propriétés chimiques différentes des lipoaminoacides et des lipopeptides GABA.

II) Fragmentation des lipopeptides et lipoaminoacides

Mécanismes de fragmentation proposés pour les lipoaminoacides : Dans notre étude, de nombreux chemin de fragmentation peuvent s'expliquer par l'apparition d'ions dipôles, c'est notamment le cas pour le fragment le plus intense des lipoaminolipides qui correspond à la perte d'H_2O.

En présence d'un ou de deux groupements COOH, ou d'un groupement CO-NH_2, le mécanisme de la perte dH_2O est toujours identique. Le cas des lipopeptides asparagine a été développé dans l'article 1. Lorsque l'acide animé est différent, le mécanisme reste équivalent mais différent car ce sont d'autres protons qui migrent (Schéma 3). En effet, quel que soit l'acide aminé considéré, il s'agit toujours d'une attaque nucléophile d'un des atomes d'oxygène chargés négativement *a priori* soit en alpha de l'amide, soit sur la chaine latérale ou de l'amide lorsque l'acide aminé possède une chaine latérale non polaire. Il en résulte la création d'un intermédiaire tétraédrique (résultant en un cycle à 6 chaînons dans le cas de la glutamine). La libération de OH^- et la formation d'un anhydride cyclique conduit à la formation d'un ion dipôle. L'anion OH^- libre, va arracher un proton pour éliminer H_2O. Ce proton peut provenir d'une position énolisable de l'anhydride ou à partir du groupement imine qui est acide

en phase gazeuse. De manière concomitante, le système à 6 chaînons s'ouvre afin de produire un nitrile et de relibérer l'acide carboxylique. A partir de là, la perte consécutive de CO_2 se produit en impliquant le transfert d'un proton, permettant l'obtention d'un nitrile déprotoné. Ces mécanismes sont cohérents avec les règles de la chimie organique. Cette perte de CO_2 est identique pour tous les lipoaminoacides.

La charge peut être localisée soit sur l'acide carboxylique, qui sera ainsi déprotoné ou en position alpha de l'amide (Figure 99). Quand la charge se situe sur le groupement carboxylique, la perte de CO_2 attendue peut avoir lieu, c'est ce qui est observée pour l'ensemble des composés asparagine, valine, leucine, isoleucine. Dans ce cas, la perte de CO_2 est induite par la charge impliquant le transfert d'un proton mobile de la position énolisable. Le transfert de ce proton est obligatoire. Ce mécanisme mène à la formation d'un ion m/z 283 dans le cas du C13Asn.

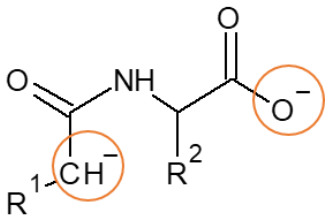


Figure 99 : Positions possibles de la charge sur les LpAA où la chaine latérale de l'aminoacide conjugué est non-polaire, menant soit à la formation d'un cétène, soit à la perte d'un groupement CO_2

Notons qu'à partir des anions $[LpAA\text{-}H]^-$ où l'AA possède une chaine latérale non polaire (Val, Leu ou Ile) qui conduit à une perte d'eau initiale réduite alors que celle du CO_2 est renforcée. Ces deux pertes compétitives sont produites selon des mécanismes basés sur le même principe que ceux décrits dans le Schéma 1 mais impliquant d'autres transferts de proton ou de cyclisation par addition/élimination. En effet, la charge de l'anion $[LpAA\text{-}H]^-$ peut être localisée ou sur le groupe carboxylate ou sur la position énolisable du groupe amide (Figure 99). Dans le premier cas, c'est la perte de CO_2 qui se produit alors qu'à partir de la seconde forme (l'énolate) c'est la perte d'eau qui se produit par un processus par étape. Notons que la formation de l'amide de l'acide gras déprotoné est fortement réduite ou absente.

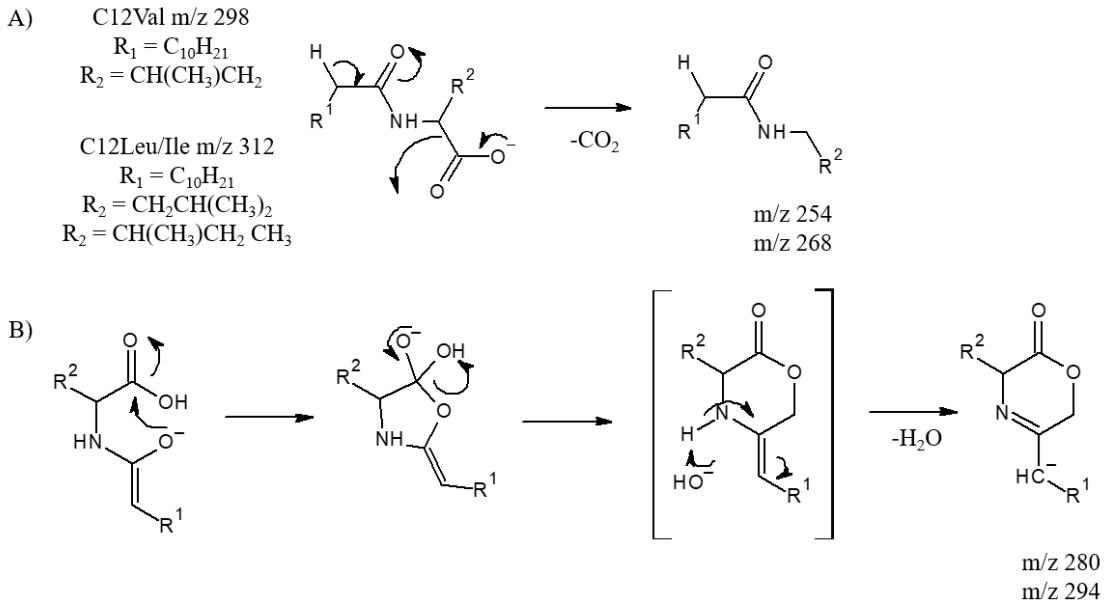


Schéma 3 : Mécanismes possibles de fragmentation par pertes directes compétitives A) majeure de CO_2 et B) mineure d'H_2O à partir de $[C12AA\text{-}H]^-$ (avec AA=Val, Leu et Ile)

Le site de déprotonation de la phénylalanine conjuguée à l'acide gras tétradécanoïque (C14Phe) étant le site énolisable de l'amide, il va induire une rupture benzylique assistée par un transfert 1-5 de proton concerté conduisant à la formation d'un ion dipôle. Cette fois, cette formation est rendue possible car le neutre libéré est l'acide cinnamique qui peut être déprotoné. Cet intermédiaire peut se dissocier soit directement pour donner un amidure (amide déprotoné) par la perte de l'acide cinnamique ($C_9H_8O_2$), soit par transfert de proton de l'acide cinnamique vers l'amide déprotoné de l'acide gras pour donner l'anion $C_9H_7O_2^-$ associé à une perte de l'amide neutre (Schéma 4).

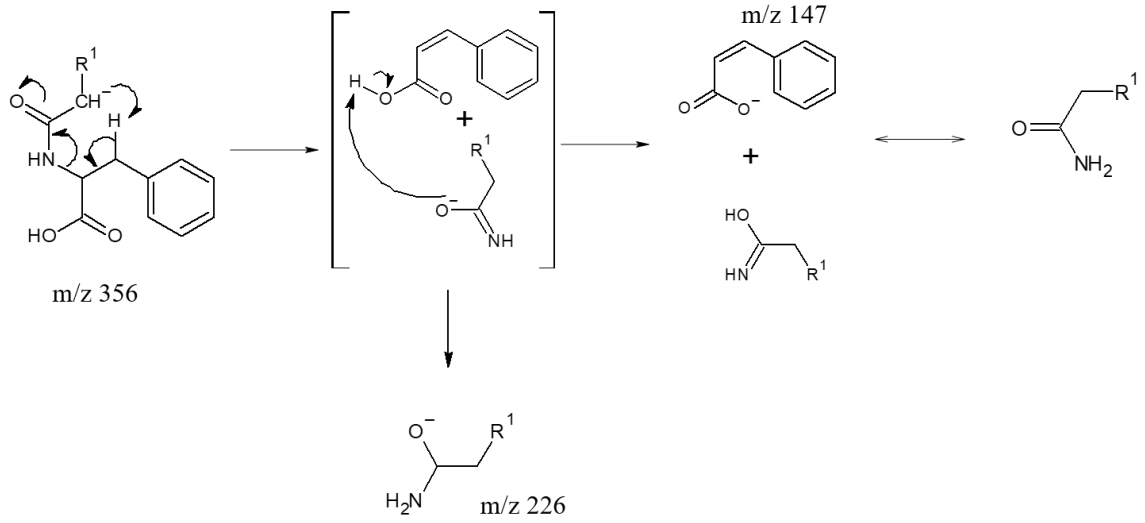


Schéma 4 : Proposition de mécanisme de fragmentation particulier de $[C14Phe\text{-}H]^-$, m/z 356, menant aux ions fragments m/z 226 et m/z 147 complémentaires.

Cas de fragmentation des lipoaminoacides sans intermédiaires dipôlaires : Cependant, les lipoaminoacides identifiés à l'aide du workflow analytique ne se décomposent pas tous selon des mécanismes de fragmentation faisant appel à des intermédiaires de type ion-dipôle. C'est le cas notamment du C12Tryptophane (C12Trp), qui présente un fragment inattendu, qui n'a pas été abordé dans la publication 1.

Dans ce cas de figure, un autre site chargé négativement peut exister en compétition avec d'autres comme par exemple un déprotomère où l'azote du noyau indole est déprotoné. A partir de ce déprotomère, la charge négative peut alors basculer et délocaliser la double liaison afin de rompre la liaison en alpha-béta de l'acide carboxylique. Cette rupture benzylique est connue pour être favorisée. Comme la localisation de la charge négative sur l'atome de carbone lié à l'atome d'azote de l'amide n'est pas autorisée selon les règles de fragmentation définies, un transfert concerté de proton pour neutraliser le carbanion se fera à partir de l'acide carboxylique voisin il en résulte la formation de l'anion $[C12Trp\text{-}H]^-$ (Schéma 5). Ici la formation d'un ion-dipôle intermédiaire n'est pas possible, car le partenaire neutre C_9H_7N ne possède pas de site acide tel qu'il est requis pour neutraliser $[C12Trp\text{-}H]^-$.

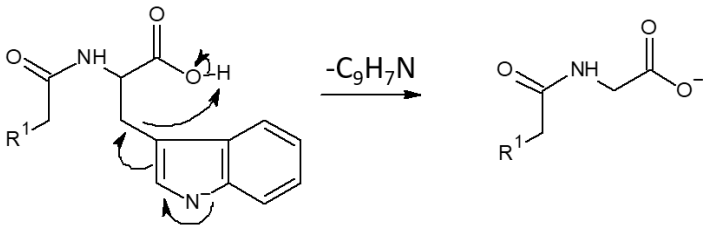
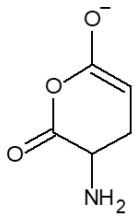
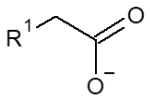


Schéma 5 : Proposition de mécanisme de fragmentation particulier de l'ion $[C12Trp\text{-}H]^-$ menant au fragment de m/z 256, où R1 correspond à une chaîne alkyle $C_{10}H_{21}$de la partie acide gras

Cas particuliers des lipoaminoacides conjugués au glutamate : Jusqu'à présent, nous n'avons abordé ni les lipoaminoacides conjugués à l'acide glutamique ni ceux liés à l'acide aspartique de façon délibérée, puisque ces deux LpAA présentent une fragmentation particulièrement inattendue. Des analyses en mode MS^n séquentiel avec un piège à ions (sous activation collisionnelle en mode résonant) ainsi que les analyses de MS^2 sur l'hybride Qq/TOF ont été nécessaires pour proposer des mécanismes de fragmentation. Celle-ci peut s'interpréter au travers de la formation d'ions dipôles pouvant se dissocier en compétition afin de donner un ion fragment acide aminé déshydraté et déprotoné et de régénérer le carboxylate de l'acide gras (Figure 100), alors que les autres LpAA régénèrent les fragments amides de l'acide gras.



m/z 128
$C_5H_6O_3N$



m/z 255
$C_{16}H_{31}O_2$

m/z 199
$C_{12}H_{23}O_2$

Figure 100 : Structure des ions m/z 128 et m/z 255 produits en compétitions à partir du $[C16Glu-H]^-$ et du $[C12Glu-H]^-$avec $R^1 = C_{14}H_{29}$ pour le C16Glu et $R^1 = C_{10}H_{21}$ pour C12Glu

Malheureusement, dans le mécanisme que nous avons proposé dans la publication n°2, il y avait une contradiction apparente qu'il fallait lever puisqu'étrangement, des différences importantes ont été mises en évidence entre les formes déprotonées de LpAsn et LpAsp, et entre celles des LpGln et LpGlu.

Dans les acides aminés conjugués avec chaine fonctionnalisée comme l'asparagine et la glutamine, l'ion fragment correspondant à l'acide gras ne peut pas être mis en évidence avec abondance. Cependant, en échangeant la fonction amide des acides aminés glutamine et asparagine par une fonction acide carboxylique (donnant naissance aux LpAsp et LpGlu), la rupture de type c^- (amide déprotoné) devient alors très abondante (minimum 27% d'abondance relative). Toutefois, il ne s'agit pas tout à fait un ion c^- mais d'un ion $[b+O]^-$ inattendu dont la structure a été précisée par le mécanisme proposé dans la publication 2. La formation de ces ions fragments impliquent la formation d'intermédiaires de type ion dipôle. Les mécanismes de formation proposés pour ces ions particuliers ont été appuyés et confirmés à l'aide de modélisation de type DFT réalisés par le Dr Yves Gimbert de l'Université de Grenoble. Comme cela ne fait pas parti de notre spécialité, nous n'en donnerons pas les aspects théoriques au cours de cette discussion.

En ce qui concerne les composés LpGlu et LpAsp, il s'agit d'un cas de figure particulier ou la molécule déprotonée conduit à la formation de deux ions complémentaires $[Lpb+O]^-$ et $[y-H_2O]^-$. Le premier semble être favorisé à plus haute énergie d'excitation et le second à plus basse énergie. Ce comportement est contraire à nos résultats d'expérience, ce qui ne veut pas dire que la formation d'un ion dipôle intermédiaire n'a pas lieu. En réalité, l'acide glutamique deshydraté et déprotoné est aussi produit par un processus de décomposition consécutive, *via* une perte d'eau, donc à un seuil d'apparition plus élevé du processus impliquant l'ion-dipôle qui est un mécanisme à multi-étapes.

Le couple d'ions fragments complémentaires (m/z 255 ; m/z 128) correspondant aux ions fragments de l'ion $[C16Glu-H]^-$ (m/z 384) et leurs comportements ont été détaillés dans la publication 2. Celui-ci est également observé pour le $[C12Glu-H]^-$ (m/z 328) et les ions fragments m/z 199 et m/z 128. Dans cette fragmentation par étapes multiples, suivant le mécanisme présenté dans la publication, un ion dipôle est directement formé. Il est constitué de l'acide gras déprotoné et de l'anhydride neutre de l'acide glutamique. Ce complexe peut se dissocier directement en carboxylate de l'acide gras (m/z 255), ou peut s'isomériser par transfert interne de proton, de l'anhydride vers le groupement carboxylate, conduisant à l'ion fragment m/z 128 (acide glutamique déshydraté et déprotoné). Ainsi, il est possible de prédire que ces deux constituants du complexe ont des acidités suffisamment proches pour avoir une durée de vie augmentée permettant l'échange de proton entre partenaires. Si l'acidité de l'anhydride est comparable à celle de l'acide gras alors, la présence de l'anhydride déprotoné avec une abondance relative non négligeable peut être observée. Effectivement, l'abondance de l'anhydride déprotoné est faible (<3%) alors que l'acide gras déprotoné a une abondance relative aux alentours de 30% à 20eV. Si l'acidité de l'acide glutamique deshydraté était trop faible l'ion fragment correspondant à l'anhydride ne serait pas détecté. Les calculs quantiques montrent que l'acidité de l'anhydride est comparable à celle de l'acide carboxylique lui-même. Toutes ces données nous amènent à conclure qu'il est normal d'observer l'anhydride de l'acide glutamique.

Enfin, si tout apparait clair concernant les différents mécanismes inattendus des LpGlu, nous nous attendions en augmentant l'énergie de collision à défavoriser davantage la production des ions fragments anhydrides (m/z 128). L'ensemble des différents profils d'évolution de l'abondance de chacun des ions fragments en fonction de l'énergie de collision conduit au graphique ERMS (energy resolved mass spectrometry). La construction de cette ensemble ERMS a été réalisé et associé à la spectrométrie de mobilité ionique pour lever en partie l'ambiguité qui pouvait subsister.

III) Les outils pour mieux comprendre les mécanismes de fragmentation

Profils d'évolution des courbes d'abondance relative en fonction de l'énergie de collision : Cette méthode d'analyse ERMS basée sur l'étude de la variation de l'abondance relative des ions parents et des ions fragments en fonction de l'énergie de collision permet une approche plus précise des processus de fragmentation qui paraissent ambigües. Cette méthode a souvent été utilisée pour étudier l'effet de l'énergie interne des ions acquise lors de collisions[370] afin d'éclaircir certains mécanismes de fragmentation, et pour déterminer les seuils d'apparition d'ions fragments[371,372]. Plus récemment cette méthode a été utilisée afin de comparer les modes d'activation comme les collisions par gaz et les collisions sur surfaces[373,374]. D'un point de vue analytique, cette méthode a surtout été employée pour distinguer des isomères[375,376,364,375,377]. C'est pour les potentialités de ces profils ERMS permettant la compréhension des mécanismes de fragmentions que nous l'avons utilisé pour étudier le comportement des ions dipôles en fonction de l'énergie tel qu'employé pour l'études des complexes ion-dipôle en mode ionisation négative[353].

De manière générale, les courbes ERMS à partir des ions précurseurs sélectionnés sont très différentes entre le mode d'activation résonant et le mode d'activation non résonant[364,338]. Les courbes ERMS construites à l'aide d'instruments où l'activation est réalisée en mode non-résonant, sont caractérisées par des profils Gaussiens (Figure 101C, m/z 198 et 101D, m/z 198). Par contre, l'ion précurseur a un profil de sigmoïde décroissante (Figure 101A, m/z 328). Enfin, des courbes monotones croissantes vers la haute énergie représentent les ions fragments qui n'atteignent pas encore leur maximum, ils correspondent à une terminaison des décompositions consécutives (Figure 101B, m/z 102).

A l'inverse, à l'aide du mode d'activation résonant dans un piège utilisé comme une cellule de collision, seuls les processus de première génération ou compétitifs sont détectés. La caractéristique du mode résonant concerne les évolutions de l'abondance relative des ions fragments qui dès la disparition de l'ion parent, deviennent toutes constantes. Cela signifie que dès que l'ion parent subit une excitation résonante trop forte, son énergie cinétique devient trop grande par rapport au puits de pseudo-potentiel et il est éjecté du piège. Ainsi les ions fragments formés dans les derniers moments se trouvent piégés et leur abondance relative devient constante. C'est le cas par exemple pour les ions m/z 310, m/z 128 et m/z 199 et m/z 284 issus du C12Glu m/z 328 (Figure 102).

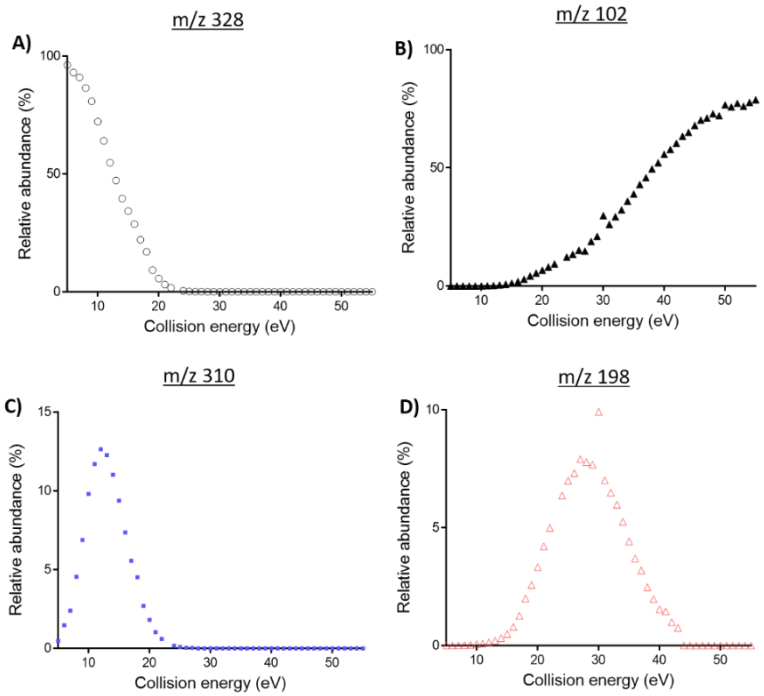


Figure 101 : Profils d'évolutions de l'abondance des ions fragments m/z B) 102, C) m/z 310 et D) m/z 198 en fonction de l'énergie de collision à partir du précurseur $[C12Glu-H]^-$ m/z 328 sous activation résonante (QqTOF Waters Xevo G2-XS)

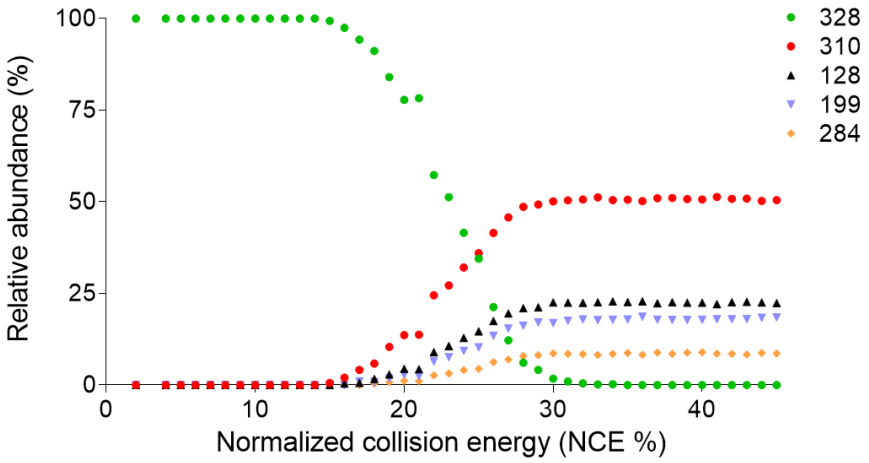


Figure 102 : Profils d'évolutions de l'abondance des ions fragments m/z 310, m/z 128, m/z 199 et m/z 284 fonction de l'énergie de collision à partir du précurseur $[C12Glu-H]^-$ m/z 328 sous activation non-résonante (LTQ, Thermo)

Toutefois avant la disparition de l'ion parent, l'évolution des ions fragments suivent une même tendance croissante, expliquée par le principe même du « chauffage lent ». C'est-à-dire que tant que l'ion parent est présent il accumule de l'énergie interne, une fois le seuil d'excitation atteint, l'ion précurseur décroit, comme pour les courbes ERMS générées à l'aide d'un mode d'activation non résonant. Cependant, les courbes ERMS des ions fragments de premières générations augmentent, jusqu'à un seuil, caractérisé par la disparition de l'ion parent et leurs profils évoluent comme une sigmoïde croissante.

Profils d'évolution des courbes d'abondance relative en fonction de l'énergie de collision pour le C12Glu : Les profils d'évolution de l'abondance des ions fragments en fonction de l'énergie cinétique pour les ions produits m/z 199 et m/z 128 sont anormalement élargis ou composites (Figure 103B, m/z 199 et Figure 103D m/z 128). Cela est vrai, lorsque les courbes ERMS sont réalisées à l'aide d'une normalisation par le courant total ionique. Lorsque les courbes ERMS sont réalisées à l'aide de l'abondance absolue de l'ion d'intérêt, alors nous ne retrouvons plus ces courbes particulières (Figure 103A m/z 199 et Figure 103C m/z 128).

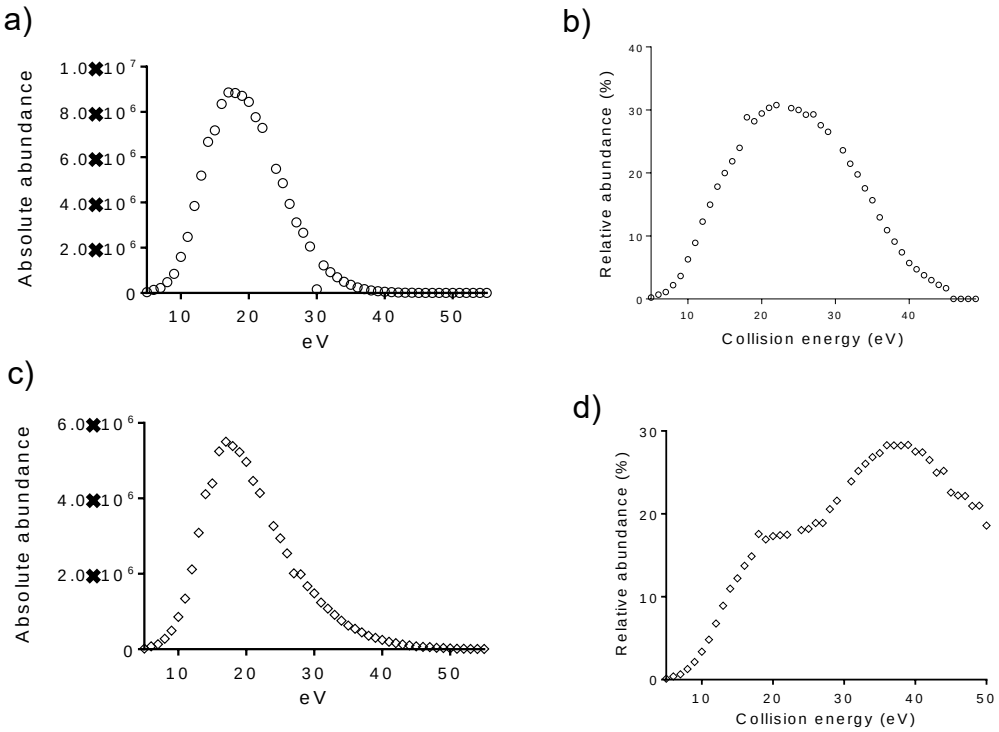


Figure 103 : Profils ERMS calculés en fonction de l'abondance absolue des ions (m/z 199 et m/z 128, A et C), en comparaison avec les courbes ERMS normalisés par le courant ionique total pour les ions m/z 199 et m/z 128 en B et D à partir du précurseur $[C12Glu\text{-}H]^-$ m/z 328 sous activation non-résonante (QqTOF Waters)

L'apport de la mobilité ionique Ces courbes ERMS ont été transposées sur un appareil de mobilité ionique couplée à la spectrométrie de masse, où nous avons pu détecter deux valeurs de CCS pour l'ion fragment m/z 128 et une seule valeur de CCS pour l'ion fragment m/z 199. Par les calculs quantiques et les expériences de recherche d'ions précurseurs, nous avons pu mettre en évidence la présence de deux chemins de mécanismes de fragmentations conduisant à deux ions fragments isomères pour le m/z 128, issus de l'ion précurseur $[C12Glu-H]^-$ et de l'ion fragment $[(C12-H_2O)-H]^-$.

Sur l'appareil cIMS les ions peuvent être fragmentés avant la cellule de mobilité ionique, puis y être injectés. Lorsqu'ils sont séparés par la cellule de mobilité, les ions sont transférés au détecteur. Ils passent consécutivement par une cellule de stockage d'ions (« post array storage »), une cellule constituée d'une succession de lentilles permettant la focalisation du faisceau d'ions (« ion guide ») qui va guider les ions, et d'une cellule de transfert (« XS transfert device). Ensuite, les molécules atteignent le « pusher », c'est-à-dire l'entrée du détecteur. Le passage des ions d'une zone à l'autre dans le spectromètre de masse n'est possible que par une variation de tension. Entre la sortie de cellule de mobilité et l'entrée du détecteur, un potentiel de tension est donc appliqué. Lorsque des ions précurseurs perdurent à la sortie de la cellule de mobilité ionique, ce potentiel peut provoquer leurs dissociations. Ainsi, il est possible de retrouver l'ensemble des ions fragments produits à partir du précurseur à la même valeur de CCS que celle du parent. A l'inverse, si à la sortie de la cellule de mobilité ionique, il n'y a plus d'ion précurseur, seul l'ion précurseur à sa valeur de CCS, et aucun autre rapport masse-sur-charge, sera retrouvé.

L'analyse du mobilogramme à 20 eV lorsque l'ion parent est encore présent à la sortie de la cellule de mobilité et qu'il peut encore être fragmenté post-IM a été réalisée. Lorsque nous nous intéressons à l'ion m/z 128 trois valeurs de CCS apparaissent : celle de l'ion fragment m/z 128, celle de l'ion fragment m/z 310 (perte de H_2O à partir du précurseur) et celle de l'ion précurseur m/z 328. L'analyse du mobilogramme à 40 eV montre la présence d'une seule valeur de CCS pour l'ion fragment m/z 128 car les mesures correspondantes aux ions (m/z 328 et m/z 310) issus d'une fragmentation post-IM sont absents. Par cette expérience, nous confirmons qu'il est possible de produire l'ion fragment m/z 128 à partir de deux composés. Ces données sont en corrélation avec la recherche d'ion précurseur effectuée sur un appareil de type QqQ (Shimadzu TQ 8060), ou la recherche d'ions parents pour l'ion fragment m/z 128 indique que celui-ci peut être formé à partir de l'ion précurseur, ou bien de la perte d'H_2O. Les calculs quantiques démontrent également qu'il est possible, à partir de ces deux précurseurs, d'obtenir deux ions fragments m/z 128 isomères.

A l'inverse, la même expérience pour l'ion m/z 199 montre qu'il est possible d'obtenir cet ion fragment uniquement à partir de l'ion précurseur. Une fois encore, ces données sont en adéquation avec nos hypothèses : la production de deux ions de structures identiques, à partir d'un même précurseur, mais par deux voies de fragmentation différentes.

La dissociation d'un ion précurseur en un ion fragment, est toujours suivie d'une conversion d'énergie interne en énergie cinétique. Celle-ci peut se disperser spatialement, c'est-à-dire que l'on retrouvera des vecteurs Δv dans toutes les directions possibles. C'est ce phénomène qui est à l'origine de la formation de profils Gaussiens. Quand l'ion précurseur se dissocie en un ion fragment et un neutre, les directions des vecteurs Δv diffèrent. Pour les particules légères, il existe une corrélation entre le rapport m/z et sa déviation. Plus le rapport m/z est faible, plus la molécule aura une énergie cinétique dispersée spatialement. Dans ce cas-là, le signal ERMS peut potentiellement s'élargir, sans pour autant perdre son profil Gaussien. Cette interprétation n'est qu'un modèle proposé pour expliquer le profil atypique de l'ion fragment m/z 199 et doit encore être démontré par des calculs physiques réalisés à l'aide du logiciel SIM-ION. Cette interprétation semble corréler avec les expériences réalisées par Cooks[378] dans les années 1980, où il démontre que les ions de faibles m/z sont davantage détectés, parce qu'ils s'éloignent davantage de l'axe, et sont donc davantage dispersés.

Ainsi quand le profil d'un signal lié à un ion produit donné, est dissymétrique comme l'ion m/z 199, cela peut être expliqué en admettant que : (i) deux (ou plus) chemins de dissociation peuvent mener à cet ion et que (ii) leurs seuils d'apparition diffèrent et (iii) donc très probablement de même avec la dispersion de leurs excès d'énergie interne (après conversion en énergie cinétique) alors chacun des chemins pourra être caractérisé par des profils différents qui se chevauchent. Si les différences sont suffisantes et que la cellule de collision est suffisamment discriminante en énergie cinétique, ces profils pourront être distincts, et avoir un « dos de chameau ».

CONCLUSION ET PERSPECTIVES

L'ensemble de mes travaux de thèse ont permis de développer une stratéigie expérimentale de chimie analytique spécifique aux lipopeptides. Celle-ci est composée de plusieurs étapes cruciales et interconnectées, ayant été validées et optimisées individuellement. Il contient notamment une extraction sur phase solide et une méthode d'analyse de chromatographie liquide couplée à la spectrométrie de masse à haute résolution. Chacune des étapes de ce workflow est essentielle pour permettre l'identification et la caractérisation de nouveaux métabolites produits par le microbiote intestinal. Ainsi, lorsque cette stratégie expérimentale de chimie analytique a été appliquée à des culots bactériens d'*Escherichia coli* Nissle 1917 et de *Lactobacillus animalis*, 49 lipopeptides potentiels ont été détectés. A l'aide d'expériences de fragmentation par mode d'excitation résonant et non résonant, et de comparaison à des composés primaires synthétiques, nous avons pu identifier et caractériser vingt-trois nouveaux lipoaminoacides (LpAA) et six lipopeptide. L'annotation de vingt-deux composés a été réalisée avec un niveau 1 d'identification, deux composés à un niveau 2 et dix-huit à un niveau 3. Le développement d'une méthode de quantification de ces composés par LC-QqQ a été réalisée. Ces composés ont ainsi été quantifiés dans différentes souches bactériennes. Nous avons pu démontrer que la production de LpAA et de lipopeptides dans le microbiote intestinal est possible *via* des bactéries gram positives comme *Lactobacillus animalis*, ou gram négatives comme *Escherichia coli* Nissle 1917. Ces résultats démontrent la redondance fonctionnelle des bactéries du microbiote pour produire ces molécules et l'importance de comparer la fonctionnalité du microbiote et pas seulement sa taxonomie afin de déterminer son rôle dans les pathologies. Les lipopeptides identifiés contenant du GABA sont en cours de développement thérapeutique pour être utilisé comme analgésiques contre les douleurs viscérales dans certaines pathologies. Les lipopeptides dopamine, sérotonine et histamine n'ont pas encore pu être identifiés à l'aide de ce workflow analytique. Lorsque les standards seront terminés d'être synthétisés, il sera nécessaire de réoptimiser certaines étapes du workflow, notamment les paramètres de détection en spectrométrie de masse et la méthode d'extraction pour accroitre nos chances d'en découvrir.

La caractérisation et l'élucidation structurale de ces nouveaux composés a permis de mettre en évidence la régénération d'ions fragments inattendus chez les lipoaminoacides acides. Pour une meilleure compréhension et étude de ces fragmentations, nous avons mis en place une nomenclature spécifique aux lipopeptides applicable en protéomique. L'étude des courbes d'évolution ERMS a permis de mettre en évidence pour la première fois la présence de courbes Gaussiennes élargies et de courbes composites. Les calculs quantiques

ainsi que la mobilité ionique cyclique ont permis de mettre en évidence la présence de deux ions fragments isomères produits lorsque nous sommes en présence d'une courbe composite. L'interprétation des courbes Gaussiennes élargies s'explique par la présence de deux chemins réactionnels issus d'un même précurseur formant un seul et même ion fragment à deux seuils différents d'énergie de collision, menant à deux chemins de fragmentation différents.

Pour aller plus loin l'utilisation d'outils pour étudier les réseaux moléculaires serait certainement pertinent. Grâce à ces outils où les fragmentations semblables sont regroupées et mises en évidence, cela pourrait ici augmenter les probabilités de découvertes de nouvelles molécules maintenant que les fragmentations de plus de vingt LpAA ont été déterminées. Au cours de ma thèse nous avons pu faire quelques essais sur les logiciels MetGEM et GNPS mais le mode de scan utilisé sur l'instrument Waters complique cette analyse. De plus, il n'est pas encore possible de rechercher un ou plusieurs fragments d'intérêt dans le réseau moléculaire sans en analyser l'ensemble, ce qui rend le processus chronophage. Ce travail ouvre également tout un axe de recherche sur les lipoaminoacides acides qui se comportent de manière inattendue et dont certaines courbes d'évolution ERMS sont des Gaussiennes élargies, et dont nous avons surement encore beaucoup à apprendre. L'ER-IMS automatisée pourrait être d'une grande utilité pour distinguer des ions fragments isomères dont les chemins de fragmentation sont proches et dépendant de l'énergie de collision.

Nous sommes certainement très loin de connaitre la richesse de production de métabolites spécialisés par les bactéries. Les activités biologiques de ces nouveaux composés doivent encore être démontrées et les voies de biosynthèse impliquées dans leur production également. Avec ce travail nous pouvons maintenant proposer la quantification absolue de plus de quarante composés potentiellement impliqués dans certaines pathologies intestinales sur des selles ou des biopsies intestinales de patients. Certains de ces composés pourront peut-être être impliqués dans la prise en charge de ces pathologies en tant que traitement ou comme biomarqueurs *via* leurs dosages.

ANNEXE 1: La théorie du quasi-équilibre

La théorie dite du quasi-équilibre (quasi-equilibrium theory, QET), a été introduite par Rosenstock[379] dans les années 1950. Cependant, elle avait déjà été exprimée lors de la théorie RRKM (Rice-Ramsperger-Kassel-Marcus) ou l'étude portait sur les constantes de vitesse de dissociation des ions qui décrivent directement l'abondance des ions fragments présents dans le spectre de masse. En d'autres termes ces constantes de vitesse k permettent de mieux rendre compte des spectres de masses qui résultent de nombreuse fragmentations des ions dans la source où la durée de résidence est courte (c'est-à-dire de l'ordre de la microseconde en moyenne). Pour simplifier si on considère qu'approximativement la moitié des espèces moléculaires se sont dissociées pendant la microseconde (temps de demi-réaction $t_{1/2}$) alors la constante de vitesse devra être $k \geq Ln2/t_{1/2}$ (soit $0{,}693 \times 10^6\ s^{-1}$). Ainsi, qualitativement k devra être au moins de l'ordre de $10^6\ s^{-1}$. Afin de rendre compte de l'abondance relative des ions fragments, donc de leurs constantes de vitesse de formation respectives, il faut connaître la température T des espèces ionisées qui se dissocient. Cette donnée est nécessaire pour appliquer la loi macroscopique d'Arrhenius ($k=Ae^{-E°/RT}$ avec A le facteur pré-exponentiel, E° l'énergie d'activation, et R la constante des gaz parfaits). Malheureusement, les conditions d'ionisation en 1950 ne permettaient pas d'avoir accès à la température T, puisque les ions étaient formés sous vide, par conséquence il n'était pas possible, du point de vue thermodynamique, de définir une température T. En effet, les ions dans ces conditions ne sont pas dans un état d'équilibre thermique résultant d'un grand nombre de collisions avec le milieu ambiant. Ainsi il a fallu trouver une autre expression de la constante de vitesse sans que la notion de température apparaisse. Il s'agit de la QET/RRKM[380] où c'est l'énergie interne (E) qui remplace la température. Cette théorie se base sur le principe d'ergodicité. En résumé, cette théorie considère que toute l'énergie interne E est distribuée « spontanément » sur l'ensemble des liaisons chimiques de l'ion parent, et que, c'est la liaison la plus fragile dans ces conditions, qui sera rompue. La connaissance de l'énergie interne (E) des ions qui se dissocient est importante afin de comprendre l'évolution des réactions de fragmentation en fonction de E. Historiquement, les ions en question étaient des ions moléculaires ($M^{+\bullet}$) produits en ionisation électronique et possédant une grande distribution d'énergie vibrationnelle répartie dans l'état fondamental et dans les états excités. Ces derniers états électroniques étaient atteints lors de l'ionisation par électron de grande énergie cinétique. Les ions moléculaires étant des systèmes isolés (au sens thermodynamique), excepté la perte d'énergie interne par radiation, une partie de l'énergie vibrationnelle reste dans les états électroniques après que l'autre partie soit redistribuée dans

l'état fondamental. Ainsi l'énergie interne E correspond à cette large distribution d'énergie vibrationnelle dans tous les états électroniques impliqués en plus de l'état fondamental.

C'est à partir de ces considérations que l'expression de k a été construite en fonction de l'énergie interne E acquise par l'ion parent pendant l'étape d'ionisation. Cette constante de vitesse dépend de l'excès d'énergie interne (vibrationnelle) au-dessus du niveau de l'état de transition (noté ‡) (E_0 étant l'énergie d'activation). Cet excès d'énergie normalisée par rapport à E correspond à la différence ($E-E_0$)/E (ou ($1-E_0/E$)). Cet excès d'énergie normalisé est réparti de manière inégal sur l'ensemble des liaisons chimiques n selon leurs propriétés respectives. Toutefois pour faire un modèle simple, il est considéré que la répartition se fera sur l'ensemble des liaisons sans compter la liaison rompue. Pour cela, il faut respecter le principe de conservation de l'énergie. Ainsi on peut considérer que cet excès d'énergie sera distribué sur (*n-1)* liaisons. L'expression qui répond au mieux à cette distribution d'énergie correspond à l'équation de proportionnalité suivante (Eq. 5) :

$$\text{Eq. 5 : } k\ \alpha\ \left[\frac{E-E0}{E}\right]^{n-1}$$

Cette expression n'exprime que la partie énergétique, c'est-à-dire l'énergie liée au second terme de l'équation d'Arrhenius ($e^{-E°/RT}$). Elle ne considère donc pas les propriétés propres de l'état de transition qui implique des liaisons rompues et/ou créées (etat de transition tendu pour un réarrangement : RR, ou lâche pour une rupture simple : RS). Ces propriétés de transitions sont liées au terme pré-exponentiel A de l'équation d'Arrhénius. En d'autres termes, cela correspond aux densité d'états (liées aux fréguences de vibration $\upsilon_j^{\ddagger}$) caractérisant ces états de transitions : faibles dans le premier cas et grandes dans le second. Ainsi, l'équation Eq :5 devient :

$$\text{Eq. 6 : } k = \upsilon\ \left[\frac{E-E0}{E}\right]^{n-1} \quad \text{avec} \quad \upsilon = \left[\frac{\Pi j\ \upsilon j\ddagger}{\Pi i \upsilon i}\right]$$

Cette présentation de ce paramètre υ sera nommé facteur de fréquence. Ce facteur de fréquence sera proportionel aux produits de ces différents états rapportés à ceux de l'état initial.

En ce qui concerne l'étude menée dans ma thèse, les ions sont formés en électro-nébulisation, leurs énergies internes sont donc distribuées seulement dans l'état fondamental des formes protonées ou déprotonées. Cela procure comme avantage de ne pas avoir à considérer les états électriques excités. En ESI, les spectres de masse sont dépendants de l'instrument et des conditions de désorption des agrégats chargés et de leur désolvatation, ce qui revient à dire que l'énergie interne est sensible aux conditions expérimentales. Tout d'abord à pression atmosphérique les agrégats chargés sont caractérisés par une

température à l'équilibre avec son environnement. Cette température « macroscopique » de la source[337] ne correspondra pas à la température microscopique des ions issus de la désolvatation[381]. Ce sera cette température qui pourrai être utilisée dans l'équation d'Arrhenius (par approximation).

Cela a été montré en utilisant des « molécules thermomètres », comme par exemple des sels de benzylpyridium, avec la méthode des ions précurseurs survivants[337,382]. Ainsi, il est possible de connaître approximativement l'énergie interne des ions après leur désolvatation. C'est ensuite, dès que la charge se situe sur un site, qu'elle induit le clivage d'une (ou de plusieurs) liaison(s). Ce clivage sera observé dans les spectres de masse si $k \geq 10^6\ s^{-1}$ (dissociations promptes dans la source) et dans les spectres de collisions (où la fenêtre temporelle pour se dissocier est de 10 µs) si $k \geq 10^5\ s^{-1}$.

Toutefois, les mesures de températures n'étant pas si simples que cela pour les ions, nous préférons largement travailler en considérant l'énergie interne. En ce qui nous concerne, on ne raisonnera que qualitativement. Par exemple, un réarrangement sera plutôt favorisé en basse énergie (faible facteur de fréquence υ et fable énergie d'activation), alors qu'une rupture simple le sera en plus haute énergie (plus fortes valeurs de υ et d'E_0). La théorie cinétique doit nous amener à une telle conclusion, au contraire une forte énergie d'activation pour RR fera que ce réarrangement ne sera jamais observé.

Ainsi lors d'un processus par étapes, passant par une isomérisation de la molécule déprotonée en complexe ion-dipôle, alors sa dissociation directe sera le processus favorisé en haute énergie. Par contre la réaction de transfert entre partenaires sera favorisée en basse énergie. Il y a compétition entre ses deux processus. Si on admet qu'il y a uniquement ces deux dissociations alors les courbes des ions fragments qui constitut l'ensemble ERMS, l'évolution du processus direct apparaitra vers les hautes énergies alors qu'en basse énegie se sera le transfert de proton qui sera observé pour disparaitre dès que l'énergie de collision sera élevée.

ANNEXE 2: Supplementary information publication 1

Solvents and reagents were used as obtained from the suppliers (Aldrich, Alfa Aesar, Acros, Fluorochem, TCI, and Iris-Biotech for 2-chlorotrityl resin). For intermediate compounds, reactions were monitored by TLC using plates precoated with silica gel 60 (Merck). Reaction components were visualized by using a 254 nm UV lamp and treatment with basic $KMnO_4$ solution. For final compounds, TLC RP-18 (Merk) using plates precoated with silica gel 60 (typically ACN/Water 90/10 + 0.1 % TFA) were used instead, and they were visualized with Hanessian's blue stain (typically as a dark blue spot). Column chromatography in normal phase were performed on Reveleris X2 by using Macherey-Nagel Chromabound® (40-64 µm) or Interchim SIHP-PuriFlash® column (30 µm) for spherical column. In reversed phase, Interchim PuriFlash® C4 or C18 (PT-15 µm, 12g or 55g) were used instead. Following abbreviations are used: NMP = N-Methyl-2-pyrrolidone, DCM = dichloromethane, DMF = dimethylformamide, TFA = trifluoroacetic acid, TIPS = triisopropyl silane, Cyc = cyclohexane, HFIP = Hexafluoroisopropanol, Fmoc = fluorenylmethoxycarbonyl, HBTU = N,N,N′,N′-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate, HOBt = 1-Hydroxybenzotriazole, NMM = N-Methylmorpholine, OxymaPure = Ethyl cyano(hydroxyimino)acetate, DIC = N,N′-Diisopropylcarbodiimide. The ^{1}H NMR spectra are reported as follow: chemical shift in ppm [multiplicity, coupling constant(s) J in Hz, relative integral]. The multiplicities are defined as follow: l = large, m = multiplet, s = singlet, d = doublet, t = triplet, q = quadruplet, p = quintuplet, or combinations thereof. LCMS analyses were carried out on a Waters Micromass with Alliance 2695 chain with a Chromolite HR C18 column (25 × 4.6 mm, Merck Inc.) monitoring at 214 nm with positive ion mode for the ion m/z value detection. Solvents for LCMS were water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (solvent B). Compounds were eluted at a flow rate of 3mL min−1 by a linear gradient of 0–100% solvent B over 2.5 min, and finally 100% solvent B for 1 min before equilibrating the column back to 0% solvent B over 1 min.

HPLC preparative purification: Samples were prepared in DMSO. The LC/MS autopurification system consisted of a binary pump Waters 2525, an injector fraction collector Waters 2676, coupled to a Waters Micromass ZQ spectrometer (electrospray ionization mode, ESI+). Purifications were carried out using a Luna® 5 µm C18 100 Å, LC Column 100 × 21.2 mm, AXIA™ Packed. A flow rate of 20 mL min−1 and a gradient of 40–60% B over 10 min were used. Eluent A: water with 0.1% TFA; eluent B: acetonitrile with 0.1% TFA. Positive ion electrospray mass spectra were acquired at a solvent flow rate of 204 µL min−1. Nitrogen was used for both the nebulizing and drying gas. The mass spectrum data were obtained in a scan mode ranging from 100 to 1000 m/z in 0.1 s intervals; 10 scans were summed up to get the final mass spectrum. Collection control trigger is set on single protonated ion with a MIT (minimum intensity threshold) of 7.105.

Example 1: *Amide bond formation with HBTU coupling.* To a solution of H-Leu(OtBu).HCl (134 mg, 0.60 mmol, 1.0 eq.) and lauric acid (180 mg, 0.90 mmol, 1.5 eq.) in DCM (6 mL, [C] = 0.1 M) were added HBTU (341 mg, 0.90 mmol, 1.5 eq.), HOBt (8 mg, 60 µmol, 0.1 eq.), and

NMM (251 µL, 2.28 mmol, 3.8 eq.). The reaction mixture was stirred overnight. Silica was added and the solvent was evaporated. The crude was purified by flash chromatography on Reveleris X2 (25g, dry sample), monitoring with ELSD detector. Cyc/EtOAc: 100/0 to 80/20 (15 CV) to obtain C12-Leu(OtBu) (221 mg, Y = 97%) was obtained as a colorless oil.

Example 2: *Deprotection of trityl amide and terbutyl ester.* To a solution of C12-Leu(OtBu).HCl (197 mg, 0.533 mmol, 1.0 eq.) in DCM (1.1 mL, [C] = 0.5 M) was slowly added TFA (1.1 mL, 14.4 mmol, 27 eq.) at RT. Resulting yellow solution was stirred for 5h at RT. Reaction mixture was evaporated under vacuum and coevaporated three times with ACN to remove traces of TFA. The resulting solid was purified by flash chromatography on Reveleris X2 (12g, dry sample), monitoring with ELSD detector. DCM/MeOH + 1% HCO_2H: 100/0 to 93/7 (15 CV). Product was coevaporated with toluene and Et_2O, dried overnight under vacuum to give (123 mg, Y = 67 %) of C12-Leu-OH as a white solid.

Example 3: *General procedure for Solid Phase Synthesis (SPS) of aminolipids*. In a SPS reactor, 2-chlorotrityl resin (max loading 1.6 mmol/g, 313 mg, 500 µmol, 2.5 eq.) was swollen in DCM (5 mL, 5 min). Solvent was evacuated and DCM was added (5 mL). Appropriate Fmoc-Amino Acid (200 µmol, 1.0 eq.) and NMM (800 µmol, 4.0 eq.) were added and reaction mixture was stirred for 3h. Liquid was evacuated under vacuum and resin was capped with DCM/MeOH/NMM (17/2/1) (3x5 mL, 5 min each), washed with DCM (3x5 mL, 1 min each), DMF (3x5 mL, 1 min each), DCM (3x5 mL, 1 min each). Then, Fmoc group was removed with piperidine/DMF (2/8) (2x5 mL, 10 min each). Resin was washed with DMF (3x5 mL, 1 min each), DCM (3x5 mL, 1 min each), NMP (3x5 mL, 1 min each). Resin was swollen in NMP (5 mL), and the appropriate fatty acid (750 µmol, 5.0 eq.), HBTU (750 µmol, 5.0 eq.), HOBt (15 µmol, 0.1 eq.), NMM (7.5 mmol, 10.0 eq.) were added. Reaction mixture was stirred for 1h30 to overnight, monitoring with Kaiser's test. Liquid was evacuated under vacuum and resin was washed with NMP (3x5 mL, 1 min each), Pyridine (3x5 mL, 1 min each), MeOH (3x5 mL, 1 min each), DCM (6x5 mL, 1 min each). Resin was cleaved with TFA/TIPS/Water (95/5/2.5) (2*5min). Resin was washed with DCM (3x5 mL, 1 min each). Filtrate was concentrated under vacuum, coevaporated with ACN several times to remove residual TFA. The white residue was then triturated two times in ACN and filtered to give the desired lipopeptide without further purification (except specified).

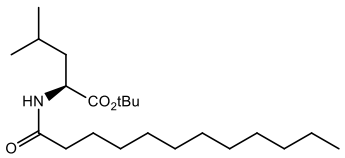
Example 4: *Synthesis of aminolipids containing a mono-unsaturated fatty acid.* Steps were identical to *example 3*, except that coupling reactions were carried out with OxymaPure (2.0 eq.)/DIC (2.0 eq.)/Fatty acid (2.0 eq.) (5 min of preactivation) during one night (> 16 h) instead of HBTU/HOBt/NMM, starting with Fmoc-Asn-OH (53 mg, 150 µmol, 1.0 eq.). Resin was cleaved with DCM/HFIP (7/3) for 1h30. After evaporation, the residue was then triturated two times in Et_2O and filtered to give the desired lipopeptide.

Example 5: *Hydrolysis of methyl ester.* A solution of C12-His(OMe) in THF/NaOH (3M) was stirred overnight. pH was adjusted to 3-4 with HCl 1M and crude reaction mixture was evaporated under vacuum. Residue was taken-up with MeOH, filtered on Celite®, washed with MeOH. Celite® was added to the filtrate, concentrated under vacuum and purified by flash

chromatography on Reveleris X2 (C4 - 12 g, Celite deposite). Water/MeOH: 40/60 to 0/100 (15CV). C12-His-OH (68 mg, Y = 57 %) was obtained as white solid.

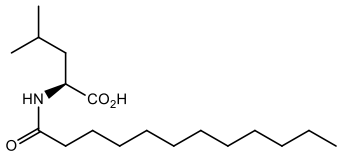
Example 6: *Synthesis of GABA aminolipids*. In a SPS reactor, flushed with argon, 2-chlorotrityl resin (max loading 1.6 mmol/g, 1.0 g, 1.6 mmol, 3.0 eq.) was swollen in DCM (12 mL, 5 min). Solvent was evacuated and DCM was added (12 mL). Fmoc-GABA-OH (1.07 g, 3.30 mmol, 3.0 eq.) - dissolved in a minimum of NMP -$HgCl_2$ (358 mg, 1.32 mmol, 1.2 eq.), NMM (726 µL, 6.60 mmol, 6.0 eq.) were added. The pink mixture was stirred overnight (>16h). Resin was capped, following example 3. Loading was determined by spectroscopy at 290 nm (Fmoc deprotection: 1*15 min) and 1.35 g of Fmoc-GABA 2-chlorotrityl resin was obtained with a loading of 0.75 mmol/g. Resin was dried overnight before use. GABA-aminolipids were obtained following the steps described in example 3.

C12-Leu(OtBu):



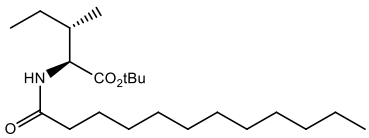
Following the example 1: Y = 99 % (221 mg). Colorless oil.
R_f (Cyc/EtOAc 80/20) = 0.35.
^{1}H NMR ($CDCl_3$, 300 Mhz): δ 5.84 (d, *J* = 8.5 Hz, 1H), 4.53 (td, *J* = 8.5, 5.3 Hz, 1H), 2.19 (lt, *J* = 8.2 Hz, 2H), 1.70 – 1.56 (m, 4H), 1.46 (s, 9H), δ 1.36 – 1.12 (m, 17H), 0.95-0.92 (m, 6H), 0.87 (t, *J* = 6.8 Hz, 3H).
^{13}C NMR ($CDCl_3$, 75 Mhz): δ 172.9, 172.7, 82.0, 51.2, 42.3, 36.9, 32.1, 29.7, 29.6, 29.5, 29.4, 28.1, 25.8, 25.1, 23.0, 22.8, 22.3, 14.3.
IR (neat) u_{max} (cm^{-1}): 3397-3250, 2962, 2923, 2855, 1737, 1646, 1536, 1460, 1368, 1250, 1146, 849.

C12-Leu-OH:



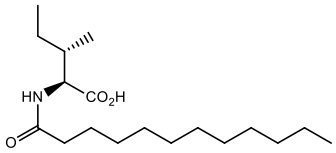
Following the example 2, starting with C12-Leu(OtBu) (197 mg, 533 µmol, 1.0 eq.) : Y = 67 % (112 mg). White solid.
^{1}H NMR ($CDCl_3$, 300 MHz): δ 5.84 (d, *J* = 7.9 Hz, 1H), 4.59 (td, *J* = 8.5, 4.7 Hz, 1H), 2.24 (t, *J* = 7.6 Hz, 2H), 1.79-1.55 (m, 4H), 1.37 – 1.11 (m, 17H), 0.96 (d, *J* = 6.3 Hz, 3H), 0.95 (d, *J* = 6.3 Hz, 3H), 0.88 (t, 3H).
^{13}C NMR ($CDCl_3$, 75 MHz,): δ 176.4, 174.3, 51.0, 41.1, 36.7, 32.1, 29.8, 29.6, 29.5, 29.3, 25.72, 25.02, 22.99, 22.84, 21.98, 14.28.
Mp (°C): 109.5-110.6.
IR (neat) u_{max} (cm^{-1}): 3324, 2957, 2922, 2853, 1695, 1615, 1554, 1467, 1272, 1238, 1168, 969.

C12-Ile(OtBu):



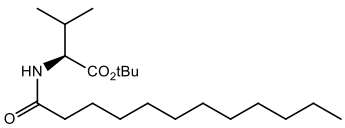
Following the example 1, starting with H-Ile-OtBu.HCl (134 mg, 0.60 mmol, 1.0 eq): Y = 73 % (131 mg). Colorless oil.
R_f (Cyc/EtOAc 80/20) = 0.35.
^{1}H NMR ($CDCl_3$, 300 MHz): δ 5.96 (d, *J* = 8.5 Hz, 1H), 4.51 (dd, *J* = 8.5, 4.5 Hz, 1H), 2.21 (t, J = 7.5 Hz, 2H), 1.91-1.81 (m, 1H), 1.69-1.54 (m, 3H), 1.53-1.37 (m, 1H), 1.46 (s, 9H), 1.37-1.21 (m, 16H), 0.91 (t, *J* = 7.6 Hz, 3H), 0.87 (d, *J* = 9.4 Hz, 3H), 0.86 (t, *J* = 6.9 Hz, 3H).
^{13}C NMR ($CDCl_3$, 75 MHz): δ 172.9, 171.5, 82.2, 56.6, 38.5, 37.0, 32.1, 29.8, 29.7, 29.5, 29.4, 28.3, 25.9, 25.6, 22.9, 15.5, 14.3, 11.9.
IR (neat) u_{max} (cm^{-1}): 3397-3230, 2962, 2924, 2855, 1737, 1646, 1635, 1460, 1367, 1250, 1146, 848.

C12-Ile-OH:



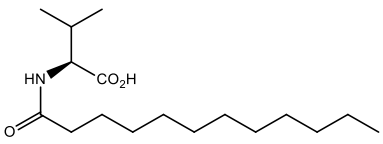
Following the example 2, starting with C12-Ile(OtBu) (132 mg, 357 µmol): Y = 82 % (92 mg). White solid.
^{1}H NMR ($CDCl_3$, 300 MHz): δ 6.01 (d, *J* = 8.4 Hz, 1H), 4.61 (dd, *J* = 8.4, 5.0 Hz, 1H), 2.25 (t, *J* = 7.6 Hz, 2H), 2.02-1.89 (m, 1H), 1.63 (p, *J* = 7.1 Hz, 3H), 1.50-1.44 (m, 1H), 1.38-1.12 (m, 17 H), 0.96 (d, *J* = 7.0 Hz, 3H), 0.94 (t, *J* = 7.7 Hz, 3H), 0.87 (t, *J* = 6.7 Hz, 3H).
^{13}C NMR ($CDCl_3$, 75 MHz): δ 175.5, 174.0, 56.6, 37.7, 36.8, 32.1, 29.8, 29.6, 29.5, 29.4, 25.8, 25.2, 22.8, 15.6, 14.3, 11.7.
Mp (°C): 106.7-107.4. IR (neat)
u_{max} (cm^{-1}): 3299, 2952, 2920, 2877, 2851, 1695, 1597, 1550, 1467, 1364, 1252, 1149, 1114, 1000, 975, 721, 692.

C12-Val(OtBu):



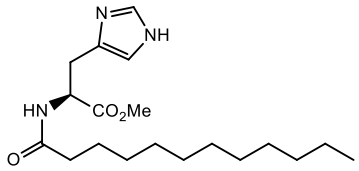
Following the example 1, starting with H-Val-OtBu.HCl (126 mg, 0.60 mmol, 1.0 eq.): Y = 99 % (213 mg). Colorless Oil.
R_f (Cyc/EtOAc 80/20) = 0.33.
^{1}H NMR (300 MHz, $CDCl_3$): d 5.92 (d, *J* = 8.6 Hz, 1H), 4.48 (dd, *J* = 8.6, 4.5 Hz, 1H), 2.22 (lt, *J* = 7.5, 2H), 2.17-2.07 (m, 1H), 1.69-1.55 (m, 2H), 1.47 (s, 9H), 1.35-1.18 (m, 16H), 0.89 (d, *J* = 6.9 Hz, 3H), 0.92 (d, *J* = 6.9 Hz, 3H), 0.87 (t, *J* = 6.8 Hz, 3H).
^{13}C NMR ($CDCl_3$, 75 MHz): δ 173.1, 171.6, 82.1, 57.2, 37.0, 32.1, 31.7, 29.7, 29.6, 29.5, 29.4, 28.2, 25.9, 22.8, 19.0, 17.8, 14.3.
IR (neat) u_{max} (cm^{-1}): 3400-3234, 2961, 2924, 2855, 1736, 1647, 1536, 1466, 1368, 1146, 848.

C12-Val-OH:



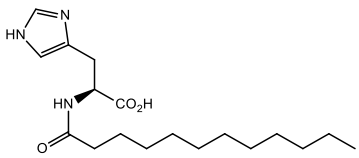
Following the example 2, starting with C12-Val(OtBu) (196 mg, 551 µmol): Y = 76 % (125 mg). White solid.
^{1}H NMR ($CDCl_3$, 300 MHz): δ 6.03 (d, *J* = 8.7 Hz, 1H), 4.57 (dd, *J* = 8.7, 4.8 Hz, 1H), 2.29 – 2.17 (m, 3H), 1.64 (p, *J* = 7.2 Hz, 2H), 1.37-1.18 (m, 16H), 0.98 (d, *J* = 9.9 Hz, 3H), 0.96 (d, *J* = 9.9 Hz, 3H), 0.88 (t, *J* = 6.7 Hz, 3H).
^{13}C NMR ($CDCl_3$, 75 MHz): δ 175.5, 174.2, 57.2, 36.8, 32.1, 31.0, 29.8, 29.6, 29.5, 29.4, 25.9, 22.8, 19.2, 17.8, 14.3.
Mp (°C): 106.7-107.4.
IR (neat) u_{max} (cm^{-1}): 3325, 2963, 2920, 2852, 1694, 1616, 1548, 1468, 1248, 972, 664.

C12-His(OMe):



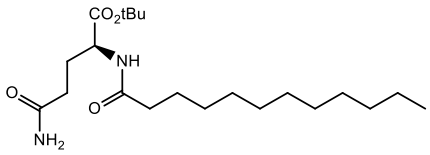
Following the example 1, starting with H-His-OMe.HCl (147 mg, 0.60 mmol, 1.0 eq.): Y = 57 % (120 mg). White wax. Crude was purified by Flash Chromatography on Reveleris X2 (40 g, silica deposite). DCM/MeOH: 100/0 to 9/1 (15 CV).
With neutral alumina TLC, R_f (DCM/MeOH 9/1) = 0.52.
^{1}H NMR ($CDCl_3$, 300 MHz): δ 7.56 (s, 1H), 7.02 (d, *J* = 7.3 Hz, 1H), 6.80 (s, 1H), 4.815 (ddd, *J* = 7.7, 5.2, 5.1, 1H), 3.69 (s, 3H), 3.14 (dd, *J* = 15.2, 5.4, 1H), 3.07 (dd, *J* = 15.2, 4.9 Hz, 1H), 2.23 (t, *J* = 7.7 Hz, 2H), 1.62 (p, *J* = 7.5 Hz, 2H), 1.37-1.18 (m, 16H), 0.88 (t, *J* = 6.5 Hz, 3H).
^{13}C NMR (MeOD, 75 MHz): d 176.3, 173.5, 136.3, 134.6, 118.1, 53.9, 52.7, 36.7, 33.1, 30.8, 30.6, 30.49, 30.46, 30.2, 30.1, 26.9, 23.8, 14.4.
IR (neat) u_{max} (cm^{-1}): 3234, 2915, 2849, 1756, 1646, 1529, 1160, 817, 659.

C12-His-OH:



Following the example 5: Y = 57 % (68 mg). White solid.
^{1}H NMR (MeOD, 300 MHz): δ 8.41 (s, 1H), 7.15 (s, 1H), 4.54 (dd, *J* = 7.5, 5.2 Hz, 1H), 3.22 (dd, *J* = 15.1, 5.2 Hz), 3.02 (dd, *J* = 15.1, 7.5 Hz, 1H), 2.23-2.17 (m, 2H), 1.55 (p, *J* = 6.8 Hz, 2H), 1.40-1.20 (m, 16H), 0.90 (t, 6.7 Hz, 3H).
^{13}C NMR (MeOD, 75 MHz): d 176.0, 175.8, 134.8, 132.8, 118.3, 54.6, 37.1, 33.1, 30.8, 30.7, 30.6, 30.5, 30.3, 29.6, 26.9, 23.7, 14.4.
Mp (°C): 163-164.5.
IR (neat) u_{max} (cm^{-1}): 3333, 3122, 2919, 2850, 1635, 1598, 1544, 1514, 1470, 1382, 860, 844.

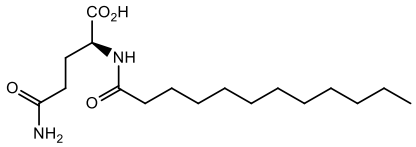
C12-Gln(OtBu):



Following the example 1, starting with H-Gln-OtBu.HCl (143 mg, 0.60 mmol, 1.0 eq.): Y = 99 % (200 mg). Crude was purified by flash Chromatography on Reveleris X2 (12g – C4, Celite

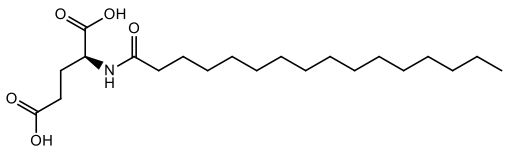
deposite®). Water/MeOH: 5/5 to 0/100 (15 CV) to obtain the desired product as a pale pink solid.
With neutral alumina, R_f (DCM/MeOH 95/5) = 0.52.
^{1}H NMR ($CDCl_3$, 300 MHz): d 6.60 (s, 1H), 6.33 (d, *J* = 7.8 Hz, 1H), 5.39 (s, 1H), 4.48 (ddd, *J* = 9.5, 7.8, 4.2 Hz, 1H), 2.33 – 2.12 (m, 5H), 1.93-1.82 (m, 1H), 1.67 – 1.57 (m, 2H), 1.46 (s, 9H), 1.36-1.20 (m, 16H), 0.87 (t, *J* = 6.7 Hz, 3H).
^{13}C NMR ($CDCl_3$, 75 MHz): 174.7, 174.0, 171.4, 82.9, 52.3, 36.8, 32.3, 32.1, 29.7, 29.6, 29.6, 29.5, 29.4, 28.1, 25.7, 22.8, 14.3. Mp (°C): 96-97.
IR (neat) u_{max} (cm^{-1}): 3394, 3333, 3205, 2917, 2850, 1727, 1648, 1523, 1415, 1392, 1365, 1159, 848.

C12-Gln-OH:



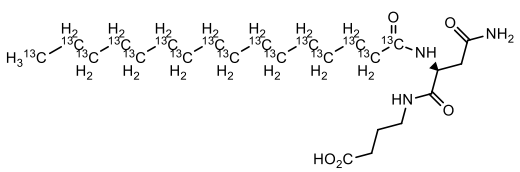
Following the example 2, starting with C12-Gln(OtBu) (189 mg, 491 µmol): Y = 60 % (97 mg). Crude was purified by flash Chromatography on Reveleris X2 (12g – C4, Celite deposite®). Water/MeOH: 70/30 to 20/80 (15 CV) to obtain the desired product as a white solid.
^{1}H NMR (MeOD, 300 MHz): δ 4.39 (dd, *J* = 9.1, 4.9 Hz, 1H), 2.34 – 2.11 (m, 5H), 2.00 – 1.88 (m, 1H), 1.62 (p, *J* = 6.8 Hz, 2H), 1.40 – 1.23 (m, 16H), 0.89 (t, *J* = 6.7 Hz, 3H).
^{13}C NMR (MeOD, 75 MHz): d 177.7, 176.5, 175.0, 53.2, 36.8, 33.1, 32.8, 30.76, 30.65, 30.5, 30.3, 28.5, 26.9, 23.7, 14.4.
Mp (°C): 132.2.
IR (neat) u_{max} (cm^{-1}): 3427, 3353, 3311, 3202, 2958, 2917, 2850, 1730, 1673, 1650, 1622, 1546, 1523, 1463, 1443, 1211, 674.

C16-Glu-OH:



Following the example 3, starting with Fmoc-Glu(tBu).H_2O (89 mg, 200 µmol, 1.0 eq.): Y = 69 % (53 mg). White solid.
^{1}H NMR (MeOD, 300 MHz): δ 4.43 (dd, *J* = 9.3, 4.9 Hz, 1H), 2.40 (t, *J* = 7.5 Hz, 2H), 2.25 (t, J = 7.5 Hz, 2H), 2.22 – 2.09 (m, 1H), 1.99-1.86 (m, 1H), 1.62 (p, J = 7.1 Hz, 2H), 1.28-1.18 (m, 24H), 0.90 (t, *J* = 6.6 Hz, 3H).
^{13}C NMR (MeOD, 75 MHz): δ 176.5, 176.3, 175.0, 52.9, 36.8, 33.1, 31.2, 30.8, 30.7, 30.5, 30.3, 27.8, 26.9, 23.8, 14.5.
Mp (°C): 118.7-119.7.
IR (neat) u_{max} (cm^{-1}): 3329, 3076, 2917, 2850, 1728, 1673, 1643, 1593, 1542, 1471, 1454, 1409, 1319, 1272, 1210, 1176, 988, 852, 780, 719.

^{13}C16-Asn-GABA-OH:

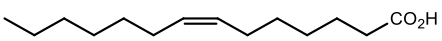


See example 4, starting with Fmoc-GABA-2-chlorotrityl resin (see example 6) (107 mg, 75 µmol, 0.70 mmol/g, 2.5 eq.) and Asn(Trt) (45 mg, 75 µmol, 1.0 eq.): Y = 79 % (27.8 mg, 97.5 % purity LC). White solid.

^{1}H NMR (500 MHz, DMSO-d_6): d 12.05 (s, 1H), 7.92 (dd, *J* = 8.1, 4.0 Hz, 1H), 7.77 (t, *J* = 5.8 Hz, 1H), 7.26 (s, 1H), 6.86 (s, 1H), 4.56 – 4.38 (m, 1H), 3.03 (q, *J* = 6.6 Hz, 2H), 2.45 (dd, *J* = 15.3, 6.1 Hz, 1H), 2.32 (dd, *J* = 15.3, 7.7 Hz, 1H), 2.18 (lt, $^{3}J_{H\text{-}H}$ = 7.5 Hz, 2H), 2.06 (dm, $^{1}J_{C\text{-}H}$ = 118 Hz, 2H), 1.58 (lp, $^{3}J_{H\text{-}H}$ = 7.1, 3H), 1.23 (dm, $^{1}J_{C\text{-}H}$ = 118 Hz, 25 H), 0.86 (ddt, $^{1}J_{C\text{-}H}$ = 125 Hz, $^{2}J_{C\text{-}H}$ = 12.0 Hz, $^{3}J_{H\text{-}H}$ = 7.3 Hz, 3H).

Mp (°C): > 181, thermal degradation.
IR (neat) u_{max} (cm^{-1}): 3412, 3307, 2945, 2910, 2844, 1699, 1665, 1649, 1607, 1563, 1531, 1435, 1413, 1215, 718.

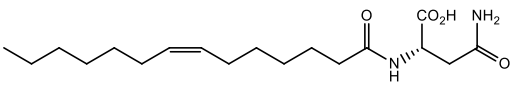
C14(n-7)-OH :



To a light brown suspension of (6-carboxyhexyl)triphenylphosphonium bromide (1.98 g, 4.2 mmol, 2.1 eq.), gently dried under vacuum with heatgun for 5 min, in degassed (bubbling for 10 min with a needle) dry THF (4.7 mL, [C] = 0.5 M) under Ar was added dropwise at 0°C KHMDS (8.0 mL, 1 M in THF, 8.0 mmol, 4.0 eq.). Resulting brown solution was stirred at RT for 1h at 0°C, and cooled at -8°C. A cooled solution (-78°C) of heptanal (282 µL, 2.00 mmol, 1.0 eq.) in dry degassed THF (2.0 mL, [C] = 1 M) under Ar was added dropwise to the ylide at -78°C. Mixture was stirred at -78°C to RT for 2h. TLC (Cyc/EtOAc: 90/10 – $KMnO_4$) showed full conversion of SM. Reaction mixture was quenched with KH_2PO_4 (0.1 M), aqueous phase was carefully acidified with HCl (1M) at pH= 2/3, and extracted with EtOAc (2x). Combined organic phase was dried over $MgSO_4$, filtered, concentrated under vacuum to give crude. Crude was purified by flash Chromatography on Reveleris X2 (dry sample, 40g column). Cyc/iPrOH: 100/0 to 95/5 (12CV). C14(n-7)-OH (305 mg, Y = 67 %) was obtained as a yellow oil, in a mixture with (E) compound (< 5% by LCMS).
^{1}H NMR ($CDCl_3$, 300 MHz): d 5.53 – 5.20 (m, 2H), 2.35 (t, *J* = 7.5 Hz, 2H), 2.02 (p, *J* = 6.8 Hz, 4H), 1.69-1.58 (m, 2H), 1.51 – 1.10 (m, 12H), 1.03 – 0.65 (m, 3H).
^{13}C NMR ($CDCl_3$, 300 MHz): d 180.1, 130.4, 129.5, 34.2, 31.9, 29.9, 29.5, 29.1, 28.98, 28.8, 27.4, 27.1, 24.7, 22.8, 14.3.

IR (neat) u_{max} (cm^{-1}): 3004, 2924, 2856, 1706, 1458, 1412, 1278, 1226, 935, 724.

C14(n-7)-Asn-OH:



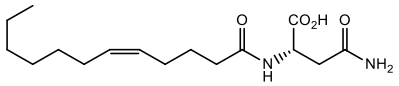
Following the example 4, starting with Fmoc-Asn-OH (53 mg, 150 µmol, 1.0 eq.), crude reaction mixture was purified by flash Chromatography on Reveleris X2 (12 g - C4, Celite deposite). Water/[ACN-iPrOH 1:1] + 0.1 % TFA, 70/30 to 40/60 (12 CV), to obtain C14(n-7)-Asn-OH
(17 mg, Y = 33 %) as a white solid.

^{1}H NMR (MeOD, 500 Mhz): d 5.40-5.33 (m, 2H), 4.73 (t, *J* = 6.0 Hz, 1H), 2.78 (dd, *J* = 16.0, 4.6 Hz, 1H), 2.74 (dd, *J* = 16.0, 6.8 Hz, 1H), 2.25 (t, *J* = 7.7 Hz, 2H), 2.09-1.99 (m, 4H), 1.63 (p, *J* = 7.1 Hz, 2H), 1.44 – 1.24 (m, 12H), 0.90 (t, *J* = 6.2 Hz, 3H).

^{13}C NMR (MeOD, 125 MHz): d 176.0, 175.0, 174.4, 130.96, 130.66, 37.8, 36.8, 33.0, 30.9, 30.6, 30.1, 29.9, 28.16, 28.05, 26.8, 23.7, 14.5.
Mp (°C): > 126, thermal degradation

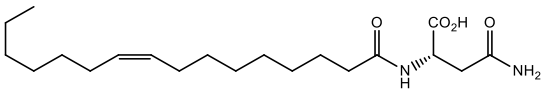
IR (neat) u_{max} (cm^{-1}): 3381, 3275, 3185, 2926, 2856, 1738, 1718, 1659, 1590, 1551, 1415, 1265, 1220, 1191, 1137, 968, 673.

C12(n-7)-Asn-OH



Following the example 4, starting with Fmoc-Asn-OH (53 mg, 150 µmol, 1.0 eq.): Y = 20 % (9.3 mg).
Off-white solid.
^{1}H NMR ($CDCl_3$ + eTFA, 500 MHz): d 7.14 (s, 1H), 6.74 (s, 1H), 5.57 – 5.34 (m, 2H), 5.39 – 5.21 (m, 2H), 4.93 – 4.73 (m, 1H), 3.13 – 2.95 (m, 2H), 2.35 (t, *J* = 7.8 Hz, 2H), 2.07 (q, *J* = 7.3 Hz, 2H), 1.98 (q, *J* = 7.1 Hz, 2H), 1.67 (p, *J* = 7.4 Hz, 2H), 1.49 – 1.06 (m, 8H), 0.87 (t, *J* = 6.7 Hz, 3H).
^{13}C NMR (125 MHz, $CDCl_3$ + eTFA): δ 176.7, 175.7, 174.2, 132.0, 127.6, 49.4, 36.4, 35.6, 31.9, 29.74, 29.71, 29.1, 27.39, 27.35, 26.5, 25.6, 22.81, 22.75, 14.2.
Mp (°C): > 121, thermal degradation
IR (neat) u_{max} (cm^{-1}): 3375, 3276, 3187, 2925, 2855, 1738, 1717, 1659, 1591, 1551, 1415, 1265, 1222, 1192, 1136, 947, 675.

C16(n-7)-Asn-OH:



Following the example 4, starting with Fmoc-Asn-OH (53 mg, 150 µmol, 1.0 eq.): crude reaction mixture was purified twice by flash Chromatography on Reveleris X2 (12 g - C4, Celite deposite). Water/[ACN-iPrOH 1:1] + 0.1 % TFA, 70/30 to 40/60 (12 CV), to obtain C16(n-7)-Asn-OH
(4.5 mg, Y = 8 %) as an off-white solid.

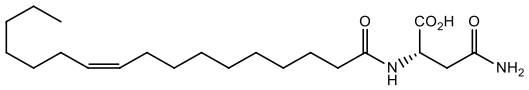
^{1}H NMR ($CDCl_3$+ eTFA, 400 MHz): d 7.25 (s, 1H), 6.59 (s, 1H), 5.40 - 5.29 (m, 2H), 4.88 (td, *J* = 9.0, 7.1 Hz, 1H), 3.20 (dd, *J* = 16.6, 4.3 Hz, 1H), 3.05 (dd, *J* = 16.6, 5.0 Hz, 1H), 2.38 (t, *J* = 7.7 Hz, 2H), 2.01 (q, *J* = 6.4 Hz, 4H), 1.62 (p, *J* = 7.2 Hz, 2H), 1.35-1.29 (m, 16H), 0.88 (t, *J* = 6.6 Hz, 3H).

^{13}C NMR ($CDCl_3$ + eTFA, 101 MHz): d 177.4, 175.7, 175.3, 130.3, 129.8, 49.4, 36.4, 36.1, 31.9, 29.9, 29.8, 29.2, 29.1, 29.1, 27.4, 27.2, 25.7, 22.8, 14.2.

Mp (°C): > 131, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3516-3035, 2926, 2855, 1662, 1608, 1429, 1189, 1138, 842, 800, 722.

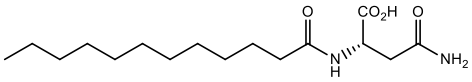
C17(n-7)-Asn-OH:



Following the example 4, starting with Fmoc-Asn-OH (53 mg, 150 µmol, 1.0 eq.): Y = 29 % (16.9 mg, LCMS purity > 91%), as an off-white solid.

^{1}H NMR (MeOD, 500 MHz): d 5.37-5.31 (m, 2H), 4.70 (t, *J* = 4.8 Hz, 1H), 2.78 (dd, J = 15.9, 4.7 Hz, 1H), 2.72 (dd, J = 15.8, 7.1 Hz, 1H), 2.23 (t, *J* = 7.8 Hz, 2H), 2.07-1.99 (m, 4H), 1.69 – 1.54 (m, 2H), 1.47 – 1.13 (m, 18H), 0.90 (t, *J* = 5.7 Hz, 3H).
^{13}C NMR (MeOD, 125 Mhz): d 176.1, 175.0, 174.5, 130.8, 50.5, 37.8, 36.9, 33.0, 30.88, 30.84, 30.54, 30.51, 30.4, 30.3, 30.1. 28.2, 26.9, 23.7, 14.5.
Mp (°C): > 128-130, thermal degradation.
IR (neat) u_{max} (cm^{-1}): 3453-3081, 2923, 2854, 1738, 1717, 1660, 1590, 1552, 1457, 1416, 1265, 1191, 967, 675.

C12-Asn-OH:



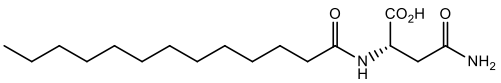
Following the example 3, starting with Fmoc-Asn(Trt)-OH (119 mg, 200 µmol, 1.0 eq.): Y = quant (63 mg). White solid.

^{1}H NMR ($CDCl_3$, 300 MHz): d 7.44 (s, 1H), 7.38 (d, *J* = 7.5 Hz, 1H), 6.80 (s, 1H), 4.94 (dt, *J* = 7.5, 4.5 Hz, 1H), 3.23 (dd, *J* = 16.9, 4.5 Hz, 1H), 3.11 (dd, *J* = 16.9, 4.7 Hz, 1H), 2.41 (t, *J* = 7.7 Hz, 2H), 1.63 (p, *J* = 7.4 Hz, 2H), 1.30-1.18 (m, 16 H), 0.87 (t, *J* = 7.2 Hz, 2H).

^{13}C NMR ($CDCl_3$ + eTFA, 75 MHz): d 177.7, 175.9 (two C), 49.4, 36.3, 36.1, 32.0, 29.7, 29.49, 29.45, 29.2, 29.0, 25.7, 22.8, 14.2.
Mp (°C): > 145 °C, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3405, 3294, 2916, 2849, 1698, 1657, 1647, 1587, 1542, 1433, ,1347, 4583, 1272, 1252, 1195, 1125, 978, 701.

C13-Asn-OH:

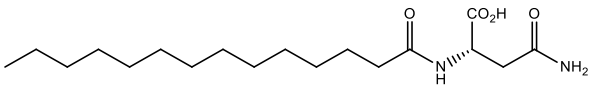


Following the example 3, starting with Fmoc-Asn(Trt)-OH (119 mg, 200 µmol, 1.0 eq.): Y = 91 % (60 mg). White solid.
^{1}H NMR ($CDCl_3$ + eTFA, 300 MHz): d 7.34 (s, 1H), 6.60 (s, 1H), 4.90 (dt, *J* = 8.1, 4.5 Hz, 1H), 3.21 (dd, *J* = 16.7, 4.3 Hz, 1H), 3.09 (dd, *J* = 16.7, 4.9 Hz, 1H), 2.39 (t, *J* = 7.7 Hz, 2H), 1.69 – 1.56 (m, 1H), 1.41 – 1.13 (m, 18H), 0.87 (t, *J* = 6.6 Hz, 3H).
^{13}C NMR ($CDCl_3$ + eTFA, 75 MHz): d 176.7, 175.5, 173.6, 49.4, 36.4, 36.2, 32.0, 29.7, 29.5, 29.2, 29.1, 25.7, 22.8, 14.2.

Mp (°C): > 143 °C, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3401, 3294, 3224, 2917, 2849, 2623-2377, 1698, 1651, 1542, 1586, 1543, 1473, 1463, 1429, 1343, 1281, 1268, 1247, 1198, 1127, 974, 896, 731, 719, 696.

C14-Asn-OH:



Following the example 3, starting with Fmoc-Asn(Trt)-OH (90 mg, 150 µmol, 1.0 eq.): Y = 62 % (35 mg). White solid.

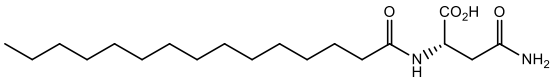
^{1}H NMR ($CDCl_3$ + eTFA, 500 MHz): d 7.43 (s, 1H), 7.32 (d, *J* = 7.5 Hz, 1H), 6.72 (s, 1H), 4.92 (dt, *J* = 8.3, 4.6 Hz, 1H), 3.32 – 2.97 (m, 2H), 2.40 (t, *J* = 7.7 Hz, 2H), 1.63 (p, *J* = 7.4 Hz, 2H), 1.39 – 1.15 (m, 20H), 0.87 (t, *J* = 6.8 Hz, 3H).

^{13}C NMR (126 MHz, $CDCl_3$) d 177.9, 176.2, 176.0, 162.6, 162.3, 161.9, 161.6, 117.7, 115.5, 113.2, 110.9, 77.4, 77.2, 76.9, 49.4, 36.2, 36.0, 32.1, 29.8, 29.82, 29.75, 29.5, 29.2, 29.0, 25.8, 22.8, 14.2.

Mp (°C): > 138-142 °C, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3408, 3297, 3206, 2954, 2920, 2851, 1699, 1664, 1639, 1563, 1539, 1471, 1435, 1410, 1323, 1254, 1213, 1191, 949, 909, 720, 681.

C15-Asn-OH :



Following the example 3, starting with Fmoc-Asn(Trt)-OH (119 mg, 200 µmol, 1.0 eq.): Y = 59 % (42 mg). White solid.

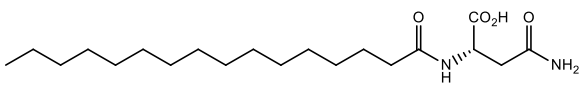
^{1}H NMR (DMSO-d_6, 300 MHz): d 7.98 (d, *J* = 7.9 Hz, 1H), 7.33 (s, 1H), 6.89 (s, 1H), 4.47 (ddd, *J* = 7.9, 7.3, 5.8 Hz, 2H), 2.59 – 2.35 (m, 2H), 2.07 (t, *J* = 7.4 Hz, 2H), 1.53 – 1.38 (m, 2H), 1.35 – 1.12 (m, 20H), 0.92 – 0.67 (m, 3H).

^{13}C NMR (DMSO-d_6, 75 MHz): d 173.1, 172.1, 171.3, 48.7, 36.8, 35.1, 31.3, 29.1, 29.1, 29.0, 28.9, 28.8, 28.6, 25.3, 22.1, 14.0.

Mp (°C): > 141, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3404, 3294, 3216, 2916, 2849, 1699, 1656, 1588, 1542, 1473, 1462, 1430, 1345, 1273, 1257, 1198, 1128, 973, 730, 719, 699.

C16-Asn-OH:

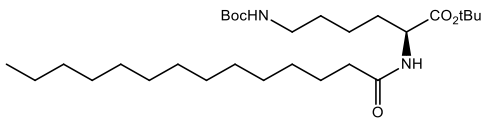


Following the example 3, starting with Fmoc-Asn(Trt)-OH (119 mg, 200 µmol, 1.0 eq.) : Y = 65 % (48 mg). White solid.
^{1}H NMR ($CDCl_3$ + eTFA, 300 MHz): d 7.42 (s, 1H), 7.36 (d, *J* = 7.7 Hz, 1H), 6.77 (s, 1H), 5.03 – 4.82 (m, 1H), 3.22 (dd, *J* = 16.8, 3.3 Hz, 1H), 3.10 (dd, *J* = 16.8, 4.1 Hz, 1H), 2.41 (t, *J* = 7.5 Hz, 2H), 1.68-1.59 (m, 2H), 1.39 – 1.17 (m, 24H), 0.87 (t, *J* = 5.8 Hz, 4H).
^{13}C NMR ($CDCl_3$ + eTFA, 75 Mhz): d 178.1, 176.1, 176.0, 49.42, 36.3, 36.0, 32.11, 29.85, 29.77, 29.71, 29.5, 29.2, 29.1, 25.8, 22.9, 14.2.

Mp (°C): > 140, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3405, 3293, 2915, 2849, 1698, 1658, 1648, 1586, 1543, 1474, 1463, 1435, 1347, 1284, 1253, 1199, 984, 702.

C14-Lys(tBu)-OtBu:



Following the example 1, starting with H-Lys(Boc)-OtBu.HCl (169 mg, 0.50 mmol, 1.0 eq.): Y = 72 % (185 mg). Crude reaction mixture was purified by flash Chromatography on Reveleris X2 (40 g, silica deposite). DCM/iPrOH 100/0 to 97.5 (12CV) to obtain C14-Lys(tBu)-OtBu as a colorless oil that partially crystallized at room temperature.

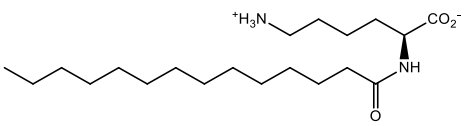
R_f (DCM/iPrOH 95/5) = 0.56

^{1}H NMR ($CDCl_3$, 300 MHz): d 6.03 (d, *J* = 7.9 Hz, 1H), 4.63 – 4.52 (m, 1H), 4.48 (td, *J* = 7.5, 5.1 Hz, 1H), 3.09 (q, *J* = 6.2 Hz, 2H), 2.25 – 2.15 (m, 2H), 1.90 – 1.73 (m, 1H), 1.62 (t, *J* = 6.5 Hz, 5H), 1.46 (s, 9H), 1.43 (s, 9H), 1.26 (d, *J* = 8.6 Hz, 22H), 0.87 (d, *J* = 6.7 Hz, 3H).

^{13}C NMR ($CDCl_3$, 75 MHz): d 173.0, 172.0, 156.2, 82.3, 79.3, 52.4, 40.3, 36.8, 32.5, 29.8, 29.79, 29.75, 29.64, 29.49, 29.43, 28.6, 28.2, 25.8, 22.4, 14.3.

IR (neat) u_{max} (cm^{-1}): 3365, 2918, 2850, 1733, 1683, 1647, 1516, 1365, 1250, 1161, 1143, 1000, 943, 851, 729, 721.

C14-Lys-OH:

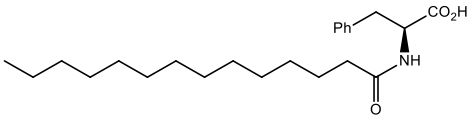


Following the example 2, starting with C14-Lys(tBu)-OtBu (165 mg, 322 µmol), using Et_3SiH (101 µL, 643 µmol, 2.0 eq.) as scavenger. Crude reaction mixture was purified by flash Chromatography on Reveleris X2 (24 g Interchim® SIHP, Celite deposite). DCM/MeOH + 0.1 %: 90/10 to remove apolar impurities then MeOH 100 % (monitoring with ELSD) to obtain Y = 84 % (96 mg) as a white powder.
^{1}H NMR ($CDCl_3$ + eTFA, 300 MHz): d 6.83 (s, 1H), 6.74 (d, *J* = 8.1 Hz, 1H), 4.74 (td, *J* = 8.6, 4.8 Hz, 1H), 3.17 (p, *J* = 6.3 Hz, 2H), 2.37 (m, 2H), 2.13 – 1.71 (m, 4H), 1.68 – 1.41 (m, 4H), 1.40 – 1.19 (m, 20H), 0.87 (t, *J* = 6.9 Hz, 3H).
^{13}C NMR ($CDCl_3$ + eTFA, 75 MHz): d 177.3, 175.1, 52.6, 40.3, 35.9, 32.0, 30.8, 29.80, 29.76, 29.71, 29.56, 29.47, 29.2, 26.4, 25.8, 22.8, 21.9, 14.2.

Mp (°C): > 178 °C, thermal degradation

IR (neat) u_{max} (cm^{-1}): 3362, 3314, 2918, 2850, 1937, 1598, 1522, 1473, 1465, 1400, 1015, 720.

C14-Phe-OH:



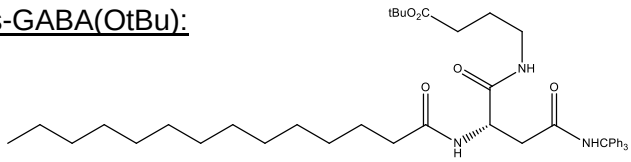
Following the example 3, starting with Fmoc-Phe-OH (200 µmol, 78 mg, 1.0 eq.): Y = 56 % (42 mg). White solid.

^{1}H NMR ($CDCl_3$, 300 MHz): d 7.37 – 7.12 (m, 5H), 5.90 (d, *J* = 7.4 Hz, 1H), 4.86 (q, *J* = 6.4 Hz, 1H), 3.25 (dd, *J* = 14.1, 5.6 Hz, 1H), 3.13 (dd, *J* = 14.1, 6.4 Hz, 1H), 2.21 – 2.14 (m, 2H), 1.55 (p, *J* = 6.6, 2H), 1.34-1.19 (m, 20 H), 0.88 (t, *J* = 6.6 Hz, 3H).

^{13}C NMR ($CDCl_3$, 300 MHz): d 174.5, 174.2, 135.7, 129.5, 128.8, 127.4, 53.3, 37.3, 36.6, 32.1, 29.8, 29.76, 29.60, 29.51, 29.44, 29.28, 25.7, 22.9, 14.3.
Mp (°C): 79.8-80.6

IR (neat) u_{max} (cm^{-1}): 3300, 2919, 2851, 1727, 1706, 1641, 1532, 1497, 1456, 1469, 1420, 1294, 1249, 1191, 1136, 925, 750, 720, 698.

C14-Asn(Trt)s-GABA(OtBu):



Following the example 1, starting with NH_2-Asn(Trt)-GABA(OtBu) (714 mg, 1.39 mmol): Y = 74 % (744 mg). White solid. Purification: DCM/AcOEt, 90/10 to 60/40.

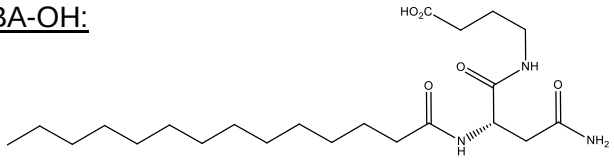
R_f (Cyc/AcOEt 50/50) = 0.2.

^{1}H NMR (DMSO-d_6, 300 MHz): d 8.53 (s, 1H), 8.01 (d, *J* = 8.1 Hz, 1H), 7.78 (t, *J* = 5.8 Hz, 1H), 7.39 – 6.90 (m, 15H), 4.47 (q, *J* = 7.7 Hz, 1H), 3.03 (q, *J* = 6.6 Hz, 2H), 2.58 – 2.52 (m, 2H), 2.24 – 2.07 (m, 4H), 1.58 (p, *J* = 7.5 Hz, 2H), 1.52 – 1.43 (m, 2H), 1.39 (s, 9H), 1.33-1.18 (m, 20 H), 0.85 (t, *J* = 6.4 Hz, 4H).
^{13}C NMR (DMSO-d_6, 75 MHz): d 172.1, 172.0, 171.1, 169.0, 144.8, 128.6, 127.4, 126.3, 79.5, 69.4, 50.1, 38.3, 37.9, 35.3, 32.1, 31.3, 29.08, 28.96, 28.90, 28.75, 27.8, 25.1, 24.7, 22.1, 14.0.

Mp (°C): 176-178.

IR (neat) u_{max} (cm^{-1}): 3286, 2922, 2852, 1724, 1652, 1636, 1519, 1493, 1448, 1367, 1251, 1148, 847, 764, 748, 697.

C14-Asn-GABA-OH:



Following the example 2, starting with C14-Asn(Trt)-GABA(OtBu) (428 mg, 590 µmol): Y = 21 % (53 mg). White solid. Purified by HPLC preparative.

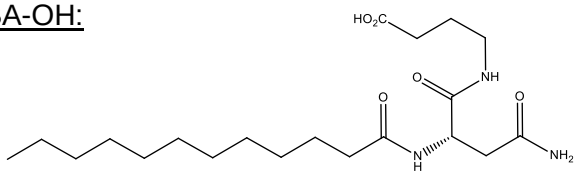
^{1}H NMR (DMSO-d_6, 300 MHz): d 7.91 (d, *J* = 8.0 Hz, 1H), 7.75 (t, *J* = 5.8 Hz, 1H), 7.26 (s, 1H), 6.84 (s, 1H), 4.47 (q, *J* = 7.3 Hz, 1H), 3.03 (q, *J* = 6.5 Hz, 2H), 2.48 – 2.24 (m, 2H), 2.18 (t, *J* = 7.5 Hz, 2H), 2.08 (t, *J* = 7.5 Hz, 2H), 1.58 (p, *J* = 7.3 Hz, 2H), 1.52 – 1.38 (m, 2H), 1.23 (ls, 20 H), 0.85 (t, *J* = 6.5 Hz, 3H).

^{13}C NMR (75 MHz, DMSO-d_6): d 174.3, 172.1, 171.5, 171.1, 49.7, 38.0, 37.4, 35.3, 31.3, 31.0, 29.10, 29.01, 28.9, 28.8, 28.7, 25.2, 24.5, 22.1, 14.0.

Mp (°C): > 197, thermal degradation.

IR (neat) u_{max} (cm^{-1}): 3408, 3297, 3206, 2954, 2920, 2851, 1699, 1664, 1639, 1563, 1539, 1471, 1435, 1409, 1323, 1254, 1213, 1191, 950, 909, 720.

C12-Asn-GABA-OH:



For the synthesis of C12-Asn-GABA-OH, see *Nat. Comm. 8, 1314 (2017)*.

ANNEXE 3 : Supplementary information, publication 2

Modelling has been performed at B3LYP/6-31+G(d,p) (Stephens, 1994) level using G16 package (Frisch , 2016) the nature of minima identified by the nature of their frequencies **A** *E(ZPE)= -742.942528*

6	0.29636	1.11206	0.38202
1	0.52596	0.96883	1.452
6	1.36572	2.11024	-0.13349
8	1.09597	3.2276	-0.53296
8	2.63017	1.68341	-0.05758
1	2.65739	0.72647	0.27111
6	-1.12342	1.68048	0.24406
1	-1.12307	2.65189	0.74837
1	-1.31747	1.88897	-0.81363
6	-2.2486	0.79607	0.82348
1	-1.90299	0.30237	1.74497
1	-3.0908	1.42818	1.12093
6	-2.83866	-0.31409	-0.10013
8	-2.03424	-0.85547	-0.94433
8	-4.03727	-0.60244	0.06765
7	0.3611	-0.19975	-0.27344
1	-0.60478	-0.52633	-0.6613
6	1.3364	-1.08354	-0.06042
8	2.40699	-0.81089	0.5517
6	1.11742	-2.4799	-0.62321
1	1.82791	-2.60501	-1.45108
1	0.10711	-2.5591	-1.03487
6	1.36572	-3.56874	0.4303
1	2.36552	-3.46752	0.8619
1	0.63415	-3.49968	1.24283
1	1.27468	-4.56398	-0.01863

TS$_{A->B}$ *E(ZPE)= -742,877372*

6	-1.314794	0.227540	0.058753
1	-0.873724	0.315579	-0.939059
6	-2.524797	-0.707286	-0.191687
8	-3.676303	-0.369566	0.134119
8	-2.130462	-1.780711	-0.773727
1	-0.531503	-1.753617	-0.933405
6	-1.415873	1.658374	0.560071
1	-1.998233	2.248970	-0.154228
1	-1.941891	1.720220	1.521492
6	0.024529	2.237548	0.697065
1	-0.001536	3.329261	0.667903
1	0.443506	1.964789	1.675407
6	1.066009	1.825805	-0.378936
8	1.464710	0.568916	-0.523087
8	1.567887	2.714689	-1.066542
7	-0.324455	-0.499404	0.892434
1	-0.108346	-0.036402	1.765000
6	0.802237	-0.857444	0.163638
8	0.500459	-1.628033	-0.894160
6	1.990088	-1.349469	0.978306
1	1.632789	-2.231197	1.526181
1	2.253369	-0.582862	1.717205
6	3.212425	-1.716675	0.133885
1	2.947595	-2.475018	-0.608121
1	3.590712	-0.839853	-0.396948
1	4.009176	-2.116760	0.772138

B *E(ZPE)=-742.884861*

6	-1.393647	0.042026	-0.216064
1	-1.293327	-0.057944	-1.313138
6	-2.713054	-0.748606	0.149827
8	-2.536670	-1.828354	0.777114
8	-3.775066	-0.219277	-0.254166
1	0.167790	-1.410239	-1.784689
6	-1.504003	1.533686	0.116713
1	-2.418881	1.929261	-0.333063
1	-1.597204	1.628872	1.205117
6	-0.291361	2.367362	-0.362103
1	-0.317127	2.438369	-1.457515
1	-0.327611	3.377671	0.056610
6	1.036751	1.743476	0.011125
8	1.338444	0.626793	-0.744062
8	1.823102	2.199842	0.813205
7	-0.251701	-0.549319	0.495768
1	-0.594647	-1.430637	0.894186
6	0.992374	-0.687965	-0.185677
8	0.993600	-1.558008	-1.302088
6	2.091245	-1.170535	0.766296
1	1.745301	-2.116964	1.200175
1	2.152154	-0.439808	1.577361
6	3.460481	-1.354184	0.107867
1	3.421271	-2.106195	-0.685504
1	3.805014	-0.412451	-0.330195
1	4.199184	-1.674240	0.852368

TS$_{B->C}$ *E(ZPE)= -742.844254*

6	-1.501893	0.269480	-0.386423
1	-1.627205	0.728737	-1.371816
6	-2.903004	-0.281548	0.101192
8	-2.860222	-1.426543	0.634792
8	-3.863831	0.499116	-0.062716
1	-0.147547	-1.567143	-1.477759
6	-1.015260	1.316987	0.620783
1	-1.871708	1.968307	0.824016

1	-0.769199	0.822048	1.569250
6	0.144268	2.206607	0.156421
1	-0.164086	2.723159	-0.765236
1	0.352214	2.987843	0.894246
6	1.480578	1.568724	-0.190626
8	1.479984	0.393917	-0.862283
8	2.527845	2.168459	-0.018221
7	-0.596233	-0.896348	-0.496358
1	-1.038815	-1.583955	0.145315
6	0.916510	-0.914626	-0.395785
8	1.128170	-1.829171	-1.350903
6	1.495360	-1.227454	0.987438
1	0.983500	-2.129264	1.343688
1	1.247795	-0.422141	1.690759
6	3.009062	-1.460633	0.952249
1	3.246956	-2.256645	0.241034
1	3.537468	-0.555462	0.640273
1	3.374956	-1.752502	1.943885

C *E(ZPE)= -742.911690*

6	0.174305	0.812153	-0.224058
1	0.953024	1.144653	-0.924657
6	-0.832844	1.964769	-0.238044
8	-1.616581	1.986772	-1.349439
8	-0.910701	2.848710	0.586268
1	0.431541	-0.894573	-1.300163
6	0.890409	0.553840	1.120003
1	0.624870	1.354871	1.819357
1	0.514994	-0.385847	1.542553
6	2.423641	0.476372	0.986057
1	2.800321	1.464402	0.678593
1	2.865265	0.273850	1.967218
6	3.011269	-0.559706	-0.029988
8	4.125165	-1.039758	0.271126
8	2.331256	-0.774521	-1.086763
7	-0.367221	-0.382838	-0.895611
1	-1.422279	1.164658	-1.843219
6	-1.372051	-1.197168	-0.407190
8	-1.519395	-2.340759	-0.834128
6	-2.348100	-0.603944	0.613831
1	-2.918344	0.193628	0.120270
1	-1.788425	-0.113185	1.417742
6	-3.302965	-1.649399	1.192386
1	-3.876950	-2.136720	0.399955
1	-2.753731	-2.433094	1.722647
1	-3.999668	-1.177838	1.894479

A'+H_2O *E(ZPE)= -819.369259*

6	-0.451947	-0.540910	0.508094
1	-0.438826	-0.272717	1.578931
6	-1.946511	-0.517222	0.111671
8	-2.564771	-1.521178	-0.212912
8	-2.552179	0.663878	0.201821
1	-1.876585	1.393118	0.459064
6	0.163579	-1.937158	0.334121
1	-0.469990	-2.633610	0.891917
1	0.082735	-2.229008	-0.718624
6	1.626084	-2.073598	0.810213
1	1.781597	-1.478318	1.723216
1	1.819656	-3.112435	1.093051
6	2.749051	-1.676515	-0.195766
8	2.484422	-0.716556	-1.012695
8	3.819657	-2.301664	-0.118602
7	0.358910	0.443623	-0.218222
1	1.272503	0.001922	-0.656499
6	0.261526	1.758287	-0.034174
8	-0.676004	2.308885	0.615266
6	1.339367	2.616234	-0.676257
1	0.859712	3.169372	-1.494602
1	2.109916	1.974978	-1.113566
6	1.953410	3.611003	0.319014
1	1.176709	4.231106	0.775189
1	2.485656	3.085654	1.119518
1	2.672027	4.263603	-0.188492
8	-5.336824	-0.684213	-0.500763
1	-5.066129	0.224221	-0.313938
1	-4.483119	-1.157196	-0.466025

VALORISATION DES TRAVAUX DE RECHERCHE

Travaux acceptés :

Amandine Hueber, Yves Gimbert, Geoffrey Langevin, Jean-Marie Galano, Alexandre Guy, Thierry Durand, Nicolas Cenac, Justine Bertrand-Michel, Jean-Claude Tabet. **New lipoamino acid from E.Coli Bacteria : origin of regenerated fatty acid carboxylate from dissociation of lipo-glutamate anion**, accépté dans le journal Aminoacids.

Amandine Hueber, Camille Petitfils, Pauline Le Faouder, Geoffrey Langevin, Jean-Marie Galano, Jean-François Martin, Jean-Claude Tabet, Nicolas Cenac, Justine Bertrand-Michel, **Discovery and quantification of new lipoamino acids in bacteria**, accepté Analytica Chimica Acta journal.

Travaux en soumission :

Amandine Hueber, Martin Green, Jakob Ujma, Keith Richardson, Yves Gimbert, Jean-François Martin, Geoffrey Langevin, Nicolas Cenac, Justine Bertrand-Michel, Jean-Claude Tabet, **Energy-Resolved Ion Mobility Spectrometry, a new concept : composite breakdown curves for distinguishing isomeric product ions**, en soumission dans le journal Analytical Chemestry.

Présentations orales :

Mise en place d'un workflow analytique pour l'identification et la caractérisation de nouveaux lipopeptides bactériens : une approche par chromatographie liquide couplée à la spectrométrie de masse à haute résolution, une approche non ciblée-ciblée. (Club jeune de la société française de spectrométrie de masse, CJ-SFSM, en ligne)

Identification et caractérisation des lipopeptides et aminolipides bactériens produits par le microbiote intestinal (Journées de l'association francophone des sciences séparatives AFSEP du sud ouest, en ligne)

Présentation poster :

Amandine HUEBER, Camille PETITFILS, Pauline LE-FAOUDER, Jean-Marie GALANO, Alexandre GUY, Thierry DURAND, Nicolas CENAC and Justine BERTRAND-MICHEL, **Experimental strategy to discover new bioactive lipopeptides produced by gut microbiota : unknown-known approach by LC-HRMS.** (Réseau Francophone de Métabolomique et de Fluxomique, Metabomeeting, Toulouse, France)

Amandine Hueber, Yves Gimbert, Martin Green, Jakub Ujma, Keith Richardson, Valérie Steiner, Geoffrey Langevin, Jean-François Martin, Nicolas Cenac, Justine Bertrand-Michel, Jean-Claude Tabet : **Unexpected "camel back" shape of Energy Resolved breakdown curves of lipo-glutamate : direct and consecutive channels evidenced by cyclic ion mobility** (American Society of Mass Spectrometry ASMS – Philadelphia, United States of America, November 2021), online presentation.

Curriculum vitæ

CONTACT

26 ans, Permis B
06.59.19.65.11
hueberamandine94@gmail.com
14, rue André Clarous
31200 Toulouse

Recommandations :
Justine BERTRAND-MICHEL
Jean-Claude TABET
Nicolas CENAC

LANGUES

Anglais ●●●●○
Allemand ●●●○○

COMPETENCES INFORMATIQUES

Pack office ●●●●○
Chemsketch ●●●○○

MassHunter
MassLynx,Unifi
LabSolution Insigh ●●●○○
Xcalibur, Chroméléon

COMPETENCES ANALYTIQUES & INSTRUMENTALES

HPLC : Waters, Agilent, Dionex, Thermo, Shimadzu ●●●●●

LR-MS QqQ : 6040 Agilent, 8060 Shimadzu ●●●●●

HR-MS : Exactive, QExactive⁺, LTQ Velos, qTOF Xevo G2-XS ●●●●●

GC-FID : Thermo, PerkinElmer ●●●○○

Elucidation structurale (fragmentation + MS^n) ●●●○○

Préparation échantillon : ELL / SPE ●●●●●

HOBBIES

Cuisine, randonnée, natation

HUEBER Amandine



Ingénieure en chimie analytique et spectrométrie de masse

EXPERIENCES PROFESSIONNELLES

2018 à nos jours METATOUL LIPIDOMIQUE TOULOUSE - **Doctorat**
Sous la direction de Justine BERTRAND-MICHEL et Nicolas CENAC

Identification et caractérisation de nouveaux lipopeptides bactériens par LC-HRMS - **QTOF Xevo G2-XS, Waters**

- **Optimisation** de méthodes d'extraction par **SPE, ELL**.
- **Développement analytique** : méthodes de **quantification LC-QqQ** Agilent 6040, **LC-QqQ** Shimadzu 8060, **GC-FID**.
- **Analyse structurale** : caractérisation **MS^2, MS^3 - QExactive⁺, LTQ Velos**

Valorisation
Oraux → Congrès jeune SFSM (10/20), AFSEP Sud-Ouest (06-20)
Poster → Congrès européen RFMF (01-20)

2019-2021 Tutorat à la faculté des Sciences de l'ingénieur - Toulouse
64h/an, Travaux pratiques de biodiversité végétale.

01/2018 - 06/2018 Institut des sciences de la vigne et du vin, Villenave d'Ornon – **Stage de Master 2 sous la direction M. Marchal**
Œnologie : Axe qualité et identité du vin → Identification et caractérisation de composés sapides provenant d'eaux-de-vie boisées de Cognac.

- Extraction par chromatographie de partage centrifuge
- Fractionnement d'extraits de Cognac par SPE et ELL,
- Identification et caractérisation par LC-HRMS Exactive + RMN

2017-2018 VégéLAB : Plant Advanced Technologies
Partenariat inclus au cursus universitaire de master :
Analyse d'un extrait issus de sous-produits de l'industrie alimentaire en vue d'une valorisation en nutraceutique.
Mise en place d'un plan d'action expérimental, veille scientifique, rédaction de rapports de résultats.

05/2017 - 07/2017 INRA Colmar - **Stage de Master 1**
Equipe métabolisme secondaire de la vigne
Extraction et **fractionnement** d'un extrait de sorgho, **SPE**, analyses par **LC-HRMS QExactive**.

FORMATIONS

2018 à nos jours : UNIVERSITE DE TOULOUSE, METATOUL LIPIDOMIQUE TOULOUSE, INSERM – **Doctorat chimie analytique**

2016-2018 UNIVERSITE DE STRASBOURG – **Master de Biologie et Valorisation des Ressources végétales**
Pharmacognosie, phytochimie, méthodes d'extractions spécifiques aux métabolites I et II des plantes, chimie analytique, bases de RMN.

2012-2016 UNIVERSITE DE STRASBOURG – **Licence de Biologie, Parcours Physiologie des Organismes**

BIBLIOGRAPHIE

1. Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* **9**, 112–124 (2008).
2. Christie, W. W. & Han, X. Lipids: their structures and occurrence. in *Lipid Analysis* 3–19 (Elsevier, 2012). doi:10.1533/9780857097866.3.
3. Carrasco-Pancorbo, A., Navas-Iglesias, N. & Cuadros-Rodríguez, L. From lipid analysis towards lipidomics, a new challenge for the analytical chemistry of the 21st century. Part I: Modern lipid analysis. *TrAC Trends in Analytical Chemistry* **28**, 263–278 (2009).
4. Fahy, E. *et al.* A comprehensive classification system for lipids. *Journal of Lipid Research* **46**, 839–861 (2005).
5. LIPID MAPS Lipidomics Gateway : Lipidomics Resources. Available at: http://www.lipidmaps.org/data/classification/LM_classification_exp.php.
6. Kihara, A. Very long-chain fatty acids: elongation, physiology and related disorders. *Journal of Biochemistry* **152**, 387–395 (2012).
7. Maier, T. Architecture of Mammalian Fatty Acid Synthase at 4.5 A Resolution. *Science* **311**, 1258–1262 (2006).
8. Guillou, H., Zadravec, D., Martin, P. G. P. & Jacobsson, A. The key roles of elongases and desaturases in mammalian fatty acid metabolism: Insights from transgenic mice. *Progress in Lipid Research* **49**, 186–199 (2010).
9. Shrago, E., Glennon, J. A. & Gordon, E. S. Comparative aspects of lipogenesis in mammalian tissues. *Metabolism* **20**, 54–62 (1971).
10. Subramanian, V., Dubini, A. & Seibert, M. Metabolic Pathways in Green Algae with Potential Value for Biofuel Production. in *The Science of Algal Fuels* (eds. Gordon, R. & Seckbach, J.) vol. 25 399–422 (Springer Netherlands, 2012).
11. Zhang, J. Y., Kothapalli, K. S. D. & Brenna, J. T. Desaturase and elongase-limiting endogenous long-chain polyunsaturated fatty acid biosynthesis: *Current Opinion in Clinical Nutrition and Metabolic Care* **19**, 103–110 (2016).
12. Bendsen, N. T. *et al.* Effect of industrially produced trans fat on markers of systemic inflammation: evidence from a randomized trial in women. *Journal of Lipid Research* **52**, 1821–1828 (2011).
13. Bassett, C. M. C. *et al.* trans-Fatty acids in the diet stimulate atherosclerosis. *Metabolism* **58**, 1802–1808 (2009).
14. Llorente-Cortés, V., Martínez-González, J. & Badimon, L. LDL Receptor–Related Protein Mediates Uptake of Aggregated LDL in Human Vascular Smooth Muscle Cells. *ATVB* **20**, 1572–1579 (2000).
15. Bassett, C. M. C. *et al.* Dietary Vaccenic Acid Has Antiatherogenic Effects in LDLr−/− Mice. *The Journal of Nutrition* **140**, 18–24 (2010).
16. Buczynski, M. W., Dumlao, D. S. & Dennis, E. A. Thematic Review Series: Proteomics. An integrated omics analysis of eicosanoid biology. *Journal of Lipid Research* **50**, 1015–1038 (2009).
17. Nicolaou, A., Mauro, C., Urquhart, P. & Marelli-Berg, F. Polyunsaturated Fatty Acid-Derived Lipid Mediators and T Cell Function. *Front. Immunol.* **5**, (2014).
18. Nicolaou, A. Eicosanoids in skin inflammation. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **88**, 131–138 (2013).
19. Nakamura, M. T. & Nara, T. Y. Essential fatty acid synthesis and its regulation in mammals. *Prostaglandins, Leukotrienes and Essential Fatty Acids* **68**, 145–150 (2003).
20. Bou Khalil, M. *et al.* Lipidomics era: Accomplishments and challenges: LIPIDOMICS ERA: ACCOMPLISHMENTS AND CHALLENGES. *Mass Spectrom. Rev.* **29**, 877–929 (2010).
21. Blée, E. Impact of phyto-oxylipins in plant defense. *Trends in Plant Science* **7**, 315–322 (2002).
22. Needleman, P., Truk, J., Jakschik, B. A., Morrison, A. R. & Lefkowith, J. B. Arachidonic Acid Metabolism. *Annu. Rev. Biochem.* **55**, 69–102 (1986).
23. Wenk, M. R. Lipidomics: New Tools and Applications. *Cell* **143**, 888–895 (2010).
24. Funk, C. D. Prostaglandins and Leukotrienes: Advances in Eicosanoid Biology. *Science* **294**, 1871–1875 (2001).

25. Gabbs, M., Leng, S., Devassy, J. G., Monirujjaman, M. & Aukema, H. M. Advances in Our Understanding of Oxylipins Derived from Dietary PUFAs. *Advances in Nutrition* **6**, 513–540 (2015).
26. Markworth, J. F. *et al.* Divergent shifts in lipid mediator profile following supplementation with n-3 docosapentaenoic acid and eicosapentaenoic acid. *FASEB j.* **30**, 3714–3725 (2016).
27. Noverr, M. C., Erb-Downward, J. R. & Huffnagle, G. B. Production of Eicosanoids and Other Oxylipins by Pathogenic Eukaryotic Microbes. *Clin Microbiol Rev* **16**, 517–533 (2003).
28. Di Marzo, V. Endocannabinoids: synthesis and degradation. in *Reviews of Physiology Biochemistry and Pharmacology* vol. 160 1–24 (Springer Berlin Heidelberg, 2006).
29. Tan, B. *et al.* Identification of endogenous acyl amino acids based on a targeted lipidomics approach. *Journal of Lipid Research* **51**, 112–119 (2010).
30. Carrasco, S. & Mérida, I. Diacylglycerol, when simplicity becomes complex. *Trends in Biochemical Sciences* **32**, 27–36 (2007).
31. Yamashita, A. *et al.* Acyltransferases and transacylases that determine the fatty acid composition of glycerolipids and the metabolism of bioactive lipid mediators in mammalian cells and model organisms. *Progress in Lipid Research* **53**, 18–81 (2014).
32. Holthuis, J. C. M. & Menon, A. K. Lipid landscapes and pipelines in membrane homeostasis. *Nature* **510**, 48–57 (2014).
33. Wood, P. L. *et al.* Targeted Lipidomics of Fontal Cortex and Plasma Diacylglycerols (DAG) in Mild Cognitive Impairment and Alzheimer's Disease: Validation of DAG Accumulation Early in the Pathophysiology of Alzheimer's Disease. *JAD* **48**, 537–546 (2015).
34. Horn, P. J. *et al.* Visualization of Lipid Droplet Composition by Direct Organelle Mass Spectrometry. *Journal of Biological Chemistry* **286**, 3298–3306 (2011).
35. Quehenberger, O. *et al.* Lipidomics reveals a remarkable diversity of lipids in human plasma. *Journal of Lipid Research* **51**, 3299–3305 (2010).
36. Himms-Hagen, J. Brown adipose tissue thermogenesis: interdisciplinary studies. *FASEB J* **4**, 2890–2898 (1990).
37. Yen, C.-L. E., Stone, S. J., Koliwad, S., Harris, C. & Farese, R. V. Thematic Review Series: Glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *Journal of Lipid Research* **49**, 2283–2301 (2008).
38. Frisardi, V., Panza, F., Seripa, D., Farooqui, T. & Farooqui, A. A. Glycerophospholipids and glycerophospholipid-derived lipid mediators: A complex meshwork in Alzheimer's disease pathology. *Progress in Lipid Research* **50**, 313–330 (2011).
39. Hishikawa, D., Hashidate, T., Shimizu, T. & Shindou, H. Diversity and function of membrane glycerophospholipids generated by the remodeling pathway in mammalian cells. *Journal of Lipid Research* **55**, 799–807 (2014).
40. Farooqui, A. A., Ong, W.-Y. & Farooqui, T. Lipid mediators in the nucleus: Their potential contribution to Alzheimer's disease. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1801**, 906–916 (2010).
41. Nagan, N. & Zoeller, R. A. Plasmalogens: biosynthesis and functions. *Progress in Lipid Research* **40**, 199–229 (2001).
42. Petrosino, S., Iuvone, T. & Di Marzo, V. N-palmitoyl-ethanolamine: Biochemistry and new therapeutic opportunities. *Biochimie* **92**, 724–727 (2010).
43. da Silva, T. F., Sousa, V. F., Malheiro, A. R. & Brites, P. The importance of ether-phospholipids: A view from the perspective of mouse models. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1822**, 1501–1508 (2012).
44. Lee, T. Biosynthesis and possible biological functions of plasmalogens. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1394**, 129–145 (1998).
45. Jiménez-Rojo, N. & Riezman, H. On the road to unraveling the molecular functions of ether lipids. *FEBS Lett* **593**, 2378–2389 (2019).
46. Boudière, L. *et al.* Glycerolipids in photosynthesis: Composition, synthesis and trafficking. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**, 470–480 (2014).
47. Chen, J., Feng, L. & Prestwich, G. D. Asymmetric Total Synthesis of Phosphatidylinositol 3-Phosphate and 4-Phosphate Derivatives. *J. Org. Chem.* **63**, 6511–6522 (1998).
48. Lemmon, M. A. Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* **9**, 99–111 (2008).

49. Vicinanza, M., D'Angelo, G., Di Campli, A. & De Matteis, M. A. Function and dysfunction of the PI system in membrane trafficking. *EMBO J* **27**, 2457–2470 (2008).
50. Hübscher, G., Dils, R. R. & Pover, W. F. R. Studies on the biosynthesis of phosphatidyl serine. *Biochimica et Biophysica Acta* **36**, 518–528 (1959).
51. Ridgway, N. D. Interactions between metabolism and intracellular distribution of cholesterol and sphingomyelin. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1484**, 129–141 (2000).
52. Kolter, T. A view on sphingolipids and disease. *Chemistry and Physics of Lipids* **164**, 590–606 (2011).
53. Hla, T. & Dannenberg, A. J. Sphingolipid Signaling in Metabolic Disorders. *Cell Metabolism* **16**, 420–434 (2012).
54. Honke, K. Biosynthesis and biological function of sulfoglycolipids. *Proc. Jpn. Acad., Ser. B* **89**, 129–138 (2013).
55. Slotte, J. P. Biological functions of sphingomyelins. *Progress in Lipid Research* **52**, 424–437 (2013).
56. Eyster, K. M. The membrane and lipids as integral participants in signal transduction: lipid signal transduction for the non-lipid biochemist. *Advances in Physiology Education* **31**, 5–16 (2007).
57. Schnaar, R. L., Suzuki, A. & Stanley, P. Glycosphingolipids. in *Essentials of Glycobiology* (eds. Varki, A. et al.) (Cold Spring Harbor Laboratory Press, 2009).
58. Yamashita, T. *et al.* A vital role for glycosphingolipid synthesis during development and differentiation. *Proceedings of the National Academy of Sciences* **96**, 9142–9147 (1999).
59. Morgan, A. E., Mooney, K. M., Wilkinson, S. J., Pickles, N. A. & Mc Auley, M. T. Cholesterol metabolism: A review of how ageing disrupts the biological mechanisms responsible for its regulation. *Ageing Research Reviews* **27**, 108–124 (2016).
60. Payne, A. H. & Hales, D. B. Overview of Steroidogenic Enzymes in the Pathway from Cholesterol to Active Steroid Hormones. *Endocrine Reviews* **25**, 947–970 (2004).
61. Tsai, M. & O'Malley, B. W. Molecular Mechanisms of Action of Steroid/Thyroid Receptor Superfamily Members. *Annu. Rev. Biochem.* **63**, 451–486 (1994).
62. Kousteni, S. Reversal of Bone Loss in Mice by Nongenotropic Signaling of Sex Steroids. *Science* **298**, 843–846 (2002).
63. Girod, J. & Brotman, D. Does altered glucocorticoid homeostasis increase cardiovascular risk? *Cardiovascular Research* **64**, 217–226 (2004).
64. Priego Capote, F., Jiménez, J. R., Granados, J. M. M. & de Castro, M. D. L. Identification and determination of fat-soluble vitamins and metabolites in human serum by liquid chromatography/triple quadrupole mass spectrometry with multiple reaction monitoring. *Rapid Commun. Mass Spectrom.* **21**, 1745–1754 (2007).
65. Feldman, D., Krishnan, A. V., Swami, S., Giovannucci, E. & Feldman, B. J. The role of vitamin D in reducing cancer risk and progression. *Nat Rev Cancer* **14**, 342–357 (2014).
66. Garabédian, M. *et al.* Prévention de la carence en vitamine D chez l'enfant et l'adolescent. II. Validation d'un abaque décisionnel non invasif prenant en compte l'exposition solaire et les apports exogènes de vitamine D. *Archives de Pédiatrie* **12**, 410–419 (2005).
67. Tissandié, E., Guéguen, Y., A.Lobaccaro, J.-M., Aigueperse, J. & Souidi, M. Vitamine D : Métabolisme, régulation et maladies associées. *Med Sci (Paris)* **22**, 1095–1100 (2006).
68. G. Popjak, J.W. Cornforth. *The biosynthesis of cholesterol*.
69. Poupon, R., Chignard, N., Rosmorduc, O., Barbu, V. & Housset, C. La fonction biliaire et sa régulation. *Med Sci (Paris)* **20**, 1096–1099 (2004).
70. Jonas, A. Lecithin cholesterol acyltransferase. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1529**, 245–256 (2000).
71. Hur, E.-M. & Kim, K.-T. G protein-coupled receptor signalling and cross-talk. *Cellular Signalling* **14**, 397–405 (2002).
72. Li, Q., Han, Y., Dy, A. B. C. & Hagerman, R. J. The Gut Microbiota and Autism Spectrum Disorders. *Front. Cell. Neurosci.* **11**, 120 (2017).
73. Mangiola, F. Gut microbiota in autism and mood disorders. *WJG* **22**, 361 (2016).
74. Landman, C. & Quévrain, E. Le microbiote intestinal : description, rôle et implication physiopathologique. *La Revue de Médecine Interne* **37**, 418–423 (2016).
75. Eckburg, P. B. Diversity of the Human Intestinal Microbial Flora. *Science* **308**, 1635–1638 (2005).

76. MetaHIT Consortium *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
77. Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the Human Infant Intestinal Microbiota. *PLoS Biol* **5**, e177 (2007).
78. Koenig, J. E. *et al.* Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Sciences* **108**, 4578–4585 (2011).
79. Rodríguez, J. M. *et al.* The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial Ecology in Health & Disease* **26**, (2015).
80. O'Hara, A. M. & Shanahan, F. The gut flora as a forgotten organ. *EMBO Rep* **7**, 688–693 (2006).
81. Dethlefsen, L., Eckburg, P. B., Bik, E. M. & Relman, D. A. Assembly of the human intestinal microbiota. *Trends in Ecology & Evolution* **21**, 517–523 (2006).
82. Neish, A. S. Microbes in Gastrointestinal Health and Disease. *Gastroenterology* **136**, 65–80 (2009).
83. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine. *Cell* **124**, 837–848 (2006).
84. Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* **95**, 6578–6583 (1998).
85. Mahowald, M. A. *et al.* Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proceedings of the National Academy of Sciences* **106**, 5859–5864 (2009).
86. Gérard, P. Microbiote intestinal et lipides : impact sur la santé humaine. *OCL* **19**, 223–227 (2012).
87. De La Cochetière, M. F. *et al.* Resilience of the Dominant Human Fecal Microbiota upon Short-Course Antibiotic Challenge. *J Clin Microbiol* **43**, 5588–5592 (2005).
88. O'Mahony, L. *et al.* Lactobacillus and bifidobacterium in irritable bowel syndrome: Symptom responses and relationship to cytokine profiles. *Gastroenterology* **128**, 541–551 (2005).
89. Kekkonen, R. A. *et al.* Probiotic intervention has strain-specific anti-inflammatory effects in healthy adults. *WJG* **14**, 2029 (2008).
90. Isolauri, E., Sütas, Y., Kankaanpää, P., Arvilommi, H. & Salminen, S. Probiotics: effects on immunity. *The American Journal of Clinical Nutrition* **73**, 444s–450s (2001).
91. Motevaseli, E., Dianatpour, A. & Ghafouri-Fard, S. The Role of Probiotics in Cancer Treatment: Emphasis on their In Vivo and In Vitro Anti-metastatic Effects. *Int J Mol Cell Med* **6**, (2017).
92. Ait Seddik, H., Bendali, F., Cudennec, B. & Drider, D. Anti-pathogenic and probiotic attributes of Lactobacillus salivarius and Lactobacillus plantarum strains isolated from feces of Algerian infants and adults. *Research in Microbiology* **168**, 244–254 (2017).
93. Pérez-Berezo, T. *et al.* Identification of an analgesic lipopeptide produced by the probiotic Escherichia coli strain Nissle 1917. *Nat Commun* **8**, 1314 (2017).
94. Biagi, G., Cipollini, I., Pompei, A., Zaghini, G. & Matteuzzi, D. Effect of a Lactobacillus animalis strain on composition and metabolism of the intestinal microflora in adult dogs. *Veterinary Microbiology* **124**, 160–165 (2007).
95. Kaźmierczak-Siedlecka, K., Roviello, G., Catalano, M. & Polom, K. Gut Microbiota Modulation in the Context of Immune-Related Aspects of Lactobacillus spp. and Bifidobacterium spp. in Gastrointestinal Cancers. *Nutrients* **13**, 2674 (2021).
96. Park, S. H. & Itoh, K. Species-specific oligonucleotide probes for the detection and identification of *Lactobacillus* isolated from mouse faeces. *Journal of Applied Microbiology* **99**, 51–57 (2005).
97. Wassenaar, T. M. Insights from 100 years of research with probiotic *E. coli*. *European Journal of Microbiology and Immunology* **6**, 147–161 (2016).
98. Kono, T. *et al.* Epithelial transient receptor potential ankyrin 1 (TRPA1)-dependent adrenomedullin upregulates blood flow in rat small intestine. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **304**, G428–G436 (2013).
99. Ochman, H. & Selander, R. K. Standard reference strains of Escherichia coli from natural populations. *J Bacteriol* **157**, 690–693 (1984).
100. Hsiao, E. Y. *et al.* Microbiota Modulate Behavioral and Physiological Abnormalities Associated with Neurodevelopmental Disorders. *Cell* **155**, 1451–1463 (2013).

101. Nowrouzian, F. L., Adlerberth, I. & Wold, A. E. Enhanced persistence in the colonic microbiota of Escherichia coli strains belonging to phylogenetic group B2: role of virulence factors and adherence to colonic cells. *Microbes and Infection* **8**, 834–840 (2006).
102. Massot, M. *et al.* Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of Escherichia coli from community subjects in the Paris area in 2010 and evolution over 30 years. *Microbiology* **162**, 642–650 (2016).
103. Philippe Gérard, Annick Bernalier-Donadille. Les fonctions majeures du microbiote intestinal.
104. LeBlanc, J. G. *et al.* Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current Opinion in Biotechnology* **24**, 160–168 (2013).
105. Gérard, P. Le microbiote intestinal: composition et fonctions. *Phytothérapie* **9**, 72–75 (2011).
106. Nicholson, J. K., Holmes, E. & Wilson, I. D. Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* **3**, 431–438 (2005).
107. Cebra, J. J. Influences of microbiota on intestinal immune system development. *The American Journal of Clinical Nutrition* **69**, 1046s–1051s (1999).
108. Guarner, F. & Malagelada, J.-R. Gut flora in health and disease. *The Lancet* **361**, 512–519 (2003).
109. Wostmann, B. S. The Germfree Animal in Nutritional Studies. *Annu. Rev. Nutr.* **1**, 257–279 (1981).
110. Smith, K., McCoy, K. D. & Macpherson, A. J. Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. *Seminars in Immunology* **19**, 59–69 (2007).
111. Gensollen, T., Iyer, S. S., Kasper, D. L. & Blumberg, R. S. How colonization by microbiota in early life shapes the immune system. *Science* **352**, 539–544 (2016).
112. de Vos, W. M. & de Vos, E. A. Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutrition Reviews* **70**, S45–S56 (2012).
113. Ducrotté, P. Flore et syndrome de l'intestin irritable. *Gastroentérologie Clinique et Biologique* **34**, 56–60 (2010).
114. Holtmann, G. J., Ford, A. C. & Talley, N. J. Pathophysiology of irritable bowel syndrome. *The Lancet Gastroenterology & Hepatology* **1**, 133–146 (2016).
115. Chassard, C. *et al.* Functional dysbiosis within the gut microbiota of patients with constipated-irritable bowel syndrome. *Aliment Pharmacol Ther* **35**, 828–838 (2012).
116. Jailwala, J. Pharmacologic Treatment of the Irritable Bowel Syndrome: A Systematic Review of Randomized, Controlled Trials. *Ann Intern Med* **133**, 136 (2000).
117. Allison, S. D. & Martiny, J. B. H. Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences* **105**, 11512–11519 (2008).
118. Pouteau, E., Meirim, I., Métairon, S. & Fay, L.-B. Acetate, propionate and butyrate in plasma: determination of the concentration and isotopic enrichment by gas chromatography/mass spectrometry with positive chemical ionization: SCFA concentration and isotopic enrichment in plasma. *J. Mass Spectrom.* **36**, 798–805 (2001).
119. Pujo, J. *et al.* Bacteria-derived long chain fatty acid exhibits anti-inflammatory properties in colitis. *Gut* **70**, 1088–1097 (2021).
120. Shaw, N. Lipid Composition as a Guide to the Classification of Bacteria. in *Advances in Applied Microbiology* vol. 17 63–108 (Elsevier, 1974).
121. Meinwald, J. & Eisner, T. Chemical ecology in retrospect and prospect. *Proceedings of the National Academy of Sciences* **105**, 4539–4540 (2008).
122. Myrron and Sasser. *Identification of Bacteria by Gas Chromatographyof Cellular Fatty Acids*.
123. Lambert, P. A. Cellular impermeability and uptake of biocides and antibiotics in Gram-positive bacteria and mycobacteria: ANTIMICROBIAL PENETRATION OF WALLS. *Journal of Applied Microbiology* **92**, 46S-54S (2002).
124. Macfarlane, G. T. & Macfarlane, S. Bacteria, Colonic Fermentation, and Gastrointestinal Health. *Journal of AOAC INTERNATIONAL* **95**, 50–60 (2012).
125. Zhang, G., Meredith, T. C. & Kahne, D. On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Current Opinion in Microbiology* **16**, 779–785 (2013).
126. Galanos, C. & Freudenberg, M. A. Bacterial endotoxins: biological properties and mechanisms of action. *Mediators of Inflammation* **2**, S11–S16 (1993).
127. Kotani, S. *et al.* Synthetic lipid A with endotoxic and related biological activities comparable to those of a natural lipid A from an Escherichia coli re-mutant. *Infect Immun* **49**, 225–237 (1985).

128. Raetz, C. R. H. *et al.* Discovery of new biosynthetic pathways: the lipid A story. *Journal of Lipid Research* **50**, S103–S108 (2009).
129. Tzeng, Y.-L., Datta, A., Kolli, V. K., Carlson, R. W. & Stephens, D. S. Endotoxin of *Neisseria meningitidis* Composed Only of Intact Lipid A: Inactivation of the Meningococcal 3-Deoxy- D -Manno-Octulosonic Acid Transferase. *J Bacteriol* **184**, 2379–2388 (2002).
130. Schumann, R. *et al.* Structure and function of lipopolysaccharide binding protein. *Science* **249**, 1429–1431 (1990).
131. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide Endotoxins. *Annu. Rev. Biochem.* **71**, 635–700 (2002).
132. Wei, J. H., Yin, X. & Welander, P. V. Sterol Synthesis in Diverse Bacteria. *Front. Microbiol.* **7**, (2016).
133. Koppel, N. & Balskus, E. P. Exploring and Understanding the Biochemical Diversity of the Human Microbiota. *Cell Chemical Biology* **23**, 18–30 (2016).
134. Wikoff, W. R. *et al.* Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences* **106**, 3698–3703 (2009).
135. Sokol, H. Microbiote, ce qu'il faut savoir. *Colon Rectum* **8**, 136–140 (2014).
136. Schoeler, M. & Caesar, R. Dietary lipids, gut microbiota and lipid metabolism. *Rev Endocr Metab Disord* **20**, 461–472 (2019).
137. Gérard, P. *Gastrointestinal tract: microbial metabolism of steroids. In:Timmis, Kenneth. N. (Ed). 'Microbiology of Hydrocarbons, Oils, Lipids ans Derivated Compounds'.*
138. Ridlon, J. M., Kang, D.-J. & Hylemon, P. B. Bile salt biotransformations by human intestinal bacteria. *Journal of Lipid Research* **47**, 241–259 (2006).
139. McGarr, S. E., Ridlon, J. M. & Hylemon, P. B. Diet, anaerobic bacterial metabolism, and colon cancer: a review of the literature. *J Clin Gastroenterol* **39**, 98–109 (2005).
140. Guarino, M. P. L. Ursodeoxycholic acid therapy in gallbladder disease, a story not yet completed. *WJG* **19**, 5029 (2013).
141. Kovatcheva-Datchary, P., Zoetendal, E. G., Venema, K., de Vos, W. M. & Smidt, H. Review: Tools for the tract: understanding the functionality of the gastrointestinal tract. *Therap Adv Gastroenterol* **2**, S9–S22 (2009).
142. Roy, C. C., Kien, C. L., Bouthillier, L. & Levy, E. Short-Chain Fatty Acids: Ready for Prime Time? *Nutr Clin Pract* **21**, 351–366 (2006).
143. Cook & Sellin. Review article: short chain fatty acids in health and disease: REVIEW: SCFAs IN HEALTH AND DISEASE. *Alimentary Pharmacology & Therapeutics* **12**, 499–507 (1998).
144. Cummings, J. H. Short chain fatty acids in the human colon. *Gut* **22**, 763–779 (1981).
145. Ganapathy, V., Thangaraju, M., Prasad, P. D., Martin, P. M. & Singh, N. Transporters and receptors for short-chain fatty acids as the molecular link between colonic bacteria and the host. *Current Opinion in Pharmacology* **13**, 869–874 (2013).
146. Charney, A. N., Micic, L. & Egnor, R. W. Nonionic diffusion of short-chain fatty acids across rat colon. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **274**, G518–G524 (1998).
147. Stumpff, F. A look at the smelly side of physiology: transport of short chain fatty acids. *Pflugers Arch - Eur J Physiol* **470**, 571–598 (2018).
148. Wong, J. M. W., de Souza, R., Kendall, C. W. C., Emam, A. & Jenkins, D. J. A. Colonic Health: Fermentation and Short Chain Fatty Acids: *Journal of Clinical Gastroenterology* **40**, 235–243 (2006).
149. Cummings, J. H. & Macfarlane, G. T. Collaborative JPEN-Clinical Nutrition Scientific Publications Role of intestinal bacteria in nutrient metabolism. *JPEN J Parenter Enteral Nutr* **21**, 357–365 (1997).
150. Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood–Ljungdahl pathway of CO2 fixation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1784**, 1873–1898 (2008).
151. Park, J. *et al.* Short-chain fatty acids induce both effector and regulatory T cells by suppression of histone deacetylases and regulation of the mTOR–S6K pathway. *Mucosal Immunol* **8**, 80–93 (2015).
152. Hetzel, M. *et al.* Acryloyl-CoA reductase from Clostridium propionicum: An enzyme complex of propionyl-CoA dehydrogenase and electron-transferring flavoprotein. *European Journal of Biochemistry* **270**, 902–910 (2003).
153. Chirakkal, H. *et al.* Upregulation of BAK by butyrate in the colon is associated with increased Sp3 binding. *Oncogene* **25**, 7192–7200 (2006).

154. Hinnebusch, B. F., Meng, S., Wu, J. T., Archer, S. Y. & Hodin, R. A. The Effects of Short-Chain Fatty Acids on Human Colon Cancer Cell Phenotype Are Associated with Histone Hyperacetylation. *The Journal of Nutrition* **132**, 1012–1017 (2002).
155. Andoh, Mitsue Shimada, Yoshio Araki, A. Sodium butyrate enhances complement-mediated cell injury via down-regulation of decay-accelerating factor expression in colonic cancer cells. *Cancer Immunology, Immunotherapy* **50**, 663–672 (2002).
156. Sauer, J., Richter, K. K. & Pool-Zobel, B. L. Physiological concentrations of butyrate favorably modulate genes of oxidative and metabolic stress in primary human colon cells. *The Journal of Nutritional Biochemistry* **18**, 736–745 (2007).
157. B. Gowda, S. G. *et al.* Chemical Labeling Assisted Detection and Identification of Short Chain Fatty Acid Esters of Hydroxy Fatty Acid in Rat Colon and Cecum Contents. *Metabolites* **10**, 398 (2020).
158. Cao, Y. & Zhang, X. Production of long-chain hydroxy fatty acids by microbial conversion. doi:10.1007/s00253-013-4815-z.
159. Kim, K.-R. & Oh, D.-K. Production of hydroxy fatty acids by microbial fatty acid-hydroxylation enzymes. *Biotechnology Advances* **31**, 1473–1485 (2013).
160. Cohen, L. J. *et al.* Functional metagenomic discovery of bacterial effectors in the human microbiome and isolation of commendamide, a GPCR G2A/132 agonist. *Proc Natl Acad Sci USA* **112**, E4825–E4834 (2015).
161. Julkowska, D., Obuchowski, M., Holland, I. B. & Séror, Simone. J. Comparative Analysis of the Development of Swarming Communities of *Bacillus subtilis* 168 and a Natural Wild Type: Critical Effects of Surfactin and the Composition of the Medium. *J Bacteriol* **187**, 65–76 (2005).
162. Götze, S. & Stallforth, P. Structure, properties, and biological functions of nonribosomal lipopeptides from pseudomonads. *Nat. Prod. Rep.* **37**, 29–54 (2020).
163. Sorensen, K. N., Kim, K.-H. & Takemoto, J. Y. PCR Detection of Cyclic Lipodepsinonapeptide-Producing *Pseudomonas syringae* pv. syringae and Similarity of Strains. *Appl Environ Microbiol* **64**, 226–230 (1998).
164. Sorensen, K. N., Kim, K. H. & Takemoto, J. Y. In vitro antifungal and fungicidal activities and erythrocyte toxicities of cyclic lipodepsinonapeptides produced by Pseudomonas syringae pv. syringae. *Antimicrob Agents Chemother* **40**, 2710–2713 (1996).
165. Tsukamoto, T., Murata, H. & Shirata, A. Identification of Non-Pseudomonad Bacteria from Fruit Bodies of Wild Agaricales Fungi That Detoxify Tolaasin Produced by *Pseudomonas tolaasii*. *Bioscience, Biotechnology, and Biochemistry* **66**, 2201–2208 (2002).
166. Nougayrede, J.-P. Escherichia coli Induces DNA Double-Strand Breaks in Eukaryotic Cells. *Science* **313**, 848–851 (2006).
167. Vizcaino, M. I., Engel, P., Trautman, E. & Crawford, J. M. Comparative Metabolomics and Structural Characterizations Illuminate Colibactin Pathway-Dependent Small Molecules. *J. Am. Chem. Soc.* **136**, 9244–9247 (2014).
168. Cane, D. E., Walsh, C. T. & Khosla, C. Harnessing the Biosynthetic Code: Combinations, Permutations, and Mutations. *Science* **282**, 63–68 (1998).
169. Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev.* **97**, 2651–2674 (1997).
170. Cane, D. E. & Walsh, C. T. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chemistry & Biology* **6**, R319–R325 (1999).
171. Putze, J. *et al.* Genetic Structure and Distribution of the Colibactin Genomic Island among Members of the Family *Enterobacteriaceae*. *Infect Immun* **77**, 4696–4703 (2009).
172. Manzoni, C. *et al.* Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. *Briefings in Bioinformatics* **19**, 286–302 (2018).
173. Holčapek, M., Liebisch, G. & Ekroos, K. Lipidomic Analysis. *Anal. Chem.* **90**, 4249–4257 (2018).
174. Wenk, M. R. The emerging field of lipidomics. *Nat Rev Drug Discov* **4**, 594–610 (2005).
175. Zehethofer, N. & Pinto, D. M. Recent developments in tandem mass spectrometry for lipidomic analysis. *Analytica Chimica Acta* **627**, 62–70 (2008).
176. Capolupo, L. *et al. Sphingolipid Control of Fibroblast Heterogeneity Revealed by Single-Cell Lipidomics*. http://biorxiv.org/lookup/doi/10.1101/2021.02.23.432420 (2021) doi:10.1101/2021.02.23.432420.

177. Carter, J. A., O'Brien, L. M., Harville, T., Jones, B. T. & Donati, G. L. Machine learning tools to estimate the severity of matrix effects and predict analyte recovery in inductively coupled plasma optical emission spectrometry. *Talanta* **223**, 121665 (2021).
178. Trufelli, H., Palma, P., Famiglini, G. & Cappiello, A. An overview of matrix effects in liquid chromatography-mass spectrometry: AN OVERVIEW OF MATRIX EFFECTS IN LC-MS. *Mass Spectrom. Rev.* **30**, 491–509 (2011).
179. Le Faouder, P. *et al.* LC–MS/MS method for rapid and concomitant quantification of pro-inflammatory and pro-resolving polyunsaturated fatty acid metabolites. *Journal of Chromatography B* **932**, 123–133 (2013).
180. Le Faouder, P. *et al.* Untargeted Lipidomic Profiling of Dry Blood Spots Using SFC-HRMS. *Metabolites* **11**, 305 (2021).
181. *Charged Aerosol Detection for Liquid Chromatography and Related Separation Techniques*. (John Wiley & Sons, Inc., 2017). doi:10.1002/9781119390725.
182. Yang, L. *et al.* Recent advances in lipidomics for disease research: Liquid Chromatography. *J. Sep. Science* **39**, 38–50 (2016).
183. Bligh, E. G. & Dyer, W. J. A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).
184. Yu, Z. *et al.* Global lipidomics identified plasma lipids as novel biomarkers for early detection of lung cancer. *Oncotarget* **8**, 107899–107906 (2017).
185. Wolf, C. & Quinn, P. J. Lipidomics: Practical aspects and applications. *Progress in Lipid Research* **47**, 15–36 (2008).
186. Riols, F. & Bertrand-Michel, J. Analysis of Oxysterols. in *Clinical Metabolomics* (ed. Giera, M.) vol. 1730 267–275 (Springer New York, 2018).
187. Folch, J., Lees, M. & Sloane Stanley, G. H. A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* **226**, 497–509 (1957).
188. Cajka, T. & Fiehn, O. Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. *TrAC Trends in Analytical Chemistry* **61**, 192–206 (2014).
189. Löfgren, L. Liquid Extraction: BUME. in *Encyclopedia of Lipidomics* (ed. Wenk, M. R.) 1–5 (Springer Netherlands, 2016). doi:10.1007/978-94-007-7864-1_98-1.
190. Ulmer, C. Z., Jones, C. M., Yost, R. A., Garrett, T. J. & Bowden, J. A. Optimization of Folch, Bligh-Dyer, and Matyash sample-to-extraction solvent ratios for human plasma-based lipidomics studies. *Analytica Chimica Acta* **1037**, 351–357 (2018).
191. Dupuy, A. *et al.* Simultaneous quantitative profiling of 20 isoprostanoids from omega-3 and omega-6 polyunsaturated fatty acids by LC–MS/MS in various biological samples. *Analytica Chimica Acta* **921**, 46–58 (2016).
192. www.agilent.com/.
193. Habib, A. *et al.* Inhibition of monoacylglycerol lipase, an anti-inflammatory and antifibrogenic strategy in the liver. *Gut* **68**, 522–532 (2019).
194. Schwudke, D., Liebisch, G., Herzog, R., Schmitz, G. & Shevchenko, A. Shotgun Lipidomics by Tandem Mass Spectrometry under Data-Dependent Acquisition Control. in *Methods in Enzymology* vol. 433 175–191 (Elsevier, 2007).
195. Han, X. & Gross, R. W. Shotgun lipidomics: Electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass Spectrom. Rev.* **24**, 367–412 (2005).
196. James, A. T. & Martin, A. J. P. Gas-liquid partition chromatography: the separation and micro-estimation of volatile fatty acids from formic acid to dodecanoic acid. *Biochemical Journal* **50**, 679–690 (1952).
197. Principe de la chromtographie gazeuse.
198. Ichihara, K. *et al.* Fatty acid analysis of triacylglycerols: Preparation of fatty acid methyl esters for gas chromatography. *Analytical Biochemistry* **495**, 6–8 (2016).
199. Řezanka, T. & Řezanková, H. Characterization of fatty acids and triacylglycerols in vegetable oils by gas chromatography and statistical analysis. *Analytica Chimica Acta* **398**, 253–261 (1999).
200. Li, M., Zhou, Z., Nie, H., Bai, Y. & Liu, H. Recent advances of chromatography and mass spectrometry in lipidomics. *Anal Bioanal Chem* **399**, 243–249 (2011).

201. O. Berdeaux, P. Juanéda et J.L. Sébédio. Analyse des acides gras conjugués et trans après dérivation. *ANALUSIS MAGAZINE* (1998).
202. Christie, W. W. Gas chromatography-mass spectrometry methods for structural analysis of fatty acids. *Lipids* **33**, 343–353 (1998).
203. Blaško, J. *et al.* Chemometric deconvolution of gas chromatographic unresolved conjugated linoleic acid isomers triplet in milk samples. *Journal of Chromatography A* **1216**, 2757–2761 (2009).
204. Teo, C. C. *et al.* Advances in sample preparation and analytical techniques for lipidomics study of clinical samples. *TrAC Trends in Analytical Chemistry* **66**, 1–18 (2015).
205. Phenomenex - brochure colonnes LC.
206. Li, M., Yang, L., Bai, Y. & Liu, H. Analytical Methods in Lipidomics and Their Applications. *Anal. Chem.* **86**, 161–175 (2014).
207. Holmes, E., Wijeyesekera, A., Taylor-Robinson, S. D. & Nicholson, J. K. The promise of metabolic phenotyping in gastroenterology and hepatology. *Nat Rev Gastroenterol Hepatol* **12**, 458–471 (2015).
208. Patti, G. J., Yanes, O. & Siuzdak, G. Metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol* **13**, 263–269 (2012).
209. Barbier Saint Hilaire, P. *et al.* Mechanistic study of competitive releases of H 2 O, NH 3 and CO 2 from deprotonated aspartic and glutamic acids: Role of conformation. *Journal of Chromatography B* **1047**, 64–74 (2017).
210. Fenn, J., Mann, M., Meng, C., Wong, S. & Whitehouse, C. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71 (1989).
211. Whitehouse, C. M., Dreyer, R. N., Yamashita, Masamichi. & Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* **57**, 675–679 (1985).
212. Karas, Michael., Bachmann, Doris. & Hillenkamp, Franz. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem.* **57**, 2935–2939 (1985).
213. Dumas, M.-E. *et al.* Analyzing the Physiological Signature of Anabolic Steroids in Cattle Urine Using Pyrolysis/Metastable Atom Bombardment Mass Spectrometry and Pattern Recognition. *Anal. Chem.* **74**, 5393–5404 (2002).
214. Murray, K. K. *et al.* Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure and Applied Chemistry* **85**, 1515–1609 (2013).
215. Lin, I.-H. & Thomson, R. Cleavage, dislocation emission, and shielding for cracks under general loading. *Acta Metallurgica* **34**, 187–206 (1986).
216. Kim, H. S., Yu, M. L., Thomson, M. G. R., Kratschmer, E. & Chang, T. H. P. Energy distributions of Zr/O/W Schottky electron emission. *Journal of Applied Physics* **81**, 461–465 (1997).
217. Hoover, H. & Washburn, H. A Preliminary Report on the Application of the Mass Spectrometer to Problems in the Petroleum Industry. *Transactions of the AIME* **142**, 100–106 (1941).
218. Gohlke, R. S. & McLafferty, F. W. Early gas chromatography/mass spectrometry. *Journal of the American Society for Mass Spectrometry* **4**, 367–371 (1993).
219. Tabet, J. Claude. & Cotter, R. J. Laser desorption time-of-flight mass spectrometry of high mass molecules. *Anal. Chem.* **56**, 1662–1667 (1984).
220. Karas, M., Bachmann, D., Bahr, U. & Hillenkamp, F. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes* **78**, 53–68 (1987).
221. Bruins, A. P. Atmospheric-pressure-ionization mass spectrometry. *TrAC Trends in Analytical Chemistry* **13**, 37–43 (1994).
222. Tanaka, K. *et al.* A new rapid and comprehensive peptidome analysis by one-step direct transfer technology for 1-D electrophoresis/MALDI mass spectrometry. *Biochemical and Biophysical Research Communications* **379**, 110–114 (2009).
223. Tabet, J.-C. *et al.* Combining Chemical Knowledge and Quantum Calculation for Interpreting Low-Energy Product Ion Spectra of Metabolite Adduct Ions: Sodiated Diterpene Diester Species as a Case Study. *J. Am. Soc. Mass Spectrom.* jasms.1c00154 (2021) doi:10.1021/jasms.1c00154.
224. Cooks, R. G. & Beynon, J. H. Metastable ions and ion kinetic energy spectrometry: The development of a new research area. *J. Chem. Educ.* **51**, 437 (1974).
225. McLafferty, F. W. Tandem mass spectrometry (MS/MS): a promising new analytical technique for specific component determination in complex mixtures. *Acc. Chem. Res.* **13**, 33–39 (1980).

226. McLafferty, F. W., Bente, P. F., Kornfeld, R., Tsai, S.-C. & Howe, I. Collisional activation spectra of organic ions. *J. Mass Spectrom.* **30**, 797–806 (1995).
227. McLafferty, F. Tandem mass spectrometry. *Science* **214**, 280–287 (1981).
228. Beynon, J. H., Cooks, R. G., Amy, J. W., Baitinger, W. E. & Ridley, T. Y. Design and Performance of a Mass-analyzed Ion Kinetic Energy (MIKE) Spectrometer. *Anal. Chem.* **45**, 1023A-1031A (1973).
229. Maquestiau, A. *et al.* The gas-phase pinacol rearrangement and related reactions in organic cations generated by chemical ionisation. *Tetrahedron* **36**, 1993–1998 (1980).
230. Yost, R. A. & Enke, C. G. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Anal. Chem.* **51**, 1251–1264 (1979).
231. Stafford, G. C., Kelley, P. E., Syka, J. E. P., Reynolds, W. E. & Todd, J. F. J. Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Processes* **60**, 85–98 (1984).
232. Berger, C. Compensator role of the electrostatic mirror in time of flight mass spectrometry. *International Journal of Mass Spectrometry and Ion Physics* **46**, 63–66 (1983).
233. Loboda, null, Krutchinsky, null, Bromirski, null, Ens, null & Standing, null. A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. *Rapid Commun Mass Spectrom* **14**, 1047–1057 (2000).
234. Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **72**, 1156–1162 (2000).
235. Wang, Y., Shi, S. D.-H., Hendrickson, C. L. & Marshall, A. G. Mass-selective ion accumulation and fragmentation in a linear octopole ion trap external to a fourier transform ion cyclotron resonance mass spectrometer. *International Journal of Mass Spectrometry* **198**, 113–120 (2000).
236. Nguyen, V. H., Afonso, C. & Tabet, J.-C. Concomitant EDD and EID of DNA evidenced by MSn and double resonance experiments. *International Journal of Mass Spectrometry* **301**, 224–233 (2011).
237. Cooper, H. J., Håkansson, K. & Marshall, A. G. The role of electron capture dissociation in biomolecular analysis: ELECTRON CAPTURE DISSOCIATION. *Mass Spectrom. Rev.* **24**, 201–222 (2005).
238. Kalli, A. & Håkansson, K. Preferential cleavage of SS and CS bonds in electron detachment dissociation and infrared multiphoton dissociation of disulfide-linked peptide anions. *International Journal of Mass Spectrometry* **263**, 71–81 (2007).
239. Baba, T., Campbell, J. L., Le Blanc, J. C. Y., Baker, PaulR. S. & Ikeda, K. Quantitative structural multiclass lipidomics using differential mobility: electron impact excitation of ions from organics (EIEIO) mass spectrometry. *Journal of Lipid Research* **59**, 910–919 (2018).
240. Gelpí, E. Interfaces for coupled liquid-phase separation/mass spectrometry techniques. An update on recent developments: Interfaces for coupled liquid-phase separation/MS. *J. Mass Spectrom.* **37**, 241–253 (2002).
241. Hiraoka, K., Fujimaki, S., Kambara, S., Furuya, H. & Okazaki, S. Atmospheric-pressure Penning ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **18**, 2323–2330 (2004).
242. Gronert, S. Mass Spectrometric Studies of Organic Ion/Molecule Reactions. *Chem. Rev.* **101**, 329–360 (2001).
243. Becker, L., Bunch, T. E. & Allamandola, L. J. Higher fullerenes in the Allende meteorite. *Nature* **400**, 227–228 (1999).
244. Biemann, K. *et al.* Search for Organic and Volatile Inorganic Compounds in Two Surface Samples from the Chryse Planitia Region of Mars. *Science* **194**, 72–76 (1976).
245. Donahue, T. M., Hoffman, J. H. & Hodges, R. R. Krypton and xenon in the atmosphere of Venus. *Geophys. Res. Lett.* **8**, 513–516 (1981).
246. Niemann, H. B. *et al.* The Galileo Probe Mass Spectrometer: Composition of Jupiter's Atmosphere. *Science* **272**, 846–849 (1996).
247. Abell, P. I. *et al.* Organic Analysis of the Returned Lunar Sample. *Science* **167**, 757–759 (1970).
248. Hersant, F., Gautier, D., Tobie, G. & Lunine, J. I. Interpretation of the carbon abundance in Saturn measured by Cassini. *Planetary and Space Science* **56**, 1103–1111 (2008).
249. Biemann, K. Organic Analysis. *Appl. Opt.* **9**, 1282 (1970).
250. Horning, E. C., Horning, M. G., Carroll, D. I., Dzidic, I. & Stillwell, R. N. New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure. *Anal. Chem.* **45**, 936–943 (1973).

251. Lundström, S. L. *et al.* HPLC/MS/MS-Based Approaches for Detection and Quantification of Eicosanoids. in *Lipidomics* (ed. Armstrong, D.) vol. 579 161–187 (Humana Press, 2009).
252. Kolakowski, B. M., Grossert, J. S. & Ramaley, L. Studies on the positive-ion mass spectra from atmospheric pressure chemical ionization of gases and solvents used in liquid chromatography and direct liquid injection. *J. Am. Soc. Mass Spectrom.* **15**, 311–324 (2004).
253. Meot-Ner, M., Elmore, D. E. & Scheiner, S. Ionic Hydrogen Bond Effects on the Acidities, Basicities, Solvation, Solvent Bridging, and Self-Assembly of Carboxylic Groups. *J. Am. Chem. Soc.* **121**, 7625–7635 (1999).
254. Chang, T. M., Prell, J. S., Warrick, E. R. & Williams, E. R. Where's the Charge? Protonation Sites in Gaseous Ions Change with Hydration. *J. Am. Chem. Soc.* **134**, 15805–15813 (2012).
255. Sunner, Jan., Nicol, Gordon. & Kebarle, Paul. Factors determining relative sensitivity of analytes in positive mode atmospheric pressure ionization mass spectrometry. *Anal. Chem.* **60**, 1300–1307 (1988).
256. Covey, T. R., Thomson, B. A. & Schneider, B. B. Atmospheric pressure ion sources. *Mass Spectrom. Rev.* **28**, 870–897 (2009).
257. Wu, J., Liu, J. & Cai, Z. Determination of triclosan metabolites by using in-source fragmentation from high-performance liquid chromatography/negative atmospheric pressure chemical ionization ion trap mass spectrometry: In-source fragmentation of triclosan and its metabolites. *Rapid Commun. Mass Spectrom.* **24**, 1828–1834 (2010).
258. Carrier, D. J., Eckers, C. & Wolff, J.-C. 'In-source' fragmentation of an isobaric impurity of lamotrigine for its measurement by liquid chromatography tandem mass spectrometry after pre-concentration using solid phase extraction. *J Pharm Biomed Anal* **47**, 731–737 (2008).
259. Tian, Q., Duncan, C. J. G. & Schwartz, S. J. Atmospheric pressure chemical ionization mass spectrometry and in-source fragmentation of lutein esters. *J. Mass Spectrom.* **38**, 990–995 (2003).
260. Damont, A. *et al.* Proposal for a chemically consistent way to annotate ions arising from the analysis of reference compounds under ESI conditions: A prerequisite to proper mass spectral database constitution in metabolomics. *J Mass Spectrom* **54**, 567–582 (2019).
261. Körner, R., Wilm, M., Morand, K., Schubert, M. & Mann, M. Nano electrospray combined with a quadrupole ion trap for the analysis of peptides and protein digests. *J Am Soc Mass Spectrom* **7**, 150–156 (1996).
262. Takáts, Z., Wiseman, J. M. & Cooks, R. G. Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *J. Mass Spectrom.* **40**, 1261–1275 (2005).
263. Chen, H., Venter, A. & Cooks, R. G. Extractive electrospray ionization for direct analysis of undiluted urine, milk and other complex mixtures without sample preparation. *Chem. Commun.* 2042 (2006) doi:10.1039/b602614a.
264. Li, H., Ha, E., Donaldson, R. P., Jeremic, A. M. & Vertes, A. Rapid Assessment of Human Amylin Aggregation and Its Inhibition by Copper(II) Ions by Laser Ablation Electrospray Ionization Mass Spectrometry with Ion Mobility Separation. *Anal. Chem.* **87**, 9829–9837 (2015).
265. Charles Bossut. *Traité élementaire hydrodynamique*. (1775).
266. Abbé NOLLET. *Essai sur l'électricité des corps*. (1774).
267. Yamashita, M. & Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **88**, 4451–4459 (1984).
268. Cech, N. B. & Enke, C. G. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom. Rev.* **20**, 362–387 (2001).
269. Thomson, B. A. & Iribarne, J. V. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *The Journal of Chemical Physics* **71**, 4451–4463 (1979).
270. Iribarne, J. V. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287 (1976).
271. Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *International Journal of Mass Spectrometry* **219**, 39–62 (2002).
272. Kebarle, P. & Verkerk, U. H. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898–917 (2009).
273. Fernandez de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica Chimica Acta* **406**, 93–104 (2000).

274. Pascoe, R., Foley, J. P. & Gusev, A. I. Reduction in Matrix-Related Signal Suppression Effects in Electrospray Ionization Mass Spectrometry Using On-Line Two-Dimensional Liquid Chromatography. *Anal. Chem.* **73**, 6014–6023 (2001).
275. Marshall, A. G. & Hendrickson, C. L. High-Resolution Mass Spectrometers. *Annual Rev. Anal. Chem.* **1**, 579–599 (2008).
276. Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. An introduction to quadrupole-time-of-flight mass spectrometry. *J. Mass Spectrom.* **36**, 849–865 (2001).
277. Schmid, D. G., Grosche, P., Bandel, H. & Jung, G. FTICR-mass spectrometry for high-resolution analysis in combinatorial chemistry. *Biotechnol Bioeng* **71**, 149–161 (2000).
278. Jennings, K. R. Metastable Transitions in the Mass Spectrum of Benzene. *The Journal of Chemical Physics* **43**, 4176–4177 (1965).
279. Haddon, W. F. & McLafferty, F. W. Metastable ion characteristics. VII. Collision-induced metastables. *J. Am. Chem. Soc.* **90**, 4745–4746 (1968).
280. McLafferty, F. W. & Schuddemage, H. D. R. Minimization of rearrangement reactions in mass spectra by use of collisional activation. *J. Am. Chem. Soc.* **91**, 1866–1868 (1969).
281. Beynon, J. H. & Cooks, R. G. Ion kinetic energy spectrometry. *International Journal of Mass Spectrometry and Ion Physics* **19**, 107–137 (1976).
282. Beynon, J. H., Caprioli, R. M., Baitinger, W. E. & Amy, J. W. The ion kinetic energy spectrum and the mass spectrum of argon. *International Journal of Mass Spectrometry and Ion Physics* **3**, 313–321 (1969).
283. Liehr, J. G., Kingston, E. E. & Beynon, J. H. Collision-induced dissociation-mass-analysed ion kinetic energy (CID-MIKE) analysis of ions generated by fast atom bombardment of isomeric bile salts. *Biol. Mass Spectrom.* **12**, 95–99 (1985).
284. Paul, W. & Steinwedel, H. Notizen: Ein neues Massenspektrometer ohne Magnetfeld. *Zeitschrift für Naturforschung A* **8**, 448–450 (1953).
285. Dawson, P. H., French, J. B., Buckley, J. A., Douglas, D. J. & Simmons, D. The use of triple quadrupoles for sequential mass spectrometry: 1—The instrument parameters. *Org. Mass Spectrom.* **17**, 205–211 (1982).
286. Ferguson, R. E., McCulloh, K. E. & Rosenstock, H. M. Observation of the Products of Ionic Collision Processes and Ion Decomposition in a Linear, Pulsed Time-of-Flight Mass Spectrometer. *The Journal of Chemical Physics* **42**, 100–106 (1965).
287. Wolff, M. M. & Stephens, W. E. A Pulsed Mass Spectrometer with Time Dispersion. *Review of Scientific Instruments* **24**, 616–617 (1953).
288. Wiley, W. C. & McLaren, I. H. Time-of-Flight Mass Spectrometer with Improved Resolution. *Review of Scientific Instruments* **26**, 1150–1157 (1955).
289. Guilhaus, M., Selby, D. & Mlynski, V. Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrom Rev* **19**, 65–107 (2000).
290. Della-Negra, Serge. & Le Beyec, Yvon. New method for metastable ion studies with a time of flight mass spectrometer. Future application to molecular structure determinations. *Anal. Chem.* **57**, 2035–2040 (1985).
291. Kingdon, K. H. A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. *Phys. Rev.* **21**, 408–418 (1923).
292. Garnier, N., Rolando, C., Høtje, J. M. & Tokarski, C. Analysis of archaeological triacylglycerols by high resolution nanoESI, FT-ICR MS and IRMPD MS/MS: Application to 5th century BC–4th century AD oil lamps from Olbia (Ukraine). *International Journal of Mass Spectrometry* **284**, 47–56 (2009).
293. Holguin, F. O. & Schaub, T. Characterization of microalgal lipid feedstock by direct-infusion FT-ICR mass spectrometry. *Algal Research* **2**, 43–50 (2013).
294. Knight, R. D. Storage of ions from laser-produced plasmas. *Appl. Phys. Lett.* **38**, 221–223 (1981).
295. A. Makarov. United States Patent : Mass spectrometer. U.S Patent 5,886,346,1999.
296. Scigelova, M. & Makarov, A. Orbitrap Mass Analyzer – Overview and Applications in Proteomics. *Proteomics* **6**, 16–21 (2006).
297. Makarov, A., Denisov, E., Lange, O. & Horning, S. Dynamic range of mass accuracy in LTQ orbitrap hybrid mass spectrometer. *J Am Soc Mass Spectrom* **17**, 977–982 (2006).

298. Makarov, A. *et al.* Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Anal. Chem.* **78**, 2113–2120 (2006).
299. Makarov, A. & Denisov, E. Dynamics of ions of intact proteins in the Orbitrap mass analyzer. *J Am Soc Mass Spectrom* **20**, 1486–1495 (2009).
300. Makarov, A., Denisov, E. & Lange, O. Performance evaluation of a high-field orbitrap mass analyzer. *J Am Soc Mass Spectrom* **20**, 1391–1396 (2009).
301. Langevin, P. Mesure de la valence des ions dans les gaz. *Radium (Paris)* **10**, 113–118 (1913).
302. The speed of positive ions in nitrogen. *Proc. R. Soc. Lond. A* **146**, 911–921 (1934).
303. Cravath, A. M. The Rate of Formation of Negative Ions by Electron Attachment. *Phys. Rev.* **33**, 605–613 (1929).
304. Van De Graaff, R. J. The Mobility of Ions in Gases. *Nature* **124**, 10–11 (1929).
305. Bradbury, N. E. & Nielsen, R. A. Absolute Values of the Electron Mobility in Hydrogen. *Phys. Rev.* **49**, 388–393 (1936).
306. Edward A. Mason, Earl W. McDaniel. *Transport Properties of Ions in Gases*. (2005).
307. Smith, D. & Adams, N. G. The Selected Ion Flow Tube (Sift): Studies of Ion-Neutral Reactions. in *Advances in Atomic and Molecular Physics* vol. 24 1–49 (Elsevier, 1988).
308. McFarland, M., Albritton, D. L., Fehsenfeld, F. C., Ferguson, E. E. & Schmeltekopf, A. L. Flow-drift technique for ion mobility and ion-molecule reaction rate constant measurements. III. Negative ion reactions of O^- with CO, NO, H_2, and D_2. *The Journal of Chemical Physics* **59**, 6629–6635 (1973).
309. Cohen, M. J. & Karasek, F. W. Plasma Chromatography --A New Dimension for Gas Chromatography and Mass Spectrometry. *Journal of Chromatographic Science* **8**, 330–337 (1970).
310. Karasek, F. W. & Fong, I. Analysis of Chlorinated Benzene Compounds by Gas Chromatography. *Journal of Chromatographic Science* **9**, 497–499 (1971).
311. Karasek, F. W., Hill, H. H. & Kim, S. H. Plasma chromatography of heroin and cocaine with mass-identified mobility spectra. *Journal of Chromatography A* **117**, 327–336 (1976).
312. Eiceman, G. A. & Karpas, Z. *Ion Mobility Spectrometry*. (CRC Press, 2005). doi:10.1201/9781420038972.
313. Dodds, J. N. & Baker, E. S. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J. Am. Soc. Mass Spectrom.* **30**, 2185–2195 (2019).
314. Gabelica, V. & Marklund, E. Fundamentals of ion mobility spectrometry. *Current Opinion in Chemical Biology* **42**, 51–59 (2018).
315. Mason, E. A. & Schamp, H. W. Mobility of gaseous Ions in weak electric fields. *Annals of Physics* **4**, 233–270 (1958).
316. Delvaux, A., Rathahao-Paris, E. & Alves, S. Different ion mobility-mass spectrometry coupling techniques to promote metabolomics. *Mass Spec Rev* mas.21685 (2021) doi:10.1002/mas.21685.
317. Guntner, A. S., Thalhamer, B., Klampfl, C. & Buchberger, W. Collision cross sections obtained with ion mobility mass spectrometry as new descriptor to predict blood-brain barrier permeation by drugs. *Sci Rep* **9**, 19182 (2019).
318. Eiceman, G. A., Krylov, E. V., Krylova, N. S., Nazarov, E. G. & Miller, R. A. Separation of Ions from Explosives in Differential Mobility Spectrometry by Vapor-Modified Drift Gas. *Anal. Chem.* **76**, 4937–4944 (2004).
319. Lindinger, W., Hansel, A. & Jordan, A. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research. *International Journal of Mass Spectrometry and Ion Processes* **173**, 191–241 (1998).
320. Shaffer, S. A., Prior, D. C., Anderson, G. A., Udseth, H. R. & Smith, R. D. An Ion Funnel Interface for Improved Ion Focusing and Sensitivity Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **70**, 4111–4119 (1998).
321. Odenkirk, M. T. & Baker, E. S. Utilizing Drift Tube Ion Mobility Spectrometry for the Evaluation of Metabolites and Xenobiotics. in *Ion Mobility-Mass Spectrometry* (eds. Paglia, G. & Astarita, G.) vol. 2084 35–54 (Springer US, 2020).
322. Buryakov, I. A., Krylov, E. V., Nazarov, E. G. & Rasulev, U. Kh. A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field. *International Journal of Mass Spectrometry and Ion Processes* **128**, 143–148 (1993).

323. Lambertus, G. R. *et al.* Silicon Microfabricated Column with Microfabricated Differential Mobility Spectrometer for GC Analysis of Volatile Organic Compounds. *Anal. Chem.* **77**, 7563–7571 (2005).
324. Purves, R. W., Guevremont, R., Day, S., Pipich, C. W. & Matyjaszczyk, M. S. Mass spectrometric characterization of a high-field asymmetric waveform ion mobility spectrometer. *Review of Scientific Instruments* **69**, 4094–4105 (1998).
325. Guevremont, R. & Purves, R. W. Atmospheric pressure ion focusing in a high-field asymmetric waveform ion mobility spectrometer. *Review of Scientific Instruments* **70**, 1370–1383 (1999).
326. Giles, K. *et al.* Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.* **18**, 2401–2414 (2004).
327. Bush, M. F., Campuzano, I. D. G. & Robinson, C. V. Ion Mobility Mass Spectrometry of Peptide Ions: Effects of Drift Gas and Calibration Strategies. *Anal. Chem.* **84**, 7124–7130 (2012).
328. Giles, K. *et al.* A Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **91**, 8564–8573 (2019).
329. Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. Naked Protein Conformations: Cytochrome c in the Gas Phase. *J. Am. Chem. Soc.* **117**, 10141–10142 (1995).
330. Paglia, G., Kliman, M., Claude, E., Geromanos, S. & Astarita, G. Applications of ion-mobility mass spectrometry for lipid analysis. *Anal Bioanal Chem* **407**, 4995–5007 (2015).
331. Groessl, M., Graf, S. & Knochenmuss, R. High resolution ion mobility-mass spectrometry for separation and identification of isomeric lipids. *Analyst* **140**, 6904–6911 (2015).
332. Comisarow, M. B. & Marshall, A. G. Fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters* **25**, 282–283 (1974).
333. Comisarow, M. B., Grassi, V. & Parisod, G. Fourier transform ion cyclotron double resonance. *Chemical Physics Letters* **57**, 413–416 (1978).
334. Scott A. McLuckey, Douglas E. Goeringer. *SPECIAL FEATURE:TUTORIAL Slow Heating Methods in Tandem Mass Spectrometry*. vols 32, Issue 5 (1998).
335. E. March. SPECIAL FEATURE: TUTORIAL An Introduction to Quadrupole Ion Trap Mass Spectrometry, JOURNAL OF MASS SPECTROMETRY.
336. Darii, E. *et al.* First Direct Evidence of Interpartner Hydride/Deuteride Exchanges for Stored Sodiated Arginine/Fructose-6-phosphate Complex Anions within Salt-Solvated Structures. *J. Am. Soc. Mass Spectrom.* **32**, 1424–1440 (2021).
337. Naban-Maillet, J. *et al.* Internal energy distribution in electrospray ionization: Internal energy distribution in electrospray ionization. *J. Mass Spectrom.* **40**, 1–8 (2005).
338. Ichou, F. *et al.* Comparison of the activation time effects and the internal energy distributions for the CID, PQD and HCD excitation modes: Theoretical comparison of CID, PQD and HCD. *J. Mass Spectrom.* **49**, 498–508 (2014).
339. https://www.waters.com/webassets/cms/library/docs/local_seminar_presentations/GE_Events/ge_2014_ms%20tech%20days_Xevo%20G2-XS%20QTof_Matthias%20Hofmann.pdf.
340. Watson, A. D. Thematic review series: Systems Biology Approaches to Metabolic and Cardiovascular Disorders. Lipidomics: a global approach to lipid analysis in biological systems. *Journal of Lipid Research* **47**, 2101–2111 (2006).
341. Colsch, B., Seyer, A., Boudah, S. & Junot, C. Lipidomic analysis of cerebrospinal fluid by mass spectrometry–based methods. *J Inherit Metab Dis* **38**, 53–64 (2015).
342. Taguchi, R. *et al.* Focused lipidomics by tandem mass spectrometry. *Journal of Chromatography B* **823**, 26–36 (2005).
343. Harkewicz, R. & Dennis, E. A. Applications of Mass Spectrometry to Lipids and Membranes. *Annu. Rev. Biochem.* **80**, 301–325 (2011).
344. Humbert, L. *et al.* Bile acid profiling in human biological samples: Comparison of extraction procedures and application to normal and cholestatic patients. *Journal of Chromatography B* **899**, 135–145 (2012).
345. Salek, R. M., Steinbeck, C., Viant, M. R., Goodacre, R. & Dunn, W. B. The role of reporting standards for metabolite annotation and identification in metabolomic studies. *GigaSci* **2**, 13 (2013).
346. Lipidomics Standards Initiative Consortium. Lipidomics needs more standardization. *Nat Metab* **1**, 745–747 (2019).
347. Liebisch, G. *et al.* Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures. *Journal of Lipid Research* **61**, 1539–1555 (2020).

348. Raczyńska, E. D., Gal, J.-F. & Maria, P.-C. The guanylated bioamine agmatine – A theoretical investigation of its structure and exceptional high basicity in the gas phase. *Computational and Theoretical Chemistry* **1109**, 10–18 (2017).
349. Jones, J. W., Thompson, C. J., Carter, C. L. & Kane, M. A. Electron-induced dissociation (EID) for structure characterization of glycerophosphatidylcholine: determination of double-bond positions and localization of acyl chains: Electron-Induced Dissociation of Glycerophosphatidylcholine. *J. Mass Spectrom.* **50**, 1327–1339 (2015).
350. Cook, S. L., Collin, O. L. & Jackson, G. P. Metastable atom-activated dissociation mass spectrometry: leucine/isoleucine differentiation and ring cleavage of proline residues: Metastable atom-activated dissociation mass spectrometry. *J. Mass Spectrom.* **44**, 1211–1223 (2009).
351. Longevialle, P. & Botter, R. Electron impact mass spectra of bifunctional steroids. The interaction between ionic and neutral fragments derived from the same parent ion. *Org. Mass Spectrom.* **18**, 1–8 (1983).
352. Debrauwer, L., Paris, A., Rao, D., Fournier, F. & Tabet, J.-C. Mass spectrometric studies on 17β-estradiol-17-fatty acid esters: Evidence for the formation of anion-dipole intermediates. *Org. Mass Spectrom.* **27**, 709–719 (1992).
353. Bowen, R. D. & Williams, D. H. Non-concrted unimolecular reactions of ions in the gas-phase: The importance of ion-dipole interactions in carbonium ion isomerizations. *International Journal of Mass Spectrometry and Ion Physics* **29**, 47–55 (1979).
354. Convert, O. *et al.* Reactions of Estradiol-2,3-quinone with Deoxyribonucleosides: Possible Insights in the Reactivity of Estrogen Quinones with DNA. *Chem. Res. Toxicol.* **15**, 754–764 (2002).
355. Johnson, R. S., Martin, S. A., Biemann, Klaus., Stults, J. T. & Watson, J. Throck. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* **59**, 2621–2625 (1987).
356. Roepstorff, P. & Fohlman, J. Letter to the editors. *Biol. Mass Spectrom.* **11**, 601–601 (1984).
357. Harrison, A. G. Characterization ofα- andγ-glutamyl dipeptides by negative ion collision-induced dissociation. *J. Mass Spectrom.* **39**, 136–144 (2004).
358. Kind, T. & Fiehn, O. Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **8**, 105 (2007).
359. Colsch, B., Fenaille, F., Warnet, A., Junot, C. & Tabet, J.-C. Mechanisms governing the fragmentation of glycerophospholipids containing choline and ethanolamine polar head groups. *Eur J Mass Spectrom (Chichester)* **23**, 427–444 (2017).
360. Colsch, B., Damont, A., Junot, C., Fenaille, F. & Tabet, J.-C. Experimental evidence that electrospray-produced sodiated lysophosphatidyl ester structures exist essentially as protonated salts. *Eur J Mass Spectrom (Chichester)* **25**, 333–338 (2019).
361. Tan, B. *et al.* Identification of endogenous acyl amino acids based on a targeted lipidomics approach. *Journal of Lipid Research* **51**, 112–119 (2010).
362. Moriarty, M. *et al.* Development of an LC-MS/MS method for the analysis of serotonin and related compounds in urine and the identification of a potential biomarker for attention deficit hyperactivity/hyperkinetic disorder. *Anal Bioanal Chem* **401**, 2481–2493 (2011).
363. Asakawa, D., Mizuno, H., Sugiyama, E. & Todoroki, K. Fragmentation study of tryptophan-derived metabolites induced by electrospray ionization mass spectrometry for highly sensitive analysis. *Analyst* **146**, 2292–2300 (2021).
364. Bennaceur, C., Afonso, C., Alves, S., Bossée, A. & Tabet, J.-C. Instrumental Dependent Dissociations of n-Propyl/Isopropyl Phosphonate Isomers: Evaluation of Resonant and Non-Resonant Vibrational Activations. *J. Am. Soc. Mass Spectrom.* **24**, 1260–1270 (2013).
365. Cheng, C. & Gross, M. L. Applications and mechanisms of charge-remote fragmentation. *Mass Spectrom Rev* **19**, 398–420 (2000).
366. Baba, T., Campbell, J. L., Le Blanc, J. C. Y. & Baker, P. R. S. Distinguishing Cis and Trans Isomers in Intact Complex Lipids Using Electron Impact Excitation of Ions from Organics Mass Spectrometry. *Anal. Chem.* **89**, 7307–7315 (2017).
367. Li, P., Hoffmann, W. D. & Jackson, G. P. Multistage mass spectrometry of phospholipids using collision-induced dissociation (CID) and metastable atom-activated dissociation (MAD). *International Journal of Mass Spectrometry* **403**, 1–7 (2016).

368. Afonso, C., Riu, A., Xu, Y., Fournier, F. & Tabet, J.-C. Structural characterization of fatty acids cationized with copper by electrospray ionization mass spectrometry under low-energy collision-induced dissociation. *J. Mass Spectrom.* **40**, 342–349 (2005).
369. Fay, L. & Richli, U. Location of double bonds in polyunsaturated fatty acids by gas chromatography-mass spectrometry after 4,4-dimethyloxazoline derivatization. *Journal of Chromatography A* **541**, 89–98 (1991).
370. McLuckey, S. A. *et al.* Energy-resolved tandem and fourier-transform mass spectrometry. *International Journal of Mass Spectrometry and Ion Physics* **44**, 215–229 (1982).
371. Loh, S. K., Hales, D. A., Lian, L. & Armentrout, P. B. Collision-induced dissociation of Fe^+_n (n =2–10) with Xe: Ionic and neutral iron binding energies. *The Journal of Chemical Physics* **90**, 5466–5485 (1989).
372. Armentrout, P. B. & Beauchamp, J. L. Experimental and theoretical studies of the reaction Ba+(D2, D)BaD+: sequential impulse model for endothermic reactions. *Chemical Physics* **48**, 315–320 (1980).
373. Wysocki, V. H., Kenttämaa, H. I. & Cooks, R. G. Internal energy distributions of isolated ions after activation by various methods. *International Journal of Mass Spectrometry and Ion Processes* **75**, 181–208 (1987).
374. Laskin, J. & Futrell, J. H. Collisional activation of peptide ions in FT-ICR mass spectrometry. *Mass Spectrom. Rev.* **22**, 158–181 (2003).
375. Carlesso, V., Fournier, F. & Tabet, J.-C. Stereochemical Differentiation of Four Monosaccharides Using Transition Metal Complexes by Electrospray Ionization/Ion-Trap Mass Spectrometry. *Eur J Mass Spectrom (Chichester)* **6**, 421–428 (2000).
376. Daikoku, S., Widmalm, G. & Kanie, O. Analysis of a series of isomeric oligosaccharides by energy-resolved mass spectrometry: a challenge on homobranched trisaccharides: Analysis of structural isomers of branched trisaccharides. *Rapid Commun. Mass Spectrom.* **23**, 3713–3719 (2009).
377. Menicatti, M. *et al.* The power of energy-resolved tandem mass spectrometry experiments for resolution of isomers: the case of drug plasma stability investigation of multidrug resistance inhibitors: Liquid chromatography/energy-resolved MS/MS for resolution of isomers. *Rapid Commun. Mass Spectrom.* **30**, 423–432 (2016).
378. Fedor, D. M. & Cooks, R. G. Angle resolved mass spectrometry with a reversed geometry spectrometer. *Anal. Chem.* **52**, 679–682 (1980).
379. Rosenstock, H. M., Wallenstein, M. B., Wahrhaftig, A. L. & Eyring, H. Absolute Rate Theory for Isolated Systems and the Mass Spectra of Polyatomic Molecules. *Proceedings of the National Academy of Sciences* **38**, 667–678 (1952).
380. Baercor, T. & Mayerfn, P. M. Statistical Rice-Ramsperger-Kassel-Marcus quasiequilibrium theory calculations in mass spectrometry. *J. Am. Soc. Mass Spectrom.* **8**, 103–115 (1997).
381. Pak, A., Lesage, D., Gimbert, Y., Vékey, K. & Tabet, J.-C. Internal energy distribution of peptides in electrospray ionization : ESI and collision-induced dissociation spectra calculation. *J. Mass Spectrom.* **43**, 447–455 (2008).
382. Drahos, Heeren, Collette, De Pauw E, & Vekey. Thermal energy distribution observed in electrospray ionization. *J Mass Spectrom* **34**, 1373–1379 (1999).