

UNIVERSITE DE STRASBOURG

THESE

Pour obtenir le grade de

Docteur de l'Université Louis Pasteur de Strasbourg

**Ecole doctorale n°269 Mathématiques, Sciences de l'Information et de l'Ingénieur
(MSII)**

Discipline : Traitement du signal et des Images

Présentée par

Clément HEUDE

Développement de nouvelles méthodes analytiques dans l'agroalimentaire par RMN

Présentée et soutenue le 14 Octobre 2015, devant le jury composé de :

Rapporteur	M. Douglas RUTLEDGE	Professeur – AgroPariTech
Rapporteur	M. François MESNARD	Professeur – Université de Picardie
Directeur de thèse	M. Karim ELBAYED	Maître de conférence – Université de Strasbourg
Directeur de thèse	M. Martial PIOTTO	Responsable des projets collaboratifs – Bruker BioSpin France

Remerciements

Ce travail a été réalisé, pour sa plus grande partie, au sein du laboratoire d'application RMN de la société Bruker BioSpin France dirigée par M. Alain Belguise. Je le remercie pour m'avoir accueilli au sein de son entreprise et d'avoir participé au financement de cette thèse. Cette thèse n'aurait pu se réaliser également sans le soutien financier de BpiFrance (ex Oséo).

Je remercie vivement Messieurs François Mesnard et Douglas Rutledge de m'avoir fait l'honneur d'accepter le travail de rapporteurs. J'adresse également mes remerciements à messieurs Bruno Kiefer et Dimitri Heintz pour avoir accepté de faire partie du jury de cette thèse.

Je remercie plus particulièrement Messieurs Karim Elbayed et Martial Piotto qui ont assuré la direction de cette thèse. Je remercie Monsieur Karim Elbayed pour les conseils et les discussions que nous avons eues ainsi que pour la confiance et la liberté qu'il a su m'accorder pour la réalisation des différents travaux de cette thèse. Je tiens à exprimer mes plus profonds remerciements à Monsieur Martial Piotto pour son encadrement, sa disponibilité, pour nos nombreuses discussions scientifiques et son soutien tout au long de cette thèse.

J'adresse mes remerciements à tous les ingénieurs du laboratoire d'application RMN de Bruker BioSpin France, à savoir, Madame Maryse Bourdonneau, Fabien Aussenac et Jérôme Coutant pour leur accueil et la bonne ambiance de travail qu'ils ont su faire régner pendant ces trois années. J'adresse un remerciement plus particulier à Monsieur Jérôme Coutant pour ses nombreux conseils dans l'utilisation du spectromètre et de certains logiciels.

Je remercie tous les stagiaires que j'ai eu la chance de co-encadrer au cours de cette thèse et qui ont grandement contribué aux manipulations. A ce titre je remercie Elise Lemasson, Mathieu Fanuel et Tangi Jezequel. Je remercie également PedroPablo Lopez « Pépé », pour sa contribution au développement des fonctions Matlab.

L'étude sur l'origine du caviar n'aurait pu se faire sans la participation et l'envoi d'échantillons authentiques. Je remercie ainsi vivement tous les producteurs de l'Aquitaine membres de la démarche IGP de m'avoir fourni ces échantillons. Je remercie particulièrement Monsieur Philippe Benoit, directeur d'exploitation de la société Sturgeon, pour toutes nos discussions et les précieuses informations au sujet de l'élevage de l'esturgeon qu'il m'a transmises.

Enfin, pour être sûr de n'oublier personne, je remercie chaleureusement toutes les personnes que j'ai pu côtoyées au sein de la société Bruker BioSpin durant ces trois années.

Sigles et Abréviations

ABVT : Azote Basique Volatil Total

AD : Analyse Discriminante

ADL : Analyse Discriminante Linéaire

ADQ : Analyse Discriminante Quadratique

ADN : Acide DésoxyriboNucléique

ACP : Analyse en Composantes Principales

ARTSI: Amplitude-corrected Referencing Through Signal Injection

ATP : Adénosine TriPhosphate

BCU : Bruker Cooling Unit

B_{eff} : Champ magnétique efficace

B_{loc} : Champ magnétique local

CAH : Classification Ascendante Hiérarchique

CLHP : Chromatographie Liquide Haute Performance

CPMG : Carr-Purcell-Meiboom-Gill

CPG : Chromatographie en Phase Gazeuse

DLUO : Date Limite d'Utilisation Optimale

DSS : Acide 4,4-diméthyl-4-silapentane-1-sulfonique

E_D : Energie dipolaire

ERETIC: Electronic REference To access In-vivo Concentration

IGP : Indication Géographique Protégée

INEPT: Insensitive Nuclei Enhanced by Polarization Transfer

FINS: Fractionnement Isotopique Naturel Spécifique

HR-MAS: High Resolution-Magic Angle Spinning

HSQC: Heteronuclear Single Quantum Correlation

Kel-F : Polychlorotrifluoroéthylène

m: moment magnétique

M₀: Aimantation à l'équilibre

M_z: Aimantation longitudinale

M_{x,y}: Aimantation transversale

NOE: Nuclear Overhauser Effect

NOESY: Nuclear Overhauser Effect Spectroscopy

OPLS-DA: Orthogonal Partial Least Square regression-Discriminant Analysis

OTMA-N : Oxyde de TriMéthylAmine

PME: Petite et Moyenne Entreprise

PLS-DA: Partial Least Square regression-Discriminant Analysis

PULCON: PULse Length-based CONcentration

RF: Radio Fréquence

RMN: Résonance Magnétique Nucléaire

RMSE : Root Mean Square Error

SIMCA: Soft Independent Modeling of Class Analogy

SM: Spectrométrie de Masse

SMRI: Spectroscopie de Masse des Rapports Isotopiques

SPE: Solid Phase Extraction

T1: Constante de temps associée à la relaxation longitudinale

T2: Constante de temps associée à la relaxation transversale

τ_c : Temps de corrélation

TMA-N : TriMéthylAmine

TMS: TétraMethylSilane

TOCSY: TOtal Correlation Spectroscopy

TSP: TriméthyleSilyl Propionate deutéré

Sommaire

Introduction.....	1
Partie I- Principes de la RMN.....	4
I) RMN : Théorie et concepts.....	4
I.1) Niveaux d'énergie et Hamiltonien.....	4
I.2) Relaxation.....	7
I.2.1) La relaxation longitudinale ou relaxation spin-réseau.....	7
I.2.2) La relaxation transversale ou relaxation spin-spin.....	8
II) Principes de la RMN HR-MAS.....	9
II.1) Les phénomènes contribuant à l'élargissement des raies RMN.....	10
II.1.1) L'interaction Zeeman.....	10
II.1.2) L'interaction dipolaire.....	11
II.1.3) Interaction électron-noyau : l'anisotropie de déplacement chimique.....	15
II.1.4) Les différences de susceptibilités magnétiques (milieu anisotrope).....	17
II.1.5) Dynamique.....	19
II.2) La rotation à l'angle magique (MAS).....	20
Partie II- Métabolomique.....	23
I) Introduction à la métabolomique.....	23
II) La RMN en métabolomique.....	24
II.1) La RMN en métabolomique.....	24
II.2) Les principales applications de la RMN dans l'agroalimentaire.....	36
II.2.1) La RMN HR-MAS dans l'agroalimentaire.....	36
II.2.2) La RMN liquide dans l'agroalimentaire.....	39
III) Méthodes Statistiques multivariées – Chimiométrie.....	41
III.1) Introduction.....	41
III.2) L'Analyse en Composantes Principales (ACP).....	42
III.2.1) Approche géométrique.....	43
III.2.2) Approche algébrique.....	47
III.3) Soft Independent Modelling of Class Analogy (SIMCA).....	48
III.4) La régression PLS discriminante (Partial Least Square regression-Discriminant Analysis) ...	53
III.5) Prétraitement des données RMN.....	59
III.6) Sélection du nombre de composantes et validation croisée.....	61
Partie III- Matériels et Méthodes.....	66
I) Matériels.....	66

I.1) Le spectromètre RMN	66
I-2) HR-MAS	66
I.2.1) Instrumentation et spécificités du matériel en HR-MAS	66
I.2.2) Protocole préparation d'échantillons HR-MAS	72
I-3) RMN liquide	73
I.3.1) Instrumentation	73
I.3.2) Protocole préparation d'échantillons de caviar	73
II) Séquences utilisées	74
II.1) Expériences 1D	74
II.1.1) ZGPR	74
II.1.2) CPMGPR1D et CPMG pseudo-2D	75
II.1.3) NOESYGPPR1D.....	77
II.1.4) Inversion-Récupération.....	79
II.2) Expériences 2D	80
II.2.1)TOCSY ET ATOCSY.....	80
II.2.2)HSQC.....	83
Partie IV- Résultats :	87
I) Etudes de matrices semi-solides par RMN HR-MAS.....	87
I.1) Méthode de détermination de la fraîcheur du poisson (Publication 1-Rapid assessment of fish freshness and quality by ¹ H HR-MAS NMR Spectroscopy-Annexe 2).....	87
I.1.1) Introduction.....	87
I.1.2) Détermination de la fraîcheur du poisson par RMN HR-MAS.....	91
I.1.3) Conclusion	102
I.2) Analyse de blocs de farce de colin d'Alaska	103
I.2.1) Détermination de la fraîcheur des blocs de farce de colin d'Alaska : la valeur K	104
I.2.2) Analyses statistiques	106
I.2.3) Conclusion	112
I.3) Etude de la texture du poisson : Temps de relaxation transversale (T2) de l'eau	113
I.3.1) Etude de la texture des blocs de farce de colin d'Alaska	113
I.3.2) Etude de la texture des bâtonnets de poisson (Croustibat)	122
I.3.3) Conclusion	125
I.4) Quantification T1/T2 de biopsies	126
I.4.1) Introduction.....	126
I.4.2) Développement méthodologique : Quantification de l'arginine au sein d'une biopsie de poisson	127

I.4.3) Biopsies de cerveau.....	130
I.4.3) Conclusion	137
I.5) Conclusion générale des travaux réalisés par HR-MAS.....	138
II) Etude de matrices liquides par RMN.....	139
II.1) Introduction	139
II.2) Etude de l'origine du caviar par RMN (Publication 2-Metabolic characterization of caviar specimens by ¹ H NMR Spectroscopy: Towards caviar authenticity and integrity- Annexe 3).....	141
II.2.1) Etude préliminaire par RMN HR-MAS	141
II.2.2) Mise en place du plan d'échantillonnage.....	144
II.2.3) Analyses statistiques	146
II.2.4) Conclusion	171
II.3) Etude de la dégradation du caviar au cours de sa conservation par RMN	172
II.3.1) Etude eppendorf	173
II.3.2) Etude longue sur des boîtes commerciales.....	178
II.3.3) Analyses microbiologiques.....	180
II.3.4) Conclusion	182
Conclusion et Perspectives	183
Liste des Tableaux	186
Liste des Figures	188
Liste des publications et communications orales.....	194
Publications	194
Communication orale.....	194
Posters.....	194
References.....	195
Annexes	209

Introduction

Dans le domaine de l'agroalimentaire, le volume des échanges commerciaux mondiaux a triplé durant les vingt dernières années. L'élargissement de l'espace Schengen, combiné à la mondialisation, favorise la provenance de matières premières et d'ingrédients issus de pays de plus en plus lointains dont les structures d'analyse et de contrôle sont moindres voire inexistantes. Dans ce contexte, émerge la question de la sécurité alimentaire. En effet, l'industrie agroalimentaire s'expose ainsi à des risques de contamination ou de fraudes, pouvant aller jusqu'à mettre en danger la santé du consommateur comme ce fut le cas de la présence de mélanine dans la poudre de lait qui a provoqué le décès de nourrissons en Chine.

De nombreuses techniques analytiques existent actuellement pour la détection de contaminants ou pour la caractérisation de l'authenticité des produits alimentaires, telles que la Chromatographie Liquide Haute Performance (CLHP), la Chromatographie en Phase Gazeuse (CPG), la spectroscopie infrarouge, la spectroscopie Raman, l'analyse de l'ADN, les méthodes immunologiques et électrophorétiques, les analyses microbiologiques, la Spectrométrie de Masse (SM) et la Résonance Magnétique Nucléaire (RMN). De plus, certaines technologies dites de « couplage » sont apparues et sont couramment utilisées couplant certaines de ces techniques telles que la GPG-SM, la CLHP-SM et la Chromatographie Liquide-RMN permettant d'obtenir des informations plus précises ou complémentaires. Des méthodes mesurant les rapports isotopiques de différents noyaux existent également : la Spectroscopie de Masse des Rapports Isotopiques (SMRI) et le Fractionnement Isotopique Naturel Spécifique par RMN (FINS-RMN). Ces techniques sont capables de déceler des falsifications entre des matériaux chimiquement similaires mais provenant d'origines différentes, ou de détecter des frelatages à l'état de traces.

Cependant, la majorité des méthodes d'analyses et de contrôle actuelles sont basées sur une approche ciblée, c'est-à-dire avec une définition en amont des contaminants recherchés, et présentent ainsi le risque de ne pas détecter certaines sources potentielles de contaminations ou d'altération de produits authentiques. C'est autour de cette problématique que le projet Agrifood GPS (Global Protection System) a vu le jour. Ce projet regroupe deux grandes entreprises industrielles (Eurofins Scientific et Bruker BioSpin), quatre instituts publics (GEPEA, IAQA, CEA LIST et IMMM) et une PME (Profilomic), et a pour objectif principal de mettre en place de nouvelles méthodes analytiques non ciblées garantissant au client ou à l'utilisateur la sécurité des produits analysés en termes d'absence de contaminants chimiques

ou microbiens et d'authenticité. Les produits en question concernent les aliments pour humains ou animaux ainsi que de l'eau agroalimentaire ou environnementale.

Dans ce projet, le laboratoire d'application de la société Bruker BioSpin basé à Wissembourg, est chargé du développement de nouvelles techniques analytiques de criblage principalement sur des échantillons de type viande ou poisson par RMN et plus particulièrement par RMN HR-MAS (High Resolution-Magic Angle Spinning) avec pour principaux axes de recherche sur ces matrices l'étude de fraudes relatives à l'authenticité et à la fraîcheur. L'authenticité d'un produit alimentaire est définie par des caractéristiques bien précises (origine géographique, procédés ou savoir-faire de fabrication, conformité à l'espèce ou à la variété, méthode de culture, non adultération ...). La notion de fraude alimentaire est employée dès lors que certains produits alimentaires sont intentionnellement commercialisés avec tromperie du consommateur. La raison des fraudes alimentaires sur l'authenticité est le plus souvent économique, dans le but de vendre un produit à plus faible valeur commerciale pour un produit à plus forte valeur ajoutée ou en remplaçant de manière illicite certains produits par des composés de qualité moindre.

La RMN est une technique spectroscopique non destructive, rapide et très reproductible permettant l'acquisition de spectres contenant une quantité d'information importante sur l'échantillon analysé. En effet, cette technique spectroscopique permet l'identification des composants majeurs et mineurs dans un mélange complexe en une seule analyse sans prédéfinition des composés recherchés et est donc totalement adaptée à la mise au point de méthodes analytiques non ciblées. Plusieurs études ont utilisé la RMN liquide Haute Résolution comme technique de criblage ou profiling combiné à des analyses statistiques multivariées dans le but de vérifier l'authenticité de produits alimentaires (vins, jus de fruits, huiles alimentaires...).^{1,2}

Toutefois, certains produits alimentaires présentent une consistance qualifiée de semi-solide, à l'interface entre l'état liquide et l'état solide, rendant leur analyse directe par RMN liquide Haute Résolution impossible du fait de leur nature hétérogène. En effet, la non-uniformité de susceptibilité magnétique de ce type d'échantillon va conduire à un élargissement de raie significatif sur le spectre RMN. Pour étudier cette famille de composés, une technique RMN a été développée au milieu des années 1990 : la RMN Haute Résolution en rotation à l'angle magique (HR-MAS). La mise en rotation de l'échantillon à cet angle va permettre de moyennner à zéro la plupart des interactions anisotropes, améliorant sensiblement la résolution et rendant ainsi possible l'utilisation des séquences employées en RMN liquide.

L'objectif de cette thèse est par conséquent de développer des nouvelles méthodes analytiques dans l'agroalimentaire principalement par RMN HR-MAS, l'étude des échantillons pouvant être analysés par RMN liquide étant essentiellement confiée à Eurofins Scientific Nantes. En effet, les offres analytiques doivent constamment être modifiées et améliorées pour suivre le rythme des fraudes. Une fois qu'une méthode est connue pour détecter une certaine forme d'adultération, les fraudeurs s'adaptent pour contourner sa détection.

Dans la première partie (**Partie I**), nous nous intéresserons aux principes de la RMN et plus particulièrement à ceux de la RMN HR-MAS. Nous présenterons notamment les différentes interactions présentes dans les échantillons semi-solides et les phénomènes responsables de l'élargissement des raies RMN, ainsi que la solution qu'apporte la rotation à l'angle magique.

Puis dans une seconde partie (**Partie II**), dédiée à la métabolomique, nous réaliserons, tout d'abord, une présentation de l'approche utilisée en RMN pour les études métabolomiques ainsi qu'une bibliographie des applications de la RMN dans l'agroalimentaire. Nous introduirons également le principe des différentes méthodes statistiques multivariées (ACP, SIMCA, PLS-DA et OPLS-DA), ainsi que les différents traitements préalables à ces analyses statistiques, que nous utiliserons par la suite pour l'étude des données RMN.

Ensuite dans une troisième partie expérimentale (**Partie III**), nous présenterons succinctement le matériel RMN, notamment les spécificités du matériel HR-MAS pour l'obtention d'un spectre RMN Haute Résolution, ainsi que les différents protocoles mis au point pour la préparation des échantillons en vue de leur analyse. Nous verrons également les différentes séquences d'impulsion et paramètres utilisés pour l'acquisition des spectres RMN.

Enfin, la **Partie IV** sera entièrement consacrée à la présentation des différents résultats obtenus sur les matrices semi-solides et liquides. Cette partie regroupera ainsi, tout d'abord, les résultats obtenus par RMN HR-MAS sur la détermination de la fraîcheur du poisson ainsi que ceux acquis lors de l'étude menée en partenariat avec la société Findus France. Dans un second temps, les résultats des travaux réalisés sur la détermination de l'origine du caviar par RMN liquide dans le cadre de la démarche IGP (Indication Géographique Protégée) des producteurs français du bordelais, et sur l'étude de la dégradation du caviar au cours de sa conservation au réfrigérateur, seront alors présentés.

Partie I- Principes de la RMN

I) RMN : Théorie et concepts

I.1) Niveaux d'énergie et Hamiltonien

Le noyau d'un atome de spin non nul peut être assimilé, en première approximation, à une particule sphérique chargée positivement tournant autour d'un axe, auquel est associé un moment cinétique nucléaire \vec{I} . La rotation de ces charges va ainsi induire un petit champ magnétique appelé moment magnétique nucléaire noté $\vec{\mu}$ colinéaire et proportionnel à \vec{I} défini par :

$$\vec{\mu} = \gamma \cdot \hbar \cdot \vec{I} \quad (1)$$

Où γ est le rapport gyromagnétique caractéristique de chaque isotope et \hbar correspond à la constante de Planck (h) divisée par 2π .

Le spin nucléaire étant quantifié, sa projection sur l'axe z du champ magnétique, notée m , ne pourra prendre que les $(2I+1)$ valeurs comprises entre $-I$ et $+I$ par pas entier de 1, où I est le nombre quantique de spin.

Si l'on considère l'exemple du proton ^1H ($I=1/2$), placé dans un champ magnétique homogène, celui-ci possède deux orientations possibles : $m=1/2$ appelé état α et $m=-1/2$ appelé état β .

En l'absence de champ magnétique extérieur, tous les états de spin possèdent la même énergie, ils sont dits dégénérés. En revanche, en présence d'un champ magnétique \vec{B}_0 , une interaction se crée entre ce champ et le moment magnétique du noyau $\vec{\mu}$, et l'énergie de cette interaction est définie par :

$$E_m = -\vec{\mu} \cdot \vec{B}_0 = -m \cdot \gamma \cdot \hbar \cdot B_0 \quad (2)$$

Cette énergie dépend donc de l'orientation du spin (Figure 1), conduisant à l'apparition de deux niveaux d'énergie pour un spin 1/2 (correspondant pour le cas du proton aux états α ($m=1/2$) et β ($m=-1/2$)) et dont la différence d'énergie entre ces deux niveaux d'énergie consécutifs ΔE vaut :

$$\Delta E = \gamma \cdot \hbar \cdot B_0 \quad (3)$$

La dégénérescence est ainsi levée et l'énergie nécessaire $h\nu$ pour provoquer une transition entre les deux niveaux d'énergie est donnée par :

$$h\nu = \gamma \cdot \hbar \cdot B_0 \quad (4)$$

soit :

$$\nu = \frac{\gamma \cdot B_0}{2\pi} \quad (5)$$

Où ν est appelée fréquence de Larmor. L'absorption de cette quantité d'énergie induit un changement d'état du moment magnétique et est réalisée sur un spectromètre RMN par l'application pendant un temps très court (quelques μs) d'un champ radio fréquence (RF).

Les deux niveaux d'énergie sont peuplés en accord avec la statistique de Boltzmann et pour un spin 1/2 :

$$\frac{N_\beta}{N_\alpha} = e^{-\frac{\Delta E}{k.T}} \quad (6)$$

Où N_α et N_β représente le nombre de noyaux qui se trouvent à l'équilibre dans l'état de spin α ou β (Figure 1). L'état de plus faible énergie est le plus peuplé et la différence de population est directement reliée à la valeur de B_0 .

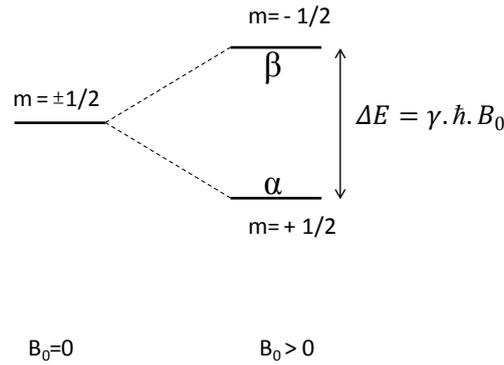


Figure 1 : Niveaux d'énergie des spins 1/2 en l'absence et en présence de champ magnétique B_0 . La présence d'un champ magnétique conduit à l'apparition de deux niveaux d'énergie pour les spins 1/2 dont la différence d'énergie vaut ΔE . Ces niveaux d'énergie sont peuplés en accord avec la statistique de Boltzmann.

Un système de spins nucléaires placé dans un champ magnétique subit des perturbations résultant de différentes interactions. Il est courant de distinguer l'interaction principale liée à la présence du champ magnétique (interaction Zeeman) des interactions internes, c'est-à-dire propres au système de spin (interactions dipolaire, d'anisotropie de déplacement chimique, quadripolaire et de couplage scalaire).

Pour un système de spins nucléaires donné, l'ensemble des Hamiltoniens à considérer sont regroupés au sein de l'Hamiltonien total (\widehat{H}) :

$$\widehat{H} = \widehat{H}_Z + \widehat{H}_D + \widehat{H}_{CS} + \widehat{H}_Q + \widehat{H}_J \quad (7)$$

Avec \widehat{H}_Z l'Hamiltonien Zeeman, \widehat{H}_D l'Hamiltonien dipolaire, \widehat{H}_{CS} l'Hamiltonien d'anisotropie de déplacement chimique, \widehat{H}_Q l'Hamiltonien quadripolaire et \widehat{H}_J l'Hamiltonien de couplage scalaire.

Pour un spin 1/2, l'interaction quadripolaire est nulle. L'importance relative des différents Hamiltoniens caractérisant les interactions internes (\widehat{H}_Z , \widehat{H}_D , \widehat{H}_{CS} et \widehat{H}_J) est intimement liée à la nature de l'échantillon étudié (solide, liquide ou semi-solide). En effet, dans le cas d'échantillons liquides, qualifiés en RMN de milieu isotrope, \widehat{H}_D est moyennée à zéro par le mouvement moléculaire rapide (mouvement brownien). En revanche, pour le cas des milieux anisotropes (solides ou semi-solides), certaines de ces interactions influencent de manière significative la largeur de raie des différents signaux du spectre RMN comme nous le verrons plus tard dans cette partie.

I.2) Relaxation

La relaxation est le nom donné au processus décrivant comment l'aimantation RMN (ensemble des spins nucléaires) retourne à sa valeur d'équilibre après avoir subi une perturbation. Pour les spins 1/2, il existe plusieurs sources de mécanismes de relaxation cependant trois principales sources se distinguent : le mécanisme dipolaire, le mécanisme d'anisotropie de déplacement chimique et le mécanisme d'échange chimique. Les composantes longitudinale (M_z) et transversale ($M_{x,y}$) sont affectées différemment par ce phénomène de relaxation comme nous allons le voir brièvement.

I.2.1) La relaxation longitudinale ou relaxation spin-réseau

Le retour de l'aimantation longitudinale (M_z) à sa valeur d'équilibre est appelée relaxation longitudinale ou relaxation spin-réseau. Les mécanismes de relaxation longitudinale sont liés aux fluctuations des champs magnétiques locaux à la fréquence de Larmor. En effet, les composantes transversales de ces champs magnétiques locaux ont la capacité d'induire des transitions entre états quantiques du système de spins, avec une probabilité proportionnelle à l'amplitude des fluctuations, et ainsi ramener le système de spins vers son état d'équilibre.

La théorie de Bloch³ stipule que la récupération de M_z suit un comportement exponentiel :

$$\frac{dM_z}{dt} = \frac{(M_0 - M_z)}{T_1} \quad (8)$$

Où M_0 représente l'aimantation à l'équilibre, et T_1 la constante de temps associée à la relaxation longitudinale.

L'aimantation longitudinale suit ainsi un comportement mono-exponentiel et peut s'écrire au temps t :

$$M_z = M_0 (1 - e^{-t/T_1}) \quad (9)$$

I.2.2) La relaxation transversale ou relaxation spin-spin

La décroissance de l'aimantation transversale ($M_{x,y}$) dans le plan xOy (plan transversal orthogonal au champ \vec{B}_0) jusqu'à atteindre une valeur nulle est appelée relaxation transversale ou relaxation spin-spin. Les mécanismes de relaxation transversale sont, quant à eux, composés de deux contributions : l'une appelée non séculaire et correspondant à l'interaction avec les composantes transversales des champs magnétiques locaux décrite précédemment, et une contribution séculaire correspondant à l'interaction avec les composantes longitudinales des champs magnétiques locaux (interaction entre dipôles magnétiques). Cette dernière interaction avec les champs magnétiques locaux va modifier très légèrement la fréquence de précession des spins : les spins ressentant un champ magnétique plus intense auront une précession plus rapide et ceux ressentant un champ magnétique plus faible auront une précession plus lente. La précession des moments individuels va ainsi être déphasée (perte de la cohérence de phase) et l'aimantation va ainsi diminuer.

Plus le mouvement moléculaire est important (cas des liquides), plus la fréquence de précession sera moyennée et donc moins l'envergure des champs magnétiques locaux sera importante, et ainsi plus les spins relaxeront lentement.

La décroissance de l'aimantation dans le plan transversal peut être décrite par une fonction mono-exponentielle caractérisée par la constante de temps des processus de relaxation naturels, notée T_2 . Cependant, la décroissance de $M_{x,y}$ est rarement uniquement due à des processus naturels et une partie de cette décroissance provient d'inhomogénéités du champ magnétique le long du volume de l'échantillon introduisant lui aussi un déphasage entre les moments magnétiques.

Le temps de relaxation apparent noté T_2^* a ainsi été introduit et décrit donc la somme de ces deux processus et peut être exprimé selon l'équation suivante :

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2 \text{ inhom}}} \quad (10)$$

Quelque soit le système de spins étudiées, $T_2 \leq T_1$, et pour les protons (^1H), la constante de temps de relaxation transversale (T_2) peut aller de la microseconde (μs) dans les solides, à quelques millisecondes (ms) pour les lipides voire de l'ordre de la seconde (s) pour les petites

molécules en solution, tandis que la constante de temps de relaxation longitudinale (T_1) est généralement de plusieurs secondes en RMN liquide.

Enfin la constante de temps de relaxation transversale (T_2) est également reliée à la largeur de raie à mi-hauteur ($\nu_{1/2}$) par la relation :

$$\nu_{1/2} = \frac{1}{\pi T_2} \quad (11)$$

Cependant, la largeur de raie sur le spectre RMN est également influencée par les inhomogénéités du champ magnétique principal qui comme nous l'avons vu précédemment vont introduire une dispersion de fréquence.

II) Principes de la RMN HR-MAS

Nous allons à présent nous intéresser à une technique particulière de la RMN, adaptée de la RMN du solide et permettant l'analyse de composés semi-solides : la RMN HR-MAS. Cette technique est relativement récente, puisque les premiers travaux publiés utilisant cette technique remontent à 1994⁴ et concernaient alors la caractérisation de composés organiques obtenus par synthèse en phase solide (milieu hétérogène). Cheng et al⁵ révélèrent en 1997 un nouveau champ d'application de cette technique, en étudiant pour la première fois des tissus biologiques par RMN HR-MAS. Les différents domaines d'application de cette technique n'ont, depuis, cessé de croître et regroupe à présent la biologie^{6,7}, la chimie organique^{8,9} ou inorganique, l'étude de polymères^{10,11} et la médecine^{12,13}. Ces différents travaux de recherche ont pu bénéficier depuis 1996, de la première sonde dédiée aux analyses par HR-MAS mise sur le marché par Bruker BioSpin, des avancées technologiques sur ces systèmes (régulation de température, amélioration des gradients, amélioration des bobines de détection) contribuant ainsi au développement de cette technique.

La largeur de raie intrinsèque des signaux RMN, de l'ordre de plusieurs centaines de hertz (Hz) pour un spectre proton, a pendant longtemps rendu l'étude de composés semi-solides chimérique. Nous allons à présent exposer les phénomènes contribuant à l'élargissement des raies et comment la technique HR-MAS permet de s'affranchir de ces problèmes.

II.1) Les phénomènes contribuant à l'élargissement des raies RMN

Dans cette section, les différents paramètres influençant la largeur de raie des composés semi-solides étudiés par HR-MAS sont présentés. La rédaction de cette section reprend les équations et les explications des travaux de Lippens & Piotto⁹. Certaines interactions sont bien connues en RMN du solide, cependant leur importance en HR-MAS est différente de celle rencontrée lors de l'analyse d'échantillons solides.

II.1.1) L'interaction Zeeman

Pour les noyaux de spin 1/2 (¹H, ¹³C, ¹⁵N, ³¹P), l'interaction Zeeman est la plus importante des interactions. Cependant, celle-ci ne contient aucune information structurale, elle fournit simplement un axe de quantification pour les spins.

Lorsqu'un spin I est placé dans un champ magnétique externe (B₀), il va interagir avec celui-ci à travers l'interaction Zeeman comme nous l'avons vu dans la section **I-1**). En effet, un spin peut être considéré comme un petit dipôle magnétique de moment magnétique μ qui précesse à la fréquence de Larmor (ω₀) autour du champ magnétique B₀. L'énergie de l'interaction (E_m) de m avec B₀ est donnée par :

$$E_m = -\vec{\mu} \cdot \vec{B}_0 = -m \cdot \gamma \cdot \hbar \cdot B_0 \quad (12)$$

En mécanique quantique, le moment magnétique nucléaire d'un noyau j est défini par :

$$m_j = \gamma_j \hbar I_j \quad (13)$$

Où γ_j représente le rapport gyromagnétique du noyau j, ħ correspond à la constante de Planck (h) divisée par 2π et I_j l'opérateur de spin du moment angulaire.

L'énergie de l'interaction entre m_j et le champ magnétique B₀, orienté suivant l'axe z, est donnée par l'Hamiltonien Zeeman :

$$\widehat{H}_z = -m_j \cdot B_0 = -\gamma \hbar B_0 \widehat{I}_z \quad (14)$$

Seules les interactions qui modifient localement le champ B_0 ont une importance pour l'élucidation structurale. Ces interactions, qui varient du Hz au kHz, sont considérées comme des perturbations au premier ordre de l'interaction Zeeman.

II.1.2) L'interaction dipolaire

L'interaction dipolaire ou interaction dipôle-dipôle, provient de l'interaction intermoléculaire ou intramoléculaire entre les moments magnétiques (également appelés dipôles magnétiques) de deux spins voisins. Cette interaction est l'une des plus importantes en RMN puisqu'elle permet notamment d'obtenir des informations sur les distances internucléaires et donc sur la géométrie de la molécule. En effet, dans un échantillon solide et dans les cas favorables, la constante de couplage dipolaire peut être directement mesurée sur le spectre. En RMN du liquide, cette information n'est pas accessible à cause du mouvement moléculaire qui moyenne cette interaction. Cependant elle est impliquée dans le processus de relaxation croisée qui gouverne l'effet Overhauser nucléaire (NOE).

La précession des spins autour du champ magnétique va modifier localement le champ magnétique B_0 en créant localement une composante tournante dans le plan perpendiculaire à B_0 (la projection de m sur le plan (x,y)) et une composante magnétique statique le long de B_0 (la projection de m le long de l'axe z). Seule cette dernière modifie le champ magnétique B_0 de manière significative¹⁴. En fonction de l'orientation du moment magnétique (m_α ou m_β) la composante statique sera dans le même sens ou opposée à B_0 .

Le champ magnétique local statique $B_j(r)$ généré à une distance r par un noyau j de moment magnétique m_j est donné par :

$$B_j(r) = \frac{\mu_0}{4\pi} \cdot \frac{[3(r \cdot m_j)r - r^2 m_j]}{r^5} \quad (15)$$

Ainsi, le champ magnétique dipolaire étant dépendant du terme $1/r^3$, il n'atteindra une valeur nulle que pour une distance infinie.

Puisque l'énergie dipolaire (E_D) est définie par $E_D = -m_k \cdot B_j(r)$, l'énergie dipolaire ressentie par un noyau k de moment magnétique m_k à une distance r d'un noyau j est donnée par :

$$E_D = \frac{\mu_0}{4\pi} \cdot \frac{[r^2(m_j \cdot m_k) - 3(r \cdot m_j)(r \cdot m_k)]}{r^5} \quad (16)$$

Cependant, seule la projection des moments magnétiques m_j et m_k sur l'axe z contribue au champ magnétique local. Par conséquent, en exprimant les moments magnétiques m_j et m_k en coordonnées polaires et en ne conservant que la composante le long de l'axe z l'équation précédente devient :

$$E_D = \frac{\mu_0}{4\pi} \cdot \frac{(m_{jz}m_{kz})}{r^3} \cdot (1 - 3 \cos^2 \theta) \quad (17)$$

Où θ représente l'angle entre le vecteur internucléaire et B_0 et m_{jz} et m_{kz} les composantes le long de l'axe z des moments magnétiques.

Pour que la composante statique le long de z soit nulle, et qu'ainsi l'énergie dipolaire soit annulée, il faut que l'orientation de j et k soit telle que le champ dipolaire ressenti par k soit perpendiculaire à m_k , c'est-à-dire lorsque $B_j(r)$ n'a pas de composantes le long de z .

L'équation précédente (17) indique que l'énergie dipolaire ne dépend que des deux paramètres r et θ et que celle-ci ne disparaît que lorsque le terme $(1 - 3 \cos^2 \theta)$ devient nul.

En se basant sur l'équation (16), il est possible d'écrire l'Hamiltonien dipolaire en substituant m_j et m_k par leur expression quantique vu précédemment :

$$H_D = \frac{\mu_0}{4\pi} \cdot \gamma_j \gamma_k \hbar \frac{[r^2(I_j I_k) - 3(r \cdot I_j)(r \cdot I_k)]}{r^5} \quad (18)$$

En développant les différents produits scalaires l'équation précédente devient⁹ :

$$H_D = \frac{\mu_0}{4\pi} \cdot \gamma_j \gamma_k \hbar (I_{jx} I_{jy} I_{jz}) \begin{pmatrix} 1 - \frac{3x^2}{r^2} & \frac{3xy}{r^2} & -\frac{3zx}{r^2} \\ -\frac{3xy}{r^2} & 1 - \frac{3y^2}{r^2} & -\frac{3yz}{r^2} \\ -\frac{3zx}{r^2} & \frac{3yz}{r^2} & 1 - \frac{3z^2}{r^2} \end{pmatrix} \begin{pmatrix} I_{kx} \\ I_{ky} \\ I_{kz} \end{pmatrix} \quad (19)$$

En mécanique quantique, il est courant d'exprimer l'Hamiltonien en coordonnées sphériques c'est-à-dire en utilisant les relations par rapport aux coordonnées cartésiennes $x = r \cdot \sin \theta \cos \phi$, $y = r \cdot \sin \theta \sin \phi$ et $z = r \cdot \cos \theta$ et en introduisant les opérateurs dits de « montée » $I_+ = I_x + iI_y$ et de « descente » $I_- = I_x - iI_y$, l'Hamiltonien devient :

$$H_D = \omega_D (A + B + C + D + E + F) \quad (20)$$

Avec

$$\omega_D = \frac{\mu_0 \gamma_j \gamma_k \hbar}{4\pi r_{jk}^3} \quad (21)$$

Et

$$A = I_{jz} I_{kz} (1 - 3 \cos^2 \theta) \quad (22)$$

$$B = -\frac{1}{4} (I_{j+} I_{k-} + I_{j-} I_{k+}) (1 - 3 \cos^2 \theta) \quad (23)$$

$$C = -\frac{3}{2} (I_{j+} I_{kz} + I_{jz} I_{k+}) \sin \theta \cos \theta e^{-i\phi} \quad (24)$$

$$D = -\frac{3}{2} (I_{j-} I_{kz} + I_{jz} I_{k-}) \sin \theta \cos \theta e^{+i\phi} \quad (25)$$

$$E = -\frac{3}{4} I_{j+} I_{k+} \sin^2 \theta e^{-2i\phi} \quad (26)$$

$$F = -\frac{3}{4} I_{j-} I_{k-} \sin^2 \theta e^{+2i\phi} \quad (27)$$

Parmi les termes de l'équation (20) seuls A et B commutent avec l'Hamiltonien Zeeman, ils sont donc les seuls à contribuer à l'énergie dipolaire et donc à la largeur de raie. L'expression de l'Hamiltonien dipolaire peut ainsi se simplifier :

$$\begin{aligned}
 H_D &= \omega_D \left[I_{jz} I_{kz} - \frac{1}{4} (I_{j+} I_{k-} + I_{j-} I_{k+}) \right] (1 - 3 \cos^2 \theta) \\
 &= \omega_D (3I_{jz} I_{kz} - I_j I_k) \cdot \frac{1}{2} (1 - 3 \cos^2 \theta)
 \end{aligned} \tag{28}$$

L'équation (28) représente l'Hamiltonien dipolaire dans le cas d'un système homonucléaire. Dans le cas d'un système hétéronucléaire (IS) celui-ci peut être encore plus simplifié du fait du couplage faible entre les deux spins et est défini par :

$$H_{D(hetero)} = \omega_D \cdot 2I_z S_z \cdot \frac{1}{2} (1 - 3 \cos^2 \theta) \tag{29}$$

Les équations (28) et (29) montrent que dans les deux cas, l'Hamiltonien dipolaire est composé d'une partie liée au spin pouvant être manipulé par des séquences d'impulsions et d'une partie spatiale $(1 - 3 \cos^2 \theta)$ dépendante de l'orientation des spins par rapport au champ magnétique qui peut être moyenné à zéro par la rotation de l'échantillon à l'angle magique.

Au sein d'un échantillon solide composé d'une grande quantité de spin, chaque spin nucléaire ressent le champ dipolaire des spins voisins (homonucléaire ou hétéronucléaire) et l'Hamiltonien dipolaire est en fait la somme de toutes ces interactions. Dans ce cas, chaque vecteur internucléaire r possède une amplitude et une orientation différente (distribution en θ et ω_D) conduisant à une large distribution des fréquences de résonance provoquant un élargissement important des pics sur le spectre RMN.

En revanche, pour le cas des échantillons liquides, où le mouvement moléculaire est important, le terme $(1 - 3 \cos^2 \theta)$ est moyennée et l'interaction dipolaire n'affecte pas les fréquences de résonance, cependant la fluctuation des champs magnétiques dipolaires (à la fréquence de Larmor) va induire des transitions entre certains niveaux d'énergie et est à l'origine de la relaxation dipolaire.

Pour ce qui est des échantillons semi-solides, les propriétés intrinsèques de ces matrices (à mobilité intermédiaire entre les liquides et les solides) combinées à l'ajout de solvant qui amène un plus grand degré de mobilité des molécules, réduisent sensiblement les interactions dipolaires présentes au sein de l'échantillon et donc le phénomène d'élargissement de raie. L'influence des interactions dipolaires en RMN HR-MAS est alors beaucoup moins notable que dans le cas de la RMN du solide.

II.1.3) Interaction électron-noyau : l'anisotropie de déplacement chimique

Sous l'influence du champ magnétique B_0 , le nuage électronique peut générer un champ magnétique local. C'est ce champ magnétique qui est à l'origine d'un paramètre RMN important à savoir le déplacement chimique. Le spin subit ainsi l'effet conjugué du champ magnétique B_0 et celui du champ magnétique induit B_{induit} provoqué par la circulation des électrons dans le voisinage du noyau.

L'Hamiltonien de déplacement chimique pour un noyau j est défini par :

$$\widehat{H}_{CS} = \gamma_j \cdot \widehat{I}_j \cdot \overline{\sigma}_j \cdot B_0 \quad (30)$$

ou $\overline{\sigma}_j$ représente le tenseur de déplacement chimique. Le champ magnétique B_0 étant orienté suivant l'axe z , il peut être réécrit :

$$\widehat{H}_{CS} = \gamma_j \cdot (\widehat{I}_x \sigma_{xz} + \widehat{I}_y \sigma_{yz} + \widehat{I}_z \sigma_{zz}) B_0 \quad (31)$$

Puisque seuls les termes qui commutent avec l'Hamiltonien Zeeman sont à considérer, il est possible d'écrire l'Hamiltonien de déplacement chimique réduit :

$$\widehat{H}_{CS} = \gamma_j \cdot \sigma_{zz} \cdot B_0 \widehat{I}_z \quad (32)$$

En exprimant l'élément σ_{zz} du tenseur dans un nouveau repère, établi de telle sorte que le tenseur $\bar{\sigma}$ soit diagonal dans ce système d'axes, appelé Système d'Axes Principaux (SAP) donné en Figure 2, σ_{zz} devient :

$$\sigma_{zz} = \sigma_{xx}^{SAP} (\cos \phi \sin \theta)^2 + \sigma_{yy}^{SAP} (\sin \phi \sin \theta)^2 + \sigma_{zz}^{SAP} (\cos \phi)^2 \quad (33)$$

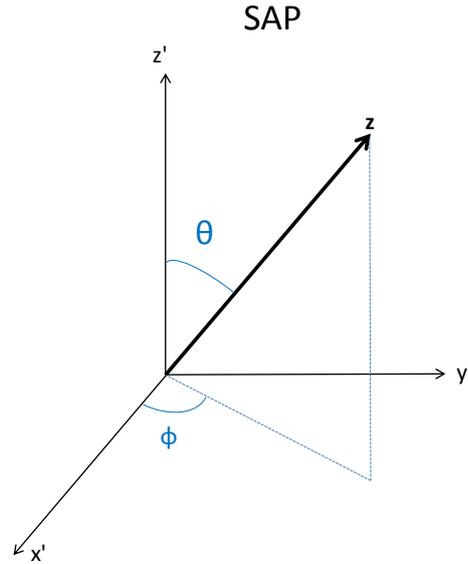


Figure 2. Représentation de l'orientation de l'axe z du laboratoire dans le système d'axes principaux (SAP)

L'équation montre clairement la dépendance spatiale (angulaire) de σ_{zz} et que le spectre RMN est une moyenne spatiale de tous les déplacements chimiques correspondant aux différentes orientations possibles des spins de l'échantillon. Si le tenseur $\bar{\sigma}$ est de symétrie axiale, autrement dit si $\sigma_{xx}^{SAP} = \sigma_{yy}^{SAP}$, σ_{zz} se simplifie et peut être réécrit sous la forme :

$$\sigma_{zz} = \sigma_{iso} - (\sigma_{iso} - \sigma_{zz}^{SAP}) \cdot \frac{1}{2} (1 - 3 \cos^2 \theta) \quad (34)$$

avec

$$\sigma_{iso} = \frac{1}{3} (\sigma_{xx}^{SAP} + \sigma_{yy}^{SAP} + \sigma_{zz}^{SAP}) \quad (35)$$

où σ_{iso} représente le déplacement chimique isotrope. L'équation précédente indique que σ_{zz} et donc le déplacement chimique est la somme d'une quantité scalaire (σ_{iso}) et d'un terme spatial qui dépend de θ . Lorsque $\theta=54.7^\circ$, ce terme spatial disparaît et σ_{zz} ne dépend plus que du déplacement chimique isotropique. En solution, le mouvement naturel rapide des molécules va moyenniser à zéro la partie spatiale et ainsi seule la partie isotropique s'exprimera, les raies du spectre RMN seront donc fines.

En revanche, pour le cas des échantillons solides (poudres cristallines) où la mobilité est restreinte voire inexistante, les différentes orientations possibles du cristallite par rapport au champ magnétique (θ) vont conduire à une distribution des fréquences de résonance et donc à un élargissement des raies.

Pour les composés semi-solides, c'est-à-dire des composés dont les propriétés se situent entre celles des liquides et des solides, la mobilité bien que restreinte sera plus importante que dans le cas des composés solides, conduisant à une distribution moins large des fréquences que celles observées en RMN solide. Ainsi en RMN HR-MAS, les effets de l'anisotropie de déplacement chimique sont bien moins importants et peuvent facilement être moyennés par la rotation à l'angle magique.

II.1.4) Les différences de susceptibilités magnétiques (milieu anisotrope)

Le principal phénomène contribuant à la largeur de raie en RMN HR-MAS provient de l'importante différence de susceptibilité magnétique, notée χ , au sein de l'échantillon. Cet effet est rarement pris en compte en RMN du solide du fait de sa plus faible importance sur la largeur de raie que l'élargissement dû aux interactions dipolaires et à l'anisotropie de déplacement chimique. En RMN du liquide, les composés étudiés possèdent une susceptibilité magnétique homogène au sein de l'échantillon ce qui réduit considérablement l'importance de ce phénomène.

Les échantillons étudiés en RMN HR-MAS sont des mélanges hétérogènes composés de parties liquides, à forte mobilité, et de parties solides, où la mobilité est beaucoup plus restreinte. Des gradients importants de susceptibilité magnétique existent ainsi à travers l'échantillon.

Lorsque placé dans un champ magnétique uniforme \vec{H}_0 , une substance isotrope et non ferromagnétique acquiert une aimantation \vec{M} défini par :

$$\vec{M} = \chi \cdot \vec{H}_0 \quad (36)$$

Où χ représente la susceptibilité magnétique de la substance isotrope et ne possède pas de dimension.

L'induction magnétique \vec{B}_0 devient alors :

$$\vec{B}_0 = \mu_0(\vec{H}_0 + \vec{M}) = \mu_0(1 + \chi) \cdot \vec{H}_0 \quad (37)$$

Où μ_0 est la perméabilité du vide.

La fréquence de résonance du noyau sera déterminé par \vec{B}_0 , et comme le montre l'équation précédente (37), \vec{B}_0 dépend de la susceptibilité magnétique. Un échantillon hétérogène peut être représenté comme une distribution de régions microscopiques, de volumes et de formes différentes, assimilables à des petites sphères de volume V_n au sein desquelles la susceptibilité magnétique (χ_n) est isotrope (Figure 3). Les différences de susceptibilité magnétique de toutes ces différentes parties de l'échantillon, vont ainsi provoquer une distribution de \vec{B}_0 qui va augmenter de manière significative la largeur des raies.

L'effet de cet élargissement lié aux différences de susceptibilité magnétique au sein de l'échantillon peut être considéré comme une interaction dipolaire, entre un dipôle macroscopique (l'échantillon) et un dipôle microscopique (un spin), notée $E_\chi(r)$.

L'énergie dipolaire ($E_\chi(r)$) ressentit par un noyau k à une distance r_n des sphères de volumes V_n et d'aimantation M_n peut être décrit par l'équation suivante :

$$E_\chi(r) = \frac{\mu_0}{4\pi} \sum_n \frac{[r_n^2(M_n m_k) - 3(r_n \cdot M_n)(r_n \cdot m_j)]}{r_n^5} \quad (38)$$

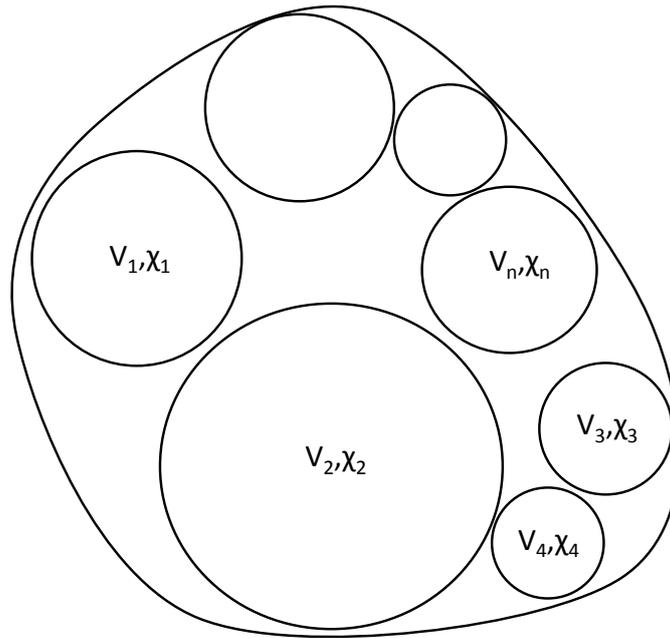


Figure 3. Schéma d'un échantillon hétérogène. Chaque partie de l'échantillon peut être assimilée à une somme de sphères de différents volumes V_n au sein desquelles la susceptibilité magnétique est homogène (χ_n).

En exprimant les moments magnétiques m_j et m_k en coordonnées polaires et en ne conservant que la composante le long du champ magnétique (de l'axe z) l'équation précédente devient :

$$E_\chi(r) = \frac{\mu_0}{4\pi} \sum_n \frac{(M_{nz}m_{nz})}{r_n^3} \cdot (1 - 3 \cos^2 \theta) \quad (39)$$

L'équation précédente (39) montre que l'effet de l'élargissement dû aux différences de susceptibilité magnétique au sein de l'échantillon peut être annulé à l'angle magique. En effet, la rotation à cet angle va moyenniser à zéro la contribution dipolaire de l'échantillon et ainsi diminuer considérablement cet effet représentant la principale source d'élargissement en RMN HR-MAS.

II.1.5) Dynamique

Lorsque toutes les interactions précédemment étudiées sont moyennées à zéro, le seul paramètre influençant alors la largeur de raie est la dynamique de la molécule. En effet, comme nous l'avons vu dans la section I.2.2) la largeur de raie est reliée au temps de relaxation spin-spin (T_2) par la relation :

$$\nu_{1/2} = \frac{1}{\pi T_2} \quad (40)$$

Par conséquent, une molécule possédant un mouvement lent (cas des grosses molécules de type protéines ou lipides) est traditionnellement caractérisée par un temps de relaxation spin-spin court et donc par des raies larges sur le spectre RMN. En revanche, les petites molécules (métabolites) sont caractérisées par des temps de relaxation transversale lents conduisant à des raies fines.

Parmi tous les phénomènes responsables de l'élargissement des raies en RMN HR-MAS précédemment cités, la dynamique est le seul phénomène immuable celui-ci étant une propriété intrinsèque des molécules. Cependant, la dynamique dépend également d'autres paramètres tels que le solvant utilisé, la température et l'environnement de la molécule.

II.2) La rotation à l'angle magique (MAS)

Nous avons vu précédemment que les différentes interactions à l'origine de l'élargissement des raies à savoir l'interaction dipolaire et l'interaction d'anisotropie de déplacement chimique pouvaient être décrites par des Hamiltoniens contenant chacun un terme spatial de la forme $(1 - 3 \cos^2 \theta)$. Les différences de susceptibilité magnétique au sein d'un échantillon donnent elles aussi lieu à une interaction proportionnelle à $(1 - 3 \cos^2 \theta)$. Ainsi, en appliquant une rotation de l'échantillon autour d'un axe dont l'inclinaison par rapport au champ magnétique statique correspond à $\theta=54,74^\circ$ (Figure 4), appelé « angle magique », le terme $(1 - 3 \cos^2 \theta)$ s'annule.

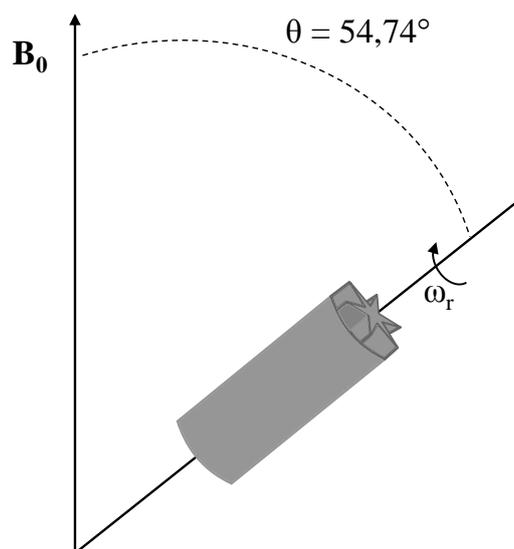


Figure 4 : Schéma illustrant la mise en rotation d'un rotor contenant l'échantillon à la vitesse ω_r autour d'un axe formant un angle de $54,74^\circ$ par rapport au champ magnétique externe, appelé « angle magique ».

D'un point de vue géométrique, la mise en rotation de l'échantillon autour d'un axe faisant un angle β par rapport au champ magnétique statique implique que le terme $(3 \cos^2 \theta - 1)$ devienne dépendant du temps. Il est alors défini par :

$$\begin{aligned}
 & (3\cos^2\theta(t) - 1) \\
 &= \frac{1}{2} (1 - 3 \cos^2 \beta) (1 - 3 \cos^2 \beta'_{ij}) \\
 &+ \frac{3}{2} \sin 2\beta \sin 2\beta'_{ij} \cos(\omega_r t + \phi_0) \\
 &+ \frac{3}{2} \sin^2 \beta \sin^2 \beta'_{ij} \cos 2(\omega_r t + \phi_0)
 \end{aligned} \tag{41}$$

Où β'_{ij} est l'angle entre le vecteur internucléaire r et l'axe de rotation, ω_r la vitesse de rotation, t le temps et ϕ_0 un facteur de phase.

Cette équation (41) est composée d'un terme indépendant du temps et de deux termes dépendant du temps. Le terme dépendant du temps s'annule pour $\beta = 54,74^\circ$, tandis que les deux autres termes sont responsables des raies satellites, situées à $\pm n$ fois la vitesse de rotation de part et d'autre du pic isotropique, appelées bandes de rotation. Si la vitesse de rotation est supérieure à la largeur spectrale d'intérêt du spectre, elles ne seront pas visibles sur le spectre RMN. Afin d'obtenir des raies fines, la vitesse de rotation de l'échantillon doit être au moins égale à la largeur de raie obtenue en conditions statiques.

Cependant, afin d'obtenir un spectre RMN HR-MAS avec une résolution proche de celle obtenue en RMN liquide, une sonde MAS solide classique ne suffit pas. Certaines spécificités sont requises (bobine particulière, adaptation d'un système de gradients à la configuration MAS, canal supplémentaire pour le « lock », méthode de réglage des « shims » particulière, rotor) comme nous le verrons plus loin dans ce manuscrit (**Partie III**).

La plupart des études menées par RMN liquide ou RMN HR-MAS sont basées sur l'utilisation de la chimométrie. En effet, les données générées par la spectroscopie RMN sont constituées d'un nombre très élevé de variables (plusieurs centaines) et d'individus (échantillons) rendant la comparaison entre les différents spectres ou groupe d'individus impossible à l'aide de l'œil humain. Il est alors nécessaire d'employer des méthodes

statistiques multivariées pour mettre en avant des similitudes ou des distinctions entre les différentes analyses.

C'est une des thématiques que nous allons développer dans la **Partie II** dédiée à la métabolomique. En effet, après avoir effectué une présentation globale de cette approche, nous décrirons en détails au cours d'une seconde section la démarche classique employée en RMN pour les études métaboliques, et réaliserons une revue de la littérature résumant l'état de l'art de la RMN dans l'agroalimentaire.

Ensuite, nous présenterons les différentes techniques statistiques utilisées au cours des travaux de cette thèse. Nous nous intéresserons également au traitement des données préalable à l'application des outils statistiques, qui est une étape extrêmement importante, ce procédé influençant grandement les résultats.

Partie II- Métabolomique

I) Introduction à la métabolomique

La métabolomique est une science très récente basée sur l'identification et la quantification exhaustive et non sélective de l'ensemble des métabolites (petites molécules) présents dans une cellule, un tissu, un organe, un organite ou un organisme à un temps donné et dans des conditions données. Le terme métabolite est réservé aux petites molécules, c'est-à-dire dont la masse moléculaire est inférieure à 1500 Da et issues du métabolisme. Les différents métabolites sont regroupés en deux classes : les métabolites primaires qui sont directement impliqués dans la reproduction, la croissance ou le développement d'un organisme ou d'une cellule, et communs à de nombreux organismes tels que les acides aminés, les sucres, certains alcools, les vitamines et les nucléotides, et les métabolites secondaires qui sont des composés plus spécifiques, a priori non impliqués directement dans ces processus tel que les phénols, certains composés azotés comme les alcaloïdes, les terpènes, certaines drogues, les xénobiotiques et les toxines. A la différence des métabolites primaires, ils peuvent provenir de métabolismes inhabituels, à la suite de réaction de défense ou d'interactions biologiques et sont spécifiques à certains organismes (plantes, champignons, bactéries...).

Le concept de métabolomique est apparu à la fin des années 1960 avec l'avènement des techniques de spectrométrie de masse couplée à la CPG. Cependant ce n'est qu'en 1971, que le premier article relatif à la métabolomique fut publié par Pauling et al¹⁵, à travers l'analyse quantitative de métabolites de fluides corporels par chromatographie tandis que Horning et al¹⁶ introduisirent la notion de profil métabolique. La même année Mamer et al¹⁷, réalisèrent également des travaux métaboliques sur l'identification des acides présents dans l'urine en utilisant la spectrométrie de masse (couplé à la chromatographie). Il fallut attendre la fin des années 1990 et le développement des approches en « omique », pour que le terme « métabolomique » apparaisse à travers le mot « métabolome » cité pour la première fois par Oliver et al¹⁸ lors de leur étude cherchant à comprendre le rôle des produits issus du gène de levure *Saccharomyces cerevisiae*. Ainsi, la métabolomique fut l'une des dernières « omique » (étude globale d'un ensemble moléculaire) à apparaître, bien après d'autres techniques bien connues comme la génomique (étude du fonctionnement d'un système biologique en tenant compte de son génome), la protéomique (étude de l'ensemble des protéines d'un organisme à un temps donné et dans des conditions données) et la transcriptomique (étude de l'ARNm issus de la transcription du génome).

Le développement des différentes techniques analytiques depuis la fin des années 1990, a conduit à l'essor de la métabolomique comme en atteste le nombre de publications : 3 articles dans *PubMed* référencés en 1999 avec le mot clé *metabolomics* plus de 12000 à présent !

Il est actuellement impossible d'étudier simultanément tous les métabolites et ceci pour deux raisons principales : l'une quantitative et l'autre qualitative. Pour la première, le nombre de molécules à caractériser peut atteindre plusieurs dizaines voire dizaine de milliers en fonction du seuil de concentration et aucune méthode analytique ne dispose aujourd'hui d'une telle sensibilité. Pour la seconde, les métabolites possèdent des structures chimiques extrêmement variées et donc des propriétés physico-chimiques hétéroclites rendant impossible la mise en place de technique d'extraction universelle.

Ainsi plusieurs techniques analytiques sont utilisées pour les études métabolomiques, les deux techniques de référence étant la spectrométrie de masse le plus souvent couplée à des techniques séparatives (SMⁿ, CLHP-SM ou CPG-SM) et la spectroscopie RMN. C'est sur cette dernière que nous allons nous focaliser bien que très souvent les études métaboliques combinent les deux techniques¹⁹, celles-ci offrant des informations complémentaires.

II) La RMN en métabolomique

II.1) La RMN en métabolomique

La RMN est une technique permettant l'observation de plusieurs noyaux (¹H, ¹³C, ¹⁵N, ¹⁹F, ³¹P) en fonction du type d'expériences RMN utilisé. Dans le cas des échantillons biologiques, le principal noyau observé est le ¹H du fait de sa présence dans la quasi-totalité des composés d'intérêt, de sa forte abondance isotopique naturelle (99.98%) et de son rapport gyromagnétique élevé, conduisant à une forte sensibilité l'aimantation observable en RMN étant principalement proportionnelle à ces deux paramètres. Chaque proton d'une molécule génère un signal spécifique sur le spectre RMN à une fréquence qui dépend de l'environnement chimique de ce proton. Plutôt que d'exprimer la position du pic sur le spectre RMN en termes de fréquence, on utilise couramment le déplacement chimique qui permet de s'affranchir du champ magnétique utilisé. La forme du pic RMN du proton étudié dépend de sa multiplicité, c'est-à-dire du nombre de protons voisins. Le spectre RMN ¹H d'un composé est donc la somme des signaux des différents noyaux ¹H de celui-ci et dans le cas de l'analyse d'un mélange de composés la somme des différents spectres de chaque composé.

La spectroscopie RMN présente plusieurs avantages dans l'établissement d'un profil métabolique par rapport à la spectrométrie de masse malgré une sensibilité inférieure²⁰. La sensibilité en RMN a toutefois été améliorée ces dernières années avec le développement de l'instrumentation que ce soit avec l'apparition d'aimants à plus hauts champs ou le développement de nouvelles sondes refroidies telles que les cryo-sondes²¹.

Le principal avantage de la RMN est que cette technique permet l'acquisition de l'empreinte métabolique globale d'un échantillon en une seule analyse, en déterminant simultanément la concentration de nombreux métabolites et ceci sans longue procédure de calibration préalable ou de traitement des données.

De plus, la RMN présente l'avantage d'être une technique ne requérant pas nécessairement une séparation préalable des composés et est non destructive, permettant ainsi l'étude directe d'échantillons intacts aussi bien liquides (urine, sérum, plasma, aliments solubles) que semi-solides (biopsies de tissus¹², certains petits organismes^{22,23}, aliments semi-solides) et la récupération des échantillons après analyse.

Cette technique est également extrêmement robuste et reproductible²⁰. Enfin, la RMN permet à travers l'utilisation du déplacement chimique d'utiliser des bases de données universelles et très riches, ne dépendant pas comme la spectrométrie de masse du type d'appareil, de méthode séparative préalable, du détecteur ou du mode de ionisation utilisé, facilitant la comparaison des résultats avec ceux des bases de données.

Les principales applications de la métabolomique par RMN sont la toxicologie ou l'évaluation de la toxicité²⁴, le diagnostic de maladies²⁵ (par exemple la détection à un stade précoce de cancers ou de certaines pathologies), la génomique fonctionnelle^{26,27} (déterminer la fonction de certains gènes à partir de leur produit d'expression), la nutriginomique²⁸ (étude des interactions entre gènes et nutriment), la microbiologie²⁹, l'écologie³⁰ ou l'étude du métabolisme des plantes³¹.

Démarche métabolomique RMN

Dans l'optique d'une étude métabolomique par RMN, un schéma analytique est à respecter et ceci qu'importe le type d'échantillon ou de technique RMN utilisé. Celui-ci est donné en Figure 5.

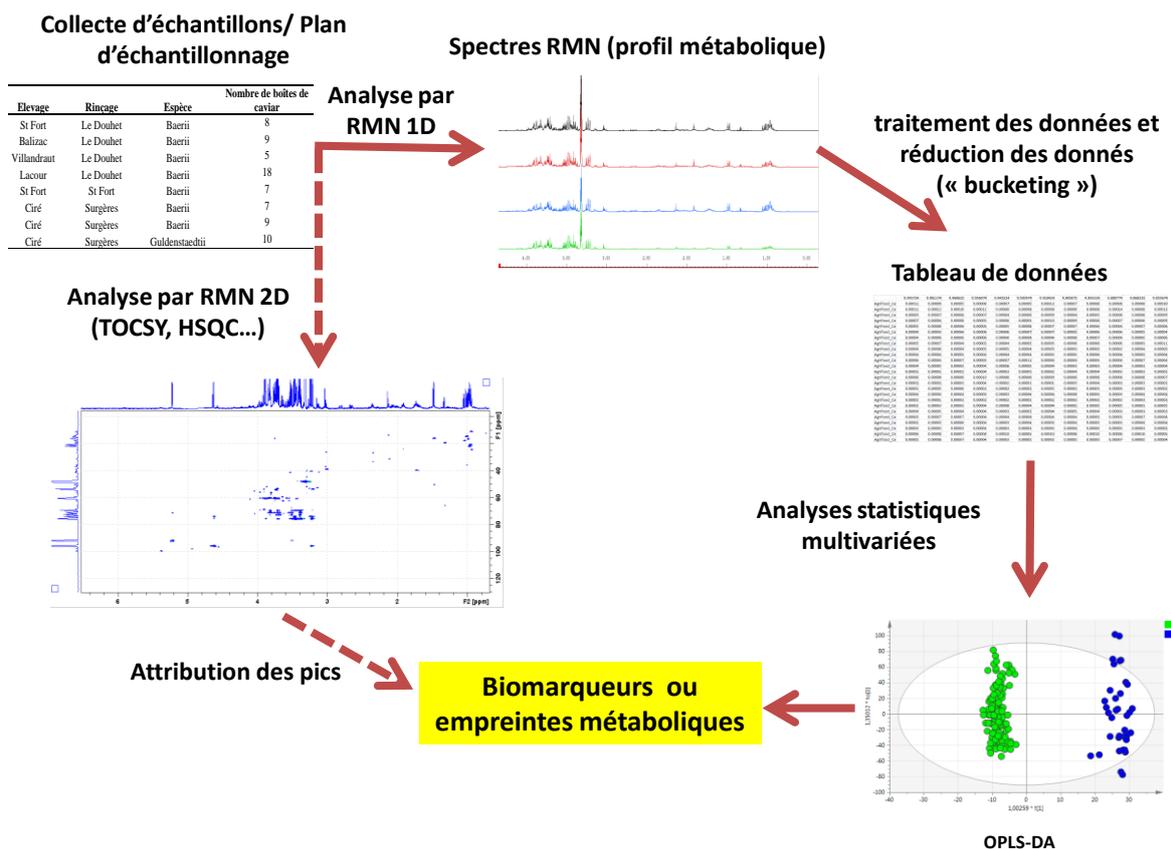


Figure 5 : Principe de la démarche en métabolomique par RMN. Les flèches en pointillés indiquent les étapes menées en parallèle des études statistiques.

La première étape consiste à travers la mise au point d'un plan d'échantillonnage, à définir la collecte d'un nombre suffisamment important et représentatif d'échantillons. Lors de cette étape, il est nécessaire de définir, dans un premier temps, l'ensemble des paramètres susceptibles d'influencer le profil métabolique, dans le but d'établir un plan d'échantillonnage permettant de démontrer quels paramètres peuvent affecter les empreintes métaboliques. Cependant, l'influence de certains paramètres ne peut pas être forcément établie en amont et la collecte d'échantillons peut être soumise à des contraintes pratiques (cas de l'analyse de biopsies humaines où tous les paramètres ne peuvent pas être maîtrisés ou connus).

La deuxième étape consiste en la préparation et en l'analyse des différents échantillons par RMN. Le protocole de préparation des échantillons biologiques en RMN est relativement simple et ceci pour les différents types d'échantillons étudiés. Pour les études par RMN liquide de fluides biologiques (urine, serum ou plasma)³², une certaine quantité du fluide en question est mélangée à une solution de tampon phosphate typiquement de pH 7,4 dans le but de contrôler les variations liées au pH de l'échantillon, du D₂O afin de fournir une fréquence

deutérium pour le « lock » et un standard interne typiquement du TSP ou du DSS pour calibrer les spectres³³.

Pour ce qui est du protocole de préparation en spectroscopie RMN HR-MAS, la démarche la plus courante consiste en l'utilisation d'inserts jetables de 33 µl en Kel-F, au sein duquel est placée une certaine quantité de solvant deutéré (classiquement environ 10 µl) et l'échantillon à analyser (10-15 mg). L'insert est alors fermé à l'aide d'un bouchon conique et d'une vis puis placé au sein d'un rotor 4mm en oxyde de zirconium ZrO₂. La principale raison de l'utilisation d'inserts jetables, plutôt que de rotors pleins, vient de la longue procédure de nettoyage nécessaire lors de l'utilisation de rotors pleins nécessaire pour éviter les contaminations entre les différents échantillons.

La plupart des profils métabolique RMN sont basés sur des analyses avec séquences d'impulsions Radio Fréquence (RF) 1D. La suppression du signal de l'eau ou du solvant utilisé est, le plus souvent, réalisée par présaturation³⁴ qui consiste à appliquer, pendant la période de récupération, une onde continue de faible amplitude à la fréquence des protons de l'eau ou du solvant. D'autres techniques de suppression du pic de l'eau sont également employées comme la séquence DANTE (Delays Alternating with Nutation for Tailored Excitation)³⁵, la WEFT (Water Eliminated Fourier Transform)³⁶ ou la suppression WATERGATE³⁷. La séquence la plus utilisée pour les études métaboliques est la 1D NOESY avec présaturation, du fait de son excellente capacité à supprimer le signal du solvant et à obtenir une bonne qualité de ligne de base³⁸. Cependant, certains échantillons biologiques particuliers comme le sérum ou le plasma contiennent des protéines et des lipides qui sont à l'origine de signaux larges et de perturbations de la ligne de base. La contribution de ces composés peut être réduite en utilisant des séquences particulières comme la séquence Carr-Purcell-Meiboom-Gill (CPMG)^{39,40}, qualifiée de filtre T₂, qui va atténuer le signal de ces molécules tout en minimisant la perte de signal pour les métabolites (Figure 6). Ces deux séquences seront détaillées d'avantage dans la **Partie III**.

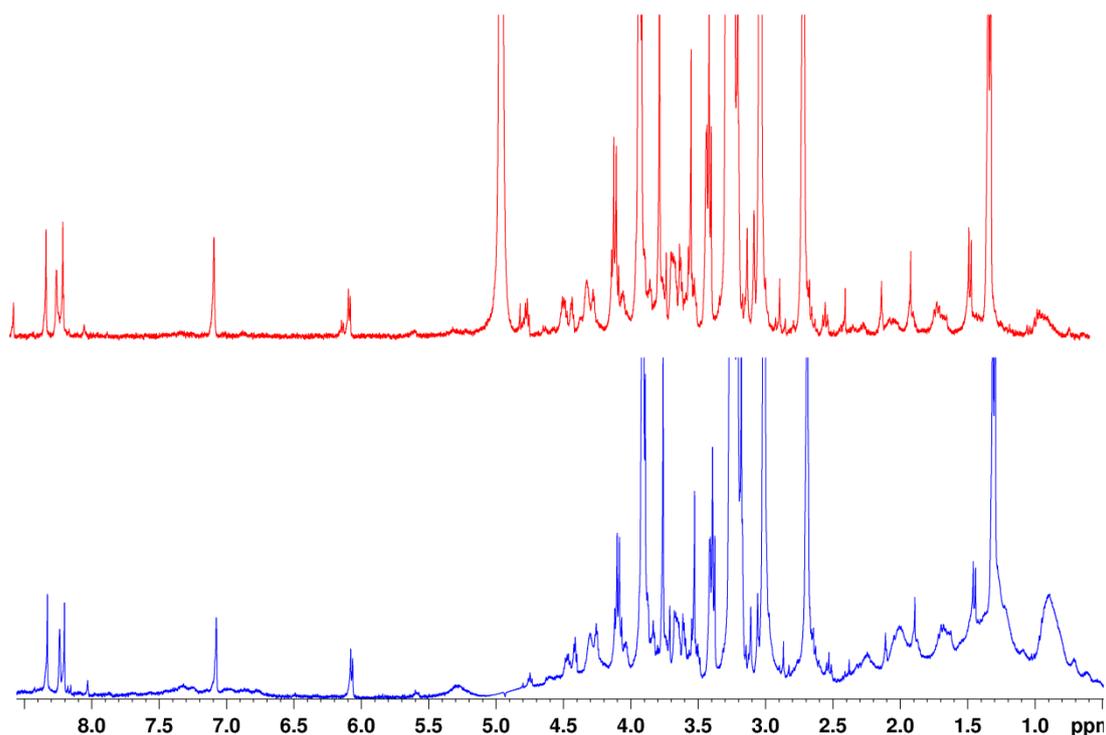


Figure 6 : Spectres ^1H enregistrés sur un spectromètre 400 MHz sur une biopsie de poisson par RMN HR-MAS à l'aide d'une séquence CPMG (spectre rouge) et NOESY 1D (spectre bleu). Pour les deux séquences une présaturation du signal de l'eau est utilisée.

D'autres séquences, cette fois-ci basées sur la différence des coefficients de diffusions moléculaires permettent également de séparer les molécules diffusant peu (grosses molécules) de celles diffusant rapidement (petits métabolites) et sont généralement utilisées pour l'étude des lipides ou de protéines liées à un ligand⁴¹.

Enfin, dans certains cas, les profils métaboliques sont enregistrées à l'aide de séquences RMN multidimensionnelles⁴²⁻⁴⁶ Cependant, ces expériences sont souvent longues et peuvent provoquer des changements biochimiques et biologiques au sein des échantillons.

Attribution des métabolites

Parallèlement à l'acquisition des premiers spectres, une étape très importante lors d'une étude métabolique est l'attribution des différents signaux présents sur le spectre RMN aux composés correspondants. En effet, les méthodes statistiques multivariées vont mettre en évidence des similitudes ou des différences entre les différents échantillons et il est capital de

savoir selon quel métabolite se font ces regroupements ou distinctions ou d'identifier d'éventuels marqueurs biologiques.

Le plus souvent la première étape de l'identification des différents métabolites consiste, lorsque cela est possible, en une analyse de l'échantillon d'intérêt par spectrométrie de masse couplée à une technique séparative afin d'identifier un maximum de composés à l'aide de leur temps de rétention, de leur rapport m/z (masse/charge) et de leur chemin de fragmentation. Il est important d'avoir à l'esprit que les composés les plus abondants par spectrométrie de masse ne sont pas nécessairement les composés les plus abondants sur le spectre RMN (cela dépend de la colonne et de la technique séparative utilisée en amont et du mode de ionisation), cependant les résultats issus de la spectrométrie de masse sont souvent un point de départ précieux et complémentaire pour l'élucidation par RMN.

L'attribution des différents métabolites observés par RMN est réalisée en comparant les spectres obtenus aux spectres de références pouvant exister soit sur les bases de données spectrales académiques telles que HMDB⁴⁷, MMCD⁴⁸ ou BMRB⁴⁹, ou à des bases de données privées comme BBIOREFCODE (Bruker BioSpin, Germany) ou Chenomx NMR suite (Chenomx Inc, Canada). Cependant, comme indiqué précédemment, les spectres 1D sont souvent sujets à des chevauchements de pics et le déplacement chimique proton ne peut pas, à lui seul, suffire à l'attribution certaine des différents composés. Des expériences RMN multidimensionnelles homonucléaires et hétéronucléaires complémentaires sont ainsi couramment enregistrées dans le but de fournir des informations additionnelles facilitant l'interprétation ou l'attribution.

La séquence 2D homonucléaire la plus utilisée est la 2D ¹H-¹H TOCSY (TOtal Correlation Spectroscopy)⁵⁰ qui permet de corrélérer tous les noyaux ¹H appartenant à un même système de spins. Un exemple de spectre ¹H-¹H TOCSY adiabatique (ATOCSY) obtenu sur un échantillon de poisson est donné en Figure 7.

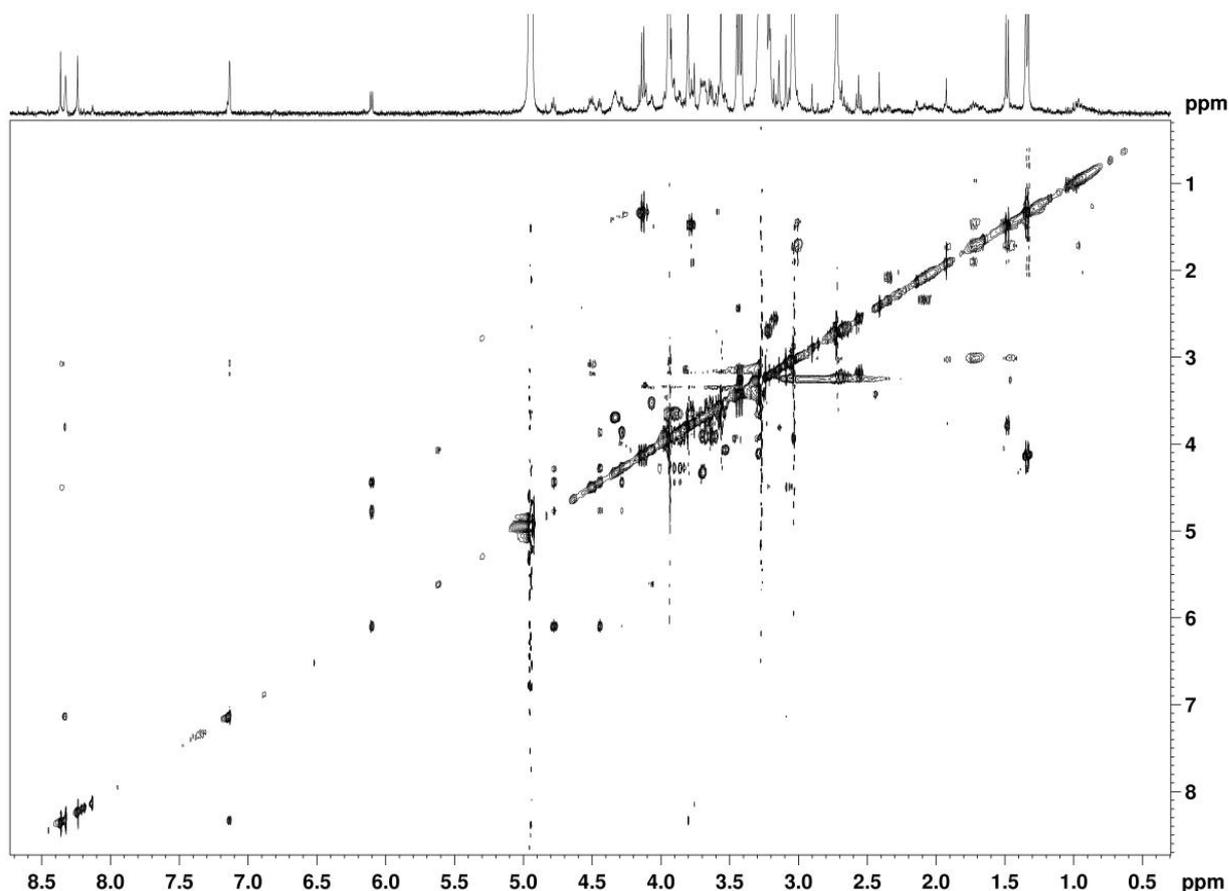


Figure 7 : Spectre ATOCSY ^1H - ^1H enregistré sur un spectromètre 400MHz par RMN HR-MAS sur une biopsie de poisson au cours de cette thèse.

Cependant, bien qu'offrant des informations additionnelles sur les pics (ou multiplets) appartenant à la même molécule et permettant ainsi de mieux comprendre le spectre RMN 1D, elle ne permet pas d'identifier dans tous les cas les molécules étudiées. Une acquisition 2D ^1H - ^{13}C HSQC (Heteronuclear Single Quantum Correlation)⁵¹ est alors utilisée pour identifier les noyaux protons directement liés à des noyaux carbonés. L'intérêt de l'utilisation de la dimension carbone est qu'elle présente une largeur spectrale plus conséquente que celle du proton et permet ainsi une meilleure discrimination des résonances sur celle-ci. Un exemple de spectre ^1H - ^{13}C obtenu sur un échantillon d'urine est donné en Figure 8.

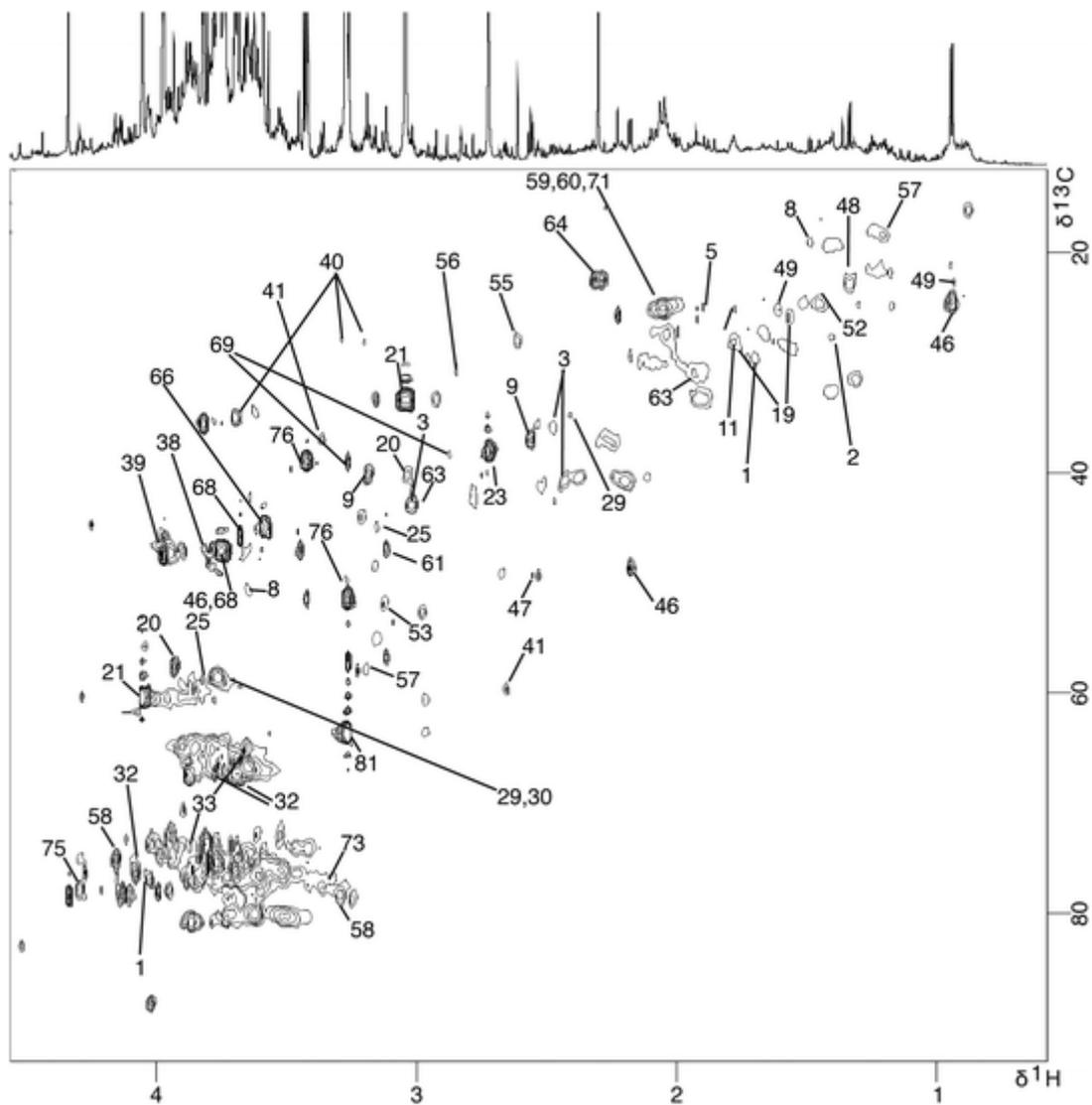


Figure 8 : Spectre HSQC ^1H - ^{13}C obtenu sur un échantillon d'urine tiré de la publication de Nicholson et al⁵².

Pour certaines études biologiques, il est également possible d'avoir recourt à l'acquisition d'HSQC ^1H - ^{15}N . Une autre expérience largement utilisée est la J-Resolved⁵³ qui permet de corréler les déplacements chimiques à la multiplicité du signal associé et la mesure directe des constantes de couplage.

Quantification RMN

La quantification des différents métabolites présents sur un spectre RMN peut être réalisée de plusieurs manières^{54,55}. La méthode la plus simple pour quantifier un composé présent sur un spectre RMN consiste à intégrer les différents signaux d'intérêt. L'une des relations fondamentales de la quantification par RMN est la relation de proportionnalité entre l'intensité d'un signal (I_n) et le nombre de noyaux responsable de cette résonance (N_n) :

$$I_n = K \times N_n$$

Où K est la constante du spectromètre et est commune à toutes les résonances du spectre.

Pour quantifier de manière absolue un composé, plusieurs techniques existent. Tout d'abord, la technique de référence interne, consistant en l'ajout d'une quantité connue d'une molécule au sein de la solution analysée comme le TMS, le TSP ou le DSS. Le standard interne ne doit pas interagir avec les autres composés présents et doit être parfaitement soluble dans le solvant utilisé.

Lorsqu'il n'est pas possible de trouver une référence interne qui n'interagit pas avec les solutés, la référence externe est alors retenue. Pour cela, un insert (capillaire) contenant la solution de référence dissoute dans un solvant deutéré est ajouté au sein du tube rempli avec la solution à analyser. Connaissant la quantité de référence au sein de l'insert, il est alors possible de remonter à la quantité absolue des différents métabolites présents dans le tube. Cette technique conduit, cependant, à une perte de signal sur bruit d'environ 30 %, liée à la difficulté d'ajuster le champ magnétique local de l'échantillon et au plus faible volume de celui-ci dans la bobine de détection.

Les étalonnages interne et externe sont également couramment utilisés pour la quantification absolue par RMN. L'étalonnage interne repose sur l'utilisation d'une droite de calibrage déterminée par l'ajout en quantité parfaitement connue d'une molécule non présente dans l'échantillon à plusieurs concentrations différentes et dans les mêmes conditions d'acquisition. La constante K du spectromètre est alors mesurée sur ce composé et appliquée à l'ensemble des métabolites à quantifier.

L'étalonnage externe consiste, quant à lui, à préparer des solutions de concentrations différentes du métabolite à quantifier afin de déterminer une courbe de calibrage pour ce

composé. La quantité absolue de cette molécule au sein de l'échantillon est alors calculée à l'aide de l'aire mesurée et en se référant à l'équation de la courbe de calibrage.

Une alternative aux méthodes d'étalonnage est la technique d'ajouts dosés. Celle-ci consiste en l'addition de quantités connues du composé à quantifier. Une courbe de calibrage est alors construite à l'aide des données enregistrées pour les différentes concentrations. La quantité absolue initiale est alors déterminée à l'aide de cette courbe et de l'aire mesurée avant les différents ajouts.

Une méthode pour quantifier de façon absolue un analyte est la technique ERETIC développée par Akoka et al⁵⁶. Celle-ci est basée sur l'utilisation d'une référence électronique appelée Electronic REference To access In-vivo Concentration (ERETIC) qui fournit un pseudo FID qui possède toutes les caractéristiques d'un vrai signal RMN dont les paramètres peuvent être contrôlés par le poste de travail. Le signal ERETIC doit être calibré par une solution de concentration connue à l'aide de l'équation suivante :

$$[ERETIC] = [REF] \times \frac{A_{ERETIC}}{A_{REF}}$$

Avec [ERETIC] et [REF] les concentrations respectives du signal électronique et de la solution de référence. La concentration absolue de n'importe quel métabolite peut alors être quantifiée à l'aide de la formule suivante :

$$[composé] = K[ERETIC] \times \frac{A_{composé}}{A_{ERETIC}}$$

Où K représente le nombre de protons et la masse moléculaire du pic d'intérêt.

Une version améliorée d'ERETIC (Amplitude-corrected Referencing Through Signal Injection, ARTSI⁵⁷) a été développée par Avizonis et al, à travers l'utilisation d'un facteur de qualité (facteur Q) de la sonde qui est calculé en fonction des propriétés diélectriques des échantillons, celui-ci étant inversement proportionnel au carré de la durée de l'impulsion 90°. Une correction de la puissance du signal ERETIC est ainsi appliquée permettant une quantification plus robuste pour les échantillons présentant des fortes variations en concentrations salines.

Wider et al⁵⁸ ont introduit une méthode (PULse Length-based CONcentration, PULCON) qui corrèle l'intensité absolue de deux spectres (un de référence et celui correspondant à l'analyse

de l'échantillon) mesurés sur un même spectromètre avec des paramètres d'acquisition similaires, en tenant compte des variations entre certaines conditions expérimentales et du fait que la durée de l'impulsion 90°(ou 360°) est inversement proportionnelle à l'intensité du signal. C'est d'ailleurs cette technique de quantification absolue que nous utiliserons lors des travaux de cette thèse. La concentration du composé inconnu est alors déterminée à l'aide de la formule suivante :

$$[composé] = f_t \times [REF] \times \frac{S_{comp} T_{comp} \theta_{360} n_{REF}}{S_{REF} T_{REF} \theta_{360} n_{comp}}$$

avec S l'intensité du signal, T la température, θ_{360} la durée de l'impulsion 360°, n le nombre de scans pour la mesure des deux échantillons et f_t le facteur tenant compte des variations d'intensité lorsque différents gains du receveur sont appliqués et/ou lorsque différentes séquences d'impulsion sont utilisées entre le spectre de référence et celui correspondant à l'échantillon.

Cependant, la simple intégration des signaux est généralement rendue compliquée, particulièrement dans le cas des spectres RMN 1D, par la présence de plusieurs dizaines voire centaines de composés conduisant à la superposition de plusieurs signaux résonnants à des fréquences très proches.

Plusieurs solutions ont ainsi été développées pour surmonter ce problème. La première approche consiste à déterminer par déconvolution, la contribution des différents spectres 1D individuels des différentes molécules présentes dans l'échantillon. Malheureusement, cette approche est restreinte au cas où la superposition des pics est limitée et quand les différents composés présents dans l'échantillon sont identifiés.

Une autre solution consiste en l'acquisition de spectres 1D de noyaux offrant une largeur spectrale plus importante (^{13}C , ^{15}N) et conduisant à une meilleure discrimination des signaux. Cependant, cette approche basée sur l'analyse de noyaux, généralement caractérisés par des rapports gyromagnétiques et des abondances naturelles nettement plus faibles, conduisant à une plus faible sensibilité, nécessite des concentrations ou des temps d'acquisition beaucoup plus importants.

Une autre approche est basée sur la RMN multidimensionnelle qui offre une meilleure discrimination à travers l'introduction d'une dimension orthogonale supplémentaire. Cependant, la RMN multidimensionnelle implique de longs temps d'acquisition liés à la

nécessité d'acquérir plusieurs spectres 1D pour différents incréments, pouvant induire des artefacts sur le spectre RMN liés aux instabilités de l'appareillage RMN pendant l'analyse et présentant l'inconvénient que le volume des différents pics dépend de plusieurs paramètres RMN propres à chaque pic (comme les temps de relaxation transversale et longitudinale, des constantes de couplage homo et hétéro nucléaires, des effets hors résonance, de certains délais de la séquence impulsionnelle...). Malheureusement, ces techniques, impliquant également de longs temps d'acquisition, ne sont pas toujours applicables aux études biologiques, certains systèmes biologiques évoluant de manière significative au cours du temps.

Le récent développement de la RMN Ultra-rapide (ULTRAFAST)⁵⁹ ouvre de nouvelles perspectives à l'utilisation de la RMN multidimensionnelle pour les études métaboliques⁶⁰ mais se heurte, pour le moment, à la nécessité de travailler sur des échantillons concentrés et est réservée, pour le moment, à l'étude de métabolites majoritaires ou enrichis en certains isotopes. L'essor de la polarisation dynamique nucléaire (DNP) permettant un gain de sensibilité extrêmement élevé (de 100 à 10000) ouvre également de nouvelles perspectives à l'utilisation de la RMN multidimensionnelle pour des études métaboliques.

Réduction du nombre de variables spectrales : « bucketing » et analyses statistiques

Une fois les données correctement traitées, la dernière étape de la démarche métabolique qui est couramment utilisée en métabolomique, appelée « bucketing » ou « binning », consiste à décomposer le spectre RMN en plusieurs petites régions de taille constante ou variable, puis à quantifier le signal RMN dans chacune de celles-ci par intégration dans le but de réduire le nombre de variables spectrales. Récemment des techniques de « bucketing » intelligentes sont apparues^{61,62}. Les différents traitements statistiques pouvant être appliqués aux « buckets » générés seront détaillés dans la **Partie III**.

Une fois les données générées et réduites, celles-ci se présentent sous la forme d'une matrice de n lignes, chaque ligne représentant un échantillon, et p colonnes représentant les différentes variables déterminées pour chacun des individus. Il est alors possible, et indispensable, d'appliquer des méthodes statistiques multivariées afin d'interpréter l'ensemble des données (ACP, PLS-DA, OPLS-DA, SIMCA...). A l'aide de ces analyses statistiques, des différences ou des similitudes entre groupes d'échantillons peuvent être mises en avant et dans les cas favorables des biomarqueurs discriminants sont identifiés.

II.2) Les principales applications de la RMN dans l'agroalimentaire

II.2.1) La RMN HR-MAS dans l'agroalimentaire

La RMN HR-MAS, initialement utilisée, pour l'étude de molécules greffées sur des supports solides ou pour les études de chimie combinatoire, s'est largement développée et est désormais appliquée à de nombreux domaines (médical^{12,63}, chimie organique⁸, polymères^{10,11}...). Cependant, le domaine sur lequel nous allons nous focaliser, est l'utilisation de cette technique pour l'étude de produits alimentaires à travers une revue bibliographique détaillée. De nombreux produits agroalimentaires ont, en effet, une consistance ni solide ni liquide et sont ainsi parfaitement adaptés à une analyse par RMN HR-MAS : cette technique permettant l'analyse brute de ces produits sans traitement ou extraction préalable (mis à part l'ajout de solvant deutéré).

Fruit/Légumes

Les premières études par RMN HR-MAS sur des fruits ou des légumes remontent au début des années 2000 avec les études de Gil et al⁶⁴ et de Sobolev⁶⁵. Gil et al se sont intéressés aux changements de composition de la pulpe de mangue en fonction de son stade de murissement et ont observé que la quantité de certains sucres, notamment le saccharose, augmentait au cours de la maturation du fruit tandis que l'acide citrique diminuait. Sobolev et al ont étudié la pulpe de tomate et déterminé des différences de composition entre deux variétés (Cilieginio et Red setter).

La RMN HR-MAS a été utilisée à plusieurs reprises dans le but d'établir un profil métabolique pour l'étude de fruits ou de légumes. Ce fut notamment le cas de Perez et al⁶⁶ qui se sont intéressés à la différenciation entre la peau et la chair de tomate, la peau étant plus riche en certains sucres et en résidus glycosyls. Ils ont également réussi à faire des différences entre l'empreinte métabolique du fruit à différents stades de murissement (vert et rouge principalement) et ont également souligné des variations significatives entre les analyses du à une relative hétérogénéité en fonction de la zone de prélèvement. Plus récemment, Iglesias et al⁶⁷ ont étudié l'impact de la saison de récolte et des traitements sur la qualité de la tomate tandis que Mallamace et al⁶⁸ ont utilisé le profil métabolique obtenu par RMN HR-MAS, ainsi que des analyses statistiques multivariées pour différencier des tomates cerises provenant de différentes origines de celles provenant de Sicile, ces dernières ayant obtenu une IGP.

Ritota et al⁶⁹ ont, quant à eux, étudié le profil métabolique du poivron italien et l'utilisation de la PLS-DA a permis la différenciation entre les deux variétés de la région du Piémont (Corno et Cuneo) mais également entre les différentes origines géographiques. Les métabolites à l'origine de ces distinctions, ont été les sucres, les acides organiques (acide malique et ascorbique) et certains acides aminés (asparagine notamment).

Ritota et al⁷⁰ ont réalisé une étude similaire sur l'ail en tentant de différencier l'ail blanc de l'ail rose, venant de plusieurs régions italiennes. A l'aide de la PLS-DA, ils ont réussi à faire des différences entre les deux types d'ail mais également à remonter à leur origine géographique.

Vermathen et al ont soumis le profil métabolique obtenu par RMN HR-MAS à des analyses statistiques multivariées (ACP et PLS-DA) dans le but de discriminer des pommes selon le producteur^{71,72}.

Fromages

Le fromage a été étudié de manière exhaustive par Shintu & Caldarelli entre 2004 et 2007. En 2004, ils ont établi le profil métabolique du Parmigiano Reggiano⁷³ et ont identifié une trentaine de métabolites. Ils ont ensuite poursuivi leurs travaux sur le Parmigiano Reggiano en s'intéressant cette fois-ci à travers l'étude du vieillissement⁷⁴ et ont mis en évidence une possibilité de classement selon leur âge à l'aide de l'ACP et par Analyse Discriminante (AD) avec, comme marqueurs, certains acides aminés et certains pics non attribués. Ils se sont ensuite intéressés à l'étude de la traçabilité de l'emmental⁷⁵, à travers l'analyse d'échantillons provenant de 7 régions différentes (dont 5 pays différents). Toujours à l'aide l'ACP et de l'AD ils ont réussi à différencier les différentes origines sur la base de teneur en acides aminés et de certains acides gras insaturés.

Récemment, Mazzei et al⁷⁶ ont mené une étude sur la traçabilité de la mozzarella dans le cadre de la mise ne place d'une IGP (Indication Géographique Protégée) de la MBC (Mozzarella di Bufala Campana). Dans le but de classifier les différents échantillons, la CAH (Classification Ascendante Hiérarchique) et l'AD ont été utilisées et ces deux techniques ont permis la différenciation des échantillons MBC des autres en fonction des teneurs en galactose, lactose, acide acétique et glycérol.

Céréales/pain

Sacco et al ont, sans doute, été les premiers à utiliser la RMN HR-MAS appliquée à la caractérisation de produits alimentaires. En effet, Sacco et al⁷⁷ ont étudié des céréales provenant de différentes régions de l'Italie et attribué plusieurs composés à l'aide de méthodes 1D et 2D (¹H et ¹³C). Quelques années plus tard, Brescia et al⁷⁸ se sont intéressés à la différenciation en fonction de l'origine géographique et de la variété, de semoule de blé du sud de l'Italie à l'aide de la RMN HR-MAS mais aussi de la SMRI (rapports isotopiques ¹³C/¹²C et ¹⁵N/¹⁴N). La combinaison de ces deux techniques avec des analyses statistiques a permis une classification correcte en fonction des deux paramètres susnommés. L'année suivante, Brescia et al⁷⁹ ont mené une étude similaire (RMN HR-MAS et SMRI ¹³C/¹²C et ¹⁸O/¹⁶O) sur différentes farines de pain et à l'aide de la PCA ont élaboré une méthode de classification en fonction de la zone de production et du type de pain. Durant cette étude, ils ont également réussi à quantifier les acides lactiques et acétiques permettant de déterminer le ratio de fermentation (acide lactique/acide acétique), responsable du phénomène d'acidification lors du traitement des céréales.

La RMN HR-MAS a également été utilisée pour caractériser certaines propriétés des céréales. Ainsi, Larsen et al⁸⁰ se sont intéressés au niveau d'hydratation de granule d'amidon, cette propriété étant directement liée à la qualité du produit. Les différents niveaux d'hydratation ont été déterminés à l'aide du signal sur bruit des acquisitions ¹H et ³¹P par RMN, des spectres à signal/bruit faible indiquant une faible mobilité pour ces granules.

Poisson /Viande

Plusieurs études ont été menées sur l'étude de produits alimentaires de type viande/poisson, ces produits présentant une consistance adéquate à une analyse par RMN HR-MAS et permettant leur analyse directement sur le produit brut.

En 2005 Sacco et al⁸¹ ont étudié de la viande d'agneau originaire de différentes régions du sud de l'Italie en essayant de les différencier en fonction de leur origine et des conditions d'élevage en combinant les résultats obtenus par SMRI (¹⁵N/¹⁴N ¹³C/¹²C) et les profils métaboliques acquis par RMN HR-MAS. A l'aide de l'AD, ils ont réussi à classer correctement les 3 zones géographiques et les 4 races différentes.

Shintu et al⁸² ont, quant à eux, appliqué la HR-MAS pour l'étude de viande de bœuf séchée dans le but de déterminer des marqueurs potentiels en fonction de la qualité et de l'origine

géographique. La teneur en acides gras ainsi que certains métabolites liés à l'élevage (proline, phénylalanine, acide glutamique et certains acides aminés) ont ainsi été présentés comme marqueurs potentiels confirmant des résultats obtenus sur d'autres études utilisant d'autres techniques analytiques.

En 2012 Ritota et al⁸³ se sont attachés à séparer les différents types de muscles sur 4 races bovines distinctes. L'utilisation de la PLS-DA et de l'OPLS-DA permet une bonne classification pour 2 espèces mais une moins efficace pour les deux autres, avec comme métabolites discriminants, certains acides aminés et la carnitine.

Aursand et al⁸⁴ ont étudié le saumon d'atlantique et ont également montré la capacité de la RMN HR-MAS pour déterminer la teneur en oméga 3, indiquant cependant que des améliorations étaient à apporter à cette technique.

Nestor et al⁸⁵ ont mené une étude sur un autre salmonidé, l'ombre chevalier, et ont pu déterminer la teneur en oméga 3, mais également, séparer l'acide ecosapentanoïque (EPA) et l'acide docosahexénoïque (DHA) indiquant que la RMN HR-MAS pouvait être une technique complémentaire à la CPG pour la détermination du profil des acides gras.

Castejon et al⁸⁶ ont pour leur part mené une étude approfondie et ont pu caractériser le profil métabolique du saumon fumé (environ 30 composés). Au cours de cette étude, ils ont également suivi l'évolution de l'empreinte métabolique (acides gras, sucres, acides aminés, acides organiques et nucléosides) lors de la conservation du poisson au congélateur sur 30 jours.

Enfin, Bankefors et al⁸⁷ ont comparé les résultats obtenus pour le profilage métabolique par RMN HR-MAS sur des tissus intacts et par RMN liquide, sur des extraits de muscles et de foie de saumon d'atlantique. Les mêmes métabolites ont été retrouvés à l'aide des deux techniques. A l'aide de l'ACP, ils ont également pu, mettre en avant, l'effet du régime alimentaire, avec ou sans zygomycètes, sur le profil métabolique.

II.2.2) La RMN liquide dans l'agroalimentaire

La RMN liquide est également présente dans l'analyse de nombreux produits agroalimentaires, soit par l'analyse directe de produits liquides ou solubles dans certains solvants, soit par l'analyse d'extraits de composés solides ou semi-solides issus de procédures

d'extraction. Une revue complète des différentes applications de la RMN liquide haute résolution dans l'agroalimentaire serait beaucoup trop exhaustive tant les domaines d'application et le nombre d'articles correspondants sont importants, nous allons donc nous concentrer sur ses principales utilisations.

Une des principales applications est l'analyse isotopique par RMN, FINS-RMN (Fractionnement Isotopique Naturel Spécifique par Résonance Magnétique Nucléaire) développée dans les années 1980 par Martin et Martin⁸⁸. Cette technique a notamment été utilisée pour détecter la chaptalisation du vin⁸⁹ qui a été par la suite approuvée par l'Union Européenne en 1990. De nombreuses applications de la FINS-RMN pour l'authentification de produits alimentaires ont été développées tels que la bière⁹⁰, la vanilline⁹¹, les jus de fruits⁹² et à l'huile de moutarde⁹³.

Récemment des techniques de « screening », basées sur une approche holistique du produit étudié et entièrement automatisées, ont été développées dans le but de caractériser certains produits alimentaires comme les jus de fruit¹. Ces méthodes sont basées sur la quantification de certains composés dans le but de détecter des ajouts de sucres ou des traitements enzymatiques, et sur l'utilisation d'analyses multivariées afin de classer les différents échantillons analysés par rapport à une base de données contenant les spectres de référence enregistrés sur des échantillons authentiques et de pouvoir remonter à leur origine géographique. Cette méthodologie a également été développée pour l'analyse d'échantillons de type vin^{94,95}, miel, poudre de lait et certaines huiles alimentaires. Une étude récente a également combinée les profils métaboliques ¹H avec des données acquises à l'aide de la FINS-RMN pour la prédiction de l'origine géographique du vin⁹⁶.

Des travaux ont également porté sur l'étude de l'origine⁹⁷ ou de la qualité⁹⁸ du thé mais également sur le café instantané⁹⁹. Certaines études métaboliques se sont intéressés à la différenciation des variétés de fruit ou légumes tels que la tomate¹⁹ ou le ginseg¹⁰⁰.

Pour ce qui est de l'analyse par RMN liquide d'échantillons de poisson ou de viande l'étude de ces composés requiert une procédure d'extraction préalable. Cependant quelques études portant sur la détermination du profil d'acides gras par RMN ¹³C d'extraits de saumon notamment¹⁰¹ ont été réalisées permettant même, dans certains cas, de différencier un saumon sauvage d'un saumon d'élevage¹⁰². Jung et al¹⁰³ ont réussi à différencier différentes viandes de bœuf provenant de différentes origines (Nouvelle Zélande, Australie, Etats-Unis et Corée) à l'aide d'analyses statistiques multivariées.

Comme nous l'avons au travers des premières sections de cette partie, la plupart des études menées par RMN liquide ou RMN HR-MAS sont basées sur l'utilisation de la chimiométrie. En effet, les données générées par la spectroscopie RMN sont constituées d'un nombre très élevé de variables (plusieurs centaines) et d'individus (échantillons) rendant la comparaison entre les différents spectres ou groupe d'individus impossible à l'aide de l'œil humain. Il est alors nécessaire d'employer des méthodes statistiques multivariées pour mettre en avant des similitudes ou des distinctions entre les différentes analyses.

C'est cette thématique que nous allons développer dans la suite de cette partie, à travers la présentation des différentes techniques statistiques utilisées au cours des travaux de cette thèse. Nous nous intéresserons également au traitement des données préalable à l'application des outils statistique, qui est une étape extrêmement importante, ce procédé influençant grandement les résultats.

III) Méthodes Statistiques multivariées – Chimiométrie

III.1) Introduction

La chimiométrie, qui est une branche de la chimie analytique, est l'application d'outils mathématiques et plus particulièrement statistiques, dans le but d'extraire le maximum d'informations pertinentes à partir de données physico-chimiques. La chimiométrie comprend notamment l'optimisation de la collecte de données à l'aide de plans d'expériences, les méthodes multivariées d'analyse de données et les méthodes de classification.

La classification entre les différentes méthodes d'analyse de données statistiques s'organise selon plusieurs distinctions.

Tout d'abord, la première d'entre elles, sépare les techniques univariées et multivariées. Une technique univariée ne s'intéresse qu'à une seule variable ou mesure physico-chimique, par exemple la température, le pH, la concentration d'un composé chimique ou alors à une variable sélectionnée parmi un jeu de données. En revanche, une technique multivariée traite simultanément plusieurs variables et montre les relations qu'il peut exister entre celles-ci. Les variables peuvent être issues de plusieurs analyses différentes ou alors correspondre à plusieurs mesures réalisées simultanément (spectre RMN par exemple).

La seconde distinction peut se faire suivant le caractère descriptif ou explicatif de la méthode statistique. Une analyse statistique descriptive (ou exploratoire) vise à structurer et simplifier

un ensemble d'observations (de variables), dans son ensemble sans privilégier l'une d'entre elles en particulier. Ces méthodes permettent ainsi de visualiser rapidement une grande quantité d'informations. Cette famille regroupe notamment l'Analyse en Composantes Principales (ACP), l'Analyse Factorielle des Correspondances (AFC) et l'Analyse des Correspondances Multiples (ACM).

Une analyse statistique explicative cherche, quant à elle, à expliquer une variable à l'aide de deux ou plusieurs variables explicatives et étudie la relation entre une variable dite dépendante et un ensemble de variables explicatives. De nombreuses méthodes statistiques en font ainsi partie telle que : l'analyse de régression multiple, l'analyse discriminante, l'analyse conjointe, l'analyse de la variance...

Une autre distinction peut se faire suivant le caractère supervisé ou non supervisé de la méthode statistique. Une méthode statistique est dite supervisée, ou non, suivant le fait qu'elle requiert ou non des connaissances supplémentaires que les données expérimentales, comme par exemple le groupe d'appartenance des échantillons lors d'une classification d'échantillons. L'ACP ou la CAH sont des exemples d'analyses statistiques non supervisées tandis que la PLS-DA ou SIMCA sont des techniques supervisées.

Au cours de cette section, nous allons ainsi introduire les différentes méthodes statistiques multivariées employées pour le traitement des données RMN enregistrées dans le cadre de cette thèse, à savoir, l'ACP, SIMCA, PLS-DA et OPLS-DA. Nous présenterons également les différentes étapes du traitement des données préalables à l'application des analyses statistiques ainsi que les différentes procédures de validation croisée qui sont des méthodes d'estimation de la fiabilité des modèles statistiques.

Les différentes équations de cette section sont issues des livres édités par Kowalski¹⁰⁴ et Marini¹⁰⁵.

III.2) L'Analyse en Composantes Principales (ACP)

L'Analyse en Composantes Principales fut introduite en 1901 par Pearson¹⁰⁶ puis développée et formalisée en 1933 par Hotelling¹⁰⁷. Un de ses principaux objectifs est de simplifier et de réduire un jeu de données à X_n variables initiales corrélées, en un nouveau jeu de données à nombre réduit de variables non corrélées appelées composantes principales (CP_n) par rotation des X_n axes initiaux. Les composantes principales (CP_n) obtenues sont des combinaisons linéaires des variables initiales, pondérées de coefficients appropriés de telle sorte que la CP_1

(la première composante principale) incorpore le maximum de variations, CP₂ la deuxième plus grande variation et ainsi de suite.

Dans le cas des données RMN, le nombre de variables générées à partir des données brutes (spectres) peut être extrêmement conséquent, pouvant aller jusqu'à plusieurs centaines voire milliers, il est donc nécessaire de réduire la dimension des données d'origine. De plus, les données initiales sont le plus souvent inter-corrélées ce qui provoque une certaine redondance rendant compliquer les calculs statistiques.

Par ailleurs, l'ACP par la réduction de la dimensionnalité des données simplifie considérablement la visualisation des données. En effet, les nouvelles coordonnées obtenues pour chaque variable appelées scores peuvent facilement être visualisées sur un graphique à deux ou trois dimensions construit à l'aide des nouveaux axes et où ces nouveaux points forment des groupes en fonction de leur similarité. La correspondance chimique de ces similarités peut être obtenue par l'interprétation des *loadings* qui représentent les coefficients de corrélation linéaires entre les variables d'origine et les facteurs (composantes principales). En d'autres mots, cela correspond aux coordonnées des variables d'origine dans le nouvel espace.

Par ailleurs, un autre intérêt de l'ACP est qu'elle permet la détection de points aberrants et de les exclure avant d'utiliser d'autres analyses statistiques, ces points pouvant conduire à une perte d'informations ou une mauvaise interprétation des résultats (l'ACP tirant le modèle de façon à ce que celui-ci passe par ces points extrêmes et modifie ainsi la direction des composantes principales).

Enfin, dans certains cas favorables, l'ACP peut servir de technique de classification des échantillons. En effet, l'ACP va regrouper les points partageants des similitudes tandis que les points différents apparaîtront éloignés. La séparation obtenue se fait sans attribution au préalable des échantillons à un groupe d'appartenance, l'ACP est donc qualifiée de méthode statistique descriptive non supervisée.

III.2.1) Approche géométrique

L'ensemble des données brutes peut être regroupé au sein d'une matrice X constituée de n lignes correspondant aux différents objets (échantillons) et p colonnes (variables). La représentation graphique de l'ensemble de cette matrice correspond à un espace à p

dimensions ou chaque objet est représenté par un point. La matrice X correspond donc à un nuage de points.

L'ACP peut être considérée comme une méthode permettant de décrire la localisation et la répartition de tout ce nuage de points à l'aide de paramètres statistiques tels que les moyennes et la variance le long de différentes directions de cet espace p .

De manière générale, l'ACP peut être comprise comme la décomposition de la matrice X en une matrice de scores T et une matrice de *loadings* P à laquelle est ajoutée une matrice E correspondant au bruit résiduel :

$$X = \mathbf{1} \cdot \bar{x} + T \cdot P + E \quad (1)$$

D'un point de vue géométrique, l'ACP est une projection orthogonale de X dans un système de coordonnées dont les axes correspondent aux différents vecteurs lignes P .

La première étape consiste en la rotation de l'axe d'origine de façon à ce que ce nouvel axe passe par l'ensemble des points du nuage de points de façon à obtenir un étalement maximum des projections des points sur cet axe c'est-à-dire de telle sorte que les déviations soient les plus petites possibles au sens des moindres carrés (Figure 9).

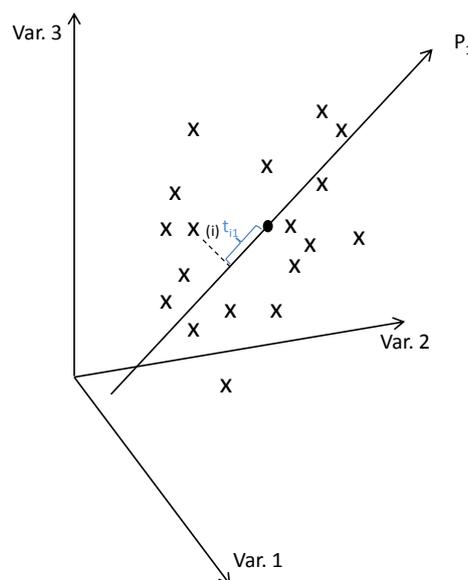


Figure 9 : Construction de la première composante principale par rapport au nuage de points des données initiales. Les scores (t_{1i}) sur la première composante principale (P_1) pour chaque observation correspondent aux projections de ceux-ci sur ce nouvel axe.

Les coefficients de direction de cet axe appelés *loadings* et dénotés p_{1k} formant le vecteur ligne P_1 correspondant à la première composante. Les points sont alors projetés sur cet axe nous obtenons les scores t_{1j} c'est-à-dire les coordonnées de ces points sur l'axe P_1 . La soustraction de t_{1jk} de x_{ik} donne les résidus e_{ik} . La première composante P_1 représente ainsi la plus grande quantité de variance expliquée contenue dans les données d'origine.

L'ensemble des résidus E est alors remplacé dans l'équation précédente par X . Cela correspond à supprimer la direction P_1 des données. La nouvelle matrice X est utilisée pour construire un nouvel axe suivant la même procédure précédemment appliquée. La deuxième composante P_2 ainsi générée, représente alors la deuxième plus importante quantité de variance expliquée avec la restriction d'être orthogonale à P_1 (Figure 10).

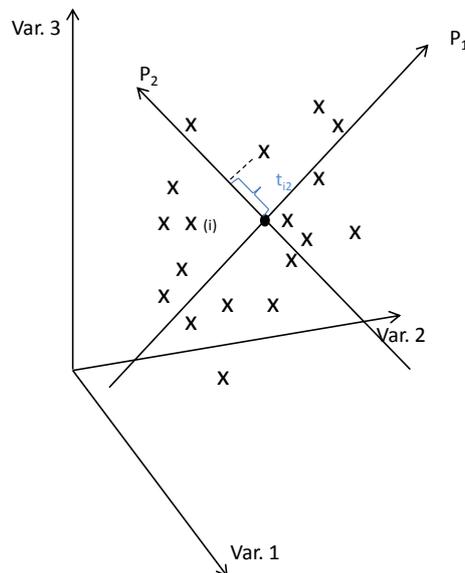


Figure 10 : Construction de la seconde composante principale (P_2) par rapport au nuage de points des données initiales. Celle-ci est perpendiculaire à la première composante et les scores pour chaque observation sont obtenus de la même manière que pour P_1 (projection sur P_2).

Cette procédure est répétée jusqu'à ce toute la variation soit expliquée c'est-à-dire jusqu'à ce que les résidus soient suffisamment petits. Un nouveau système de coordonnées a ainsi été construit avec P axes et où les n objets peuvent être représentés par des points dans ce nouvel espace.

Il est possible de représenter graphiquement l'ensemble des résultats de l'ACP de deux manières différentes. La première consiste à représenter les scores dans un repère P_1, P_2 (Figure 11a). Sur cette représentation des groupes d'objets peuvent être observés et les

distances entre les points correspondent à l'approximation des distances entre les points dans l'espace.

La deuxième représentation correspond à celle des *loadings* en fonction de P_i et P_j (Figure 11b). Les variables avec peu d'importance sur l'explication des données sont situées au centre de ce graphique (coefficients de corrélation faibles) tandis que celles possédant un fort coefficient de corrélation avec la composante principale sont éloignées du centre graphique. Plus ces *loadings* sont éloignés du centre, plus leur importance est significative.

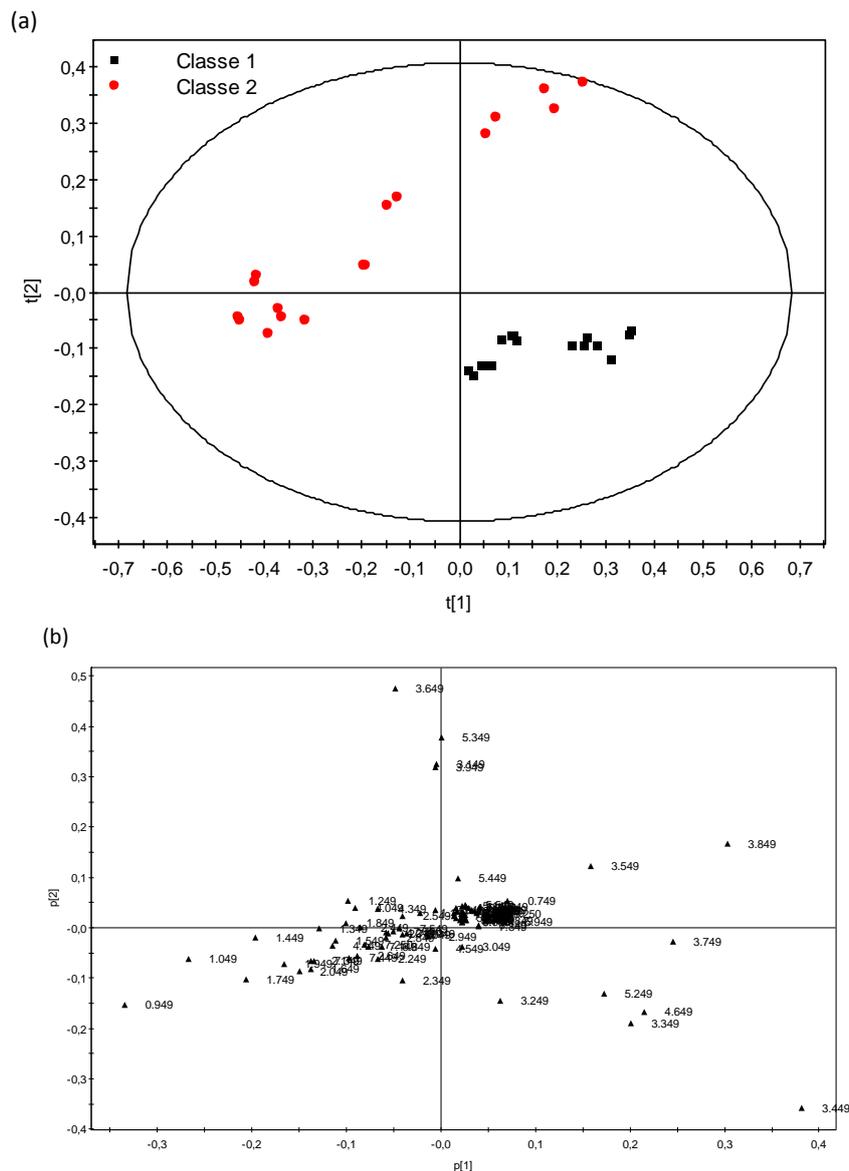


Figure 11 : Exemple de (a) graphe des scores et (b) graphe des *loadings* ; en fonction de CP_1 et CP_2 obtenus à l'aide de l'ACP sur des échantillons de caviar. Sur le graphe des scores, une distinction est faite entre les deux classes correspondant à deux espèces d'esturgeon différentes. Le graphe des *loadings* permet de déterminer quelles sont les variables initiales (origine chimique) responsables de la discrimination entre les deux groupes. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

III.2.2) Approche algébrique

D'un point de vue algébrique, l'ACP peut être formulée comme un problème de maximisation mathématique sous certaines contraintes. Nous avons vu précédemment que les composantes principales sont des combinaisons linéaires des variables d'origine telles que :

$$t_a = X \cdot p_a \quad (2)$$

Où p_a , les vecteurs des *loadings*, sont soumis à certaines restrictions : $p_a^T p_a = 1$ (contrainte de normalisation), $p_a^T p_b = 0$ (contrainte d'orthogonalité) et de telle sorte que la variance expliquée par la première composante $var(p_a)$ soit maximale. Ainsi l'expression à maximiser est, pour $a=1 \dots j$:

$$(X p_a)^T (X p_a) = p_a^T X^T X p_a = p_a^T cov(X) p_a \quad (3)$$

Où cov représente la covariance et la solution peut être reformulée sous la forme d'un problème de vecteurs propres/valeurs propres pour chaque valeur de a :

$$cov(X) p_a = \lambda_a p_a \quad (4)$$

Cela signifie que les valeurs inconnues pour les *loadings* correspondent aux vecteurs propres de la matrice de covariance X et λ sont les valeurs propres correspondantes. En d'autres mots, le calcul des composantes principales consiste en la diagonalisation de la matrice de covariance de X .

Par conséquent, les composantes principales sont rangées par ordre décroissant de variance et en considérant la propriété algébrique de conservation de la trace qui pour n'importe quelle matrice B à laquelle est associée sa matrice diagonale D implique:

$$trace(B) = trace(D) \quad (5)$$

La somme des valeurs propres équivaut alors à la variance totale de la matrice X :

$$\sum_a \lambda_a = var(X) \quad (6)$$

La proportion de variance P_{var} expliquée pour la $j^{\text{ème}}$ composante peut alors s'écrire :

$$P_{var} = \frac{\lambda_j}{\sum_a \lambda_a} \quad (7)$$

Les *loadings* peuvent ainsi être obtenus par plusieurs algorithmes de calculs des vecteurs propres/valeurs propres. Les deux algorithmes les plus couramment utilisés sont NIPALS (« Non-linear Iterative PARTial Least Squares »)¹⁰⁸ et SVD (« Singular Value Decomposition »)¹⁰⁹. Lorsque le nombre de variables mesurées est largement supérieur au nombre d'individus, l'algorithme NIPALS est le plus adapté, celui-ci pouvant être stoppé dès qu'un certain critère est atteint comme le pourcentage désiré de variance expliqué par exemple.

III.3) Soft Independent Modelling of Class Analogy (SIMCA)

SIMCA (Soft Independent Modelling of Class Analogy) est une méthode statistique descriptive supervisée de classement de données développée par Wold et al en 1976¹¹⁰ basée sur l'ACP qui requiert un ensemble de données de calibration où l'appartenance de chaque objet à une classe est définie en amont. L'hypothèse de base de cette méthode est, que la principale variabilité systématique caractérisant les échantillons d'une classe, peut être capturée par un modèle construit à l'aide d'une ou plusieurs composantes principales de dimensionnalité déterminée par validation croisée sur le jeu de données de formation. SIMCA génère autant de modèles différents qu'il n'y a de classes.

A partir de la matrice X_A correspondant aux données collectées sur les échantillons du groupe A, SIMCA suppose que l'information importante en termes de similarité à travers les échantillons de cette classe, peut être capturée par un modèle ACP à C composantes selon l'équation suivante :

$$X_A = T_A P_A + E_A \quad (8)$$

Où T_A et P_A sont respectivement les matrices des scores et des *loadings* obtenus sur les C composantes des échantillons du groupe A et E_A la matrice contenant les résidus non modélisés de X_A .

Une fois la décomposition calculée par l'ACP, les résultats sont utilisés pour définir le modèle de la classe en question.

Initialement, dans la première version de l'algorithme SIMCA, seulement l'information contenue dans les résidus est utilisée pour vérifier si un échantillon appartient ou non à une certaine classe. La distance entre un objet p et le modèle de la classe A est reliée à la déviation standard résiduelle selon :

$$s_{0,A} = \sqrt{\frac{\sum_{i=1}^{n_A} \sum_{j=1}^v e_{ij,A}^2}{(v - C)(n_a - C - 1)}} \quad (9)$$

Où v représente le nombre de variables, n_A le nombre d'échantillons de la classe A , C le nombre de composantes principales du modèle et e_{ij} la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne de la matrice des résidus E_A .

$s_{p,A}$ peut être considérée la mesure de l'étendue, en moyenne, de la distance entre les différents points du modèle et le modèle construit. C'est ce concept de distance au modèle qui constitue la base de la classification des échantillons inconnus. Ainsi, afin de vérifier si un échantillon inconnu, caractérisé par un vecteur de mesures x_n appartient ou non à une classe, celui-ci est projeté sur l'espace en composantes principales défini dans l'équation (8) :

$$t_{n,A} = x_n P_A \quad (10)$$

Où $t_{n,a}$ est le vecteur contenant les scores de l'échantillon inconnu sur les C composantes principales du modèle de la classe en question. Ensuite, la représentation du modèle de l'échantillon inconnu $x_{n,a}$ est obtenu par une projection des scores au sein de l'espace original des variables :

$$x_{n,a} = t_{n,a} P_A \quad (11)$$

Il est alors possible de calculer le vecteur des résidus, noté $e_{n,A}$ pour l'échantillon inconnu par rapport à la représentation en composantes principales de la classe A comme étant la différence entre x_n et $x_{n,A}$:

$$e_{n,A} = x_n - x_{n,A} \quad (12)$$

Une fois le vecteur des résidus calculé, celui-ci est alors utilisé pour définir la distance de l'échantillon inconnu au modèle qui est exprimé comme nous l'avons vu précédemment par une déviation standard $s_{n,A}$ définie par :

$$s_{n,A} = \sqrt{\frac{\sum_{j=1}^y e_{nj,A}^2}{(v - C)}} \quad (13)$$

Où $e_{nj,A}$ représente le $j^{\text{ème}}$ élément du vecteur des résidus $e_{n,A}$. Pour vérifier si la distance de l'échantillon au modèle est comparable ou trop importante par rapport à la variabilité de la classe étudiée, un test F (test de Fisher), est appliqué pour comparer $s_{0,A}$ et $s_{n,A}$. Le test de Fisher mesure le rapport entre les deux variances et le compare à une certaine valeur théorique issue de la table de Fisher. Si la valeur est inférieure à celle théorique l'hypothèse nulle est acceptée, auquel cas celle-ci est rejetée.

Généralement, une valeur de 95 % est utilisée pour le test de Fisher (5% d'erreur), et une limite pour les valeurs des distances entre les échantillons inconnus et le modèle de la classe A est définie par :

$$s_{crit,A} = \sqrt{F_{0.95,v-C,n_A-C-1} \cdot s_{0,A}} \quad (14)$$

Ainsi, si la distance de l'échantillon au modèle de la classe A $s_{n,A}$ est plus petite que la valeur du seuil , alors celui-ci est accepté par la classe A sinon il est rejeté. Si plusieurs classes sont présentes, cette procédure est alors appliquée à l'ensemble des modèles.

Les mêmes concepts sont utilisés pour les différentes améliorations existantes dans la littérature apportées à cette première version de SIMCA. Wold proposa ainsi, une première amélioration en 1977¹¹¹, tenant compte, non seulement de la distance orthogonale d'un objet

au modèle (résidus), mais également de la distance entre celui-ci et les autres observations dans l'espace des composantes principales (distance en termes de scores).

Nous allons à présent nous intéresser à une version alternative, qui est celle que nous allons utiliser dans cette thèse, où le degré d'appartenance d'un individu est évalué à l'aide d'une procédure empruntée aux contrôles de processus statistiques multivariés (Multivariate Statistical Process Control)¹¹². Après avoir calculé le modèle statistique selon l'équation (8) le critère d'acceptabilité de cette méthode alternative est défini à l'aide d'une distance orthogonale (notée Q) et d'une distance en termes de scores (notée T²) de la façon suivante.

Tout d'abord, pour mesurer la distance d'un échantillon dans l'espace des scores, une variable statistique notée T², qui représente le carré de la distance de Mahalanobis¹¹³ est introduite :

$$T_{k,A}^2 = t_{k,A}^T \Lambda_A^{-1} t_{k,A} \quad (15)$$

Où $t_{k,A}$ est le vecteur colonne rassemblant les scores du k^{ème} échantillon sur les composantes principales du modèle de la classe A, Λ_A est une matrice diagonale possédant autant d'éléments que de valeurs propres des C composantes principales et $T_{k,A}^2$ est la valeur T² pour le k^{ème} échantillon par rapport au modèle de la classe A.

De plus, la somme des carrés résiduels, notée Q, est utilisée pour définir la distance orthogonale :

$$Q_{k,A} = \sum_{j=1}^v e_{kj,A}^2 \quad (16)$$

Où $Q_{k,A}$ est la valeur Q pour le k^{ème} échantillon par rapport au modèle de la classe A.

Ensuite, les deux limites pour T² et Q sont calculées basées sur 95 % de leurs distributions correspondantes :

$$T_{0.95,A}^2 = F_{0.95,C,n_A-C} \frac{C(n_A^2 - 1)}{n_A(n_a - C)} \quad (17)$$

Où $T_{0.95,A}^2$ est la limite critique de T^2 pour la classe A, $F_{0.95,C,n_A-C}$ le 95^{ème} centile de la distribution F correspondant aux C et $(n_A - C)$ degrés de liberté.

Et la limite de Q est calculée en utilisant l'approximation de Jackson et Mudholkar¹¹⁴ :

$$Q_{0.95,A} = \theta_{1,A} \left[1 - \frac{\theta_{2,A} h_{0,A} (1 - h_{0,A})}{\theta_{1,A}^2} + \frac{z_{0.95} \sqrt{(2\theta_{2,A} h_{0,A}^2)}}{\theta_{1,A}} \right]^{\frac{1}{h_{0,A}}} \quad (18)$$

$$\text{Avec } h_{0,A} = 1 - 2\theta_{1,A}\theta_{3,A}/3\theta_{2,A}^2 \quad (19)$$

$$\text{Et } \theta_{k,A} = \sum_{l=C+1}^{rang(X_A)} \lambda_{l,A}^k \quad (20)$$

Où $Q_{0.95,A}$ est la limite critique de Q pour la classe A, $z_{0.95}$ est le 95^{ème} centile de la distribution normale standardisée, $\lambda_{l,A}^k$ est la valeur propre e la lème composante principale pour la classe A et la somme de l'équation (20) est calculée pour toutes les composantes non incluses dans la définition du modèle de la classe A.

Une fois que ces deux valeurs limites ont été déterminées, la distance d'un échantillon k au modèle de la classe A, notée $d_{k,A}$ est exprimée selon :

$$d_{k,A} = \sqrt{(T_{red_{k,A}}^2)^2 (Q_{red_{k,A}})^2} = \sqrt{\left(\frac{T_{k,A}^2}{T_{0.95,A}^2}\right)^2 + \left(\frac{Q_{k,A}}{Q_{0.95,A}}\right)^2} \quad (21)$$

Où $T_{red_{k,A}}^2$ et $Q_{red_{k,A}}$ correspondent respectivement au T^2 réduit et Q réduit qui sont les valeurs pour l'échantillon k lorsqu'il est projeté sur le modèle de la classe A, normalisés par leur valeurs limites correspondantes pour la classe A. Une valeur de $\sqrt{2}$ est couramment fixée comme critère d'acceptation ou de rejection d'un échantillon à un modèle donné. Ainsi, si $d_{k,A} \leq \sqrt{2}$, l'échantillon k est considéré comme appartenant au modèle de la classe A auquel cas celui-ci est rejeté.

Un exemple de modèle SIMCA est présenté en Figure 12. Sur celle-ci une valeur de 1 est fixée comme critère d'acceptation.

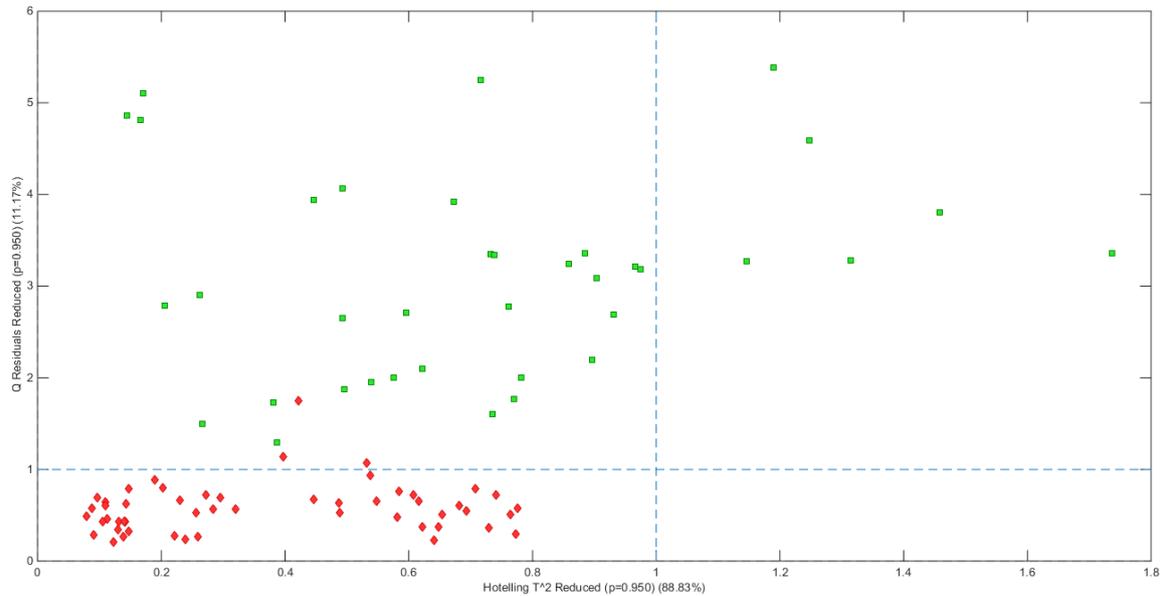


Figure 12 : Exemple de modèle SIMCA à 5 composantes principales, obtenu à l'aide des données acquises sur des échantillons de caviar. En rouge les points utilisés pour la construction du modèle et en vert ceux utilisés pour le classement des échantillons inconnus. L'axe des abscisses représente les valeurs de $T_{red,k,A}^2$ tandis que celui des ordonnées représente les valeurs de $Q_{red,k,A}$. Les limites du modèle sont délimitées par les axes en pointsillés bleu (valeur limite de 1 dans le cas présent).

III.4) La régression PLS discriminante (Partial Least Square regression-Discriminant Analysis)

Tout comme l'ACP, la régression des moindres carrés partiels (PLS) est une méthode statistique de réduction des données, cependant elle diffère de l'ACP par son caractère supervisé. La régression PLS requiert tout comme SIMCA, une prédéfinition à un groupe d'appartenance (classe) des différents objets du jeu de données.

La régression PLS discriminante (PLS-DA) est une extension de la régression PLS et peut facilement s'adapter au cas de la classification supervisée surtout lorsque le nombre de variables explicatives est nettement supérieur au nombre d'individus et/ou lorsque les variables sont corrélées. En effet, dans ce cas précis, l'utilisation de techniques de classement discriminante telle que l'Analyse Discriminante Linéaire (ADL) ou l'Analyse Discriminante Quadratique (ADQ) est inappropriée.

La régression PLS étant initialement introduite pour la construction de modèles de calibration, il est nécessaire de reformuler le problème de classification sous la forme d'une équation de régression. Ceci est accompli à travers l'introduction d'une matrice factice Y , qui codifie

l'appartenance des échantillons à une classe. Cette matrice Y possède autant de colonnes que de nombre de classes définies et autant de lignes que de nombre d'échantillons. Pour chaque observation dans le jeu de données de formation, son appartenance à une classe donnée est codée de manière binaire (1 pour sa classe, 0 sinon). Par exemple, si 4 classes sont définies et que l' $i^{\text{ème}}$ échantillon appartient à la classe 3 la $i^{\text{ème}}$ ligne (y_i) de la matrice Y vaut :

$$y_i = [0 \ 0 \ 1 \ 0] \quad (22)$$

En utilisant ce codage binaire pour l'appartenance des classes, il est alors possible de transformer le problème de classification en un problème de régression à travers une fonction associant la matrice de données X contenant les variables mesurées sur les échantillons et la matrice factice Y telle que :

$$Y = f(X) \quad (23)$$

En supposant qu'il existe une relation linéaire entre ces deux matrices, l'équation précédente devient :

$$Y = XB \quad (24)$$

où B est une matrice de coefficients de régression. L'approche PLS peut alors être utilisée pour calculer le modèle et ceci même lorsque le nombre d'échantillons est inférieur au nombre de variables. La méthode de classification correspondante est appelée régression PLS discriminante (PLS-DA).

La régression PLS suppose que la matrice indépendante X et la matrice dépendante Y puissent être projetées sur un espace à dimensions réduites et qu'une relation linéaire existe entre les scores des deux blocs. D'un point de vue mathématique, X et Y sont décomposées en matrice de scores et de *loadings* selon les équations suivantes :

$$X = TP^T + E_X \quad (25)$$

$$Y = UQ^T + E_Y \quad (26)$$

où T et U sont les matrices contenant respectivement les scores de X et de Y, P et Q sont celles contenant les *loadings* de X et les *loadings* de Y, et E_X et E_Y celles contenant les résidus.

De plus, la dépendance linéaire entre les scores de X et de Y supposée précédemment implique :

$$U = TC \quad (27)$$

où C est une matrice diagonale des coefficients. En se basant sur les deux équations précédentes (26) et (27), il est possible de calculer une matrice contenant les coefficients de régression, notée B, qui permet de prédire les valeurs de la matrice dépendante Y_n pour les échantillons inconnus à laquelle sont associées les variables mesurées X_n :

$$Y_n = X_n B \quad (28)$$

Il est important de souligner qu'à la différence de la matrice factice Y qui est codée en binaire, les valeurs prédites sur les échantillons inconnus Y_n seront des nombres réels. Ainsi, si 4 classes sont impliquées dans le problème de classification comme ce fut le cas de notre exemple précédent, quand le modèle est testé sur un échantillon inconnu, le résultat se présentera sous la forme d'un vecteur de valeurs réelles à 4 dimensions qui pourra prendre des valeurs telles que [0.05 -0.16 0.88 0.09]. La classification de l'échantillon est alors réalisée en assignant cet échantillon à la catégorie (classe) correspondant à la valeur la plus élevée des valeurs prédites, dans notre exemple à la classe 3 (0.88).

Puisque la PLS est une technique basée sur le nombre de composantes, lors de la construction du modèle, il est nécessaire d'estimer le nombre de variables latentes opportun afin d'obtenir un modèle fiable et d'éviter la surestimation des résultats. Cette procédure est le plus souvent réalisée par validation croisée comme nous le verrons plus tard dans cette **Partie** (III-6)).

Une fois le modèle calculé, il est essentiel d'interpréter les résultats afin d'identifier les variables initiales qui sont les plus significatives pour la discrimination. En effet, il est important d'identifier les éventuels marqueurs chimiques et de s'assurer que cette différence ne provient pas d'erreurs de prétraitement des données ou de sur-apprentissage (ou sur-ajustement).

La méthode la plus courante pour estimer la contribution des variables individuelles au modèle dans le cadre de la PLS-DA, est l'inspection de l'importance des variables dans la projection (Variable Importance in Projection, notée VIP). VIP est un index qui a été spécialement introduit pour exprimer à quel point un prédicteur est significatif dans la définition des vecteurs latents du modèle. D'un point de vue mathématique, le score VIP d'une variable j est défini tel que :

$$VIP_j = \sqrt{N_{var} \frac{\sum_{k=1}^F (c_k^2 t_k^T t_k) (w_{jk} / \|w_k\|)^2}{\sum_{k=1}^F (c_k^2 t_k^T t_k)}} \quad (29)$$

Où t_k est le vecteur des scores des échantillons le long de la $k^{\text{ème}}$ variable latente, c_k est le coefficient de la $k^{\text{ème}}$ variable latente, N_{var} le nombre de variables expérimentales et w_{jk} est le poids de la $j^{\text{ème}}$ variable pour la $k^{\text{ème}}$ variable latente et w_k le vecteur poids pour la $k^{\text{ème}}$ variable latente.

L'avantage d'utiliser les scores VIP pour estimer la contribution des variables initiales au modèle PLS-DA est qu'il peut être démontré que la moyenne des scores VIP au carré est égale à 1. Ainsi, tous les scores supérieurs à 1 correspondent aux variables significatives. Plus cette valeur est élevée, plus cette variable est importante dans la construction du modèle.

L'interprétation des résultats peut être améliorée par l'inspection des coefficients de régression du modèle PLS-DA qui, soigneusement examinés, peuvent indiquer si les valeurs des différentes variables mesurées sur les échantillons provenant d'une certaine classe, sont supérieures ou inférieures de ceux enregistrés pour les autres classes.

En effet, comme l'équation (24) l'indique, les coefficients de régression sont les poids linéaires associés avec les variables indépendantes pour la prédiction des différentes composantes de Y . Ainsi plus leur valeur est élevée, plus la contribution du prédicteur à la composante spécifique de Y est importante. En effet, la $k^{\text{ème}}$ colonne de B contient les coefficients des variables X pour la prédiction de $k^{\text{ème}}$ composante de Y . Cependant, certains facteurs doivent être pris en compte dans l'interprétation des coefficients de régression. Tout d'abord, l'amplitude des éléments de B est associée à celle de X . Ainsi, la valeur absolue des coefficients de régression associée à une variable peut être plus élevée que les autres juste à cause d'une faible amplitude de cette variable mesurée impliquant un coefficient élevé afin d'être consistant avec y et non pas à cause de sa contribution. L'interprétation de ces

coefficients est rendue beaucoup plus intuitive lorsque les données ont été préalablement réduites.

Analyse Discriminante par projections orthogonales des moindres carrés partiels (Orthogonal Partial Least Square-Discriminant Analysis, OPLS-DA)

L'OPLS-DA, introduite par Trygg et Wold¹¹⁵, est une modification récente de la PLS-DA dont l'objectif est d'améliorer l'interprétation des modèles PLS et de réduire la complexité du modèle. Comme nous l'avons vu précédemment, l'utilisation de classes dans la PLS-DA permet à l'algorithme PLS d'améliorer la séparation des groupes dans l'espace des scores. L'O-PLS fournit une méthode pour éliminer la variation systématique des données d'entrée X qui n'est pas corrélée au jeu de réponse Y, en d'autres mots cela revient à supprimer la variabilité dans X qui est orthogonale à Y. Cette technique analyse la variation gênante de chaque composante PLS. La variation corrélée à X est séparée de celle non-corrélée, cette dernière pouvant alors être analysée et étudiée séparément. La suppression de la variation non corrélée des données, préalablement à la modélisation, n'est pas seulement intéressante d'un point de vue prédictif, elle facilite également l'interprétation des modèles prédictifs.

La méthode O-PLS proposée par Trygg et al¹¹⁵, consiste en une modification de l'algorithme NIPALS originel. Celle-ci peut être vue comme une méthode de prétraitement pour éliminer la variation orthogonale systématique d'un jeu de données X (Figure 13).

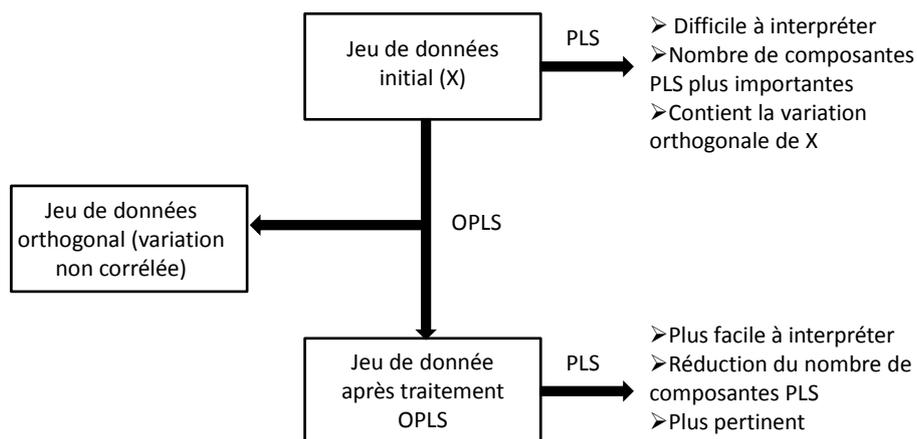


Figure 13 : Schéma de l'apport de l'OPLS vis-à-vis de la PLS inspiré de la publication de Trygg et Wold¹¹⁵. La variation systématique non corrélée à Y est séparée du jeu de données facilitant l'interprétation et réduisant la complexité de modèles PLS. Celle-ci peut alors être étudiée et analysée séparément et les sources de variation orthogonale peuvent être identifiées.

Le modèle O-PLS est constitué de deux variations modélisées, celle prédictive de Y ($T_p P_p^T$) et celle orthogonale de Y ($T_0 P_0^T$). En revanche, seule la variation prédictive de Y est utilisée pour la modélisation de Y. Le modèle OPLS est ainsi défini par :

$$\text{Modèle de X : } X = T_p P_p^T + T_0 P_0^T + E \quad (30)$$

$$\text{Modèle de Y : } Y = T_p C_p^T + F \quad (31)$$

Où T_p et T_0 sont les matrices contenant les scores, P_p et P_0 celles contenant les *loadings*, E et F représentent les matrices contenant les résidus de X et Y respectivement. L'O-PLS peut alors être appliquée à la PLS-DA et être utilisée pour la discrimination (OPLS-DA). La variation entre les classes et la variation interclasses sont alors séparées par l'OPLS-DA ce qui n'est pas le cas lors d'une PLS-DA, facilitant l'interprétation du modèle OPLS-DA. En effet, les scores et les loadings prédictifs de l'OPLS-DA utilisés pour estimer Y sont alors composés de la variation directement corrélée avec Y exempte de celle non corrélée, améliorant l'interprétation par rapport à une analyse PLS-DA classique¹¹⁶.

De plus, le risque de sur-apprentissage est grandement réduit par l'utilisation de méthodes de validation appropriée et/ou de critères sur les valeurs propres, dans le but de déterminer le nombre de composantes orthogonales. Pour la validation croisée, il est courant d'employer une technique de validation croisée intégrale. L'approche basée sur les valeurs propres consiste, quant à elle, à analyser le rapport $\|p - [w^T p / (w^T w)] w\| / \|p\|$, qui devient nul pour les composantes PLS corrélées lorsqu'il n'y a plus de variation orthogonale au sein de X. Un graphique de ce ratio en fonction du nombre de composantes orthogonales fournit une bonne indication du nombre de composantes à extraire. Ce graphique est alors interprété de la même manière que pour le diagramme d'éboullis que nous présenterons dans la section III-6) de cette partie pour le cas de l'ACP.

Enfin, il est important de signaler que l'OPLS-DA ne fournit pas d'avantage prédictif par rapport à la PLS-DA ; en effet, en l'absence de variation non corrélée à Y au sein de X, les modèles résultants de ces deux méthodes statistiques seront quasi-similaires.

III.5) Prétraitement des données RMN

Dès lors que le nombre d'observations ou de variables devient trop important, ou lorsque les différences entre les individus sont faibles, l'utilisation de méthodes statistiques multivariées devient indispensable dans le but de mettre en évidence des différences ou des corrélations entre les différents échantillons.

Cependant, ces méthodes statistiques sont extrêmement sensibles au traitement préalable à leur analyse statistique. En effet, afin que les statistiques ne soient pas erronées et afin que les différences ou similarités mises en évidence par ces analyses soient bien réelles et non issues du prétraitement, il est indispensable que tous les spectres RMN aient été traités suivant le même protocole.

Tous les spectres doivent ainsi tout d'abord être précisément phasés de telle sorte que tous les pics soient en parfaite absorption, puis une correction de ligne de base est appliquée afin que le milieu du bruit correspondent à une intensité nulle. Effectivement, la correction de la ligne de base est une étape cruciale puisqu'elle affecte directement la valeur absolue de l'intégrale (aire). Les distorsions de la ligne de base peuvent provenir de l'instabilité ou d'imperfections de l'appareillage¹¹⁷, ou de la nature de l'échantillon (macromolécules, échantillons hétérogènes, échanges chimiques).

Les déplacements chimiques de certains composés sur un spectre RMN sont extrêmement sensibles au pH et des variations même très légères du pH peuvent entraîner des décalages en fréquence et ainsi fausser les aires mesurées pour chaque intervalle. En effet, les déplacements chimiques de certains signaux sont affectés par les champs électriques locaux provenant des groupes chargés ou polaires. Les charges positives ont tendance à déblindier les protons voisins.

Comme nous le verrons dans la **Partie III**, pour corriger ces variations locales, une solution de tampon, couramment utilisée pour les analyses métaboliques, est ajoutée dans le but de corriger les variations de pH entre les échantillons. Cependant, au vu du grand nombre d'analyses réalisées il a été nécessaire de commander plusieurs fois du tampon et des variations même infimes entre les différentes solutions entraînent des décalages pour les pics les plus sensibles comme le montre la Figure 14a. Pour s'affranchir de ce problème, des fonctions Matlab (MathWorks) ont été développées, basées sur la fonction « icoshift »¹¹⁸. Cette fonction découpe les spectres en intervalles réguliers (0.25ppm par exemple) et réaligne les différents pics en fonction d'une référence choisie (spectre moyen ou spectre de référence)

avec comme critère de performance la maximisation du coefficient de corrélation croisée obtenu par transformation de Fourier. Certains intervalles ont également dû être redéfinis pour permettre un réalignement optimal. Les résultats après réalignement sont donnés en Figure 14b).

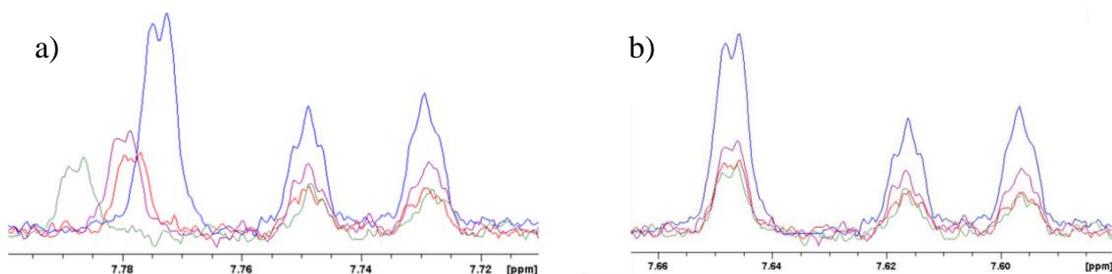


Figure 14 : Spectres ^1H NOESY 1D de la phase aqueuse du caviar enregistrés à 400 MHz a) avant réalignement et b) après réalignement à l'aide la fonction Matlab « icoshift ». L'utilisation de cette fonction permet de corriger les légères variations de pH des différentes solutions de tampon employées pour la réalisation des travaux de cette thèse et de permettre

Les fonctions Matlab développées ont été conçues de telle sorte qu'à l'issue du réalignement, les nouveaux spectres réalignés puissent être à la fois sauvegardés au sein un fichier de traitement différent de celui contenant les données initiales (PROCNO), mais également utilisés sous Matlab dans le but d'appliquer des traitements ou des méthodes statistiques directement sur ces nouvelles données.

Afin de rendre les données RMN utilisables pour une analyse statistique, le spectre RMN est ainsi découpé en petites régions, appelé « buckets » de taille constante ou variable et l'aire sous la courbe jusqu'à la ligne de base (intégrale) pour chacun de ces intervalles est mesurée. La taille de ces intervalles peut aller du point (0.001 ppm) jusqu'à 0.05 ou 0.10 ppm. Il est également possible d'utiliser des techniques de découpage automatiques du spectre, de telle sorte que chaque intervalle représente un pic ou un groupe de pics.

Ensuite, afin de corriger les variations d'intensité du spectre dues aux différences de concentration des échantillons liées aux imprécisions lors du prélèvement des différents volumes, la valeur de l'intégration de chaque bucket est divisée par l'intensité totale de tout le spectre en excluant la zone correspondant à la suppression du solvant. Des algorithmes plus élaborés comme PQN¹¹⁹ (Probabilistic Quotient Normalization) permettent également de corriger ces variations dans la préparation en calculant le facteur de dilution le plus probable à travers l'ensemble d'un spectre vis-à-vis d'un spectre de référence.

La dernière étape du prétraitement des données consiste à centrer les variables, c'est-à-dire à réaliser une translation du centre de gravité du nuage de données sur l'origine. Cette opération consiste à soustraire à chaque variable la moyenne de celle-ci, calculée sur l'ensemble des données, et ceci pour toutes les observations. Cette étape de centrage des données est réalisée de façon systématique en ACP puisque la première composante principale doit passer par l'origine et par l'ensemble du nuage de points afin de décrire un maximum de variance. Si l'origine ne se trouve pas au centre du nuage de points, la première composante principale décrira moins bien la longueur du nuage de données.

Une étape complémentaire au centrage des données est souvent réalisée à ce stade du prétraitement de celles-ci, il s'agit de la mise à l'échelle des données. En effet, la longueur d'une variable étant égale à son écart-type, plus la variable possède un écart-type élevé, plus elle apporte de l'inertie en projection et donc plus elle « attire les axes ». Ainsi, pour éviter d'accorder une plus grande importance aux variables exprimées avec de plus grandes valeurs, il est courant de les mettre à l'échelle en divisant la valeur des variables centrées par la valeur de l'écart-type pour chaque variable, la racine carrée de l'écart-type (Pareto) ou bien encore par transformation logarithmique. Cette procédure permet d'accorder la même importance à chaque variable et donc de permettre aux signaux de faible intensité d'avoir la même influence que ceux de plus grande intensité. Etant donné que les signaux RMN sont tous mesurés dans la même unité de mesure cette réduction est à appliquer au cas par cas puisque la mise à l'échelle des données peut ainsi donner de l'importance à des signaux à faible signal/bruit dont la valeur mesurée est moins précise et plus sensible aux variations et différences de la ligne de base.

III.6) Sélection du nombre de composantes et validation croisée

De manière générale, quelle que soit la méthode statistique utilisée et présentée dans les premières sections de ce chapitre (ACP, OPLS-DA ou SIMCA), une question se pose : Combien de composantes doivent être utilisées ? En effet, le nombre de composantes doit être déterminé de telle sorte que toute l'information contenue dans le jeu de données initial soit utilisée tout en faisant attention à ne pas réaliser de sur-apprentissage.

La réponse diffère en fonction de la méthode statistique employée, nous allons donc présenter les différentes techniques ou critères utilisés pour déterminer le nombre optimal de composantes en séparant le cas de l'ACP de celui des techniques de classification.

Une décomposition classique de l'ACP consiste en la distinction de la variance expliquée et du bruit. Cependant, il est plus juste de la présenter sous la forme « information pertinente » + « autre variance expliquée » + bruit. Ainsi, l'établissement du nombre de composantes principales correspondant à l'information pertinente est crucial et extrêmement utile.

Un des règles les plus simples est de retenir un nombre de composantes correspondant à un pourcentage donné de la variance expliquée (critère de Jolliffe¹²⁰), par exemple 80-90%. Dans ce cas, le type de données et de prétraitement doivent être pris en compte. Par exemple, si les données n'ont pas été centrées la première composante principale expliquera un grand pourcentage de la variance de celles-ci.

Un autre règle est basée sur l'évaluation numérique des valeurs propres et consiste à ne garder que les valeurs propres (λ_a) qui sont supérieures à la moyenne des valeurs propres (critère de Kaiser¹²¹). Ce critère admet que dans le cas d'une parfaite indépendance entre les variables, les composantes principales sont les mêmes que les variables originales et de variance égales (l'ACP représente une rotation de ces axes) et que seules les valeurs supérieures à la moyenne sont jugées informatives. Dans le cas d'une ACP normée la moyenne est de 1 et toutes les composantes possédant une valeur inférieure à 1, seront écartées.

Enfin, un dernier moyen de déterminer le nombre de composantes principales optimales est la représentation graphique du diagramme d'éboulis (« scree-plot ») correspondant à la représentation des valeurs propres (λ_a) en fonction de chaque composante principale. Le principe consiste à rechercher s'il existe un point d'inflexion, également appelé « coude », sur la courbe de décroissance de l'inertie et à ne conserver que les composantes jusqu'à ce « coude » (critère de Cattell¹²²). Un exemple de diagramme d'éboulis est donné en Figure 15.

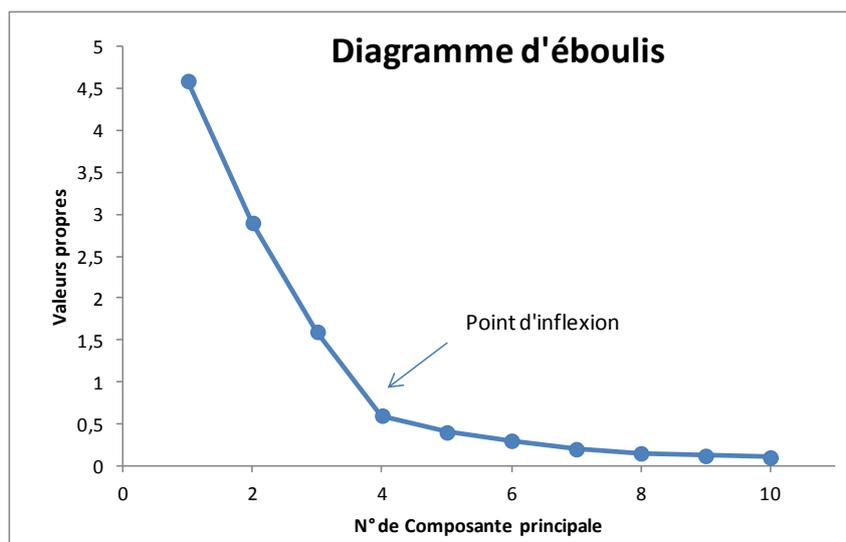


Figure 15 : Exemple de diagramme d'éboulis, représentant les valeurs propres (λ_a) en fonction de chaque composante principale. Dans le cas présent, le coude est obtenu au niveau de la 4^{ème} composante principale.

Différentes considérations sont à prendre dans le cas où l'ACP est utilisée comme technique de classification (SIMCA) ou dans le cas des régressions (OPLS-DA) et la détermination du nombre de composantes optimal est alors réalisée par validation croisée.

Avec la validation croisée, les données initiales sont utilisées à la fois pour la construction du modèle et pour l'évaluation de celui-ci.

Tout d'abord, le jeu de données initial est séparé en deux matrices : une matrice de calibration et une matrice de validation. Ainsi, certains individus sont alors exclus du jeu de données de calibration et le modèle est calibré sur les individus restants. Ensuite, les valeurs pour les objets exclus sont prédites et les résidus de prédiction sont calculés. Cette procédure est alors répétée avec un autre sous ensemble de jeu de données de calibration et ainsi de suite jusqu'à ce que tous les objets aient été exclus une fois. Une fois cette étape réalisée, tous les résidus de prédictions ainsi obtenus sont combinés pour calculer la variance résiduelle de validation et la racine carrée de l'erreur quadratique (« Root Mean Squared Error », notée RMSE) qui permettent d'estimer le score de validation croisée. Le modèle optimal (et donc le nombre de composantes optimal) est défini comme celui ayant obtenu le meilleur score de validation croisée.

Plusieurs schémas de validation croisée peuvent être employés :

- La validation croisée intégrale (« leave-one-out ») qui exclut seulement un objet à la fois. Les autres objets sont utilisés pour construire le modèle qui est par la suite appliqué à l'échantillon écarté. Cette procédure est répétée jusqu'à ce que tous les échantillons aient été exclus. Cette option est principalement utilisée pour les jeux de données peu importants.
- La validation croisée segmentée qui décompose en un certain nombre de segments aléatoires la matrice de données initiale, par exemple 10% des données dans chaque segment. Un des segments est exclu (qui servira à tester le modèle) et le modèle est construit à l'aide des autres segments. La procédure est alors répétée sur un autre segment, le segment précédemment exclu étant réincorporé aux données de construction du modèle, et jusqu'à ce que tous les segments aient été testés. Ce schéma de validation croisée est employé pour les jeux de données conséquents pour lesquels la validation croisée conduirait à un temps de calcul trop important.
- La validation croisée segmentée systématique qui exclut un certain nombre d'objets de façon logique. Un exemple typique de l'utilisation de cette méthode est le cas des répliques d'un échantillon. La matrice de données initiale est décomposée en plusieurs segments dont le nombre dépend du nombre de répliques.
- La validation croisée catégorique. Les données sont séparées en plusieurs matrices en fonction d'un paramètre spécifique (saison, fournisseur, région, année, opérateur, envoi...). Cette procédure permet d'évaluer la robustesse du modèle en fonction de paramètres particuliers.

Aucune de ces procédures de validation n'est préférable à une autre : le choix de la méthode employée dépend du jeu de données et est à réaliser au cas par cas. Il est notamment important de s'assurer que des spectres issus du même échantillon (duplicat ou triplicat) font partie du même groupe de calibration.

Maintenant que nous avons réalisé une présentation de la métabolomique et détaillé les procédures de prétraitement des données RMN en vue de leur analyse par les différentes méthodes statistiques décrites au sein de cette partie (ACP, SIMCA et OPLS-DA), nous allons présenter succinctement, dans la **Partie III**, les différentes séquences et paramètres RMN utilisées pour l'acquisition des spectres ainsi que les protocoles de préparation des échantillons de poisson par RMN HR-MAS et de caviar par RMN liquide. Une fois cette

nouvelle partie traitée, nous pourrions alors nous intéresser aux résultats obtenus lors de la réalisation des travaux de cette thèse, sur des matrices semi-solides par RMN HR-MAS puis à ceux de l'étude du caviar par RMN liquide (**Partie IV**).

Partie III- Matériels et Méthodes

I) Matériels

I.1) Le spectromètre RMN

L'ensemble de mes travaux de thèse ont été réalisés sur un spectromètre Bruker Avance III 400 opérant à la fréquence 399.90 MHz pour les noyaux ^1H , strictement réservé à l'analyse des échantillons pour le projet Agrifood GPS. L'ensemble du dispositif utilisé au sein de la société est donné en Figure 16.

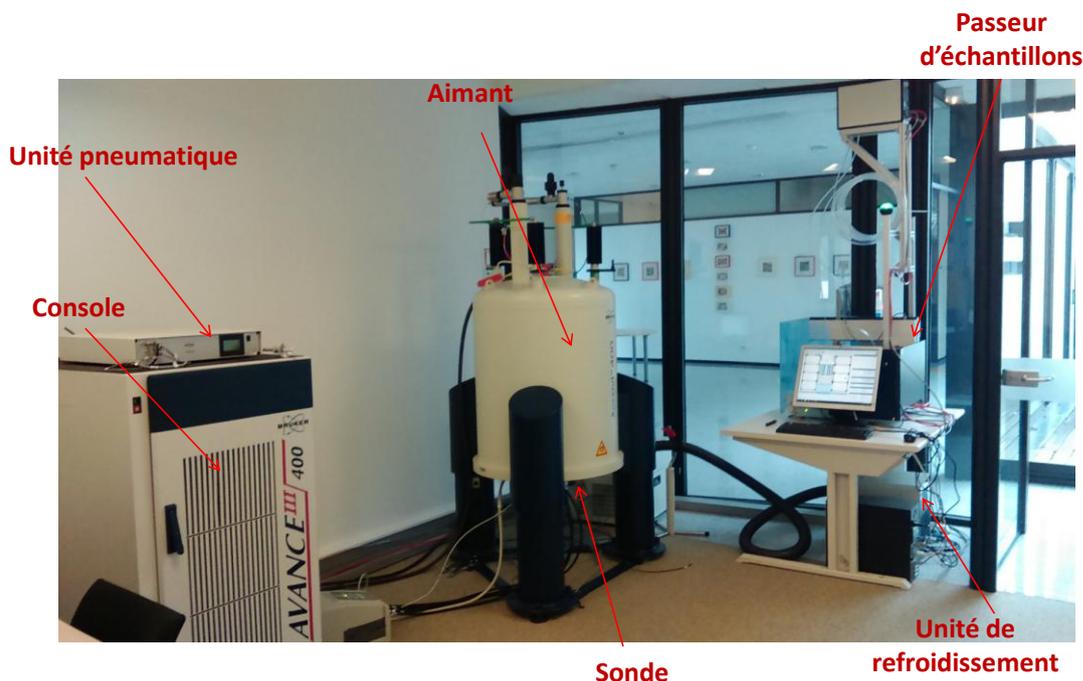


Figure 16 : Vue du spectromètre Agrifood (AVANCE III 400 MHz) utilisé pour l'acquisition de l'ensemble des données du projet. Celui-ci est équipé d'une unité pneumatique permettant la mise en rotation de l'échantillon, d'un passeur d'échantillon ainsi que d'une unité de refroidissement garantissant le maintien des échantillons à la température souhaitée.

I-2) HR-MAS

I.2.1) Instrumentation et spécificités du matériel en HR-MAS

I.2.1.a) Le rotor HR-MAS

Le rotor est un des éléments importants dans une expérience RMN puisque c'est à l'intérieur de celui-ci qu'est placé l'échantillon et il doit être conçu de manière à pouvoir supporter des

vitesses de rotation pouvant atteindre 15 kHz. Les rotors utilisés en RMN HR-MAS sont construits sur la base de ceux utilisés en RMN solide. Ces rotors en ZrO_2 dont le diamètre est de 4 mm existent sous différents volumes effectifs pour l'échantillon, les plus utilisés étant les rotors de 12 μl et 50 μl . Cependant, ces rotors présentent le désavantage de devoir être nettoyés à chaque changement d'échantillon. Afin de remédier à ce problème des rotors de 96 μl ont été développés dans le but de permettre l'insertion d'un insert jetable hermétique (fermé à l'aide d'un capuchon et d'une vis) de 33 μl en Kel-F à l'intérieur de celui-ci évitant ainsi toute contamination ou fuite de solvant. C'est ce type de rotor que nous avons utilisé pour l'ensemble de nos analyses et une vue schématique de l'ensemble du dispositif est donnée en Figure 17.

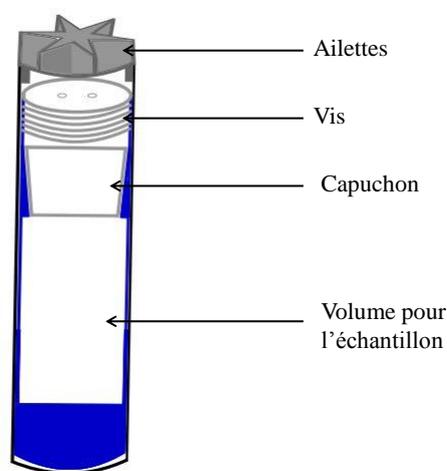


Figure 17 : Schéma du dispositif HR-MAS avec insert jetable (représenté en bleu) de 33 μl en Kel-F situé à l'intérieur d'un rotor. Ce dispositif est composé d'un capuchon et d'une vis permettant de rendre l'insert hermétique. Le rotor est alors fermé à l'aide d'un autre capuchon (ailettes) permettant la mise en rotation rapide du rotor.

I.2.1.b) La sonde HR-MAS

La sonde HR-MAS possède un rôle essentiel dans les performances d'un système HR-MAS. L'élément fondamental de la sonde est le stator qui contient la bobine radiofréquence dans laquelle est placé le rotor. La bobine utilisée est un solénoïde et non une bobine en « selle de cheval » également appelée bobine de Helmholtz, utilisée pour les sondes liquides. La bobine solénoïde présente, en effet, une meilleure sensibilité et une meilleure distribution du champ B_1 . De plus, la bobine solénoïde est mécaniquement plus résistante et ainsi supporte plus facilement les conditions sévères imposées par la mise en rotation de l'échantillon à haute

vitesse. Le stator est conçu de telle manière que celui-ci soit aligné avec le champ magnétique statique B_0 lors de l'insertion ou l'éjection du rotor, et qu'il puisse être basculé mécaniquement à l'angle magique pendant l'expérience RMN.

De plus, comme pour les sondes haute résolution liquide, il est nécessaire que la sonde soit équipée d'un canal ^2H . Le rôle de ce canal est de fournir une fréquence de référence pour le verrouillage champ/fréquence (« lock ») qui permet de corriger la déviation du champ magnétique principal au cours du temps par une bobine additionnelle qui produit un champ magnétique ΔB colinéaire à B_0 qui compense les variations de celui-ci.

Enfin, la principale amélioration des sondes HR-MAS par rapport à une sonde MAS classique est l'introduction d'une bobine gradient sur le stator qui permet de générer des gradients de champ magnétique linéaires dans la direction de l'angle magique¹²³ et ainsi l'utilisation de l'ensemble des séquences monodimensionnelles ou multidimensionnelles utilisées en RMN liquide haute résolution. En effet, les gradients de champ magnétique ont un rôle capital dans ces expériences du fait de leur excellente efficacité pour la sélection des chemins de cohérence et pour la suppression de certains artefacts, notamment la suppression du « bruit t_1 » pour les expériences multidimensionnelles.

L'ensemble des spectres des différents échantillons étudiés en HR-MAS ont été enregistrés à l'aide d'une sonde 4 mm Bruker HR-MAS DVT muni de gradient Z et composée de 2 canaux d'observation (^1H et ^{13}C) et d'un canal ^2H pour le « lock ». L'intérêt d'une sonde DVT vis-à-vis d'une sonde HR-MAS classique est qu'elle permet une régulation en température permanente indépendante de la rotation (du drive), contrairement aux sondes classiques qui ne régulent plus la température au sein de la bobine lorsque la rotation cesse. Ainsi, même sans mise en rotation, les échantillons sont maintenus à la température souhaitée. Le refroidissement de l'air injecté au sein de la sonde est assuré par une unité de refroidissement BCU II (Bruker Cooling Unit).

I.2.1.c) L'ajustement de l'homogénéité du champ magnétique en HR-MAS

Afin de corriger les inhomogénéités du champ magnétique B_0 , celui-ci est équipé de bobines additionnelles spéciales appelées « shims », qui produisent des champs magnétiques inhomogènes qui viennent se combiner à B_0 afin d'obtenir un champ homogène. Ces inhomogénéités proviennent des caractéristiques de la bobine supraconductrice (longueur,

forme, connexions), des perturbations extérieures, de la sonde (structure et constituants internes) et de l'échantillon. A la différence de la RMN liquide haute résolution, la technique HR-MAS ne dispose pas encore de procédures de méthodes de réglage automatique des « shims » tel que le « gradient shimming », cette étape doit donc être réalisée manuellement par l'utilisateur. En effet, l'ajustement de l'homogénéité du champ magnétique d'une sonde RMN HR-MAS diffère de celui d'un sonde liquide classique du fait de la rotation rapide de l'échantillon et de l'axe de rotation qui ne se trouve pas aligné avec le champ magnétique principal¹²⁴. Les « shims » dans un système d'axe orienté suivant l'angle magique peuvent être exprimés en fonction de ceux utilisés dans le système d'axe du laboratoire. Les principaux « shims » à régler en HR-MAS sont donnés en Tableau 1¹²⁴.

Cependant, pour un échantillon à l'angle magique et dans le cas d'une sonde HR-MAS avec un stator positionné dans le plan (x,z), seul l'ensemble des « shims » B_z^{MAS} , $B_{z^2}^{MAS}$ et $B_{z^3}^{MAS}$ sont suffisants pour ajuster convenablement l'homogénéité du champ magnétique d'un échantillon une fois que les autres « shims » ont été préalablement réglés lors de l'installation de la sonde¹²⁵.

D'un point de vue pratique, le niveau de « lock » peut être un bon moyen de régler les différents « shims », surtout pour les échantillons ne contenant pas d'eau tel que le chloroforme qui est le plus couramment utilisé. Cependant pour les échantillons contenant de l'eau, il est préférable de réaliser cette procédure en ajustant en temps réel l'allure d'un pic de référence tel que le TSP ou le TMS ou celui un métabolite de référence : des gradients de température pouvant apparaître avec la rotation du rotor¹²⁵.

Premier ordre
$B_z^{MAS} = \frac{1}{\sqrt{3}} B_z^{LAB} - \frac{\sqrt{2}}{\sqrt{3}} B_x^{LAB}$
$B_x^{MAS} = \frac{1}{\sqrt{3}} B_x^{LAB} + \frac{\sqrt{2}}{\sqrt{3}} B_z^{LAB}$
$B_y^{MAS} = B_y^{LAB}$
Deuxième ordre
$B_{z^2}^{MAS} = B_{(x^2-y^2)}^{LAB} - 2\sqrt{2} B_{zx}^{LAB}$
$B_{zx}^{MAS} = -\frac{1}{3} B_{zx}^{LAB} + \frac{\sqrt{2}}{6} B_{z^2}^{LAB} - \frac{\sqrt{2}}{6} B_{(x^2-y^2)}^{LAB}$
$B_{zy}^{MAS} = \frac{1}{\sqrt{3}} B_{zy}^{LAB} - \frac{\sqrt{2}}{\sqrt{3}} B_{xy}^{LAB}$
$B_{(x^2-y^2)}^{MAS} = \frac{2\sqrt{2}}{3} B_{zx}^{LAB} + \frac{1}{3} B_{z^2}^{LAB} + \frac{2}{3} B_{(x^2-y^2)}^{LAB}$
$B_{xy}^{MAS} = \frac{1}{\sqrt{3}} B_{xy}^{LAB} - \frac{\sqrt{2}}{\sqrt{3}} B_{zy}^{LAB}$
Troisième ordre
$B_{z^3}^{MAS} = \frac{2}{3\sqrt{3}} B_{z^3}^{LAB} - \frac{1}{\sqrt{6}} B_{z^2x}^{LAB} + \frac{5}{\sqrt{3}} B_{z(x^2-y^2)}^{LAB} - \frac{5}{3\sqrt{6}} B_{x^3}^{LAB}$
$B_{z^2x}^{MAS} = \frac{\sqrt{2}}{3\sqrt{3}} B_{z^3}^{LAB} - \frac{\sqrt{3}}{2} B_{z^2x}^{LAB} + \frac{5}{6\sqrt{3}} B_{x^3}^{LAB}$
$B_{z^2y}^{MAS} = -\frac{10\sqrt{2}}{3} B_{xyz}^{LAB} + \frac{1}{6} B_{z^2y}^{LAB} + \frac{5}{6} B_{y^3}^{LAB}$
$B_{z(x^2-y^2)}^{MAS} = \frac{1}{3\sqrt{3}} B_{z^3}^{LAB} - \frac{\sqrt{2}}{3\sqrt{3}} B_{x^3}^{LAB}$
$B_{xyz}^{MAS} = -\frac{1}{3} B_{xyz}^{LAB} + \frac{\sqrt{2}}{12} B_{z^2y}^{LAB} - \frac{\sqrt{2}}{12} B_{y^3}^{LAB}$
$B_{x^3}^{MAS} = \frac{\sqrt{2}}{3\sqrt{3}} B_{z^3}^{LAB} + \frac{1}{2\sqrt{3}} B_{z^2x}^{LAB} + \frac{2\sqrt{2}}{\sqrt{3}} B_{z(x^2-y^2)}^{LAB} + \frac{5}{6\sqrt{3}} B_{x^3}^{LAB}$
$B_{y^3}^{MAS} = 2\sqrt{2} B_{xyz}^{LAB} + \frac{1}{2} B_{z^2y}^{LAB} + \frac{1}{2} B_{y^3}^{LAB}$

Tableau 1 : Principaux « shims » (trois premiers ordres) exprimés dans le repère orienté à l'angle magique jusqu'au troisième ordre en tant que combinaisons linéaires des « shims » dans le repère du laboratoire.¹²⁴

I.2.1.d) Le passeur d'échantillon HR-MAS

Quel que soit le domaine d'application de la RMN, la plupart des études, en particulier celles métaboliques, nécessite le passage d'un grand nombre d'échantillons. En fonction des besoins ou des objectifs des laboratoires, la possibilité de pouvoir réaliser de multiples analyses lorsque l'utilisateur est absent ou occupé, est devenue une nécessité. Pour cela,

l'automatisation des différentes étapes entre la préparation de l'échantillon et la fin de son analyse (transfert au sein du spectromètre, réglage de l'accord en fréquence et en impédance de la sonde, ajustement de l'homogénéité du champ magnétique principal, détermination des durées d'impulsion, acquisition et traitement des données) est requise. Ainsi, des passeurs d'échantillons automatiques, réalisant ces opérations ont été conçus par la société Bruker. Afin de piloter ce robot dans les différentes tâches que celui-ci doit effectuer un logiciel (IconNMR) a également été développé.

Dans le cas de la RMN HR-MAS, le passeur assure l'insertion des rotors depuis un compartiment réfrigéré vers le spectromètre à travers un tuyau, qui est guidé par une unité pneumatique. Une fois l'échantillon analysé, celui-ci effectue le chemin inverse. Cependant, à la différence des sondes RMN liquides, celles employées en HR-MAS ne sont pas équipées de dispositifs permettant le réglage automatique de la syntonisation de la sonde. De plus, comme nous l'avons vu précédemment l'homogénéité du champ magnétique ne peut être ajustée automatiquement en RMN HR-MAS. Ainsi, l'utilisation d'un tel dispositif est restreinte à l'analyse d'échantillons possédant des propriétés similaires et préparés dans des conditions équivalentes.

Une photographie du passeur d'échantillons utilisé pour la réalisation d'une partie des travaux de cette thèse est donnée en Figure 18.

Le passeur d'échantillons que nous avons utilisé en RMN HR-MAS est un Sample-Pro (Bruker BioSpin) permettant le refroidissement et la régulation en température des rotors avant leur insertion au sein du spectromètre. Ce système assure également le changement de rotors entre les expériences. La vitesse de rotation est fixée à 5000 Hz pour l'ensemble des acquisitions HR-MAS afin de maintenir les bandes de rotation à l'extérieur de la zone d'intérêt du spectre RMN. Tous les spectres RMN ont été enregistrés à une température de 4°C afin de minimiser la dégradation des échantillons de poisson pendant l'acquisition (éviter l'évolution de la valeur K notamment, comme nous le verrons plus tard dans ce manuscrit).

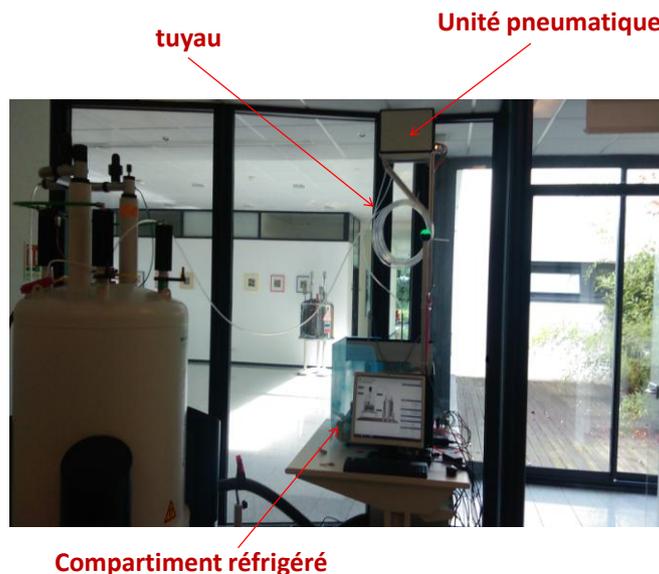


Figure 18 : Photographie du passeur d'échantillons HR-MAS Sample Pro. Celui-ci est composé d'un compartiment réfrigéré au sein duquel les rotors sont stockés jusqu'à leur analyse. Une unité pneumatique assure le passage via un tuyau des rotors jusqu'au spectromètre au sein de la sonde.

I.2.2) Protocole préparation d'échantillons HR-MAS

Tous les échantillons de poisson ont été préparés suivant le même protocole. Tout d'abord, 10 à 12 mg de muscle de poisson ont été prélevés à l'aide d'une biopsie punch de 2mm (Miltex) puis placés au sein de l'insert jetable refroidit dans de la glace pillée. Pour les analyses portant sur la fraîcheur du poisson, les prélèvements ont été systématiquement réalisés au niveau de la queue du poisson, cette partie n'étant composée que de muscle. Par la suite, 10 μ l d'oxyde de deutérium (D_2O) contenant 0,75 % de trimethylsilyl propionate deutéré (TSP) afin de fournir une fréquence deutérium pour le « lock » et un signal de référence à 0 ppm pour ajuster l'homogénéité du champ magnétique. L'insert est alors placé au sein du rotor 4 mm en ZrO_2 puis fermer à l'aide d'un capuchon muni d'ailettes (Figure 19). Le rotor est ensuite placé au sein du passeur d'échantillon refroidit à 4°C avant d'être inséré automatiquement au sein de la sonde.



Figure 19 : Photographie du dispositif HR-MAS utilisé pour la préparation des différents échantillons étudiés à l'aide de cette technique. De gauche à droite sur la Figure : le rotor, l'insert jetable de 33 μ l et le capuchon (ailettes).

I-3) RMN liquide

I.3.1) Instrumentation

L'ensemble des spectres de caviar 1D ont été enregistrés sur le spectromètre Bruker Avance III 400 opérant à la fréquence 399.90 MHz pour les noyaux ^1H équipé d'une sonde 5mm Bruker BBI (optimisée pour l'observation ^1H) muni de gradient Z. Afin d'attribuer les résonances des différents métabolites présents dans la phase aqueuse du caviar, les spectres 2D homonucléaires ^1H - ^1H et hétéronucléaires ^1H - ^{13}C ont été enregistrés sur un spectromètre Bruker AVANCE III 600 opérant à la fréquence proton 600.13 MHz équipé d'une sonde cryogénique Bruker QCI composée de 4 canaux d'observation $^1\text{H}/^{19}\text{F}$, ^{13}C et ^{15}N et d'un canal ^2H pour le « lock ». Les tubes RMN utilisés sont des tubes RMN 5mm PP 507 (Wilmad LabGlass).

I.3.2) Protocole préparation d'échantillons de caviar

Tous les échantillons de caviar, que ce soit pour l'étude sur la détermination de l'origine du caviar ou pour l'étude sur la dégradation du caviar, ont été préparés suivant le même protocole de préparation. Le protocole de préparation développé en vue de l'analyse du caviar par RMN est basé sur la méthode d'extraction Bligh & Dyer¹²⁶: cette méthode d'extraction liquide-liquide permettant une extraction totale des lipides. En effet, cette méthode utilise à la fois des solvants aqueux (eau) et organiques (méthanol et chloroforme) afin d'obtenir un système bi-phasique. Les molécules présentes dans le caviar vont ainsi migrer au sein de la phase organique ou aqueuse en fonction de leur polarité : les molécules apolaires (lipides) se retrouvent ainsi dans la phase organique tandis que les molécules polaires (métabolites) se placent, quant à elles, au sein de la phase aqueuse. Les protéines ainsi que les résidus de caviar se trouvent, quant à eux, à l'interface entre les deux phases. Notre étude se basant sur l'établissement d'un profil métabolique du caviar, seule la phase aqueuse contenant les molécules d'intérêt sera prélevée en vue de l'analyse par RMN.

Tout d'abord, 150 mg de caviar sont prélevés et insérés au sein d'un eppendorf de 1,5 ml (Eppendorf) maintenu dans de la glace pilée pendant toute la procédure d'extraction afin de minimiser les réactions métaboliques et sa dégradation. Les œufs de caviar sont alors soigneusement cassés par centrifugation puis manuellement. Nous ajoutons alors successivement 187,5 μl de chloroforme deutéré et 375 μl de méthanol deutéré. L'eppendorf est alors passé au vortex pendant 1 minute. Un deuxième ajout de solvant, cette fois ci de 187,5 μl de chloroforme deutéré et de 187,5 μl de D_2O est réalisé. L'eppendorf est de

nouveau agité au vortex pendant 1 minute. L'épendorf est ensuite laissé à décanter pendant 15 minutes et le caviar, se trouvant à l'interface des deux phases, est alors ôté. Une nouvelle centrifugation est alors appliquée puis 150 µl de la phase aqueuse sont collectés et mélangés à 450 µl d'une solution de tampon deutéré calibrée à pH 7,30 (Bruker BioSpin GmbH, Karlsruhe). Les 600 µl sont alors placés dans un tube qui sera conservé au congélateur à -32 °C jusqu'à l'analyse et centrifugé avant celle-ci. La phase aqueuse étant composée d'un mélange de nombreux métabolites, les tubes ne sont analysés que le lendemain de la préparation afin de permettre une bonne homogénéisation du tube RMN.

II) Séquences utilisées

Maintenant que nous avons présenté le matériel utilisé ainsi que les différents protocoles de préparation des échantillons en vue d'une analyse par RMN, nous allons développer, en détails, les différentes séquences RMN utilisées.

II.1) Expériences 1D

II.1.1) ZGPR

Cette séquence (Figure 20) est la plus basique des expériences RMN 1D. En effet, celle-ci est composée, tout d'abord, d'une irradiation continue faible puissance à la fréquence de résonance de l'eau, appelée présaturation pendant le temps de répétition (TR), suivie d'une simple impulsion RF 90° afin de rendre l'aimantation z initiale observable dans le plan x,y où se trouve le détecteur. Le FID est alors enregistré pendant un certain temps appelé temps d'acquisition (AQ).

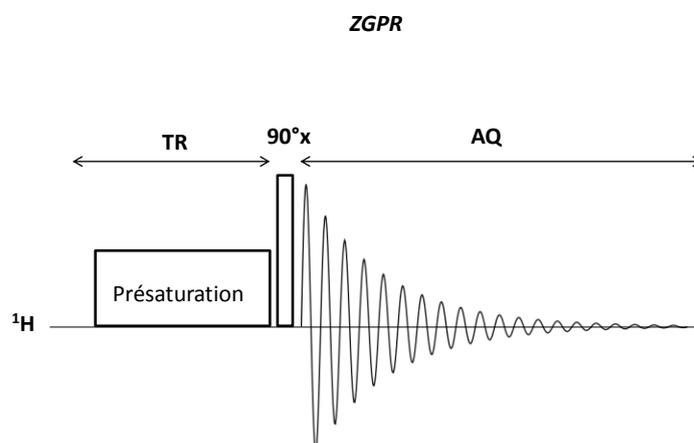


Figure 20 : Séquence d'impulsions RMN bruker *zgpr*. Celle-ci est composée d'une présaturation pendant le temps de répétition puis d'une impulsion 90°.

II.1.2) CPMGPR1D et CPMG pseudo-2D

L'expérience CPMG développée par Carr-Purcell-Meiboom-Gill et basée sur des échos de spin successifs, est l'une des séquences 1D les plus utilisées pour les études par RMN HR-MAS du fait de sa capacité à supprimer les signaux larges correspondant aux macromolécules et de corriger les imperfections de l'impulsions 180° permettant ainsi d'obtenir un spectre phasé.

Cette séquence (Figure 21) est, tout d'abord constituée d'une présaturation pendant le temps de répétition (TR) puis d'une impulsion de 90° afin de basculer l'aimantation dans le plan transversal. Elle est ensuite composée d'un écho de spin τ - 180° - τ qui permet de refocaliser les aimantations à l'issue de celui-ci. L'écho de spin permet de refocaliser les déplacements chimiques tout en permettant à la relaxation transversale de s'établir pendant cette période. La valeur du délai τ est fixée en tenant compte de deux paramètres. Tout d'abord, ce délai doit être réglé de telle manière que $1/\tau \gg J$ afin de minimiser la distorsion de phase introduite par le couplage scalaire pendant l'expérience. Ensuite, il est nécessaire de synchroniser ce délai avec la vitesse de rotation de l'échantillon : $\tau = N \times T_R$ afin que les imperfections des impulsions (B_1) appliquées soient ressenties toujours de la même manière par l'échantillon.

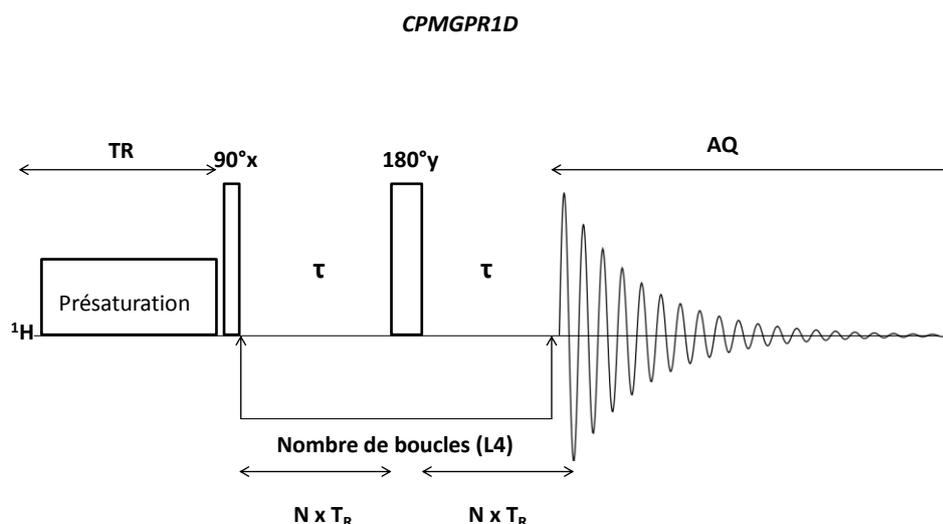


Figure 21 : Séquence d'impulsions CPMG (bruker *cpmgpr1d*), avec une présaturation du signal de l'eau pendant le temps de répétition, synchronisée sur la période du rotor. Le temps d'écho total employé pour l'ensemble des acquisitions a été fixé à 75,6 ms ($\tau = 200 \mu\text{s}$ et $L4 = 188$).

Le grand intérêt de la séquence CPMG est de pouvoir atténuer voire supprimer les signaux correspondant aux macromolécules en se basant sur la différence de temps de relaxation transversale entre ces dernières et les métabolites. Le principe consiste à fixer un temps d'écho total ($T_{echo} = 2 \cdot \tau \cdot L4$) suffisamment long pour qu'un maximum de macromolécules relaxent mais suffisamment court pour que les signaux correspondant aux métabolites ne soient pas trop atténués. Le temps d'écho total a ainsi été fixé pour les études réalisées par RMN HR-MAS à 75,6 ms ($\tau = 200 \mu\text{s}$ et $L4 = 188$). Les paramètres d'acquisition utilisés pour l'enregistrement des spectres ^1H par RMN HR-MAS sont donnés ci-dessous (Tableau 2).

Paramètres d'acquisition ^1H CPMG HR-MAS	
Echantillonnage	65536
Fenêtre spectrale (ppm)	20,5643
Temps d'acquisition (s)	4
Gain du receveur	128
Délai de relaxation (s)	2
Temps d'écho fixe (μs)	200
Nombres de boucles CPMG	188
Nombre de scans fictifs	32
Nombre de scans	512
Durée impulsion ^1H 90° (μs)	5,9
Vitesse de rotation (Hz)	5000
Champ RF présaturation (Hz)	50

Tableau 2 : Paramètres d'acquisition de la séquence 1D ^1H CPMG utilisés en RMN HR-MAS. La durée de l'impulsion 90° est la valeur moyenne observée sur l'ensemble des échantillons.

La séquence CPMG a également été utilisée sous sa forme pseudo-2D (pas de création d'une dimension indirecte) pour mesurer les temps de relaxation transversale (T_2). Afin de réaliser cette mesure, il est nécessaire de définir une liste de compteurs (*vclist*) qui contient les différents nombres de boucles pour chaque incrément de l'expérience pseudo-2D. Pour chaque incrément l'intensité des signaux d'intérêt va être déterminée et à l'aide de la courbe de décroissance de ces signaux, il est alors possible de remonter à la valeur du temps de relaxation transversale à l'aide de l'équation :

$$I(t) = I_0 \cdot e^{-t/T_2}$$

Cette formule est utilisée dans le cas d'un comportement mono-exponentiel de décroissance du signal dans le plan transversal. Cependant, pour certains cas, comme celui que nous verrons pour la relaxation transversale de l'eau au sein de tissus musculaires, des fonctions bi-

exponentielle ou tri-exponentielle sont utilisées pour modéliser le phénomène de relaxation transversale.

II.1.3) NOESYGPPR1D

La séquence NOESY 1D est la plus utilisée pour les études métaboliques par RMN du fait de son excellente capacité à supprimer le signal du solvant, et du peu de restrictions matériel et de réglages requis. En effet, la suppression de ce signal est réalisée à la fois par une présaturation pendant TR mais également pendant le temps de mélange, par le cycle de phase et par l'utilisation de gradients de nettoyage. Un cycle de phase consiste à incrémenter les phases de certaines impulsions RF et du récepteur, pour les différentes répétitions de la séquence RMN, de telle sorte que la combinaison des différentes accumulations permette de supprimer les artéfacts tandis que les signaux d'intérêt s'additionnent. Les cycles de phase les plus connus étant EXORCYCLE¹²⁷ et CYCLOPS¹²⁸.

Cette séquence (Figure 22) diffère de la séquence WEFT (Water Eliminated Fourier Transform)¹²⁹ par l'utilisation de deux impulsions 90° successives plutôt que d'une impulsion 180° afin de pouvoir appliquer un cycle de phase.

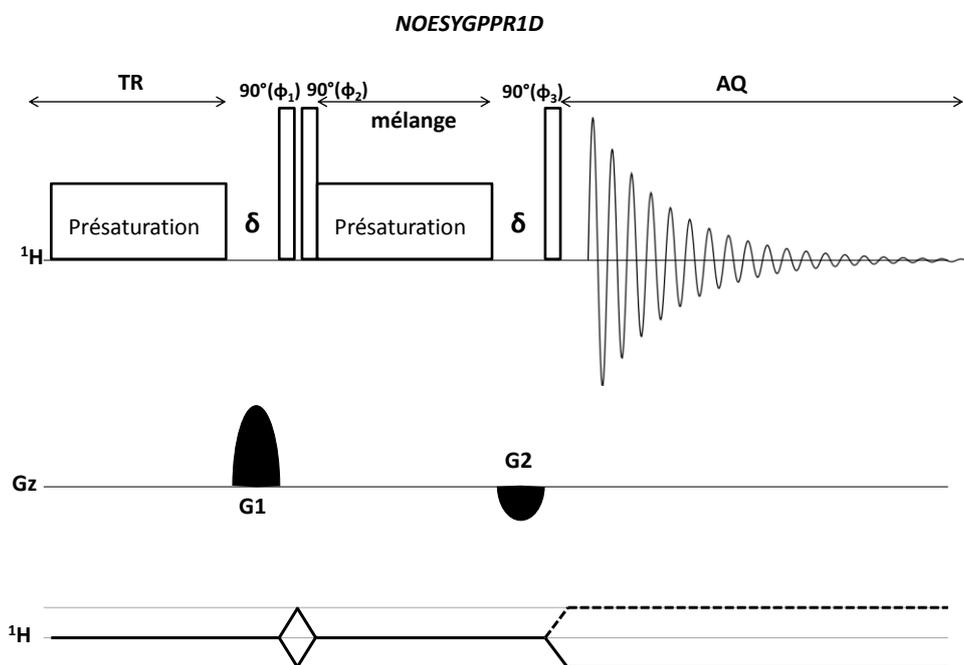


Figure 22 : Séquence d'impulsions 1D ^1H NOESY (bruker *noesygppr1d*), avec une présaturation du signal de l'eau et l'utilisation de gradients, employée pour l'ensemble des analyses réalisées sur la phase aqueuse du caviar. Les phases des impulsions utilisées pour le cyclage de phase sont $\phi_1 = (0\ 2)$, $\phi_2 = (0)$ et $\phi_3 = (0\ 0\ 2\ 2\ 1\ 1\ 3\ 3)$.

L'aimantation initiale (I_z) est, tout d'abord, basculée dans le plan transversal par la première impulsion 90° ($-I_y$). La seconde impulsion 90° va alors transférer de nouveau cette aimantation le long du champ magnétique ($-I_z$). Cette aimantation va alors évoluer pendant le temps de mélange sous l'influence du déplacement chimique et du couplage avant d'être rendue observable par la dernière impulsion 90° (I_y).

L'explication précédente de la séquence présente le cas où toutes les impulsions 90° sont effectuées le long de x. Cependant, comme vu précédemment un cycle de phase est appliqué sur les différentes impulsions mais également sur la phase du détecteur afin de permettre la suppression des artefacts et l'addition des signaux d'intérêt par combinaison des différentes acquisitions. La suppression des différents artefacts est également réalisée par l'utilisation des gradients.

Les paramètres d'acquisition utilisés pour l'enregistrement des spectres de caviar sont présentés dans le Tableau 3.

Paramètres d'acquisition ^1H NOESYGPPR1D	
Caviar	
Echantillonnage	65536
Fenêtre spectrale (ppm)	20,5643
Temps d'acquisition (s)	4
Gain du receveur	90,5
Délai de relaxation (s)	4
Temps de mélange (ms)	10
Nombre de scans fictifs	8
Nombre de scans	512
Durée impulsion ^1H 90° (μs)	10,5
Forme gradient	Sinusoidale 100 points
Force des gradients	GP1 : 50% GP2:-10%
Champ RF présaturation (Hz)	50

Tableau 3 : Paramètres d'acquisition de la séquence RMN 1D ^1H NOESY (bruker *noesygppr1d*) utilisés pour l'étude la phase aqueuse du caviar. La durée de l'impulsion 90° est la valeur moyenne observée sur l'ensemble des échantillons.

II.1.4) Inversion-Récupération

La séquence d'inversion-récupération (Figure 23) est la méthode la plus couramment utilisée pour mesurer les temps de relaxation longitudinale (T_1). Tout comme l'expérience CPMG pseudo-2D, cette séquence nécessite la création d'une liste contenant les différents délais (*vdlist*) τ pour chaque incrément et un spectre sera enregistré pour chacun de ces incréments. L'aimantation initiale qui est alignée le long du champ magnétique (I_z) est, tout d'abord, inversée par l'impulsion 180° ($-I_z$). Cette aimantation va ensuite évoluer et relaxer pendant le délai variable τ suivant l'équation $-I_z \exp^{-\tau/T_1}$. Cette aimantation est enfin rendue observable par l'impulsion 90° (I_y) et enregistrée pendant AQ.

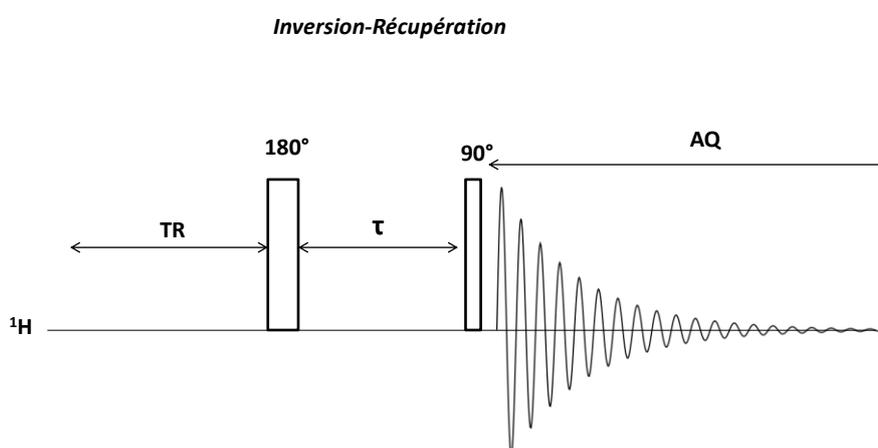


Figure 23 : Séquence d'impulsions Inversion-Récupération utilisée pour les mesures de temps de relaxation longitudinale. La durée de τ varie pour chaque incrément de la liste contenant les différents délais (*vdlist*).

Les différentes valeurs de temps de relaxation longitudinale (T_1) sont déterminées à l'aide de l'équation suivante qui décrit l'évolution de l'aimantation pendant cette séquence :

$$I(t) = I_0 \cdot [1 - 2e^{-t/T_1}]$$

Le facteur 2 provient du fait que les pics sont, au départ, négatifs sur le spectre puis passent par une valeur nulle avant de devenir de plus en plus positif au fur et à mesure que τ augmente. Il est possible d'obtenir une estimation rapide de la valeur de T_1 . Pour cela, il est nécessaire de situer le délai au bout duquel l'aimantation est nulle (τ_0). La valeur de l'estimation rapide de T_1 peut alors être déterminée suivant l'équation :

$$T_1 = \frac{\tau_0}{\ln 2}$$

II.2) Expériences 2D

II.2.1) TOCSY ET ATOCSY

La TOCSY (TOtal Correlation Spectroscopy)⁵⁰ est une expérience 2D homonucléaire, généralement utilisée pour l'étude des noyaux ^1H , qui permet d'observer les corrélations entre des spins qui sont connectés par une chaîne de couplage continue. Cette expérience diffère de la COSY où seuls les spins directement couplés sont visibles sur le spectre. Ainsi, si un spin A est couplé à un spin B et B est couplé à un spin C, il est possible d'observer sur le spectre TOCSY un pic de corrélation entre A et C même s'il n'y a pas de couplage entre ces deux spins.

La partie clé de l'expérience TOCSY est la période de mélange isotrope (« spin-lock »)¹³⁰ notée τ_{mix} sur la Figure 24 qui correspond à la période de mélange. En effet, le taux de transfert de l'aimantation z d'un spin à un autre dépend du couplage entre ces spins, de l'efficacité de la séquence de mélange et de la durée du « spin-lock », et c'est ce transfert qui est à l'origine de l'apparition du pic de corrélation sur le spectre RMN. Pendant le « spin-lock » les aimantations sont verrouillées selon une certaine direction de l'espace qui dépend de la direction du champ B_1 .

Cependant cette technique est extrêmement sensible aux effets hors résonance, il est donc nécessaire d'utiliser des périodes de mélange appropriées et efficaces. Plusieurs séquences d'impulsions existent et sont communément utilisées pour le mélange isotrope les plus courantes étant WALTZ-16, DIPSI-1 et DIPSI-2¹³¹, MLEV 16 et MLEV 17¹³².

La séquence (Figure 24) est tout d'abord composée d'une première partie, similaire à la séquence COSY (90° - t_1 - 90°), qui commence par une première impulsion 90° x qui bascule l'aimantation initiale (I_z) dans le plan transversal ($-I_y$) puis celle-ci évolue pendant t_1 sous l'influence des déplacements chimiques et des constantes de couplage puis une nouvelle impulsion 90° x est appliquée. A ce stade, juste avant la période de mélange isotrope, et en utilisant le formalisme des opérateurs produits 4 termes sont présents :

$$- \cos(\pi J_{12} t_1) \cos(\Omega_1 t_1) I_{1z} \quad [\text{A}]$$

$$- \sin(\pi J_{12} t_1) \cos(\Omega_1 t_1) 2I_{1x} I_{2y} \quad [\text{B}]$$

$$+ \cos(\pi J_{12} t_1) \sin(\Omega_1 t_1) I_{1x} \quad [\text{C}]$$

$$- \sin(\pi J_{12} t_1) \sin(\Omega_1 t_1) 2I_{1z} I_{2y} \quad [\text{D}]$$

Etant donné que le « spin-lock » est appliqué le long de z, seul le terme [A] nous intéresse. A l'issue de τ_{mix} [A] peut être réécrit :

$$\sum_{k=1}^k -I_{kz} a_{ik}(\tau_{\text{mix}}) \cdot \cos(\pi J_{ik} t_1) \cos(\Omega_1 t_1)$$

Où $a_{ik}(\tau_{\text{mix}})$ correspond au coefficient de mélange pour le transfert de l'aimantation de I_i vers I_k .

La dernière impulsion 90° transforme le terme (I_{kz}) précédent en aimantation observable (I_{ky}).

Il est intéressant de noter que dans le cas de la TOCSY, les pics diagonaux et les pics de corrélation sont en phase dans chaque dimension permettant de traiter le spectre de telle manière que tous les pics soient en absorption.

L'intensité des pics de corrélation due au couplage direct possèdent une intensité maximum pour $\tau_{\text{mix}} = \frac{1}{2J}$. Cependant, les pics de corrélation liés au transfert à travers deux ou plus couplages nécessitent des temps de mélanges plus importants. La valeur du temps de mélange généralement utilisée est de 80-120 ms.

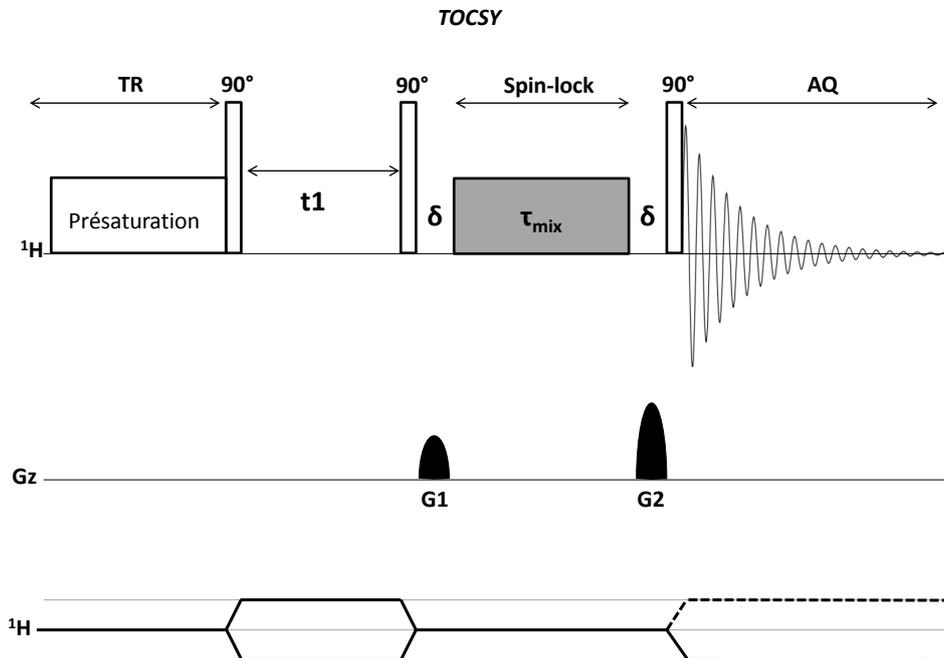


Figure 24 : Séquence d'impulsions TOCSY utilisée pour l'acquisition des spectres 2D homonucléaires ^1H - ^1H permettant la mise en évidence des couplages d'un système de spins facilitant l'identification des composés. En RMN HR-MAS le mélange isotropique (« spin-lock ») est composée d'impulsions adiabatiques (ATOCSY).

La séquence d'impulsions utilisée pour le mélange isotropique en RMN liquide est la séquence DIPSI-2 tandis qu'en HR-MAS nous avons utilisé une séquence employant des impulsions adiabatiques de 180° , le terme ATOCSY est alors employé (Adiabatic Total Correlation Spectroscopy)^{133,134}.

Les impulsions adiabatiques¹³⁵ créent un champ magnétique $\overrightarrow{B_{1eff}}$ qui bascule progressivement de $+z$ à $-z$, tandis que l'aimantation précesse dans le même temps autour de celui-ci. Ce type d'impulsion modulée en fréquence et en amplitude est relativement insensible aux inhomogénéités du champ B_1 et aux effets hors résonance et est couramment utilisé en RMN HR-MAS pour leur plus grande efficacité¹³⁶. De nos jours, les impulsions adiabatiques sont également couramment employées pour des expériences de découplage large bande homonucléaires¹³⁷ et hétéronucléaires¹³⁸.

Dans les expériences multidimensionnelles, la modulation dans la dimension t_1 est typiquement de la forme cosinus ou sinus. A cause de la propriété algébrique du cosinus, où $\cos(\theta) = \cos(-\theta)$ il est impossible de distinguer une modulation en $+\Omega$ de $-\Omega$. Lorsque la modulation est sinusoïdale, la situation est légèrement différente puisque les pics à $+\Omega$ et à $-\Omega$ seront toujours indifférenciables mais également de signes opposés puisque $\sin(-\theta) = -\sin(\theta)$. Cependant, plusieurs méthodes existent pour remédier à ce problème, celles-ci étant basées sur la possibilité de générer des données modulées à la fois en cosinus et en sinus. La première est la SHR (*States-Haberkorn-Ruben*)¹³⁹ communément appelée « States », qui permet à partir de deux signaux S_{\cos} et S_{\sin} modulés en amplitude d'obtenir un signal complexe dont la partie réelle est en absorption pure. La seconde est la méthode Redfield ou TPPI (*Time Proportional Phase Incrementation*)¹⁴⁰ implique l'incrément de la phase d'une des impulsions de la séquence utilisée à chaque incrément de t_1 , d'où la dénomination de cette technique. Cette approche ajoute une fréquence constante au déplacement chimique de tous les pics du spectre et en fixant de manière appropriée celle-ci, et donne l'impression que tous les déplacements sont positifs. Enfin, une méthode hybride basée sur la combinaison des deux techniques précédemment présentées est également employée, appelée States-TPPI¹⁴¹, celle-ci permettant d'obtenir les mêmes résultats mais également le décalage des pics axiaux hors de la fenêtre spectrale.

Les différents paramètres utilisés pour l'enregistrement des spectres TOCSY ^1H - ^1H en RMN liquide et HR-MAS sont données ci-dessous (Tableau 4).

Paramètres d'acquisition TOCSY et ATOCSY ¹ H- ¹ H	HR-MAS	Liquide
Mode d'acquisition en F1	States-TPPI	States-TPPI
Mélange isotropique	impulsions adiabatiques	DIPSI-2
Echantillonnage F2	16k	16k
Nombre d'incrément	256	1024
Fenêtre spectrale (ppm)	13,5	13,5
Temps d'acquisition F2 (s)	1,22	1
Gain du receveur	128	128
Délai de relaxation (s)	1,5	2
Temps de l'incrément (μs)	150	150
Temps de mélange (ms)	80	80
Nombre de scans fictifs	64	8
Nombre de scans	64	8
Durée impulsion 1H 90° (μs)	5,9	13
Champ RF présaturation (Hz)	50	50

Tableau 4 : Paramètres d'acquisition TOCSY et ATOCSY ¹H-¹H utilisés pour l'identification des composés présents dans les échantillons de poisson (HR-MAS) et de la phase aqueuse du caviar (Liquide).

II.2.2)HSQC

L'expérience HSQC (Heteronuclear Single Quantum Correlation)⁵¹ est largement utilisée pour enregistrer des spectres de corrélation à une liaison entre un hétéro noyau (¹³C, ¹⁵N) et un proton. Au cours de celle-ci, les transferts d'aimantation sont réalisés à l'aide de séquences INEPT¹⁴² (Insensitive Nuclei Enhanced by Polarization Transfer) et l'aimantation de l'hétéro noyau évolue dans un état simple quantum pendant la période d'évolution t1. Le plus souvent, les séquences HSQC sont acquises en détection inverse, qui consiste à obtenir dans la dimension F1 (indirecte) les paramètres RMN (couplage, déplacement chimiques...) d'un noyau peu sensible et de détecter le signal RMN d'un noyau plus sensible grâce à un transfert d'aimantation. Le transfert de polarisation permet un gain de sensibilité équivalent à:

$$Gain_{détection\ inverse} = \frac{\gamma_I^{5/2}}{\gamma_S^{5/2}} \times \frac{(1 - e^{-\frac{TR}{T_1^I}})}{(1 - e^{-\frac{TR}{T_1^S}})}$$

Où I représente le noyau sensible (généralement ¹H ou ¹⁹F) et S le noyau à faible abondance et faible rapport gyromagnétique (¹³C ou ¹⁵N). Le gain de sensibilité permet l'obtention d'un spectre équivalent en signal/bruit à celui obtenu en détection directe mais avec un nombre d'accumulations bien inférieur. Cependant, il est nécessaire d'échantillonner une gamme de

fréquences plus large dans la dimension F1 que dans le cas d'une détection directe (environ 6 kHz pour le proton contre environ 30 kHz pour le ^{13}C).

La séquence HSQC utilisée pour l'attribution des différents métabolites (par RMN liquide et HR-MAS) est donnée en Figure 25. Cette séquence est, tout d'abord, composée d'une présaturation du solvant par irradiation continue pendant TR puis d'un motif INEPT afin de transférer l'aimantation des spins I vers les spins S. Le motif INEPT correspond à une première impulsion 90° qui va basculer l'aimantation dans le plan transversal, puis d'un écho de spin pendant lequel les aimantations ne vont évoluer que sous l'influence du couplage (les déplacements chimiques sont refocalisés par les impulsions 180° sur les deux noyaux) et enfin de deux impulsions 90° simultanées sur les deux noyaux afin de transférer l'aimantation des spins I vers les spins S.

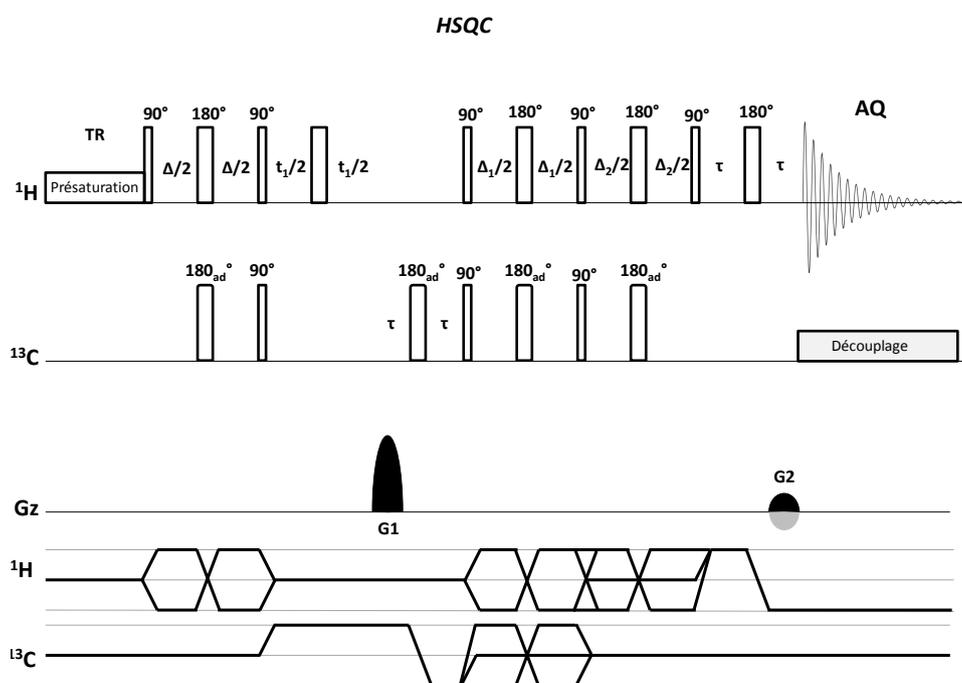


Figure 25 : Séquence d'impulsions HSQC utilisée, avec présaturation du signal de l'eau et l'utilisation de gradients, pour mettre en évidence les corrélations hétéronucléaires ^1H - ^{13}C permettant l'identification des composés. L'abréviation **ad** en indice de certaines impulsions indique que celle-ci est une adiabatique.

A la fin du transfert INEPT, l'aimantation est composée d'un terme en anti phase ($2I_zS_y$) et d'un terme en phase (I_y). Cependant seul le terme en anti phase nous intéresse puisque c'est celui qui a été transféré sur S et celui-ci se présente sous la forme :

$$-\sin(2\pi J_{IS}\Delta) 2I_zS_y$$

S'en suit une période d'évolution permettant l'incrémentation de t_1 (dimension indirecte) qui correspond à un écho de spin où les couplages sont refocalisés et où l'aimantation évolue sous l'influence des déplacements chimiques de S (l'impulsion 180° sur I refocalisant les déplacements protons). Seul le terme qui a été transféré sur S nous intéresse et à l'issue de la période d'évolution l'aimantation est décrite par :

$$\cos(\Omega_S t_1) \sin(2\pi J_{IS} \Delta) 2I_z S_y - \sin(\Omega_S t_1) \sin(2\pi J_{IS} \Delta) 2I_z S_x$$

La suppression des signaux non désirés se fait par l'application de gradient de champ pulsés (g-HSQC¹⁴³) et les meilleurs résultats sont obtenus lorsque le premier gradient est appliqué pendant la période d'évolution de l'hétéronoyau et le second juste avant la période de détection¹⁴⁴ conduisant à la formation d'un écho de gradient hétéronucléaire ce qui est le cas de la séquence que nous avons utilisé. Les deux impulsions 90° simultanées sur les deux noyaux permettent de re-transférer l'aimantation sur les spins I en vue de leur détection. Ce transfert transforme le terme $2I_z S_y$ en $-2I_y S_z$ qui est une aimantation en antiphase sur le spin I et le terme $2I_z S_x$ en $2I_y S_x$ qui est une cohérence multi-quantum non observable. Il est impossible d'appliquer un découplage à un terme en antiphase puisque cela conduirait à la suppression du multiplet. Un écho de spin est alors appliqué afin de transformer le terme en antiphase en terme en phase afin de pouvoir appliquer un découplage pendant l'acquisition. A la fin de l'écho de spin, le terme en antiphase est devenu un terme en phase :

$$- \cos(2\pi J_{IS} \Delta_1) \cos(\Omega_S t_1) \sin(2\pi J_{IS} \Delta) 2I_y S_z + \sin(2\pi J_{IS} \Delta_1) \cos(\Omega_S t_1) \sin(2\pi J_{IS} \Delta) I_x$$

Si le signal est observé sous découplage des noyaux S, seul le terme I_x contribue et conduit ainsi à un pic à la fréquence Ω_I dans la dimension ω_2 et à la fréquence Ω_S dans la dimension ω_1 ce qui correspond au pic de corrélation. L'intensité de ce pic est maximale lorsque Δ vaut $\frac{1}{4J_{IS}}$ puisque celle-ci dépend d'un facteur $\sin(2\pi J_{IS} \Delta)$. De la même façon il est facile de démontrer que la valeur optimale pour Δ_1 est $\frac{1}{4J_{IS}}$.

La dernière partie de la séquence est un schéma proposé par Schleucher et al¹⁴⁵ engendrant une amélioration de la sensibilité qui dépend des valeurs fixées pour les différents délais et dont les facteurs de gains varie en fonction du groupement (CH, CH₂ ou CH₃). Généralement, c'est la sensibilité des CH qui est la plus critique et les délais sont fixés de telle sorte que le gain soit maximum pour ces groupements et, dans ce cas, leur intensité est augmentée de 50 % contre 20 à 30 % pour les CH₂ et CH₃.

Les paramètres utilisés lors de l'acquisition des spectres HSQC sur le caviar et en HR-MAS sont donnés en Tableau 5.

Paramètres d'acquisition HSQC ¹H-¹³C	HR-MAS	Liquide
Mode d'acquisition en F1	Echo-Anti Echo	Echo-Anti Echo
Echantillonnage F2	2048	2048
Nombre d'incrément	256	512
Temps de l'incrément (µs)	3	3
Fenêtre spectrale 1H (ppm)	13,5	13,5
Fenêtre spectrale 13C (ppm)	180	180
Temps d'acquisition F2 (s)	0,4	0,13
Gain du receveur	128	128
Délai de relaxation (s)	1,5	2
Temps de l'incrément (µs)	150	150
Nombre de scans fictifs	128	32
Nombre de scans	256	32
Durée impulsion 1H 90° (µs)	5,9	13
Découplage	GARP-4	bi_p5m4sp_4sp (adiabatique)
Constante J(CH) en Hz	145	145

Tableau 5 : Paramètres d'acquisition HSQC ¹H-¹³C utilisés pour l'identification des composés présents dans les échantillons de poisson (HR-MAS) et de la phase aqueuse du caviar (Liquide)

Partie IV- Résultats :

I) Etudes de matrices semi-solides par RMN HR-MAS

Comme nous l'avons vu précédemment, la RMN HR-MAS est un outil analytique totalement approprié pour l'étude et la caractérisation d'échantillons agroalimentaires^{74,76,83}. En effet, de nombreux produits présentent une consistance semi-solide, rendant leur analyse par RMN liquide impossible, sans étape préliminaire d'extraction et de solubilisation. La RMN HR-MAS permet de s'affranchir de ces processus fastidieux et coûteux, préalables à l'analyse, ceux-ci pouvant également conduire à une perte d'information.

La première section est ainsi consacrée aux différents résultats obtenus à l'aide la RMN HR-MAS principalement sur des échantillons de poissons.

Ainsi, nous présenterons, tout d'abord, la méthode de détermination de la fraîcheur du poisson mise au point à l'aide de cette technique analytique, à travers l'utilisation de deux indicateurs chimiques : la valeur K et le TMA-N. Par la suite, nous nous intéresserons à l'étude de blocs de farce de colin d'Alaska d'un partenaire industriel (Findus France). Des analyses statistiques (ACP et OPLS-DA) ont ainsi été appliquées aux spectres 1D enregistrés par RMN HR-MAS, afin d'essayer de comprendre les différences observées entre les minerais de poissons. Nous présenterons, ensuite, une étude de la texture des différents blocs réalisée à travers l'étude du temps de relaxation transversale (T_2) de l'eau. Enfin, une méthode développée dans le but d'améliorer la quantification absolue des différents métabolites détectés lors d'une analyse CPMG sur des échantillons de poissons dans un premier temps, puis sur des biopsies humaines de cerveau, dans un second temps sera exposée.

I.1) Méthode de détermination de la fraîcheur du poisson (Publication 1-Rapid assessment of fish freshness and quality by ¹H HR-MAS NMR Spectroscopy-Annexe 2)

I.1.1) Introduction

La fraîcheur du poisson est un paramètre critique pour l'industrie agroalimentaire puisque la valeur commerciale du poisson ainsi que ses qualités gustatives sont directement reliées à son état de fraîcheur. Afin de préserver la qualité et la fraîcheur des produits de la mer, des conditions de stockage strictes doivent être respectées depuis l'instant où le poisson est pêché jusqu'à sa consommation. De mauvaises conditions de conservation peuvent en effet conduire à une dégradation de ces produits mais également, dans les cas extrêmes, à certains cas

d'intoxication alimentaire. Ainsi, pour garantir la qualité de ces produits, il est essentiel d'utiliser des méthodes analytiques permettant une détermination rapide et fiable de ces différents paramètres.

Immédiatement après la pêche et la mort du poisson, l'altération du poisson débute que ce soit par des procédés autolytiques ou microbiens. La quantité de bactéries et la teneur de certains produits de décomposition peuvent ainsi être utilisées comme indicateurs *post-mortem* de la qualité du poisson. Le schéma de dégradation du poisson conservé sous glace est composé de quatre phases :

- Une première phase où le poisson est extrêmement frais, également qualifiée d'ultra-fraîcheur, durant laquelle la dégradation est principalement autolytique. Ce stade dure environ deux jours pour la majorité des espèces et durant celle-ci la chair a une saveur douce et délicate.
- Une seconde phase caractérisée par une perte progressive de saveur des tissus. La chair possède alors un goût neutre mais la texture reste agréable. Les processus de détérioration sont majoritairement autolytiques. Cette étape dure également quelques jours et varie en fonction des espèces.
- Lors de la troisième phase, une altération importante des qualités gustatives et de la texture des tissus a lieu, cette dernière devenant soit molle et aqueuse, soit dure et sèche. Un certain nombre de substances volatiles à l'odeur désagréable (ammoniacale) apparaissent suivant les espèces et le type d'altération (aérobie, anaérobie). Les changements sont essentiellement bactériens et la principale réaction microbienne est la réduction du OTMA-N en TMA-N.
- Une dernière phase où le poisson est impropre à la consommation, celui-ci étant considéré comme trop altéré et putride.

Actuellement, plusieurs méthodes analytiques officielles sont recommandées dans le but d'évaluer la fraîcheur et la qualité du poisson se basant sur des techniques sensorielles, microbiologiques, chimiques, physiques et physico-chimiques^{146,147}. Du point de vue des méthodes chimiques, trois méthodes sont couramment utilisées : l'Azote Basique Volatil Total (ABVT)¹⁴⁸, la teneur en triméthylamine (TMA-N)¹⁴⁸ et la mesure des catabolites de nucléotides à travers la valeur K^{149,150}.

L'Azote Basique Volatil Total est un indicateur de la dégradation du poisson par dégradation autolytique (dégradation de certaines protéines par des enzymes présentes dans le poisson)

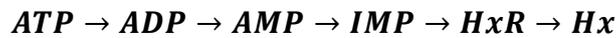
mais également par l'action de certaines bactéries (dégradation de l'oxyde de triméthylamine) et consiste en la détermination de la teneur en ammoniacque (NH_3), monométhylamine, diméthylamine (DMA), triméthylamine (TMA-N) et d'autres amines volatiles¹⁴⁸. La méthode officielle de référence (règlement CE n°2074/2005), permettant de mesurer l'ABVT, implique tout d'abord une première étape de déprotéinisation de l'échantillon par un acide (acide perchlorique ou acide borique) suivie d'une distillation de l'extrait acide à la vapeur d'eau et enfin une titration avec un acide fort (par exemple l'acide chlorhydrique) afin de déterminer la quantité de bases distillée. D'autres méthodes utilisant des cellules à microdiffusion de Conway¹⁵¹ ou la distillation directe¹⁵² existent également.

L'analyse du TMA-N réside dans la mesure de la réduction de l'oxyde de triméthylamine (OTMA-N) par des bactéries, durant la conservation du poisson, en triméthylamine qui est un composé volatil responsable de la forte odeur du poisson. La réduction de l'OTMA-N est principalement associée à des bactéries marines telles que *Photobacterium*, *Alteromonas*, *Vibrio* et *S. putrefaciens* mais également à certains microbes intestinaux de la famille des *Enterobacteriaceae*.

Une méthode officielle existe pour le dosage de la triméthylamine (AOAC 971.14) et consiste, tout d'abord, en une extraction du TMA-N de l'échantillon à l'aide d'un alcène (toluène ou benzène) puis d'un ajout d'acide picrique à l'extrait précédemment obtenu qui va engendrer la formation d'un complexe coloré. Des méthodes non officielles sont également utilisées telles que la micro diffusion ou la distillation d'un extrait acide (méthodes identiques à celles citées pour la détermination de l'ABVT sauf qu'une étape additionnelle consistant en l'ajout de formaldéhyde pour bloquer les amines primaires et secondaires est nécessaire) ou des méthodes enzymatiques¹⁵³.

La valeur K est un indicateur chimique international qui mesure le processus autolytique de dégradation de l'ATP (Adénosine-5'-Triphosphate) qui est le vecteur de l'énergie permettant de transporter l'énergie de certaines réactions ayant lieu au sein d'un organisme vers les cellules qui en ont besoin. La *rigor mortis* survient lorsque le niveau d'ATP est inférieur à 1mmole/g, cette molécule jouant également le rôle d'agent plastifiant. En effet, la contraction musculaire est contrôlée par le calcium (Ca^{2+}) et l'enzyme ATP-ase présente dans les cellules musculaires. Le ramollissement du muscle à la fin de la *rigor mortis* coïncide avec le début des processus autolytiques.

A partir de la mort du poisson et donc de l'arrêt de la respiration cellulaire, la production d'ATP cesse et les stocks d'ATP sont dégradés une fois la *rigor mortis* finie par des enzymes endogènes selon l'équation suivante :



Avec les abréviations suivantes : **ADP** : Adénosine-5'-DiPhosphate, **AMP** : Adénosine-5'-MonoPhosphate, **IMP** : Inosine-5'-MonoPhosphate, **HxR** : inosine et **Hx** : Hypoxanthine.

La valeur K est déterminée en mesurant le rapport des concentrations de ces six nucléotides selon la formule suivante :

$$K = \left[\frac{(HxR + Hx)}{(ATP + ADP + AMP + IMP + HxR + Hx)} \right] \times 100$$

Aucun lien direct n'a pu être établi entre la dégradation des catabolites de nucléotides et la cause de l'altération de la fraîcheur. Cependant, l'hypoxanthine est un composé responsable de l'arrière-goût amer du poisson détérioré et l'IMP est quant à lui responsable du goût recherché de poisson frais seulement retrouvé dans les produits extrêmement frais. Ce dernier est également considéré comme étant à l'origine de la saveur *unami*, qui est l'un des 5 goûts de base au Japon¹⁵⁴.

Dans la littérature, la valeur K est considérée comme le meilleur indicateur chimique de la fraîcheur du poisson¹⁵⁵ pour la détermination de l'ultra-fraîcheur.

A la différence des deux autres méthodes précédemment décrites, la détermination de la valeur K ne présente pas de méthodes officielles. Cependant, la technique de détection la plus communément utilisée pour la mesure de la concentration des différents nucléotides est la CLHP. Néanmoins, avant l'analyse par CLHP, une étape préliminaire d'extraction des amines est requise à l'aide d'acide perchlorique ou d'acide chloroacétique. Les amines extraites sont alors dérivées en utilisant des d'halogénures d'acides afin de permettre leur détection par CLHP sur une colonne chromatographique en phase inverse. La méthode enzymatique, basée sur la mesure de la consommation de l'oxygène durant la dernière étape de dégradation de l'ATP, catalysée par l'enzyme *xanthine oxidase*, est également largement employée. L'évaluation de la quantité d'oxygène est réalisée par des électrodes d'oxygène¹⁵⁶ ou, plus récemment, par un dispositif microfluidique électrochimique¹⁵⁷.

Quelques études portant sur la détermination de la fraîcheur ou de la qualité du poisson ont été réalisées par RMN haute résolution. Tout d'abord, Sitter et al¹⁵⁸ ont démontré qu'à l'aide de spectre RMN liquide haute résolution ¹H enregistrés sur des extraits de flétan d'Atlantique, il était possible de déterminer la teneur des différents composés impliqués dans la valeur K par simple intégration ou déconvolution des signaux RMN correspondant. Au cours de cette étude, les auteurs ont également démontré que la valeur K n'était pas affectée par la congélation préalable à l'analyse du poisson, une propriété que Ehira et Uchiyama¹⁵⁵ avaient déjà observé sur d'autres espèces de poissons. Howell et al¹⁵⁹ ont, quant à eux, montré que la spectroscopie RMN ¹H pouvait être utilisée pour estimer le niveau de TMA-N au sein d'extraits de poisson et que les mesures élevées au sein de cabillauds et haddocks indiquaient qu'une dégradation bactérienne importante avait eu lieu. Enfin, Erikson et al¹⁶⁰ ont récemment publié une revue des applications de la RMN (RMN liquide, Imagerie par Résonance Magnétique et RMN bas champ) dans le but d'optimiser les processus de transformation du poisson.

I.1.2) Détermination de la fraîcheur du poisson par RMN HR-MAS

Un spectre RMN ¹H HR-MAS d'un échantillon de biopsie non transformée de poisson (bar) est présenté en Figure 26. La première étape de l'étude du poisson par RMN HR-MAS a ainsi consisté en l'attribution des différents métabolites observés dans ce type d'échantillons. Pour cela, nous avons utilisé, en premier lieu, des tables de déplacement chimique de métabolites détectés par RMN HR-MAS sur des échantillons de saumon et des bases de données de spectre de référence (HMDB et Bruker BioSpin BBIORFCODE), combiné à des expériences 2D (¹H-¹H ATOCSY et ¹H-¹³C HSQC). Le spectre ¹H-¹³C HSQC enregistré sur une biopsie de poisson (bar) est donné en Figure 27. L'ensemble des 23 métabolites ainsi attribués, ainsi que les corrélations observées sur les analyses 2D RMN sont répertoriés dans le Tableau 6. Comme nous pouvons le constater, la résolution obtenue, sur ce type de matrice par l'utilisation de la rotation à l'angle magique, est très proche de celle obtenue en RMN liquide haute résolution, la texture semi-solide naturelle du muscle de poisson (entre 60 et 80 % d'eau) introduisant un degré de mobilité des molécules suffisant au sein de l'échantillon correspondant parfaitement au domaine d'application de la RMN HR-MAS.

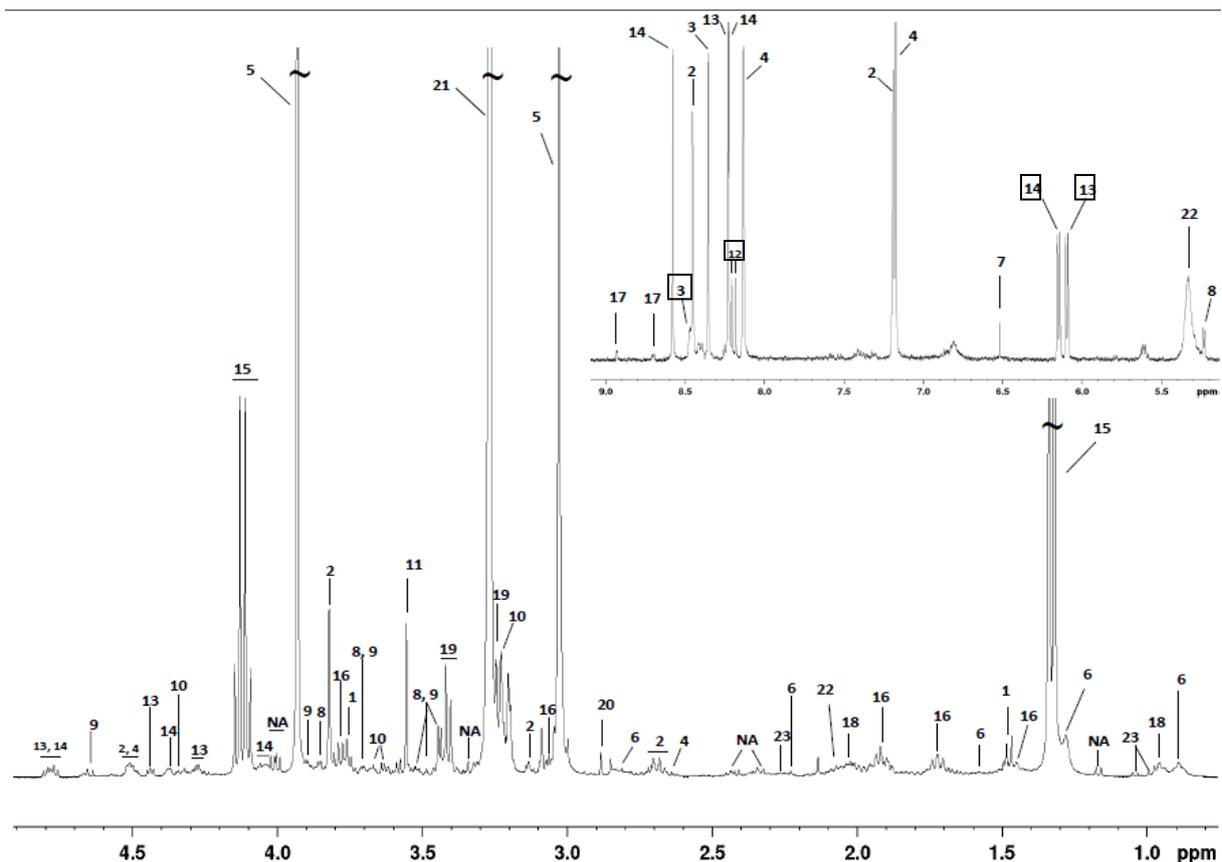


Figure 26 : Spectre 400 MHz ^1H CPMG HR-MAS enregistré sur une biopsie de bar. Le spectre correspond à l'analyse d'un poisson datant de 6 jours, la teneur en ATP, ADP et AMP est donc faible. L'attribution des différents métabolites réfère aux Tableau 6. Les signaux utilisés pour la détermination de la valeur K sont indiqués par un carré noir. L'abréviation NA est utilisée pour signaler les composés non attribués.

Le spectre RMN caractéristique obtenu sur ces échantillons peut être divisé en deux régions : une première de 5,5 à 9,5 ppm majoritairement composée de l'ATP et de ses produits de dégradation et une seconde de 0,5 à 4,7 ppm contenant les signaux des métabolites (acides aminés et sucres principalement) ainsi que les signaux correspondant au OTMA-N, au TMA-N, au DMA et quelques lipides résiduels.

Dans le but de déterminer la valeur K, il a été nécessaire d'établir les signaux à intégrer dans le but de quantifier les différents nucléotides, en prenant pour chaque métabolite un signal ne présentant pas de chevauchement avec d'autres pics.

Du point de vue de la RMN, l'ATP, l'ADP et l'AMP résonnent tous à la même fréquence sur le spectre ^1H et ne peuvent ainsi pas être différenciés. Ceci est sans conséquence pour la détermination de la valeur K puisque l'ensemble de ces trois composés se situe au dénominateur de la formule. Le singulet à 8,49 ppm, correspondant au CH₂-Purine, a ainsi été choisi comme signal d'intérêt. Pour les autres nucléotides, chacun d'entre eux possède au

moins un signal caractéristique et l'utilisation des signaux correspondant au CH1-Ribose de l'IMP et de l'inosine (HxR) résonnant respectivement sous la forme d'un doublet à 6,14 et 6,09 ppm, a été retenue. Pour l'hypoxanthine (Hx), nous avons choisi, pour l'intégration, le singulet résonnant à 8,18 ppm et correspondant au (CH-8). Le pH de l'échantillon variant avec la fraîcheur et les nucléotides possédant des signaux très sensibles au pH, un chevauchement entre certains signaux est parfois apparu. Dans ces cas-là, une déconvolution utilisant un mélange de fonctions Lorentzienne et Gaussienne a alors été appliquée (commande *dcon* sous TopSpin) afin d'obtenir l'intégral du pic d'intérêt. Une fois l'ensemble des signaux intégrés, la valeur K a pu être déterminée directement en faisant le rapport des intégrales obtenues pour chaque composé en tenant compte du nombre de protons associé à chaque signal. Les différents signaux utilisés pour la détermination de la valeur K sont encadrés sur la Figure 26.

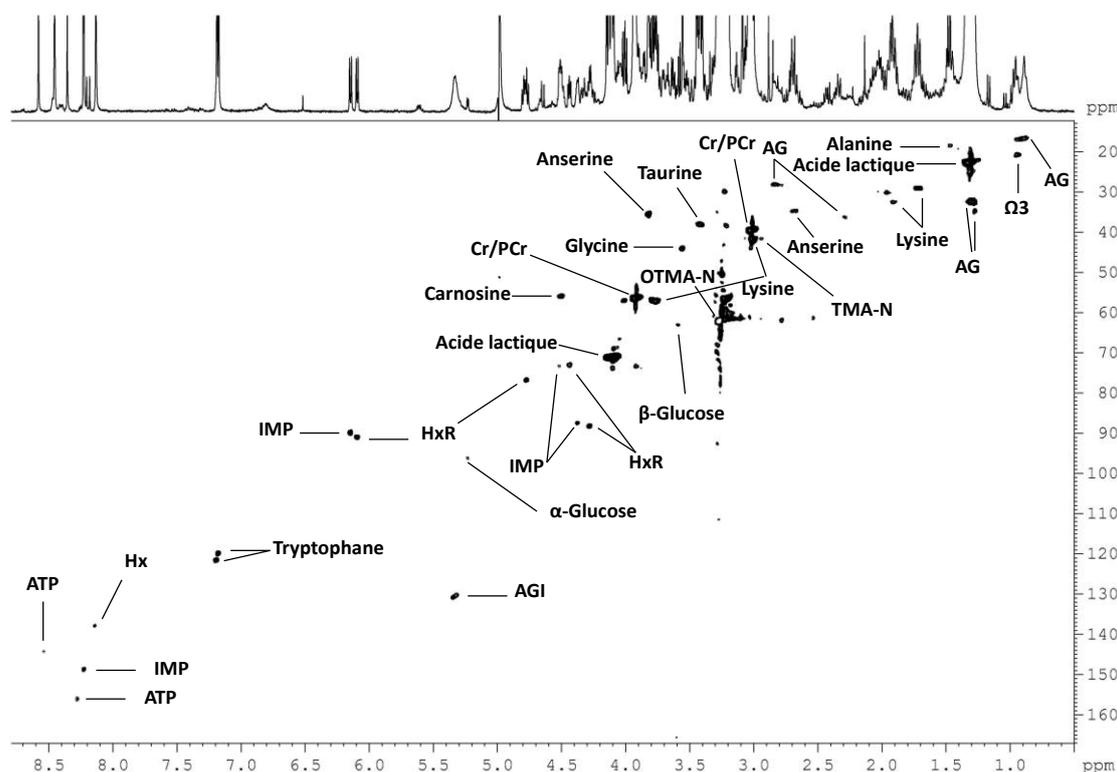


Figure 27 : Spectre 400 MHz HSQC ^1H - ^{13}C enregistré sur un échantillon de bar. La correspondance des différentes tâches avec son métabolite est indiquée.

Pour ce qui est du TMA-N, l'autre indicateur chimique de mesure de la fraîcheur mesurable par RMN, celui-ci résonne sous la forme d'un singulet correspondant au signal des CH_3 à 2,89 ppm. Dans le but de quantifier ce composé de manière absolue au sein de chaque échantillon,

la méthode PULCON/ERETIC^{56,58} a été alors appliquée. Cette technique de quantification est basée sur l'acquisition, dans un premier temps, d'un spectre de référence contenant une quantité connue d'un standard choisi (dans notre cas, une solution contenant 198,6 nmol de saccharose). Ensuite, un facteur de correction prenant en compte les variations de durée des impulsions RF induites par des concentrations en sels différentes entre le spectre de référence et les spectres d'échantillons de poisson est retenu de façon à quantifier les métabolites de manière précise.

Afin d'évaluer le potentiel de la RMN ¹H HR-MAS pour la détermination de la fraîcheur vis-à-vis des méthodes classiques, nous avons mené une étude portant sur quatre espèces distinctes (daurade, bar, truite et rouget) et sur deux poissons différents pour chaque espèce (duplicat). Sur les quatre espèces étudiées, deux avaient déjà fait l'objet de publications pour la valeur K (daurade et bar) tandis que pour la détermination du TMA-N, deux autres publications ont été utilisées pour comparer nos résultats à ceux issus de la littérature (rouget et bar).

Pour chaque étude, tous les poissons ont été conservés au réfrigérateur dans des boîtes en polystyrène remplies de glace pilée (0°C) afin de reproduire les conditions classiques de conservation du poisson utilisées dans l'industrie ; la glace pilée étant remplacée tous les jours pendant la durée de l'expérimentation. Pour l'étude portant sur la daurade et sur le bar, les poissons ayant été achetés au Marché d'Intérêt National de Strasbourg, une durée de deux jours est à considérer pour le T₀ (décès du poisson) ; ce délai correspondant au temps nécessaire entre la pêche à bord du bateau et l'arrivée du poisson au laboratoire. En revanche, ce n'est pas le cas pour les deux autres espèces, les truites et les rougets ayant été obtenus directement auprès de producteurs locaux.

Une première étude a ainsi portée sur le bar (*Dicentrarchus Labrax*) : une analyse par RMN HR-MAS étant réalisée chaque jour pendant deux semaines. Les résultats obtenus sur ce poisson sont donnés en Figure 28 pour la valeur K et en Figure 29 pour la mesure de TMA-N. Une valeur K de 78% a ainsi pu être déterminée, après 10 jours de conservation à 0°C, ce qui est en accord avec les résultats de Chang et al¹⁶¹ qui avaient trouvé une valeur de 73 % sur la même période de temps à l'aide de techniques enzymatiques. Pour le bar, les quantités d'ATP, d'ADP, d'AMP et d'IMP décroissent rapidement au cours du stockage tandis que dans le même temps la quantité d'inosine augmente et l'hypoxanthine apparaît au bout de deux jours. Pour ce qui est de l'évolution du TMA-N, celle-ci est caractérisée par une

évolution lente les premiers jours (du jour 2 au jour 6) puis augmente significativement jusqu'à la fin de la période de conservation. Au début de la période de stockage, la concentration du TMA-N est de $1,2 \pm 0,2$ mg/100g au bout de 2 jours puis atteint $2,9 \pm 0,5$ mg/100g au bout de 6 jours. Ces valeurs sont légèrement supérieures à celles déterminées par Cakli et al¹⁶² qui ont relaté des valeurs respectivement de 1,02 et 1,81 mg pour des périodes équivalentes. A la fin de l'étude, après 13 jours, la concentration de TMA-N mesurée par RMN est de $8,8 \pm 0,5$ mg/100g ce qui est en accord avec la valeur publiée $8,71$ mg/100g¹⁶².

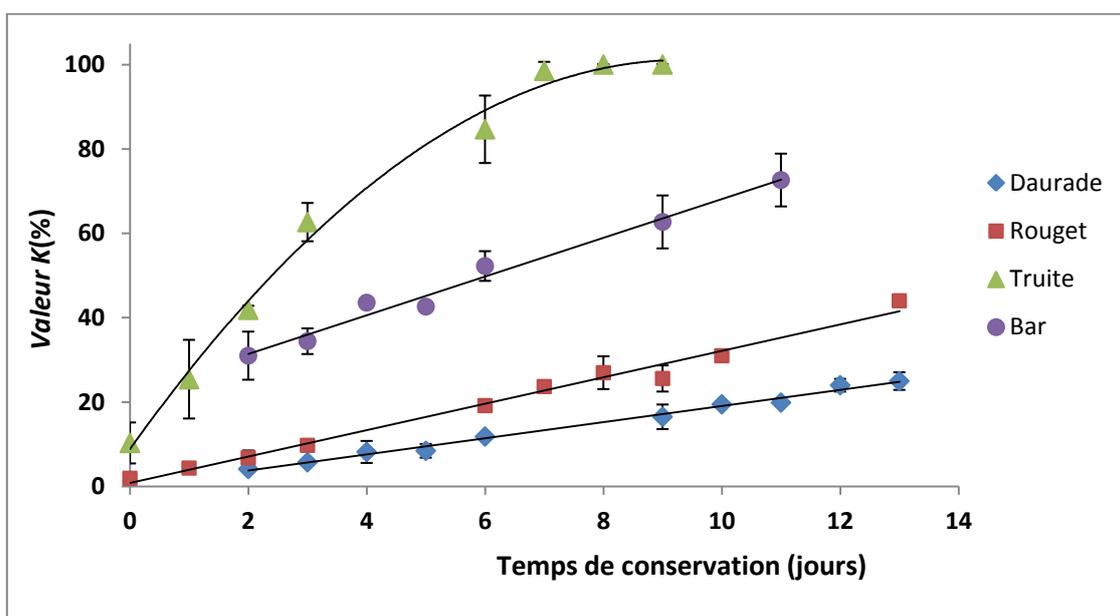


Figure 28 : Evolution de la valeur K (en %) au cours de sa conservation au réfrigérateur dans de la glace à 0°C pour les différentes espèces de poissons étudiées. Les barres représentent l'écart-type observé sur les mesures de la valeur K sur les deux poissons analysés en duplicat pour chaque espèce.

Une seconde expérimentation a été menée sur la daurade (*Sparus Aurata*) également conservée à 0°C pendant deux semaines. La valeur K atteint une valeur finale de 25% après 13 jours (Figure 28) ce qui concorde avec les valeurs déterminées par CLHP par Lougovois et al¹⁶³ qui ont déterminé une valeur de 27 % pour la même variété de poisson et ceux de Grigorakis et al¹⁶⁴ (22%) également par CLHP. Lougovois et al¹⁶³ ont signalé que la valeur K pour la daurade évoluait linéairement durant le stockage avec une pente de 2.1 ce qui est en accord avec celle obtenue dans le cas de notre étude (1.9 avec un R² de 0.991). L'ATP, l'ADP, l'AMP et l'IMP décroissent de manière continue mais à la différence du bar cette évolution est beaucoup plus lente et l'hypoxanthine n'apparaît qu'après 10 jours. La dégradation lente de l'IMP en inosine et la lente formation d'hypoxanthine sont

caractéristiques de la daurade. En effet, ces observations ont déjà été montrées dans la littérature^{164,165} et qualifient, ainsi, la daurade comme poisson conservateur d'IMP. La teneur en TMA-N n'a pu être déterminée de manière précise pour cette dernière, du fait de la présence inhabituelle de certains lipides au sein des poissons utilisés pour cette étude.

Une troisième étude a également été conduite cette fois-ci sur le rouget (dans les mêmes conditions que les deux premières espèces de poissons, sur deux semaines avec une analyse réalisée chaque jour) à la différence que celle-ci a pu débuter dès le T₀ et bien que nous n'ayons pas trouvé d'étude comparable dans la littérature. L'évolution de la valeur K pour le rouget est relativement similaire à celle de la daurade (Figure 28) : celle-ci semble suivre un modèle linéaire et est caractérisée par une dégradation lente des nucléotides du début de la chaîne réactionnelle de l'ATP (ATP, ADP, AMP et IMP) et donc une faible production d'inosine et d'hypoxanthine. Après 13 jours, une valeur K finale de 45 % a ainsi pu être déterminée. En revanche, les résultats obtenus pour la mesure de la teneur en TMA-N (Figure 29) ont pu être comparés à ceux obtenus par Cakli et al¹⁶². Une concentration de 1,5 mg±0.2 mg/100g, après un jour de conservation, a ainsi été évaluée après 1 jour de stockage. La quantité de TMA-N augmente, alors, doucement au début de la conservation comme pour le cas du bar, atteignant une valeur 8,4±0.6 mg/100g après 8 jours puis croître de manière plus importante jusqu'à la fin de l'étude pour finalement atteindre une valeur de 14 ±0.3 mg/100g. Ces valeurs sont en accord avec celles rapportées par Cakli et al, à savoir 1,50 mg/100g au stade initial et 8,01 mg/100g après 8 jours.

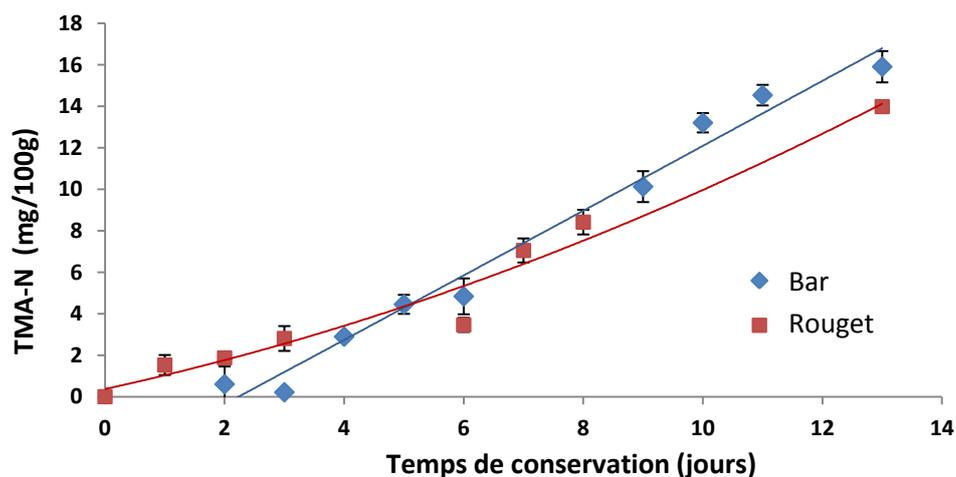


Figure 29 : Evolution de la quantité de TMA-N (en mg/100g) pour le bar et le rouget au cours de sa conservation au réfrigérateur dans de la glace à 0°C. Les barres représentent l'écart-type observé sur les deux poissons analysés en duplicat pour chaque espèce.

L'évolution générale du TMA-N, caractérisée à la fois pour le bar et pour le rouget par une augmentation lente aux prémices de l'étude puis par une teneur qui augmente de manière plus significative par la suite, s'explique par la prolifération des bactéries responsables de la production du TMA-N au fur et à mesure de l'expérimentation. En effet, plus le nombre de bactéries est important plus la dégradation de l'OTMA-N est conséquente et donc plus la quantité de TMA-N croît au cours du temps.

Enfin, notre méthodologie a été appliquée à une dernière espèce de poisson à savoir la truite marron. Pour cette variété, la valeur K (Figure 28) évolue de manière linéaire et de manière beaucoup plus rapide que pour les autres poissons étudiés puisque la valeur K atteint sa valeur maximale au bout de 6 jours. En effet, pour la truite marron, la dégradation de l'ATP en IMP et d'IMP en inosine est extrêmement rapide puisque, après 2 jours seulement, près de 60 % de l'ATP, de l'ADP, de l'AMP et l'IMP ont été catabolisés. L'évolution des différents produits de dégradation de l'ATP pour ce type de poisson est donnée en Figure 30. La teneur en TMA-N n'a pas pu être quantifiée dans la truite marron, celle-ci appartenant à la famille des salmonidés connus pour être des faibles producteurs de TMA-N. Effectivement, la quantité initiale d'OTMA-N est extrêmement faible conduisant à une faible quantité de TMA-N produite, et est ainsi considérée comme peu fiable comme estimateur de la fraîcheur¹⁶⁶.

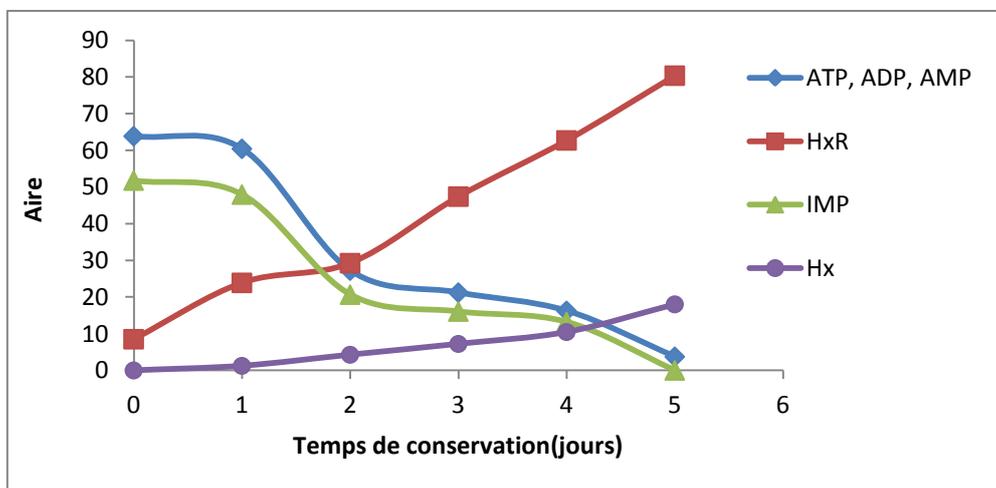


Figure 30 : Evolution des différents produits de dégradation de l'ATP pour la truite marron au cours de sa conservation au réfrigérateur dans de la glace à 0°C. Pour cette espèce, la dégradation de l'ATP, ADP, AMP et IMP est très rapide tandis que les quantités d'inosine (HxR) et hypoxanthine (Hx) croissent tout au long de l'étude.

Au vue de l'ensemble des résultats, il apparaît de manière évidente que l'évolution de la valeur K est dépendante de l'espèce de poisson étudié^{167,168}. En effet, d'une variété à l'autre,

l'activité de l'enzyme *5-nucleotidase* varie significativement¹⁶⁹, cependant, aucune explication biochimique n'a pu être trouvée pour expliquer ces différences. Pour certains cas extrêmes, comme la plie, la dégradation de l'ATP, de l'ADP, de l'AMP et de l'IMP est tellement rapide qu'après seulement un jour, l'ensemble de ces nucléotides est consommé et la valeur K atteint donc sa valeur maximale (100%)¹⁷⁰. Une étude a d'ailleurs été conduite sur ce type de poisson et les résultats obtenus ont confirmé cette évolution de la valeur K.

De plus, les faibles différences observées entre les poissons d'une même espèce sur les données expérimentales peuvent s'expliquer par plusieurs facteurs tels que la variabilité naturelle existante entre les poissons, des petites variations dans les conditions de conservation (durée ou température de stockage légèrement différentes), mais également les conditions utilisées pendant la préparation et l'analyse des échantillons (dans notre cas maintenu à 4°C pendant ces deux étapes). En outre, de faibles changements peuvent être introduits pour la valeur K par la période de l'année pendant laquelle le poisson a été pêché. Pour certaines espèces, la valeur K est 5% plus élevée en été qu'en automne¹⁷¹.

Enfin, de manière générale et ce pour la plupart des espèces de poissons, une valeur K inférieure à 20% indique un très haut niveau de fraîcheur tandis qu'une valeur de 60% est couramment utilisée comme limite d'acceptation pour le consommateur. Cependant, il est important d'avoir à l'esprit que ces valeurs limites dépendent des espèces étudiées et que ces valeurs peuvent être différentes pour les poissons préservateurs d'IMP (daurade, rouget) ou pour les producteurs d'hypoxanthine (plie). Dans ces cas-là, il est nécessaire de coupler ou d'utiliser d'autres méthodes d'estimation de la fraîcheur notamment organoleptiques pour déterminer ces limites. Au vu de celles-ci, il apparaît évident que pour la daurade ainsi que pour le rouget, la limite pour la consommation ne peut être fixée à 60%. Cependant, en utilisant cette limite de rejection pour la truite, une conservation maximale 3 jours après la pêche peut être établie. Chang et al ¹⁶¹ ayant confirmé à travers leur étude qu'une valeur de 60 % pour la valeur K pouvait être utilisée comme limite de rejection pour le bar, à l'aide des résultats obtenus par RMN HR-MAS celle-ci peut être fixée à 7 jours après la pêche pour un poisson conservé à 0°C.

Parallèlement aux différentes études précédemment décrites, l'effet d'une congélation longue a également été étudié sur la daurade. Les études portant sur la queue du poisson, la partie restante non utilisée du poisson, a ainsi été conservée au congélateur pendant 6 mois à -36 °C immédiatement après la découpe de la queue. Comme attendu, la valeur K initiale déterminée

dès la réception du poisson n'a été que très peu affectée par la procédure de congélation (4,38% après 6 mois comparé à une valeur initiale de 4,1 %). Cependant, un nouveau signal est apparu sur le spectre enregistré 6 mois plus tard et ce composé a facilement été attribué au DMA. L'apparition du DMA provient de la dégradation de l'OTMA-N par l'enzyme *OTMA-N demethylase* pendant la conservation au congélateur. Cette observation est consistante avec les données issues de la littérature : le DMA ayant été déjà signalé comme potentiel indicateur de la congélation de certains poissons conservés à -40 °C^{86,172}. La teneur en DMA a ainsi pu être quantifiée de manière absolue à l'aide de la méthode PULCON/ERETIC et une concentration de 1,0 mg/100g de muscle de daurade a alors pu être mesurée. Le DMA semblerait donc être un moyen de prouver qu'un poisson a été congelé pendant une longue période. Cependant, du fait de sa forte volatilité (température d'ébullition de 7°C pour le produit pur), sa détection sur des poissons décongelés complique considérablement son investigation.

N°	Métabolite	Groupe	Déplacement chimique		
			¹ H (ppm)	HSQC	TOCSY
				¹³ C (ppm)	¹ H (ppm)
1	Alanine	βCH ₃	1.47	18.7	
		αCH	3.78	53.1	
		β-CH ₂ (Ala)	2.69	34.64	
		β-CH ₂ (His)	3.08	28.8	
		αCH ₂ (Ala)	3.20		
2	Anserine	αCH ₂ (His)	3.22		
		N-CH ₃	3.82	35.4	
		αCH(His)	4.5	55.83	
		CH-5(His)	7.19	122.5	
		CH-2(His)	8.46	138.3	
3	ATP, ADP, AMP		4.25		
			4.31		
			4.42		
			4.66		
			4.82		
4	Carnosine	CH-8	8.32	155.2	
		CH-2	8.48	142.6	
			2.66		3.21
			3.03		4.5
			3.21	38.4	
5	Créatine/Phosphocréatine		4.50	57.60	
			7.17	119.8	
			8.14	137.65	
		N-CH ₃	3.02	37.5	
		α-CH ₂	3.92	53.9	
6	Acides Gras	CH ₃	0.90		1.30
		(n)CH ₂	1.27		0.90, 2.25
		(2)CH ₂	1.58		1.26
		(1)CH ₂	2.25		1.30
7	Acide fumarique	CH = CH	6.53	Biblio	
		CH-4	3.42	72.4	
		CH-5	3.54		5.23
		CH-3	3.72		5.23
		CH ₂ -6,6'	3.77		5.23
8	α-Glucose	CH-5	3.85	74.0	
		CH ₂ -6,6'	3.86	63.4	
		CH-1	5.23	94.8	
		CH-2	3.24		4.64
		CH-4	3.42	72.4	
9	β-Glucose	CH-3	3.47		4.64
		CH-5	3.51		4.64
		CH ₂ -6,6'	3.73		4.64
		CH ₂ -6,6'	3.89	63.4	
		CH-1	4.64	98.5	
10	GPCCho	N-CH ₃	3.23	56.4	
		CH ₂	3.62	64.7	
		βCH ₂	3.67	64.7	
		CH	3.92	73.2	
11	Glycine	αCH ₂	4.32	62.1	
		αCH	3.56	44.10	
		CH-8	8.19		
12	Hypoxanthine	CH-2	8.21		
		CH ₂ -5(Rib)	3.86		4.26, 4.43
		CH-4(Rib)	4.27	88.2	
13	Inosine (HxR)	CH-3(Rib)	4.43	73.1	
		CH-2(Rib)	4.77	76.7	
		CH-1(Rib)	6.09	90.9	

		CH-8(Purin)	8.22	148.8	
		CH ₂ -5(Rib)	4.02	66.3	
		CH-4(Rib)	4.36	87.4	
		CH-3(Rib)	4.5	73.3	
14	IMP	CH-2(Rib)	4.78		4.5, 6.13
		CH-1(Rib)	6.13	89.9	
		CH-3(Purin)	8.21	148.8	
		CH-8(Purin)	8.55	142.6	
15	Acide lactique	CH ₃	1.33	22.65	
		CH	4.12	71.05	
		γ-CH ₂	1.47	24.0	
		δ-CH ₂	1.72	29.1	
16	Lysine	β-CH ₂	1.90	32.6	
		ε-CH ₂	3.01	41.9	
		αCH	3.76	57.1	
17	Niacine	CH-5	7.53	Biblio	
		CH-4	8.26	Biblio	
		CH-6	8.70	Biblio	
		CH-2	8.92	Biblio	
18	Ω3	CH ₃	0.96		2.06, 5.31
		CH ₂	2.06		2.82, 5.31
19	Taurine	-CH ₂ -NH ₃ ⁺	3.26	50.48	
		-CH ₂ -SO ₃ ⁻	3.42	38.17	
20	TMA-N	N-CH ₃	2.89	47.2	
21	OTMA-N	N-CH ₃	3.27	62.2	
22	Acides gras insaturés(AGI)	(2)CH ₂	2.04		1.30, 5.33
		CH	5.33	130.3	1.30, 2.06, 2.80
23	Valine	γCH ₃	0.98		2.27
		γ'CH ₃	1.04	20.6	
		βCH	2.27		0.98, 1.04

Tableau 6 : Attribution des différents métabolites détectés dans une biopsie de poisson. Un blanc dans une case de la colonne des déplacements chimiques du ¹³C indiquent que le composé n'a pu être détecté sur le spectre ¹H-¹³C HSQC pour des raisons de sensibilité.

II.1.3) Conclusion

Les travaux décrits au cours de cette section démontrent, pour la première fois, que la RMN HR-MAS est une méthode analytique rapide et fiable de détermination de la qualité et de la fraîcheur du poisson à travers l'utilisation de deux indicateurs chimiques internationaux à savoir la valeur K et l'analyse du TMA-N. En effet, les résultats obtenus par RMN HR-MAS sont en accord avec ceux issus de la littérature obtenus à l'aide d'autres techniques analytiques et présentent une très bonne répétabilité. La méthodologie mise au point par RMN HR-MAS permet l'analyse directe d'échantillons maintenus à 4 °C pendant toute la durée de l'expérience, sans longues étapes de préparation et d'extraction, garantissant qu'un faible biais est introduit sur la mesure des indicateurs. De plus, les faibles quantités de poisson requises (entre 10 et 15 mg) ainsi que la durée relativement courte de l'expérience RMN (temps de préparation et d'acquisition) sont compatibles avec une analyse industrielle de routine.

I.2) Analyse de blocs de farce de colin d'Alaska

A l'initiative d'un partenaire industriel du projet Agrifood GPS (Findus France), une étude portant sur l'analyse de blocs de farce de colin d'Alaska (*Theragra chalcogramma*), a été conduite sur des échantillons provenant des différents fournisseurs de cette société, dans le but de comprendre les différences observées lors des analyses organoleptiques réalisées par leur service qualité au sein de leur usine de Boulogne-sur-Mer. Ces minerais de poisson sont utilisés sur l'ensemble des lignes de production pour la fabrication des différents produits de la mer proposés par Findus.

Pour cela, 24 blocs distincts provenant des 6 fournisseurs avec qui la société Findus travaille, nous ont été envoyés. L'échantillonnage réalisé couvre ainsi 8 mois différents de pêche sur deux années de pêche distinctes (2013 et 2014). L'ensemble du plan d'échantillonnage est donné en Tableau 7.

N°	Lot	Fournisseur	Mois de pêche-Année
1	13879-1	A	févr-13
2	13927-1	A	févr-13
3	14508-1	A	juin-14
4	14421-3	A	févr-14
5	14581-1	A	juil-14
6	13936	B	janv-13
7	13945-1	B	mars-13
8	13817-1	B	avr-13
9	14534-1	B	août-14
10	14572-5	B	juin-14
11	14552-5	B	sept-14
12	13900-1	C	janv-13
13	13900-5	C	janv-13
14	14505-3	C	mars-14
15	14542-1	C	juil-14
16	13806	D	févr-13
17	13878-1	D	févr-13
18	13947-5	D	mars-13
19	14507-1	D	juil-14
20	14521-1	E	août-14
21	14522	E	juin-14
22	14526	E	juil-14
23	14426-4	F	févr-14
24	14389	F	févr-14

Tableau 7 : Plan d'échantillonnage conçu par la société Findus France pour la collecte d'échantillon de blocs de farce de colin d'Alaska dans le cadre du projet Agrifood GPS.

L'ensemble de ces échantillons a, alors, été analysé en triplicat (trois analyses correspondant à trois zones de prélèvement différentes) par RMN ^1H HR-MAS suivant le protocole de préparation décrit précédemment (**Partie III**). Les paramètres d'acquisition ainsi que la séquence RMN employée (CPMG avec présaturation de l'eau) pour l'acquisition des différents spectres 1D, sont ceux présentés dans la **Partie III**.

Dans un premier temps, la méthode de détermination de la fraîcheur de la valeur K préalablement présentée a alors pu être appliquée à ce type d'échantillon une fois l'intégralité de ceux-ci étudiés par RMN. Dans un deuxième temps, les informations fournies par les différents spectres ^1H HR-MAS ont ensuite été soumises à des analyses statistiques multivariées (ACP, OPLS-DA) dans le but de différencier les échantillons selon les fournisseurs ou les mois de pêche.

I.2.1) Détermination de la fraîcheur des blocs de farce de colin d'Alaska : la valeur K

La méthode de détermination de la fraîcheur par RMN HR-MAS à travers la valeur K a été mise en application sur l'ensemble des blocs de farce de colin d'Alaska. Le TMA-N étant d'avantage un indicateur de l'activité microbienne et celui-ci étant mesuré lors des analyses sur la fraîcheur réalisées par Findus (ABVT), il a été décidé de concentrer nos mesures sur la valeur K. L'ensemble des mesures de la valeur K moyenne pour chaque lot est présentée en Tableau 8. Les échantillons sont classés par ordre de fraîcheur (du plus frais au moins frais). Mis à part pour l'échantillon 13806 du fournisseur D, les valeurs K mesurées sont comprises entre 65 % et 100%.

N°FINDUS	Valeur K	N°FINDUS	Valeur K
13806 (D)	23,4	14552-5(B)	85,43
14426-(F)	65,64	14521-1 (E)	88,87
13900-1 (C)	65,96	14522 (E)	91,27
13927-1 (A)	70,79	14508-1 (A)	91,97
13947-5 (D)	71,29	14542-1 (C)	92,03
14389-11 (F)	71,65	14581-1 (A)	94,9
13900-5 (C)	75,44	14534-1(B)	95,62
13879-1 (A)	78,46	13817-1 (B)	95,78
13936 (B)	84,01	14505-3 (C)	96,83
13878-1 (D)	84,54	14572-5 (B)	97,62
13945-1 (B)	84,76	14421-3 (A)	98,93
14507-1 (D)	85,22	14526 (E)	100

Tableau 8 : Moyennes de la valeur K (en%) déterminées pour chaque échantillon de farce de colin d'Alaska de la société Findus. La lettre entre parenthèses représente le code utilisé pour chaque fournisseur.

De ces résultats, certaines tendances en fonction des fournisseurs peuvent être dégagées. Tout d'abord, les échantillons provenant des fournisseurs D et surtout F apparaissent plus frais de manière générale que ceux provenant des autres prestataires. En revanche, les échantillons provenant du fournisseur B ressortent comme faisant partie des échantillons les moins frais. Cette observation n'est pas surprenante puisque le fournisseur B est le seul à utiliser un système particulier de pêche à savoir celui d'un bateau-mère (« mothership »). Ce système de pêche particulier est composé d'un bateau principal appelé bateau-mère, au sein duquel le poisson est stocké et transformé, et de plusieurs petits bateaux qui vont chaque jour pêchés et collectés les poissons en mer puis les ramènent au bateau-mère. Le bateau-mère ne retourne au port que lorsque l'ensemble de la cale est remplie cela pouvant prendre plusieurs semaines ou quelques mois. Ce système de pêche est une alternative au système de pêche classique (utilisé par les autres fournisseurs) où chaque chalutier part le matin et rentre chaque soir au port pour vendre sa marchandise. Ainsi, la différence de fraîcheur observée sur les échantillons du fournisseur B peut sans doute s'expliquer par une période plus longue entre la pêche et la transformation que pour les autres prestataires.

I.2.2) Analyses statistiques

Afin de comprendre les différences observées lors des tests organoleptiques entre les différents blocs de poissons, des analyses statistiques multivariées ont été appliquées aux données acquises par RMN HR-MAS.

I.2.2.a) Prétraitement des données

Tout d'abord, le FID est multiplié par une fonction mathématique, également appelée fonction d'apodisation, de type exponentielle décroissante avec un paramètre d'élargissement de raies égal pour cette fonction à 1 Hz. L'utilisation d'une telle fonction engendre une augmentation du signal/bruit en minimisant l'influence des derniers points du FID composés principalement de bruit et présente l'avantage de ne pas affecter le début du FID celle-ci valant 1 pour $t=0$. Ensuite, pour chaque spectre RMN, la phase a été ajustée manuellement puis une correction de ligne de base est alors appliquée. Ils sont alors référencés sur le pic de gauche du doublet de l'acide lactique à 1,3376 ppm. L'alignement des différents pics ainsi obtenu étant très bon sur les échantillons de poissons, il n'a pas été nécessaire d'utiliser de fonction de réalignement. Comme nous l'avons vu lors de la **Partie II**, afin de rendre les données utilisables pour une analyse statistique, nous avons alors découpé les spectres RMN en petites régions (« buckets ») de 0,03 ppm, en excluant la zone spectrale correspondant au signal de l'eau. La masse de muscle de poisson prélevée pouvant varier (entre 10 et 12 mg), une normalisation par rapport à l'intensité globale de l'ensemble des « buckets » a alors été effectuée. Enfin, pour chaque variable, les données ont été centrées et mises à l'échelle par division de chaque variable par la racine carrée de son écart-type, appelée mise à l'échelle Pareto. Le nombre de composantes pour les modèles OPLS-DA a été déterminé par validation-croisée.

I.2.2.b) Résultats des Analyses Statistiques

Comme nous l'avons vu précédemment, lorsque le nombre de spectres et de variables devient conséquent, il devient extrêmement difficile de pouvoir isoler par simple comparaison les zones spectrales spécifiques à un groupe ou une origine d'échantillons. Les méthodes d'analyses statistiques multivariées non supervisées, telles que l'ACP, permettent alors de réduire le nombre de variables et de détecter les marqueurs (variables significatives) responsables des groupes d'échantillons. Comme vu dans la **Partie II**, l'origine chimique est alors révélée par l'interprétation du graphe des *loadings*, à travers les valeurs des coefficients de corrélation entre les variables initiales et les composantes principales.

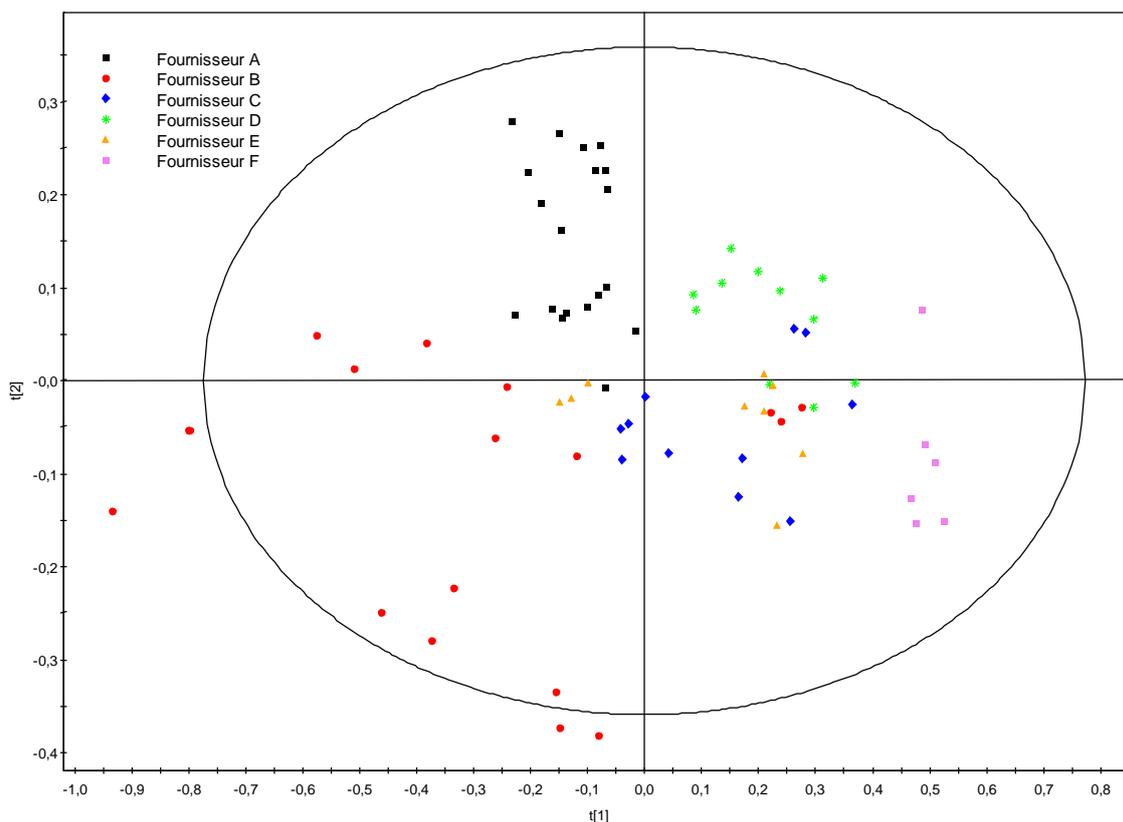


Figure 31 : Graphe des scores (CP_1 vs CP_2) de l'ACP appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. La variance totale expliquée par ces deux composantes est de 64,37% (52,98% pour CP_1 et 11,39% pour CP_2). Chaque couleur représente un Fournisseur différent. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

Les résultats obtenus à l'aide de l'ACP (CP_1 vs CP_2) sur l'ensemble des blocs de farce de colin d'Alaska sont donnés en Figure 31. Ces deux composantes principales représentent 64,37% de la variance totale (52,98 % pour CP_1 et 11,39 % pour CP_2). Comme nous pouvons le voir sur cette figure, des tendances en fonction de certains fournisseurs peuvent être extraites. Tout d'abord, les échantillons du fournisseur F se différencient des autres par une contribution positive selon CP_1 . En regardant les coefficients de corrélations de la première composante principale sur le graphe des *loadings*, celle-ci est principalement corrélée aux régions spectrales correspondant au TMA-N et à l'OTMA-N ces deux zones possédant respectivement des coefficients de corrélation de -0,878 et 0,92 sur CP_1 . Cette différenciation pour ce groupe d'échantillon est ainsi principalement liée à une teneur plus élevée en OTMA-N sur le spectre RMN (scores positifs sur CP_1) et une quantité moins élevée de TMA-N. Une teneur plus importante en OTMA-N, combinée à un niveau plus faible en TMA-N, atteste d'une plus faible dégradation bactérienne pour ces échantillons et donc d'une meilleure qualité de ces derniers. Cette observation est cohérente avec les résultats obtenus pour la

valeur K sur ces échantillons, où ceux-ci apparaissent parmi les plus frais du point de vue de cet indicateur. Une autre tendance selon la première composante principale est observée sur les échantillons du fournisseur B (points bleus sur la Figure 31). En effet, mis à part pour l'échantillon n° 13936, les spécimens de ce prestataire possèdent des scores négatifs sur CP₁. Ainsi les blocs de poisson provenant de ce vendeur sont caractérisés par des teneurs plus faibles en OTMA-N et plus élevées en TMA-N, indicatif d'une dégradation bactérienne plus importante. Là encore ces résultats sont en accord avec ceux issus de la détermination de la fraîcheur à travers la valeur K, puisque ces blocs faisaient partie des échantillons les moins frais. Enfin, une dernière tendance est observée pour les minerais correspondant au fournisseur A (points noirs sur la Figure 31), cette fois-ci, sur la deuxième composante principale (CP₂). Celle-ci est principalement corrélée à la région spectrale de l'acide lactique à 1,33ppm (coefficient de corrélation de -0,84) ainsi que d'un composé inconnu à 2,71 ppm (coefficient de 0,81). Ainsi les échantillons du fournisseur A sont définis par des teneurs plus faibles en acide lactique (coordonnées positives sur CP₂) pouvant s'expliquer par un technique de pêche différente des autres fournisseurs provoquant moins de stress *ante mortem* pour ces poissons.

Les différents points correspondant aux analyses réalisées sur trois zones différentes du bloc de colin d'Alaska d'un même lot sont relativement proches sur l'ACP pour la plupart des échantillons, cependant, dans certains cas leur position relative est un peu plus dispersée. Les variations standards observées sur la majorité des points semblent ainsi correspondre à la reproductibilité expérimentale tandis que les variations plus importantes sont probablement dues à une certaine hétérogénéité des minerais de poisson, ces blocs étant composés de filets de colin provenant de différents individus.

Cependant, bien que l'ACP puisse être utilisée comme technique statistique de classement non supervisée des échantillons nous avons également appliqué une méthode statistique supervisée de classement, l'OPLS-DA, dans le but de déterminer si nous pouvions classer l'ensemble des blocs de farce de colin d'Alaska en fonction des fournisseurs, dans un premier temps, puis en fonction du mois ou de la saison de pêche, dans un deuxième temps.

Ainsi, nous avons, tout d'abord, effectué une analyse OPLS-DA à l'aide du logiciel SIMCA-P (version 13, Umetrics) sur la même matrice de données que celle utilisée pour l'ACP à la différence qu'une matrice factice Y est introduite contenant les groupes d'appartenance des échantillons, celle-ci possédant autant de colonnes que de classes fixées. Dans le cas présent,

celle-ci est composée de 6 classes correspondant aux 6 fournisseurs (A, B, C, D, E et F). Les résultats de l'OPLS-DA sont donnés en Figure 32. Comme nous pouvons le voir, l'utilisation de l'OPLS-DA améliore, de manière significative, la discrimination des fournisseurs A, B et F par rapport à celle obtenue à l'aide de l'ACP. En revanche, l'application de cette méthode multivariée supervisée ne permet pas la distinction entre les autres prestataires (C, D et E) comme l'atteste le paramètre de prédiction ($Q^2=0,554$).

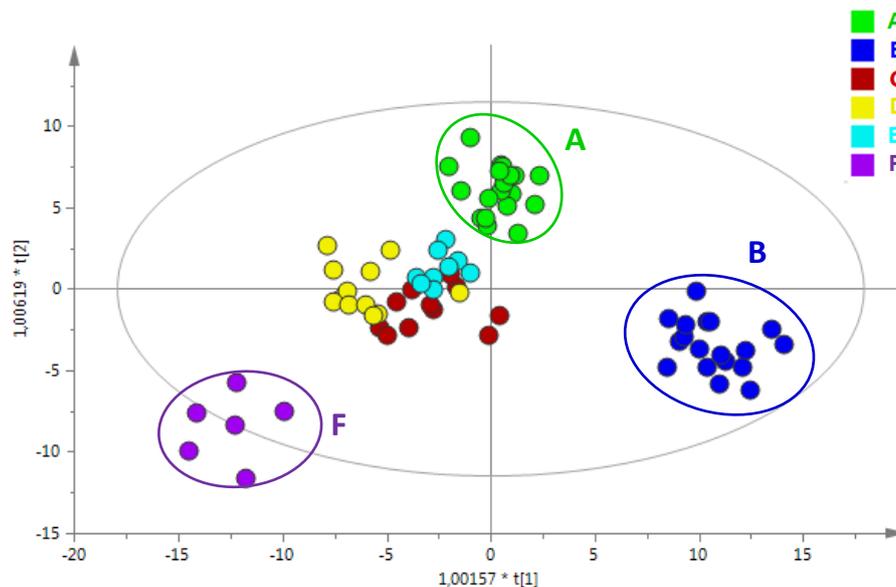


Figure 32 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. Chaque classe représente un fournisseur différent. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95% Paramètres du modèle : $R^2Y=0,807$ et $Q^2=0,554$.

De même que pour l'étude en fonction des fournisseurs, l'OPLS-DA a également été appliquée à la matrice de données initiales, cette fois-ci en regroupant les échantillons non pas par leur origine (prestataire) mais en fonction du mois durant lequel les poissons utilisés pour la constitution des blocs ont été pêchés. Ainsi pour cette étude, la matrice Y est composée de 7 groupes correspondant à 7 mois de pêche (Janvier, Février, Mars, Avril, Juin, Juillet et Août). Les données enregistrées sur l'échantillon correspondant au mois de septembre n'ont pas été utilisées pour la construction de l'OPLS-DA du fait du nombre insuffisant de données pour cette période de pêche (un seul échantillon). Comme nous pouvons le voir sur la Figure 33, les résultats obtenus à l'aide de l'OPLS-DA ($R^2Y= 0,803$ et $Q^2=0,48$) ne permettent de séparer les différents échantillons en fonction du mois de pêche. Cependant, il est tout de même important de noter qu'une tendance se dégage sur le premier axe de l'OPLS-DA, à

savoir une certaine chronologie en fonction de la période de pêche. En effet, sur cet axe, les différents groupes sont de manière générale disposés chronologiquement : les échantillons des mois correspondant à la première saison de pêche (Janvier à Avril) possédant des scores positifs tandis que ceux correspondant aux mois de la seconde saison de pêche (Juin à Octobre) ont des scores négatifs sur celle-ci.

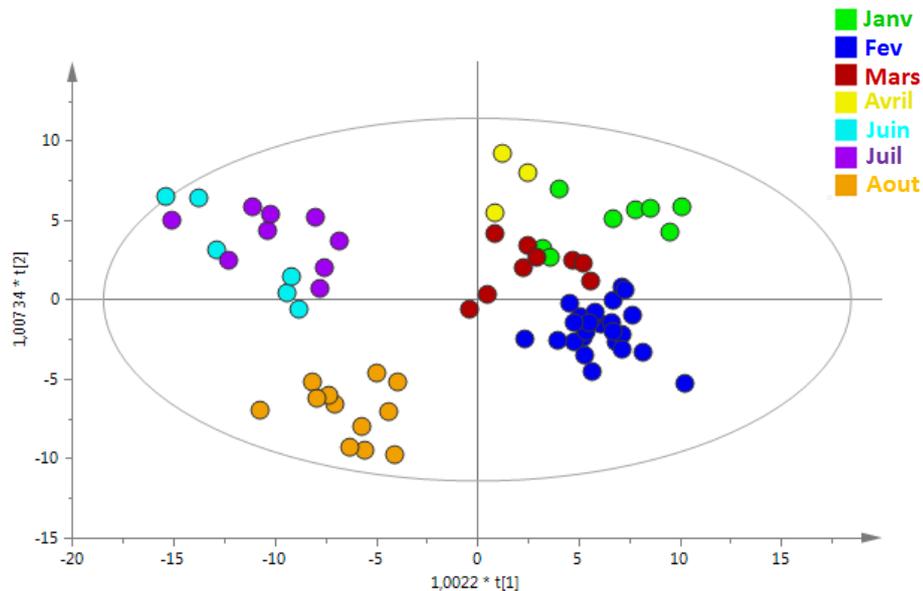


Figure 33 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. Chaque classe représente un mois de pêche différent. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%. Paramètres du modèle : $R^2Y=0,803$ et $Q^2=0,48$.

Une analyse OPLS-DA a alors été réalisée en regroupant les échantillons par saison de pêche en deux groupes : le premier rassemblant toutes les analyses des mois de Janvier à Avril et le second contenant tous les spécimens pêchés entre juin à septembre. Le graphe des scores de ce modèle est présenté en Figure 34.

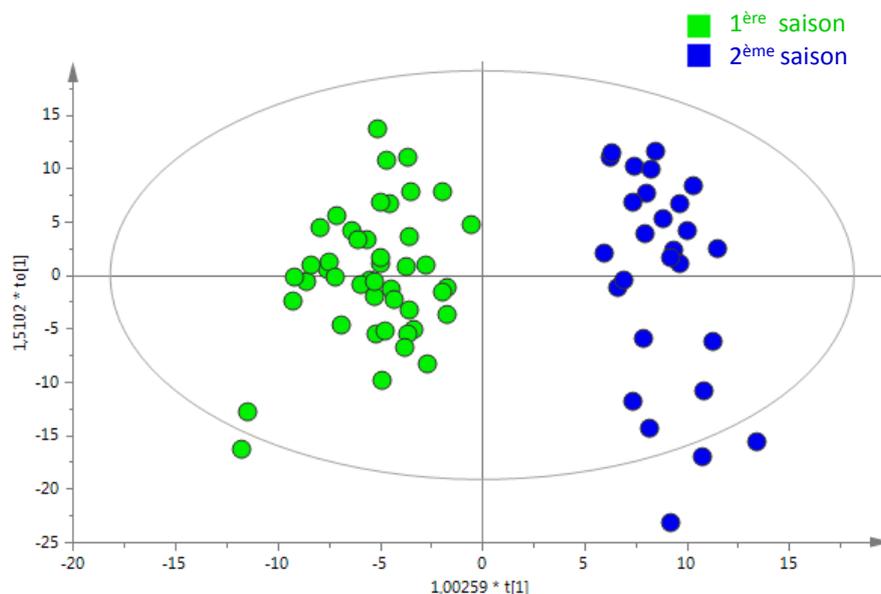


Figure 34 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. Chaque classe représente une saison de pêche différente. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%. Paramètres du modèle : $R^2Y=0,903$ et $Q^2=0,831$.

Comme nous pouvons le voir celui présente une excellente séparation entre les deux classes d'échantillons comme le confirme la valeur élevée du Q^2 (0,831). Au vu de ces résultats, il semblerait donc que la tendance observée sur le précédent modèle soit réelle et que la saison de pêche influence directement le profil du colin d'Alaska déterminé par RMN. En regardant les scores des VIP, la différence entre les deux groupes d'échantillon se fait principalement à l'aide des variables correspondant aux signaux des acides gras (saturés et insaturés) ainsi que dans une moindre mesure à ceux de certains métabolites tels que la glycine et la tyrosine.

La composition du poisson est directement reliée à son alimentation, à ses flux migratoires ainsi qu'à la proximité de la pêche avec la période de reproduction. En effet, des poissons ayant réalisés de longs flux migratoires en vue de leur reproduction présenteront des teneurs plus faibles en acides gras et protéines. Une fois la reproduction assurée, les poissons vont alors chercher des endroits riches en nourriture et la teneur en acides gras ainsi qu'en protéines va alors remonter de manière significative. La ponte des œufs ayant lieu entre Février et début Mai les échantillons correspondant à la première saison de pêche vont ainsi être caractérisés par des teneurs plus faibles en lipides tandis que ceux de la deuxième partie de l'année (seconde saison de pêche) présenteront, quant à eux, des quantités d'acides gras

nettement plus importantes. C'est cette tendance que nous observons sur le profil métabolique du colin d'Alaska déterminé par RMN.

II.2.3) Conclusion

Au cours de cette section, nous avons vu que la méthode de détermination de la fraîcheur du poisson par RMN HR-MAS, développée précédemment, a pu être appliquée à l'analyse de blocs de farce de colin d'Alaska et que différents niveaux de fraîcheur ont pu être observés entre certains fournisseurs.

Le profil métabolique enregistré par RMN HR-MAS a également été soumis à des analyses statistiques qui ont permis de dégager certaines tendances en fonction des fournisseurs, confirmant les résultats sur la détermination de la fraîcheur, les principales différences entre les échantillons provenant des teneurs en OTMA-N et TMA-N. Pour un fournisseur, une teneur en acide lactique plus faible pour le fournisseur A probablement liée à une technique de pêche différente de celle employée par les autres fournisseurs. L'utilisation de l'OPLS-DA en établissant des classes en fonction des fournisseurs a confirmé les résultats de l'ACP, celle-ci différenciant les fournisseurs A, B et F, cependant, celle constituant les différentes classes en fonction du mois de pêche, n'a pu dégager de tendances. De plus, l'OPLS-DA a également permis de mettre en évidence l'influence de la saison de pêche (teneur en lipides) sur l'empreinte métabolique du colin d'Alaska.

Au vu des résultats obtenus par RMN HR-MAS, les différences observées en fonction des fournisseurs lors des analyses organoleptiques sont, probablement, en partie liées à des niveaux de fraîcheur distincts résultant de techniques de pêches et de conservation légèrement différentes (temps avant la transformation du poisson).

Nous avons alors décidé de mener une expérimentation portant sur la texture du poisson, à travers l'étude du temps de relaxation transversale de l'eau, afin de déterminer si nous pouvions corrélérer les différences de textures observées lors des tests organoleptiques avec ce paramètre RMN. C'est cette thématique que nous allons développer dans la prochaine section.

I.3) Etude de la texture du poisson : Temps de relaxation transversale (T₂) de l'eau

I.3.1) Etude de la texture des blocs de farce de colin d'Alaska

La société Findus, a décidé, fin 2013, de participer au projet Agrifood GPS en tant que partenaire industriel. Jusqu'à présent, la plupart des analyses réalisées sur les produits bruts utilisés pour la conception des différents produits alimentaires étaient réalisées au sein de leur site de Boulogne-sur-Mer et consistaient pour la plupart en des tests visuels ou organoleptiques. Certaines analyses physico-chimiques de base (ABVT notamment) et microbiologiques étaient effectuées au sein de leur laboratoire sur ce même site tandis que celles ne pouvant être réalisées sur place (analyses ADN par exemple) étaient sous-traitées auprès de la société Eurofins Scientific.

Ainsi, afin de comprendre les différences observées sur la texture de certains blocs de farce de colin d'Alaska, plusieurs échantillons nous ont été transmis (les même que ceux étudiés précédemment) afin d'étudier ces spécimens par RMN HR-MAS et d'essayer de comprendre les différences observées lors des analyses organoleptiques sur la texture de certains échantillons au sein de l'usine de Boulogne-sur-Mer. Par exemple, le poisson provenant du fournisseur F possède une excellente texture lorsque celui-ci est contrôlé par dégustation mais présente des soucis de transformation lorsque celui-ci est utilisé sur la ligne de production. De plus, la société Findus soupçonne certains fournisseurs d'injecter certains additifs interdits tels que les carbonates pour gonfler le poids du poisson, ceux-ci permettant la rétention d'eau. Afin d'étudier la texture des différents échantillons fournis par la société Findus France, nous avons étudié le temps de relaxation transversale de l'eau (T₂) : ce paramètre RMN étant directement corrélé aux proportions et constantes de relaxation transversale des différents types d'eau présents au sein des muscles de poisson.

Plusieurs théories existent pour expliquer le comportement de la relaxation de l'eau se trouvant dans les tissus musculaires. Cependant toutes s'accordent sur le fait que celle-ci peut être décrite par une fonction multi-exponentielle, correspondant à différentes structures de l'eau¹⁷³.

La première théorie met en évidence que l'existence de deux composantes reflète l'eau intra et extra cellulaire, où la membrane cellulaire joue le rôle de barrière physique¹⁷⁴. La seconde théorie¹⁷⁵ suggère que l'eau intracellulaire est plus ou moins structurée en fonction des changements de structures dans le système actomyosine, par exemple le degré de contraction.

L'actomyosine étant une protéine formée par l'assemblage d'actine et de myosine et entrant en jeu dans le processus de contraction des fibres musculaires. La dernière théorie¹⁷⁶ explique le comportement multi-exponentiel de la relaxation par l'échange rapide entre l'eau libre et la surface macromoléculaire présentant des pores de différentes tailles (théorie de l'hétérogénéité spatiale).

De plus, bien que la décroissance de la relaxation transversale soit reconnue comme étant multi-exponentielle dans les tissus musculaires, un autre débat existe entre un comportement bi-exponentiel (T_{21} et T_{22})¹⁷⁴ ou tri-exponentiel (T_{2b} , T_{21} et T_{22})¹⁷³. Pour le premier cas, la constante de temps courte (T_{21}) représente l'eau étroitement liée aux macromolécules tandis que la constante de temps de relaxation longue (T_{22}) représente l'eau libre à l'intérieur des tissus. Dans le deuxième cas, T_{2b} est associée au comportement de T_{21} dans le cas d'une décroissance bi-exponentielle, tandis que T_{21} reflète l'eau localisée à l'intérieur de la structure des protéines (l'eau se situant dans la structure tertiaire ou quaternaire des protéines) et T_{22} représente l'eau extra myofibrillaire contenant la fraction sarcoplasmique des protéines, c'est-à-dire relative au sarcoplasme qui est le cytoplasme entourant les fibrilles des fibres musculaires.

D'un point de vue expérimental, pour chacun des 24 blocs de farce de colin d'Alaska 6 mesures des temps de relaxation transversale ont été réalisées correspondant à 3 zones de prélèvement différentes analysées en duplicat. Nous avons décidé d'effectuer autant de mesures à cause de la composition relativement hétérogène des échantillons : ceux-ci étant composés de plusieurs poissons hétéroclites. De plus, les différents prélèvements ont été réalisés sur 3 faces différentes et cela légèrement en profondeur : la texture de l'échantillon à la surface du bloc pouvant être influencée par le processus de découpe du minerai.

La séquence utilisée pour la détermination des différentes constantes de temps de relaxation transversale ainsi que leur proportion relative, est la séquence CPMG pseudo 2D présentée en **Partie III**.

Une étude préliminaire a été menée sur plusieurs échantillons d'un même bloc et d'une même zone de prélèvement dans le but de déterminer les conditions d'acquisition optimales. Tout d'abord, nos mesures portant sur le temps de relaxation transversale de l'eau, des essais avec et sans ajout de solvant ont été réalisés. L'addition de solvant modifiant légèrement les proportions d'eau libre/eau liée ainsi que leurs valeurs associées de temps de relaxation transversale, nous avons décidé que l'ensemble de l'étude serait réalisée sans ajout de solvant

deutérimé et donc sans verrouillage champs/fréquence afin de ne pas biaiser les résultats de nos analyses.

Durant cette étape préliminaire à l'analyse de l'ensemble des échantillons, l'effet de la dégradation par la mise en rotation de l'échantillon a également été exploré. Ainsi des acquisitions ont été effectuées à différentes vitesses de rotation : 5000Hz, 1500 Hz et 500Hz. De plus, des acquisitions ont été réalisées sur des rotors laissés en rotation au sein du spectromètre pendant une heure afin d'étudier l'impact de ce paramètre sur la structure des tissus de poisson. Les résultats obtenus pour les différentes vitesses de rotation sur l'échantillon n° 13947-5 sont donnés dans le Tableau 9.

Vitesse de rotation (Hz)	T ₂₁ (ms)	T ₂₂ (ms)	pT ₂₁ (%)	pT ₂₂ (%)
5000	19,6	162	50,5	49,5
5000*	19,5	167	36,2	63,8
1500	18,1	153	66,1	33,9
1500*	18,9	160	63,2	36,8
500	17,8	154	71,6	28,4
500*	17,4	152	71,4	28,6

Tableau 9 : Mesures des constantes de temps de relaxation transversale de l'eau et des proportions relatives en eau libre et eau liée déterminées sur l'échantillon 13947-5 à différentes vitesses de rotation. Les mesures marquées d'un astérisque représentent les analyses réalisées sur des échantillons ayant passé une heure en rotation au sein du spectromètre.

Comme nous pouvons le voir, les teneurs déterminées en eau libre (pT₂₂) pour des vitesses de rotation importantes (1500 Hz et surtout 5000 Hz), augmentent au cours du temps attestant d'une dégradation des tissus. En revanche, pour une vitesse de rotation faible (500Hz) les proportions en eau libre et en eau liée ne varient que très peu entre les deux mesures témoignant de la préservation de la structure de l'échantillon.

Au vue de la dégradation observée sur les échantillons soumis à des vitesses de rotation de 5000Hz et 1500Hz, nous avons pris la décision de travailler à une vitesse de 500 Hz et d'utiliser l'algorithme SLOW-MAS (BL4 SLOW) de l'unité pneumatique permettant une mise en rotation de l'échantillon douce et stable à basse vitesse. L'inconvénient de travailler à des vitesses lentes en HR-MAS, est l'apparition de nombreuses bandes de rotation sur le spectre RMN pour les pics présents en quantité significative. Cependant, étant donné que, pour cette étude, seul le pic de l'eau nous intéresse et non pas l'intégralité de la fenêtre

spectrale, et que les bandes de rotation ne se superposent pas au pic de l'eau situé vers 4,7 ppm, leurs présences n'affectent pas nos analyses.

Une fois l'ensemble des paramètres d'acquisition mis au point, les spectres CPMG pseudo 2D ont été acquis pour l'intégralité des 24 blocs de farce de colin d'Alaska puis traités sous le logiciel TopSpin (version 3.2) avant d'être analysés à l'aide du Dynamics Center (version 2.2.4) qui est un logiciel Bruker spécialement conçu pour le traitement des données de relaxation, diffusion, cinétique ou de polarisation croisée.

Préalablement à l'analyse des données sous le Dynamics Center, l'intégralité des spectres a, ainsi, été phasée manuellement et une correction de ligne de base manuelle a également été appliquée ; ces paramètres influençant directement les mesures des intégrales nécessaires pour la détermination des constantes de temps de relaxation transversale ainsi que les proportions relatives en eau libre et eau liée. De plus, les incréments correspondant à des temps de CPMG long et donc à l'application de nombreuses impulsions RF consécutives, entraînent un faible échauffement de l'échantillon provoquant un léger décalage du pic de l'eau. Pour remédier à ce problème, une correction par un décalage des données de quelques points à travers la commande sous TopSpin *ls* (décalage vers la gauche) ou *rs* (décalage vers la droite) a alors été appliquée pour les 3 derniers points de la liste correspondant respectivement à des enregistrements réalisés avec 720, 1440 et 2880 boucles CPMG.

Afin de mesurer la décroissance de l'aimantation dans le plan transversal, la sélection du pic de l'eau sous le Dynamics Center a été réalisée manuellement en définissant un seuil d'intensité sur le premier incrément entraînant seulement la sélection du pic d'intérêt situé à 4,7 ppm. Pour déterminer la quantité de signal des différents délais de la liste, nous avons utilisé l'aire de ce pic et non l'intensité; cette dernière pouvant être influencée par l'évolution de l'homogénéité du champ magnétique au cours de l'expérience. Enfin, la fonction bi-exponentielle suivante a été définie pour le calcul des différentes constantes de temps de relaxation transversale (T_{21} et T_{22}) ainsi que les proportions relatives en eau libre (pT_{22}) et eau liée (pT_{21}) :

$$f(t) = I_{0 \text{ liée}} e^{-t/T_{21}} + I_{0 \text{ libre}} e^{-t/T_{22}}$$

Les valeurs $I_{0 \text{ liée}}$ et $I_{0 \text{ libre}}$ permettent de remonter à la quantité initiale d'eau liée et d'eau libre et le rapport de ces intensités sur l'intensité totale de calculer leur proportion relative. Des essais ont été effectués avec une fonction tri-exponentielle : certaines publications décrivant la

relaxation transversale de l'eau par trois composantes comme nous l'avons vu précédemment. Cependant, l'ajustement des courbes obtenue avec la fonction bi-exponentielle étant d'excellente qualité et le logiciel de traitement n'arrivant que très rarement à calculer trois composantes, nous avons décidé d'utiliser la première solution. Un exemple de courbe de décroissance bi-exponentielle du signal dans le plan transversal de l'eau obtenue sur l'échantillon 13878-1 est donné en Figure 35.

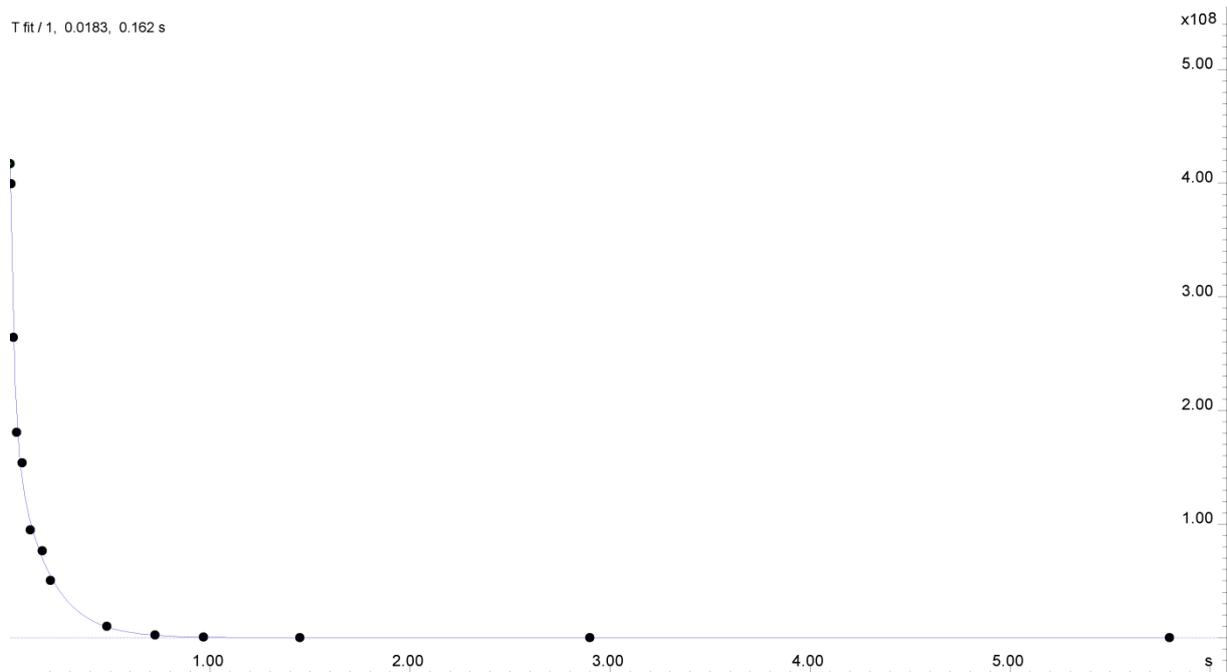


Figure 35 : Exemple de courbe d'ajustement de la décroissance du signal de l'eau dans le plan transversal obtenue avec une fonction bi-exponentielle sur l'échantillon 13878-1 sous le logiciel Dynamics Center. La corrélation entre les mesures et la courbe d'ajustement est excellente.

L'ensemble des résultats obtenus (T_{21} , T_{22} , pT_{21} et pT_{22}) pour les différents échantillons sont regroupés sous forme de graphe pour faciliter leur visualisation (Figure 36-39).

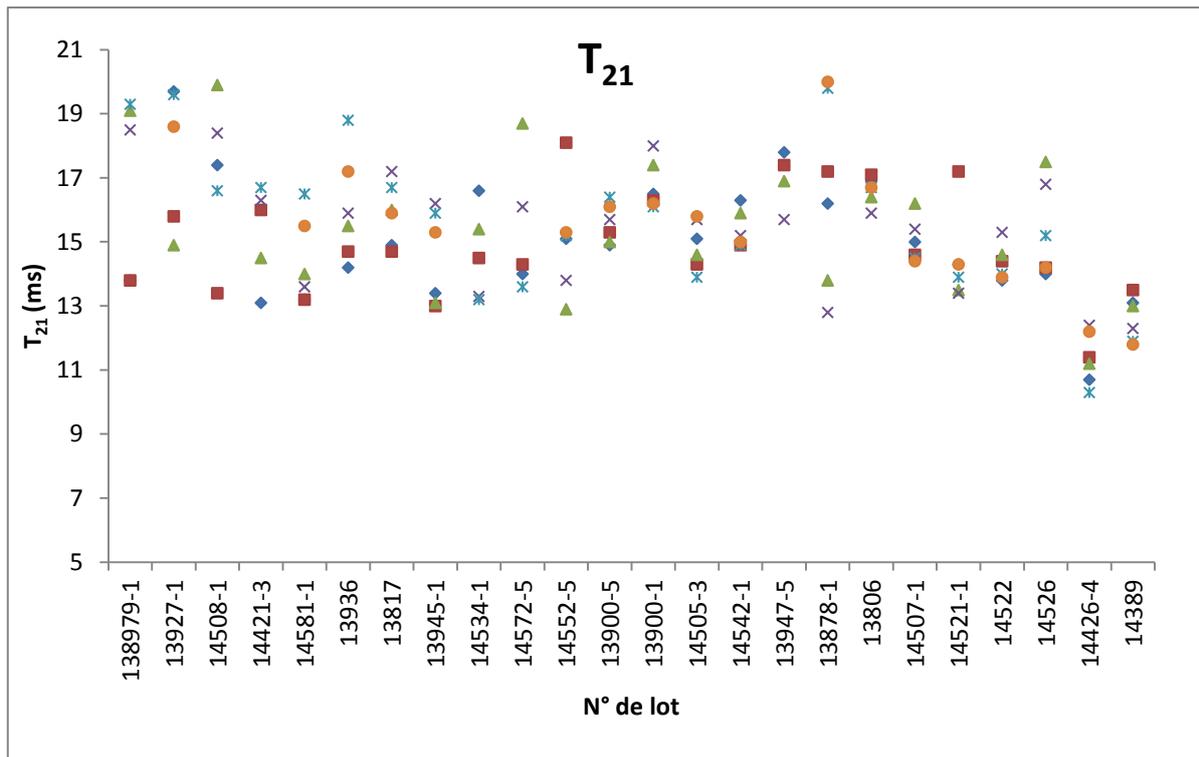


Figure 36 : Graphe regroupant les mesures des constantes de temps de relaxation transversale (T_{21}) de l'eau liée sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.

Tout d'abord, comme nous pouvons le voir sur les différentes figures, la plupart des échantillons présentent un profil « standard » avec des teneurs moyennes en eau liée comprises entre 69 et 72 % et donc des teneurs en eau libre entre 28 et 31% avec respectivement des constantes de temps de relaxation moyennes d'environ 15-16 ms pour T_{21} et d'environ 140 ms pour T_{22} .

Cependant, certains échantillons possèdent des profils bien particuliers du point de vue de nos analyses et tout particulièrement celles portant sur les blocs du fournisseur F (14426-4 et 14389). En effet, les valeurs moyennes déterminées pour T_{21} et T_{22} sur ces deux spécimens sont très significativement différentes que celles mesurées sur les autres blocs : 11,4 ms et 12,6 ms pour T_{21} , et 95 ms et 97,7 ms pour T_{22} . Comme nous l'avons vu dans la **Partie I**, plus le mouvement moléculaire est important au sein d'un échantillon, moins l'amplitude des champs magnétiques locaux est importante ce qui conduit à une augmentation de la constante de temps de relaxation transversale. En ayant cette propriété en tête, les échantillons du fournisseur F présentant des valeurs de T_{21} et T_{22} plus faibles que les autres spécimens fournis

par Findus, la mobilité au sein de ces blocs est ainsi plus restreinte témoignant d'une texture plus ferme.

De plus, les proportions moyennes calculées en eau liée pour ces deux blocs sont légèrement supérieures à celles définies sur les autres minerais de poisson (72,8% et 73,7%). Ces échantillons présentant une meilleure texture lors des analyses organoleptiques, il semblerait qu'une teneur en eau liée plus importante ainsi que des constantes de temps de relaxation transversale plus courtes, attestent d'une meilleure qualité du point de vue de la texture.

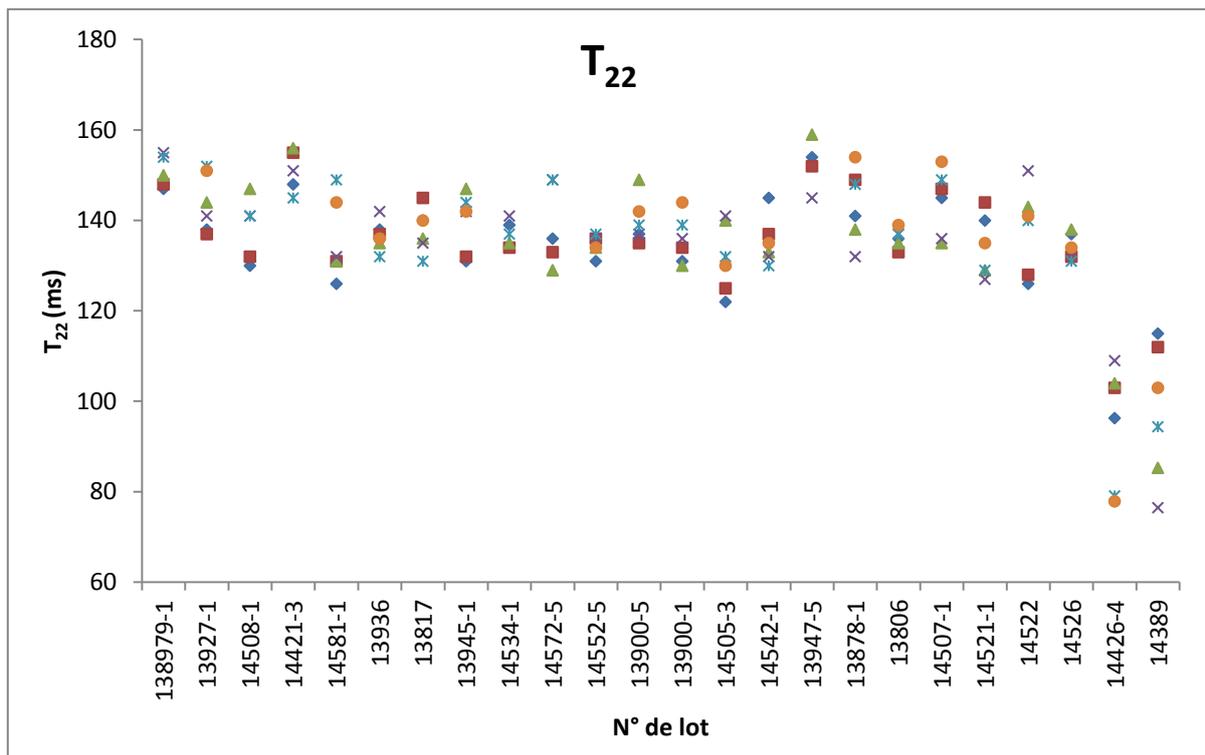


Figure 37 : Graphe regroupant les mesures des constantes de temps de relaxation transversale de l'eau libre (T_{22}) sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.

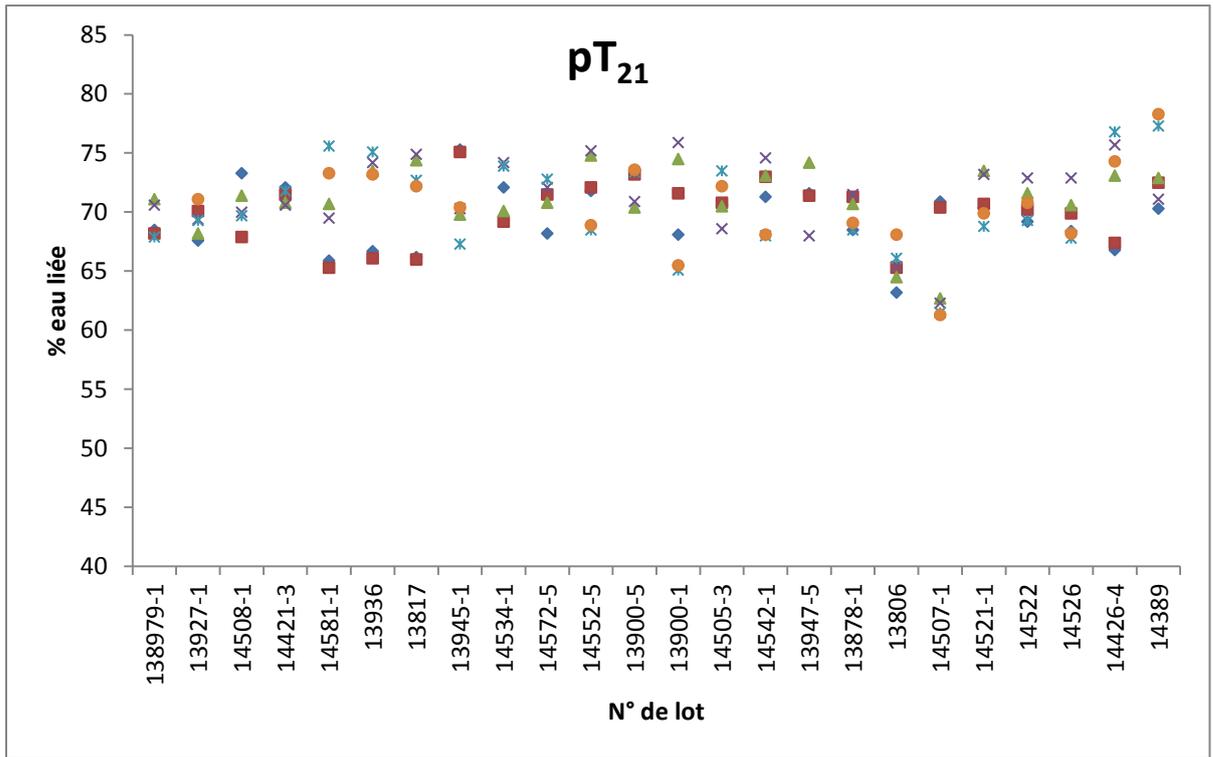


Figure 38 : Proportions en eau liée (pT₂₁) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.

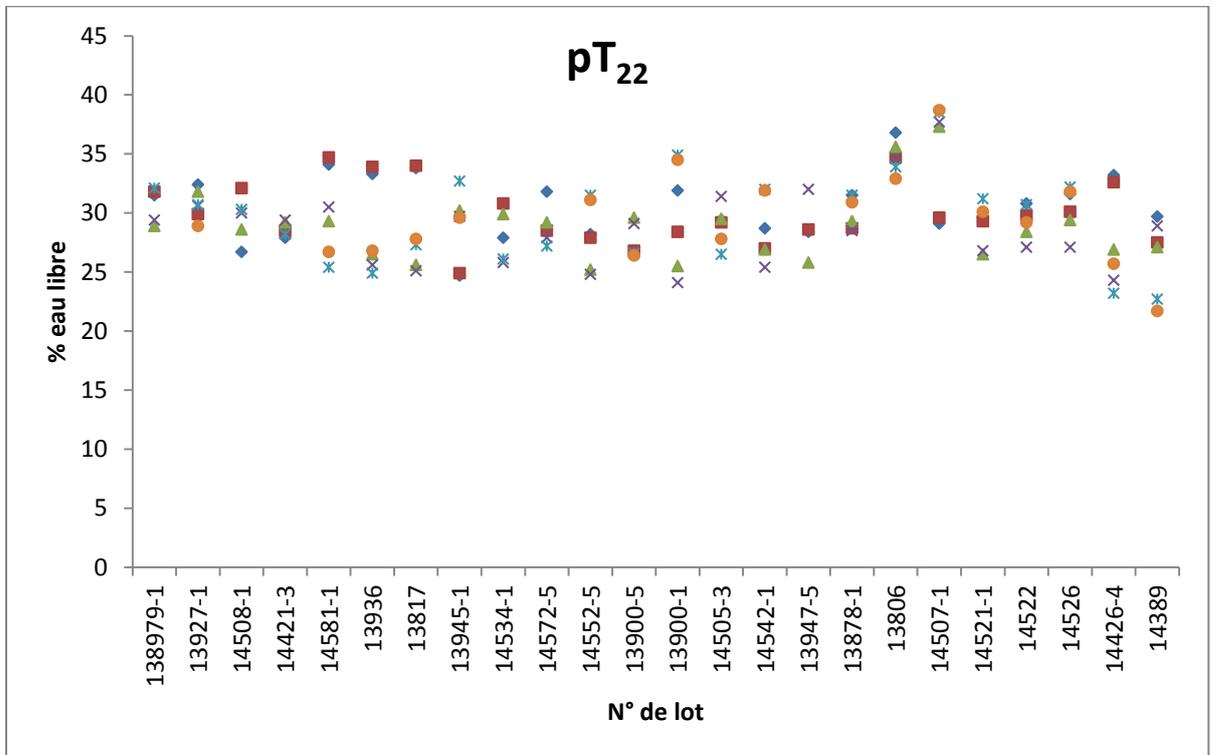


Figure 39 : Proportions en eau libre (pT₂₂) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.

Les variations de texture observées sur les différents échantillons peuvent probablement s'expliquer par certains paramètres affectant l'établissement de la *rigor mortis*. Immédiatement après la mort du poisson, les tissus musculaires sont totalement détendus et la texture est souple et élastique. Cette étape dure classiquement quelques heures puis le muscle se contracte et est alors en état de *rigor mortis*. Habituellement, cette situation dure environ une journée puis la rigor disparaît, provoquant la détente du muscle et un assouplissement de sa texture. Cependant, la chair du poisson ne retrouvera jamais son élasticité *pre-rigor mortis*. Le rapport entre l'apparition et la disparition de la *rigor* est espèce dépendant et est influencé par plusieurs facteurs tels que : la technique de pêche, la température de stockage, la manutention du poisson, la taille et la condition physique du poisson¹⁷⁷.

Tout d'abord, de manière générale, une température plus élevée conduit à une mise en place et une disparition plus rapide de la *rigor mortis*, sauf pour les poissons tropicaux pour lesquels l'effet est inversé. En fait, cela dépend de la différence entre la température de l'eau et celle du stockage : plus celle-ci est importante et plus le délai entre la mort et la *rigor mortis* est court¹⁷⁸.

Par ailleurs, la technique utilisée pour tuer le poisson (étourdissement, hypothermie, égorgement) agit également sur l'apparition de la *rigor mortis*. Par exemple, le fait de tuer le poisson par de l'eau glacée accélère considérablement son commencement par rapport à une mort par étourdissement¹⁷⁹.

Enfin, le moment de l'étape de filetage du poisson par rapport à la *rigor mortis* (avant ou pendant) a une importance capitale. Si les filets sont levés précédemment à la *rigor*, les tissus pourront se contracter et ceux-ci rétréciront à l'apparition de la *rigor*. Si cette étape est réalisée pendant la *rigor*, la chair sera complètement ferme et le rendement sera médiocre. Une étude menée par Skjervold et al¹⁸⁰ sur des saumons de 4 à 5 kg a révélé qu'un filetage *pre-rigor* par rapport à un *post-rigor* présentait le meilleur score de couleur, le moins de gaping (rétrécissement des filets), la texture la plus ferme et le moins de pertes liées à l'exsudation.

Au vu de tous ces paramètres, il semblerait que la différence de texture observée lors des tests organoleptiques menés par Findus sur les blocs de colin d'Alaska du fournisseur F, puisse être corrélée à un filetage plus rapide après la mort du poisson que celui des autres producteurs. En effet, les échantillons provenant du fournisseur F apparaissent comme plus

frais d'après notre méthode de détermination de la fraîcheur (valeur K et TMA-N). La valeur K ne variant que très légèrement une fois le poisson congelé, le temps passé entre la pêche et la congélation est donc plus court, en supposant que les conditions de stockage soient les mêmes pour tous les fournisseurs (dans de la glace pilée à 0°C). Il est également possible que ce distributeur utilise une technique de pêche différente modifiant la position de l'étape de traitement vis-à-vis de la *rigor mortis*.

I.3.2) Etude de la texture des bâtonnets de poisson (Croustibat)

Une étude similaire à celle portant sur la texture des blocs de farce de colin d'Alaska a été menée, cette fois-ci, sur des produits finis de la société Findus à savoir sur des bâtonnets de poisson commercialisés sous la marque Croustibat. Sur un lot particulier de ce type d'échantillon, la société Findus a reçu de nombreuses réclamations clients portant sur la texture des bâtonnets. Plusieurs échantillons de ce lot nous ont alors été transmis ainsi que des spécimens de lots conformes dans le but de les étudier par RMN avec la même méthodologie que celle utilisée sur les blocs de minerai de colin d'Alaska (même conditions expérimentales et même processus de traitement des données).

Les bâtonnets Croustibat sont conçus à partir des blocs étudiés au préalable, produit de base de la majorité des produits Findus. Le poisson est maintenu et travaillé congelé tout le long de la ligne de production. Le bloc est tout d'abord découpé en entrée de ligne par une machine selon le type de produit final désiré. Le poisson est alors plongé dans une solution d'eau et de farine afin de permettre à la chapelure de tenir et d'être uniforme. Le poisson qui est alors à -25°C, passe alors dans une nouvelle machine où la chapelure va être pulvérisée. Lors de ce processus la température est d'environ 200°C pendant une vingtaine de secondes, le poisson va ainsi remonter à environ -15°C pour le cœur du bâtonnet et -7°C pour les bords. Le poisson est alors aspergé par de l'huile pour la cuisson et refroidit de nouveau à -25°C. Les bâtonnets sont alors mis en boîte de manière automatisée puis stockés sur des palettes dans une chambre froide jusqu'à son expédition.

Au vu des différents traitements appliqués sur la ligne de production, nous avons décidé de réaliser les prélèvements pour les biopsies de poisson au cœur du bâtonnet, celui-ci étant moins affecté par les différents processus. Nous avons alors réalisé 4 mesures par lot, correspondant à deux prélèvements au centre de l'échantillon Croustibat sur deux bâtonnets différents. L'ensemble des résultats obtenus (T_{21} , T_{22} , pT_{21} et pT_{22}) sont regroupés sous forme de graphe comme précédemment (Figure 40-43). Les lettres A et B à côté du numéro de lot

correspondent aux deux bâtonnets tandis que les échantillons non conformes au niveau de la texture (sujet à des retours clients) sont annotés de l'abréviation NC à la suite du numéro de lot.

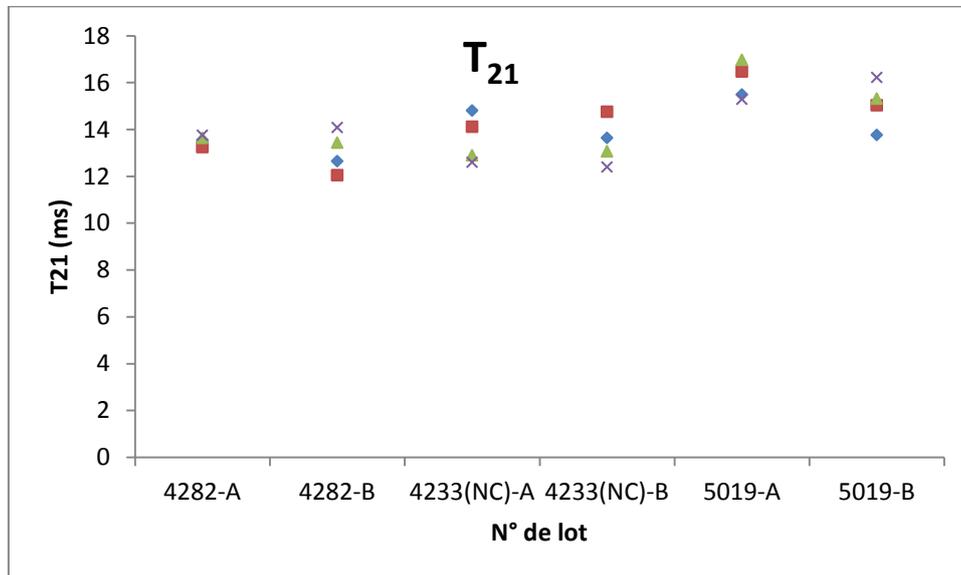


Figure 40 : Graphe regroupant les mesures des constantes de temps de relaxation transversale de l'eau liée (T_{21}) sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.

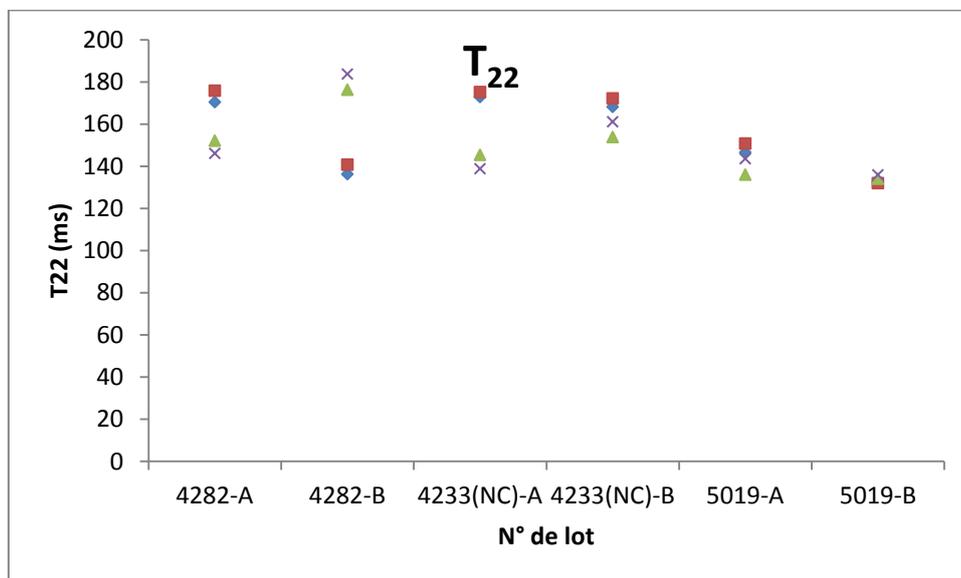


Figure 41 : Graphe regroupant les mesures des constantes de temps de relaxation transversale de l'eau libre (T_{22}) sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.

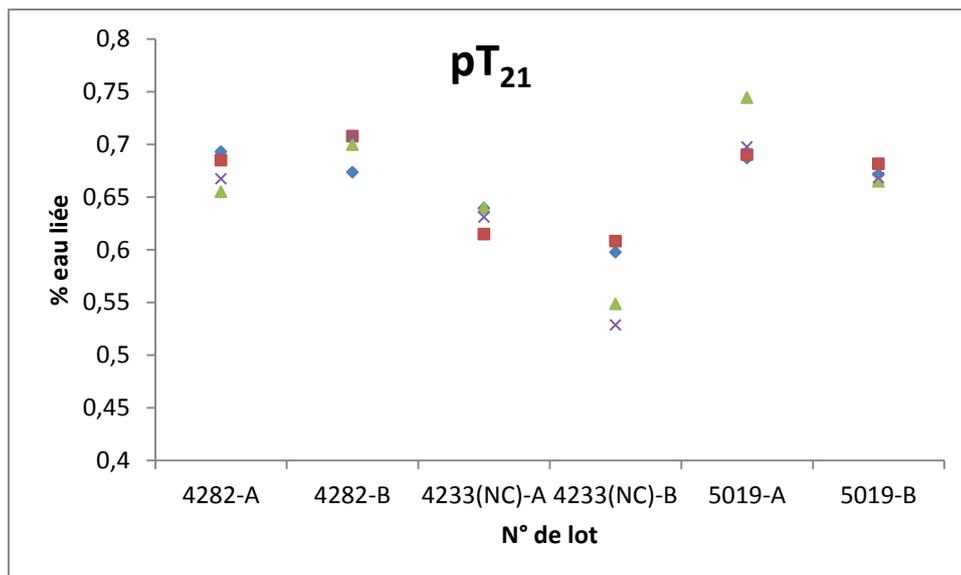


Figure 42 : Proportion en eau liée (pT₂₁) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.

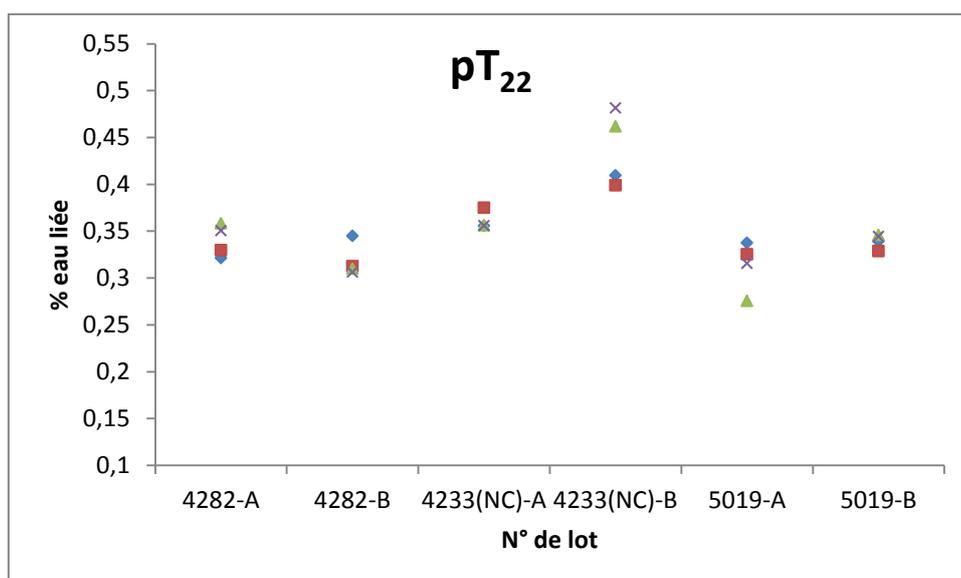


Figure 43 : Proportion en eau libre (pT₂₂) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.

Comme nous pouvons le voir sur ces résultats, les échantillons non conformes ne présentent pas de profils particuliers au niveau des valeurs de temps de relaxation transversale (T₂₁ et T₂₂) contrairement aux blocs de farce de colin d'Alaska du fournisseur F. Cependant, ces bâtonnets présentent des proportions en eau liée plus faibles surtout pour le bâtonnet B (63 %

pour A et 57% pour B) que celles déterminées sur les échantillons conformes (68,5% pour le lot 4282 et 69% pour le lot 5019).

Au vu de ces résultats et de ceux obtenus sur les blocs de farce de colin d'Alaska, il semblerait que les proportions en eau libre et eau liée influencent directement la texture de l'échantillon. En effet, les échantillons possédant les teneurs en eau liée les plus importantes correspondent aux spécimens présentant la meilleure texture au niveau des tests organoleptiques. De même, une teneur en eau liée plus faible semble attester d'une texture moins bonne, les bâtonnets Croustibat ayant donné lieu à des retours clients présentant des teneurs en eau liée nettement plus faibles que celles observées sur les autres échantillons.

I.3.3) Conclusion

A la lumière de l'ensemble des résultats de cette étude, il semblerait que nous soyons capables d'établir avec notre méthode un lien entre un paramètre RMN, le temps de relaxation transversale T_2 , et la texture de l'échantillon. En effet, les échantillons de farce de colin d'Alaska présentant des différences lors des tests organoleptiques, ou les bâtonnets Croustibat sujet à des réclamations client, présentent des profils particuliers vis-à-vis des autres échantillons analysés soit au niveau de la valeur des constantes de temps de relaxation transversale (T_{21} et T_{22}) soit au niveau des proportions en eau liée et eau libre (pT_{21} et pT_{22}).

L'avantage de notre technique vis-à-vis de la RMN Bas champ est notre capacité à différencier le temps de relaxation de l'eau et des lipides permettant ainsi une mesure précise des teneurs en eau liée et eau libre. Pour cela, il est nécessaire d'appliquer une vitesse de rotation faible pour éviter de dégrader la structure des tissus musculaire et d'utiliser une liste de délais appropriée.

Pour clore la présentation des résultats obtenus par RMN HR-MAS au cours de cette thèse, nous allons aborder une section dédiée à la quantification absolue des signaux présents sur le spectre RMN lors d'une acquisition CPMG 1D ^1H à l'aide de cette technique. En effet, la perte d'aimantation lors de l'utilisation de cette séquence RMN, liée à la relaxation transversale pendant l'écho de spin ainsi que celle associée à une relaxation longitudinale incomplète, est propre à chaque métabolite. La méthodologie mise au point permet de déterminer rapidement les pertes liées pour chaque signal et ainsi affiner leur quantification absolue.

I.4) Quantification T1/T2 de biopsies

I.4.1) Introduction

Comme nous l'avons vu dans la **Partie II**, la RMN est une technique couramment employée pour l'établissement du profil métabolomique. Le plus souvent, les données collectées par RMN sont soumises à des analyses statistiques multivariées prenant en compte la teneur de l'ensemble des composés présents sur le spectre RMN. Cependant, pour certaines études, notamment médicales, la détermination précise de la concentration absolue d'un ou plusieurs métabolites¹⁸¹ est nécessaire dans le but de comparer les résultats obtenus à ceux d'autres techniques analytiques ou d'autres laboratoires sur des échantillons similaires. De nombreuses études médicales sont réalisées à l'aide de la RMN et plus particulièrement à l'aide de la RMN HR-MAS ; cette méthode garantissant la préservation de l'échantillon en question ne requérant pas de protocole de préparation lourd ou d'extraction altérant celui-ci. Les analyses par HR-MAS sont le plus souvent acquises à l'aide de la séquence CPMG, celle-ci réduisant de manière significative la contribution des macromolécules et des lipides présents au sein des biopsies. Les spectres ainsi obtenus présentent alors une excellente ligne de base et une meilleure résolution, les signaux liés aux macromolécules étant nettement atténués. La qualité de leur atténuation dépend directement du temps de CPMG et peut varier en fonction des matrices utilisées.

Malheureusement, l'utilisation de cette séquence conduit à une perte d'aimantation liée, à la fois à la relaxation transversale durant le motif CPMG, mais également, dans une moindre mesure, à une relaxation longitudinale incomplète. Pour déterminer de manière précise la quantité d'aimantation perdue, il est nécessaire de connaître de façon rigoureuse les valeurs de T_1 et de T_2 de chaque métabolite. En effet, chaque molécule possède des propriétés de relaxation propre à elle-même et dépendante également de leur environnement (en solution ou au sein de matrices où la mobilité est plus ou moins réduite). La perte d'aimantation liée au T_2 est exprimée en fonction de l'équation suivante :

$$\% \text{ perte d'aimantation}(T_2) = \left[1 - e^{-\left(\frac{t_{cpmg}}{T_2_{composé}}\right)} \right] \times 100$$

La perte de signal liée à une relaxation longitudinale incomplète est, quant à elle définie par l'équation :

$$\% \text{ perte d'aimantation}(T1) = [1 - e^{-\left(\frac{TR+AQ}{T1_{composé}}\right)}] \times 100$$

De plus, malgré l'utilisation de la séquence CPMG, il est assez fréquent que des chevauchements de pics aient lieu entre les métabolites ou entre une petite molécule et des lipides résiduels. Dans ces cas-là, il est souvent nécessaire d'employer la déconvolution pour déterminer de manière précise la teneur de chaque composé du recouvrement. Le développement des logiciels de traitement des données RMN et des algorithmes de calcul de ces outils informatiques, rend désormais possible l'application de courbes d'ajustement bi-exponentielle (voire tri-exponentielle dans les cas les plus favorables) pour mesurer les constantes de temps de relaxation de deux molécules se chevauchant. Dans certaines circonstances, il peut arriver que plus de deux signaux se superposent. Il est alors nécessaire de trouver un des pics du multiplet caractérisant la molécule d'intérêt présentant le plus faible recouvrement avec les autres composés.

Dans ce contexte, nous avons, tout d'abord, essayé de développer une méthode par RMN HR-MAS déterminant les constantes de temps de relaxation longitudinale et transversale des composés présents sur un échantillon de poisson, dans le but d'améliorer la quantification des métabolites. Nous avons, par la suite, appliqué notre technique à des biopsies de cerveau.

I.4.2) Développement méthodologique : Quantification de l'arginine au sein d'une biopsie de poisson

Tout d'abord, nous avons élaboré une solution de référence d'arginine dans du D₂O de concentration 8,31 mmol/l. Cette valeur précise de la préparation d'arginine a été mesurée à l'aide la méthode PULCON/ERETIC^{56,58} et d'un spectre de référence, sur le spectromètre 600 MHz. Le rôle de cette solution de D₂O est double : fournir une fréquence de verrouillage champ/fréquence et servir de référence interne pour nos essais de quantification. Nous avons alors enregistré un spectre de référence pour la quantification PULCON/ERETIC sur un insert contenant 20 µl de notre solution de référence, soit 166,14 nmol, à l'aide d'une séquence CPMG extrêmement courte (2 boucles pour un temps total d'écho de 0,8436 ms), afin d'obtenir une ligne de base optimale tout en évitant toute perte d'aimantation liée à la relaxation transversale durant le temps de CPMG.

Une fois la teneur exacte en arginine déterminée, nous avons alors préparé un insert jetable contenant une biopsie de poisson (8,2mg) ainsi que 8 μ l de la solution de référence (66,46 nmol). Nous avons légèrement modifié les quantités de solvant et de poisson insérées au sein de l'insert, par rapport aux conditions classiques (10 μ l de solvant et 10 à 12 mg de poisson) afin de s'assurer que l'intégralité de la solution reste au sein de l'insert et ainsi ne pas biaiser les mesures quantitatives. Nous avons cependant pris soin de conserver le même rapport solvant/biopsie que celui employé dans les conditions standards. Trois spectres ont alors été enregistrés : le premier correspondant à une acquisition classique 1D CPMG avec les paramètres standards utilisés pour les études par RMN HR-MAS (188 boucles pour un temps total de CPMG de 75,6 ms), le second correspondant à une CPMG pseudo-2D dans le but de mesurer les valeurs des constantes de temps de relaxation transversale (T_2) et le dernier à une inversion-récupération afin de déterminer les constantes de temps de relaxation longitudinale (T_1) des différents métabolites.

Dans le but de tester notre méthode dans le cas de chevauchement de pics sur le spectre RMN, nous avons décidé de mener nos essais sur le triplet de l'arginine situé à 3,77 ppm, celui-ci présentant un chevauchement avec un autre composé (Figure 44). Pour cela, nous avons étudié le pic central du triplet situé à 3,776 ppm et le pic de droite situé à 3,763 ppm.

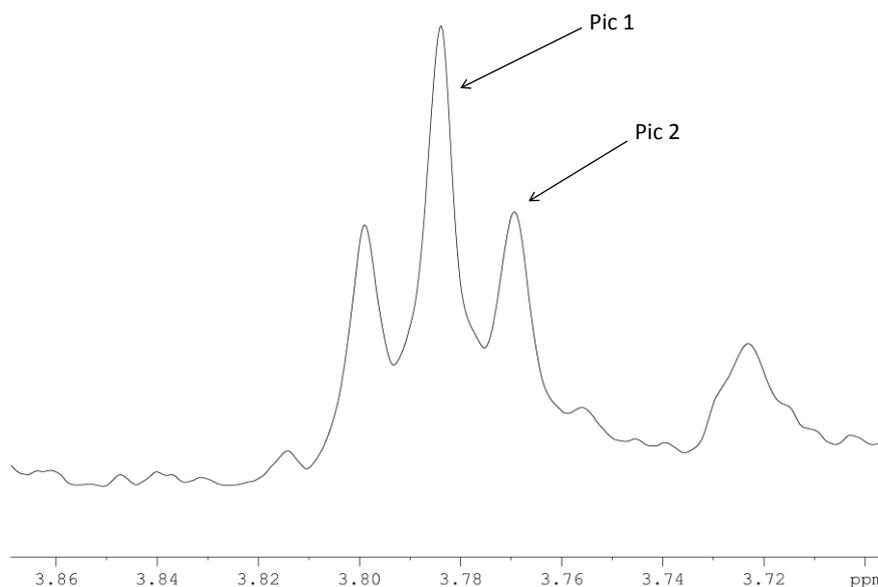


Figure 44 : Fenêtre spectrale d'un spectre ^1H 1D CPMG enregistré sur biopsie de poisson. Le triplet de l'arginine situé à 3,77 ppm présente un chevauchement entre deux composés.

Nous avons alors appliqué la méthode de quantification PULCON/ERETIC sur ce multiplet et mesuré une teneur absolue de 76,43 nmol. Cependant, cette quantité correspond aux quantités d'arginine et d'une autre molécule. Nous avons alors utilisé les valeurs des aires obtenues par déconvolution et nous avons ainsi déterminé les proportions (contribution) relatives de ces deux composés (76 % pour l'arginine et 24 % pour l'autre métabolite).

	Pic1 (3,776 ppm)	Pic 2 (3,763 ppm)
n PULCON/ERETIC (nmol)	76,43	76,43
proportion d'arginine	0,7629	0,765
n ₀ PULCON/ERETIC (nmol)	58,31	58,47
T ₂ (ms)	690	679
Perte T ₂ (%)	10,38	10,54
T ₁ (s)	1,27	1,26
Perte T ₁ (%)	0,97	0,97
n ₀ ajusté (nmol)	65,69	65,99
Valeur théorique (nmol)	66,40	

Tableau 10 : Résultats obtenus pour la quantification et la mesure des constantes de temps de relaxation (longitudinale et transversale) des deux signaux de l'Arginine sur une biopsie de poisson à l'aide de la méthode de quantification PULCON/ERETIC. La valeur en gras représente la quantité absolue d'arginine déterminée pour chaque signal. Celle-ci tient compte de la perte de signal liée aux processus de relaxation.

Nous avons alors traité les données des séquences permettant la détermination des constantes de temps de relaxation T₁ et T₂ (Inversion-Récupération et CPMG pseudo-2D). Les valeurs de T₁ et T₂ calculées pour les deux pics sont données dans le Tableau 10. A l'aide de ces constantes, les pertes d'aimantation liées aux différents processus de relaxation ont été estimées et appliquées à la teneur mesurée sur le spectre 1D dans le but de remonter à la quantité initiale (Tableau 10), et des valeurs de 65,69 et 65,99 nmol respectivement pour le pic 1 et pour le pic 2 ont été déterminées. Celles-ci sont très proches de la valeur théorique de 66,40 nmol correspondant aux 8 µl de solution de référence (C= 8,31 mmol.l⁻¹). En effet les erreurs entre les valeurs théoriques et expérimentales sont très faibles (1,07 % pour le pic 1 et 0,62 % pour le pic 2).

Les résultats obtenus lors de cette phase de développement étant prometteurs, nous avons alors décidé de mener des études complémentaires sur des biopsies de cerveau humain utilisées dans le cadre du projet EXTEMPO RMN.

I.4.3) Biopsies de cerveau

Le projet EXTEMPO RMN est un projet collaboratif entre Bruker BioSpin, l'université de Strasbourg, le CHU de Strasbourg-Hautepierre et Ariana Pharma, et est soutenu par Bpi France (anciennement Oséo). L'objectif de ce projet est de développer une méthode innovante, par RMN HR-MAS, de diagnostic de tissus cancéreux en temps réel, au cours de l'acte chirurgical.

L'intérêt de l'application de notre méthode à des biopsies était de pouvoir fournir à ce projet une technique de quantification absolue plus précise des différents métabolites analysés à l'aide de la séquence CPMG par RMN HR-MAS.

Ainsi, nous avons, tout d'abord, appliqué notre méthode à un premier type de tumeur de cerveau : l'oligodendrogliome. Dans ce but, le même protocole de préparation que celui employé pour la biopsie de poisson, à savoir environ 8mg de tumeur et 8 μ l de la solution de référence d'arginine. Comme dans le cas de l'analyse du tissu musculaire de poisson, le triplet de l'arginine présentait un chevauchement de pics, nous avons donc appliqué une déconvolution pour mesurer la proportion des deux composés. Les différents spectres (Inversion-Récupération, CPMG pseudo-2D et CPMG 1D) ont été enregistrés sur un spectromètre 500 MHz, à l'aide des mêmes séquences, cependant, certains paramètres ont dû être adaptés. Tout d'abord, la liste contenant le nombre de boucles pour la CPMG pseudo-2D a été ajustée ; le délai τ étant synchronisé sur la vitesse de rotation du rotor et celle-ci étant différente de celle employée précédemment (3500 Hz à l'hôpital d'Hautepierre pour la biopsie de poisson, 5000 Hz pour l'analyse du poisson). De plus, le temps de CPMG étant de 93,716 ms pour l'analyse des échantillons de cerveau (contre 75,6 ms précédemment), le calcul de la perte d'aimantation liée à la relaxation transversale a été adapté à ce temps de filtre T_2 . Les temps de relaxation longitudinale et transversale (T_1 et T_2) ont alors été déterminés pour chaque signal à l'aide d'ajustements bi-exponentiels certains pics présentant des chevauchements notamment à cause de macromolécules résiduelles. Dans le cas où un seul composé est présent, l'algorithme de calcul utilise une fonction d'ajustement mono-exponentielle.

Les différents résultats obtenus sur l'oligodendrogliome sont donnés dans le Tableau 11. Les résultats de quantification absolue de l'arginine déterminées pour les deux pics (67,88 et 67,69 nmol) sont conformes à la valeur théorique attendue (66,40 nmol). Les erreurs entre les

valeurs théoriques et expérimentales sont légèrement plus élevées que celles mesurées sur des biopsies de poisson (2,22 % pour le pic 1 et 1,94 % pour le pic 2).

	Pic1 (3,776 ppm)	Pic 2 (3,763 ppm)
n PULCON/ERETIC (nmol)	79,47	79,47
proportion d'arginine	0,734	0,765
n0 PULCON/ERETIC (nmol)	58,29	58,00
T2 (ms)	657	647
Perte T2 (%)	13,29	13,48
T1 (s)	1,27	1,26
Perte T1 (%)	0,97	0,97
n0 ajusté (nmol)	67,88	67,69
Valeur théorique (nmol)		66,40

Tableau 11 : Résultats obtenus pour la quantification et la mesure des constantes de temps de relaxation (longitudinale et transversale) des deux signaux de l'Arginine sur une biopsie de cerveau (oligodendrogliome) à l'aide de la méthode de quantification PULCON/ERETIC. La valeur en gras représente la quantité absolue d'arginine déterminée pour chaque signal. Celle-ci tient compte de la perte de signal liée aux processus de relaxation.

Une fois que nous nous étions assuré que notre méthodologie fonctionnait effectivement sur la nouvelle matrice, nous avons alors mesuré les valeurs de T_1 et de T_2 sur 17 métabolites présents sur les spectres des différents types de biopsies de cerveau, dans le but de déterminer la perte d'aimantation pour ces composés. Pour cela, nous avons étudié deux échantillons de cortex, deux d'oligodendrogliomes et deux glioblastomes. L'intérêt d'analyser différents types de tumeurs est de mesurer des variations éventuelles des valeurs de T_1 et de T_2 entre les différents échantillons. En effet, les valeurs des constantes de temps de relaxation sont directement liées à l'environnement des métabolites et des variations peuvent exister entre les différents types de biopsies.

L'ensemble des mesures de temps de relaxation longitudinale et transversale réalisées sur les différentes tumeurs de cerveau sont données respectivement en Figure 45 et 46.

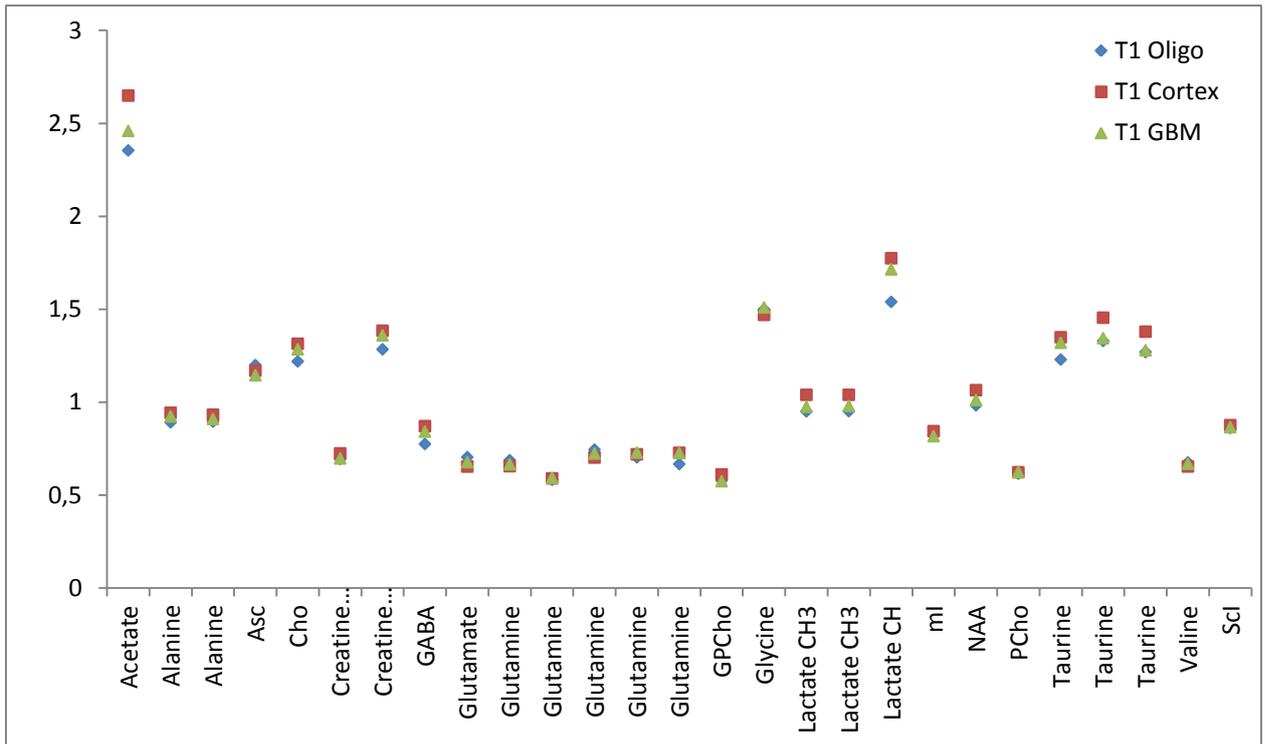


Figure 45 : Mesures des constantes de temps de relaxation longitudinale (T1) des principaux métabolites présents dans les différents types de biopsies de cerveau (♦) oligodendrogliome, (■) cortex et (▲) glioblastome.

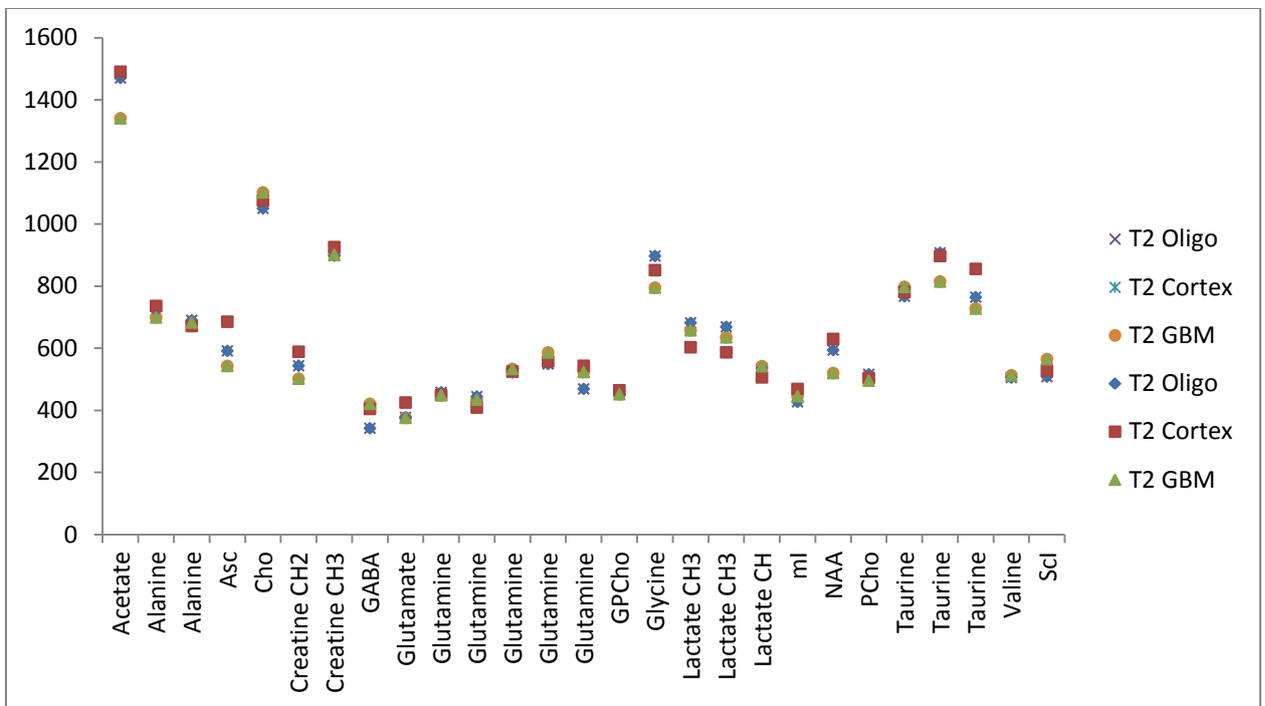


Figure 46 : Mesures des constantes de temps de relaxation transversale des principaux métabolites présents dans les trois types de biopsies de cerveau étudiés sur deux échantillons différents représentés par deux symboles distincts (X et ♦) oligodendrogliome, (■ et X) cortex et (▲ et ●) glioblastome. Les valeurs mesurées sur les mêmes types de biopsies sont extrêmement proches.

Pour les mesures de T_2 , deux acquisitions CPMG, correspondant à deux listes de nombres de boucles différentes, ont été effectuées pour chaque échantillon ; la première étant la liste de 8 points classiquement utilisée, tandis que la seconde est constituée de 16 points. Les résultats obtenus à l'aide des deux listes sont très proches. Cependant certaines valeurs sont légèrement plus élevées pour la seconde acquisition, ce qui semble attester d'une légère dégradation des tissus liée au temps passé par l'échantillon au sein du spectromètre (effet de la mise en rotation prolongée).

Comme nous pouvons le voir sur ces deux figures, les valeurs déterminées par RMN HR-MAS sur les différentes matrices sont homogènes et celles-ci peuvent alors être utilisées pour mesurer les pertes d'aimantation liées aux différents processus de relaxation au cours de l'expérience CPMG (Tableau 12 et 13). La perte totale d'aimantation, lors de la séquence CPMG employée pour l'étude des différents types de tumeur, est alors la combinaison de la perte liée à une relaxation longitudinale incomplète pendant le temps de récupération et de celle associée à la relaxation transversale pendant le motif CPMG de 93,716 ms (Tableau 14).

	$\delta^1\text{H}$ (ppm)	T_2 OLIGO _{moy} (ms)	T_2 CORTEX _{moy} (ms)	T_2 GBM _{moy} (ms)	OLIGO: perte d'aimantation (%)	CORTEX: perte d'aimantation (%)	GBM: perte d'aimantation (%)
Acetate	1,924	1470	1490	1340	6,18	6,10	6,75
Alanine	1,483	721	736	699	12,19	11,96	12,55
Alanine	1,468	689	671	683	12,71	13,04	12,83
Asc	4,523	591	685	543	14,66	12,79	15,87
Cho	3,21	1050	1075	1101	8,54	8,35	8,16
Creatine	3,935						
CH ₂		543	588	501	15,84	14,73	17,05
Creatine	3,03						
CH ₃		897	925	901	9,92	9,63	9,88
GABA	1,904	342	404	421	23,97	20,69	19,98
Glutamate	2,054	377	425	376	21,99	19,80	22,09
Glutamine	2,13	458	449	449	18,52	18,84	18,85
Glutamine	2,16	444	409	434	19,03	20,50	19,41
Glutamine	2,424	522	525	533	16,44	16,36	16,13
Glutamine	2,44	549	557	586	15,69	15,49	14,78
Glutamine	2,47	469	543	523	18,12	15,85	16,41
GPCho	3,236	459	464	452	18,47	18,28	18,74
Glycine	3,558	897	851	795	9,93	10,43	11,12
Lactate CH ₃	1,337	681	603	657	12,85	14,40	13,29
Lactate CH ₃	1,323	668	587	635	13,08	14,77	13,73
Lactate CH	4,127	535	506	542	16,07	16,91	15,89
ml	4,055	427	469	445	19,71	18,13	18,99
NAA	2,018	594	630	520	14,60	13,83	16,50
PCho	3,228	516	504	495	16,62	16,97	17,24
Taurine	3,432	767	781	797	11,50	11,31	11,09
Taurine	3,418	907	897	814	9,82	9,93	10,87
Taurine	3,405	764	855	727	11,54	10,38	12,09
Valine	1,036	505		513	16,94		16,71
Scl	3,344	508	528	565	16,85	16,26	15,30

Tableau 12 : Valeurs de T_2 mesurées sur les différents dans les biopsies de cerveau ainsi que la perte de signal associée à la relaxation transversale. OLIGO : Oligodendrogliome GBM : Glioblastome

	$\delta^1\text{H}$ (ppm)	T_1 OLIGO _{moy} (s)	T_1 CORTEX _{moy} (s)	T_1 GBM _{moy} (s)	OLIGO: perte d'aimantation (%)	CORTEX: perte d'aimantation (%)	GBM: perte d'aimantation (%)
Acetate	1,924	2,355	2,650	2,460	15,22	18,77	16,49
Alanine	1,483	0,892	0,944	0,924	0,69	0,91	0,82
Alanine	1,468	0,896	0,933	0,910	0,71	0,86	0,77
Asc	4,523	1,200	1,170	1,145	2,49	2,26	2,08
Cho	3,21	1,220	1,315	1,285	2,64	3,43	3,17
Creatine	3,935						
CH ₂		0,695	0,725	0,698	0,17	0,22	0,17
Creatine	3,03						
CH ₃		1,285	1,385	1,360	3,17	4,07	3,84
GABA	1,904	0,776	0,872	0,842	0,33	0,62	0,52
Glutamate	2,054	0,705	0,654	0,678	0,18	0,11	0,14
Glutamine	2,13	0,688	0,656	0,666	0,16	0,12	0,13
Glutamine	2,16	0,583	0,591	0,594	0,05	0,05	0,06
Glutamine	2,424	0,745	0,703	0,724	0,26	0,18	0,22
Glutamine	2,44	0,704	0,720	0,729	0,18	0,21	0,23
Glutamine	2,47	0,668	0,728	0,727	0,13	0,23	0,22
GPCho	3,236	0,580	0,612	0,575	0,05	0,07	0,04
Glycine	3,558	1,500	1,470	1,510	5,20	4,90	5,31
Lactate CH ₃	1,337	0,951	1,040	0,976	0,94	1,41	1,06
Lactate CH ₃	1,323	0,952	1,040	0,981	0,95	1,41	1,09
Lactate CH	4,127	1,540	1,775	1,715	5,62	8,23	7,54
ml	4,055	0,824	0,845	0,818	0,46	0,52	0,44
NAA	2,018	0,984	1,065	1,012	1,10	1,56	1,25
PCho	3,228	0,616	0,624	0,625	0,07	0,08	0,08
Taurine	3,432	1,230	1,350	1,320	2,72	3,75	3,48
Taurine	3,418	1,330	1,455	1,345	3,57	4,75	3,70
Taurine	3,405	1,270	1,380	1,280	3,05	4,02	3,13
Valine	1,036	0,677	0,655	0,670	0,14	0,11	0,13
Scl	3,344	0,862	0,877	0,866	0,58	0,64	0,60

Tableau 13 : Valeurs de T_1 mesurées sur les différents dans les biopsies de cerveau ainsi que la perte de signal associée à la relaxation transversale. OLIGO : Oligodendrogliome GBM : Glioblastome

	$\delta^1\text{H}$ (ppm)	Perte totale OLIGO (%)	Perte totale CORTEX (%)	Perte totale GBM (%)
Acide acétique	1,924	20,45	23,72	22,13
Alanine	1,483	12,80	12,76	13,27
Alanine	1,468	13,33	13,79	13,50
Asc	4,523	16,78	14,76	17,62
Cho	3,21	10,95	11,50	11,07
Créatine CH ₂	3,935	15,98	14,92	17,20
Créatine CH ₃	3,03	12,78	13,31	13,34
GABA	1,904	24,22	21,18	20,39
Acide glutamique	2,054	22,14	19,89	22,20
Glutamine	2,13	18,65	18,93	18,95
Glutamine	2,16	19,07	20,54	19,46
Glutamine	2,424	16,66	16,51	16,31
Glutamine	2,44	15,85	15,66	14,97
Glutamine	2,47	18,23	16,04	16,59
GPCho	3,236	18,51	18,34	18,78
Glycine	3,558	14,61	14,82	15,84
Acide lactique CH ₃	1,337	13,67	15,60	14,21
Acide lactique CH ₃	1,323	13,91	15,97	14,67
Acide lactique CH	4,127	20,78	23,75	22,23
ml	4,055	20,08	18,56	19,35
NAA	2,018	15,54	15,17	17,54
PCho	3,228	16,68	17,04	17,31
Taurine	3,432	13,91	14,64	14,19
Taurine	3,418	13,03	14,20	14,17
Taurine	3,405	14,24	13,99	14,84
Valine	1,036	17,06		16,82
Sci	3,344	17,33	16,80	15,80

Tableau 14 : Tableau récapitulatif de la perte totale d'aimantation au cours de la séquence CPMG utilisée pour l'étude de biopsies de cerveau par RMN HR-MAS au sein de l'hôpital d'Hautepierre pour chaque métabolite sur les différents types de tumeurs. OLIGO : Oligodendrogliome GBM : Glioblastome

En observant les résultats du Tableau 14, nous nous apercevons que la perte de signal pour les différents métabolites varie de 11 à 24 %. Le principal phénomène, à l'origine de la perte d'aimantation pour l'ensemble des composés, est la relaxation transversale, sauf pour l'acide acétique où l'atténuation est principalement liée à une relaxation longitudinale incomplète.

Les signaux les plus atténués correspondent au CH de l'acide lactique, au myo-inositol, à l'acide glutamique, à l'acide acétique et l'acide γ -aminobutyrique. En revanche, les signaux de la choline, de l'alanine, de la taurine font partie de ceux qui sont le moins modérées par la séquence RMN acquise dans ces conditions. Il est important de noter que les signaux des CH₃ de l'acide lactique de la créatine/phosphocréatine sont moins affectés que ceux correspondant à d'autres groupements de ces molécules (CH₂ ou CH). Les valeurs ainsi déterminées pour chaque métabolite peuvent alors être appliquées aux valeurs de quantification absolue obtenue par la méthode PULCON/ERETIC sur le spectromètre de l'hôpital d'Haute-pierre, à l'aide d'un spectre de référence correspondant à une acquisition sans ou avec un CPMG court (quelques boucles), dans le but de mesurer de manière plus précise la concentration absolue de chaque composé. En effet, les pertes calculées tiennent compte à la fois des propriétés de chaque groupement moléculaire mais également de l'environnement de celles-ci (matrice étudiée).

I.4.3) Conclusion

La méthodologie mise au point sur un échantillon de poisson, puis appliquée à l'étude de biopsies de cerveau, permet la mesure rapide de la perte d'aimantation associée aux différents processus de relaxation ayant lieu lors d'une acquisition CPMG 1D ¹H pour chaque métabolite. Les pertes de signal sont déterminées en calculant de manière simultanée les constantes de relaxation de l'ensemble des composés présents sur le spectre RMN à l'aide des séquences appropriées (Inversion-Récupération pour les calculs de T₁ et CPMG pseudo 2D pour les mesures de T₂). De plus, lorsque certains pics présentent des chevauchements (maximum 2 molécules impliquées), l'intensité relative calculée à l'aide du rapport des aires déterminés par déconvolution permet d'évaluer les proportions des métabolites à l'origine du recouvrement.

Cependant, les constantes de temps déterminées et donc les pertes associées, dépendent de nombreux paramètres (champ du spectromètre, sonde utilisée, paramètres d'acquisition, matrice d'intérêt) et doivent donc être mesurées à l'aide de la méthodologie développée à chaque nouvelle étude.

I.5) Conclusion générale des travaux réalisés par HR-MAS

Les différentes études menées au cours de cette thèse dans le cadre du projet Agrifood GPS par RMN HR-MAS démontrent, tout d'abord, que cette technique est appropriée pour l'étude du poisson. En effet, les spectres obtenus par RMN HR-MAS sur ce type de matrice présentent une excellente résolution proche de celle obtenue en RMN liquide. Ceux-ci sont principalement composés d'acides aminés, de sucres, de certains acides organiques ainsi que de certains métabolites comme les produits de dégradation de l'ATP, l'OTMA-N et le TMA-N. Ces composés ont permis de développer une méthode relativement rapide (40 minutes) de détermination de la fraîcheur, mesurant deux indicateurs chimiques en une seule analyse.

En outre, l'application d'analyses statistiques multivariées aux profils métaboliques enregistrés par RMN HR-MAS sur les blocs de farce de colin d'Alaska, a permis de mettre en avant certaines variations entre les différents fournisseurs, notamment pour le prestataire A caractérisé par des teneurs en acide lactique plus faibles.

De plus, la RMN HR-MAS est également une technique de choix pour l'étude de la texture du poisson comme nous l'avons vu à travers les travaux menés sur les échantillons de la société Findus, les spécimens présentant des profils particuliers lors des analyses organoleptiques, apparaissant également différents sur la méthode développée par RMN HR-MAS à basse vitesse.

Une nouvelle étude sur la texture devrait être menée en partenariat avec deux fournisseurs de la société Findus dans le but d'établir si un lien peut être établi entre le temps passé par le poisson entre sa pêche et sa transformation, et la valeur K. L'influence de cette durée sur les mesures de temps de relaxation transversale de l'eau sera également étudiée en parallèle.

Par ailleurs, la méthodologie mise au point pour la détermination de la concentration absolue des métabolites au sein de biopsies lors d'une acquisition CPMG 1D ^1H constitue un nouvel outil rapide affinant les mesures de quantification réalisées par RMN HR-MAS, la séquence CPMG étant l'une des plus utilisées en à l'aide de cette technique.

Maintenant que les différents résultats obtenus par RMN HR-MAS sur des matrices semi-solides ont été présentés, nous allons, à présent, aborder ceux ayant été réalisées par RMN liquide haute résolution. Cette dernière partie est composée de deux études : la première, portant sur la détermination de l'origine géographique du caviar, tandis que la seconde est dédiée à l'investigation la dégradation du caviar au cours de sa conservation au réfrigérateur.

II) Etude de matrices liquides par RMN

II.1) Introduction

Le caviar est un mets gastronomique et consiste en des œufs d'esturgeon non fécondés traités avec du sel ou du borax (tétraborate de sodium). Il fait partie des produits de la mer les plus onéreux du fait de sa rareté, de son mode d'élevage, mais également, de la difficulté rencontrée lors des différentes étapes de préparation. Tout d'abord, les œufs sont extraits de la roque de l'esturgeon femelle celle-ci représentant entre 10 et 20 % de la masse du poisson. Les œufs ainsi récupérés sont alors tamisés à plusieurs reprises en différentes tailles et triés en fonction de leur couleur, leur fermeté et leur grosseur. Ils sont, ensuite, soigneusement nettoyés et rincés avant de passer à l'étape de salage. Le but du salage est, principalement, de préserver la qualité du caviar. La quantité ainsi que la méthode utilisée pour cette étape est propre à chaque producteur où savoir-faire. Les œufs d'esturgeon sont alors placés dans des grosses boîtes de plusieurs kilogrammes et placées au réfrigérateur (conservés entre -4 et -2 °C).

Jusqu'à la fin des années 1990, la production mondiale de caviar était principalement composée de caviar provenant d'esturgeons sauvages de la mer Caspienne (Russie et Iran) tels que le Beluga (*Huso Huso*), l'Osciètre (*Acipenser Guldenstaedtii*) et le Sevruga (*Acipenser Stellatus*). Cependant, les stocks de ces trois espèces, les plus commercialisés jusqu'alors, ont décliné de manière drastique atteignant un niveau inquiétant, poussant la Convention sur le Commerce International des Espèces de faune et de flore sauvages menacées d'extinction (CITES), a déclaré ces poissons comme espèces en danger de disparition. La considérable diminution des stocks d'esturgeons sauvages s'explique par les dommages environnementaux causés par plusieurs paramètres telles que la construction de barrages sur certaines rivières et la pollution des eaux et des sédiments, ainsi qu'à une pêche à outrance liée à la demande toujours plus importante de caviar¹⁸².

En raison de la pénurie d'esturgeons sauvages, l'élevage de ceux-ci, commencé dans les années 1980 en Europe de l'Ouest¹⁸³, en Italie et en France, s'est accrue au fur et à mesure des années et représente, à présent, une des principales sources de production. Les premières espèces d'esturgeon utilisées au sein de ces fermes d'élevage étaient l'esturgeon Adriatique pour l'Italie (*Acipenser nacarii*), et l'esturgeon européen d'Atlantique (*Acipenser sturio*) pour la France. Cependant, en France, afin d'éviter d'éventuels dommages sur la population des esturgeons natifs, les études ont rapidement été menées sur un autre modèle biologique

d'esturgeon, l'esturgeon Sibérien (*Acipenser baerii*). En effet, ce poisson originaire du lac Baïkal est extrêmement bien adapté à l'élevage celui-ci vivant en eau douce et s'acclimatant parfaitement aux conditions retrouvées en France. A présent, les principales espèces retrouvées dans les fermes d'élevage sont l'esturgeon blanc (*Acipenser transmontanus*) en Italie et aux Etats-Unis, l'esturgeon japonais (*Acipenser shrenckii*) en Chine, l'esturgeon sibérien (*Acipenser baerii*) et l'esturgeon russe (*Acipenser guldenstaedtii*) en France, en Italie, en Allemagne et en Chine et l'esturgeon adriatique (*Acipenser naccarii*) en Espagne et en Italie¹⁸³.

Les caviars d'élevage italiens et français sont ainsi reconnus pour la haute qualité de leur production. En France, quatre producteurs du bordelais ont décidé, conjointement, d'obtenir une Indication Géographique Protégée (IGP) dans le but de protéger leur savoir-faire dans la production de caviar. En effet, le caviar est un produit à haute valeur ajoutée et son prix est directement relié à sa qualité et sa rareté. La haute valeur ajoutée de ce produit ainsi que la mondialisation des échanges commerciaux provoquent de plus en plus l'émergence de fraudes sur ce produit et tout particulièrement sur l'origine de celui-ci.

Dans ce contexte, nous avons pris contact avec le principal producteur français de caviar, à savoir la société Sturgeon ainsi que l'ensemble des producteurs du bordelais, dans le but d'essayer de développer une méthode physico-chimique permettant la différenciation entre le caviar issu de leurs productions et celui provenant d'autres pays. En effet, à l'heure actuelle des méthodes par analyse PCR (Réaction en chaîne par polymérase) de l'ADN existent pour distinguer certaines espèces d'esturgeon¹⁸⁴⁻¹⁸⁶. Cependant, celles-ci ne permettent pas la détermination de l'origine géographique de celui-ci: le génotype de l'esturgeon étant extrêmement complexe. D'autres techniques analytiques sont capables de discriminer le caviar d'élevage du caviar sauvage à l'aide du profil d'acides gras du caviar déterminé par CPG¹⁸⁷. Par ailleurs, des études récentes ont évalué et comparé la valeur nutritionnelle d'œufs d'esturgeon d'élevage en Chine¹⁸⁸ et en Corée du Sud¹⁸⁹.

II.2) Etude de l'origine du caviar par RMN (Publication 2- Metabolic characterization of caviar specimens by ¹H NMR Spectroscopy: Towards caviar authenticity and integrity- Annexe 3)

II.2.1) Etude préliminaire par RMN HR-MAS

Tout d'abord, une étude préliminaire a été réalisée par RMN HR-MAS. Pour cela, un premier protocole de préparation consistant en l'insertion de trois à quatre œufs d'esturgeon au sein d'un insert jetable auquel 10 µl de d'une solution de D2O contenant 0,75 % de TSP sont ajoutés, a été utilisé. Cependant, plusieurs problèmes sont apparus lors de l'acquisition des spectres CPMG 1D ¹H. Tout d'abord, le caviar est un produit relativement gras, contenant entre 15 et 20 % de lipides. Comme nous l'avons vu précédemment lors de la **Partie I**, les signaux des lipides, qui sont relativement larges comparés à ceux des métabolites, masquent une grande partie de l'information contenue dans les spectres RMN. Pour remédier à ce problème, nous avons essayé d'appliquer une multi présaturation au lieu d'une simple présaturation, en saturant cette fois-ci le signal de l'eau mais également celui des lipides. Pour cela nous avons utilisé une impulsion sélective modulée en phase avant la séquence CPMG, permettant de saturer simultanément plusieurs fréquences (12 au total donc celle de l'eau). Un exemple de spectres obtenus sur un même échantillon à l'aide d'une acquisition proton classique et d'une séquence CPMG avec une multiprésaturation (*cpmgps1d*) est donné en Figure 47.

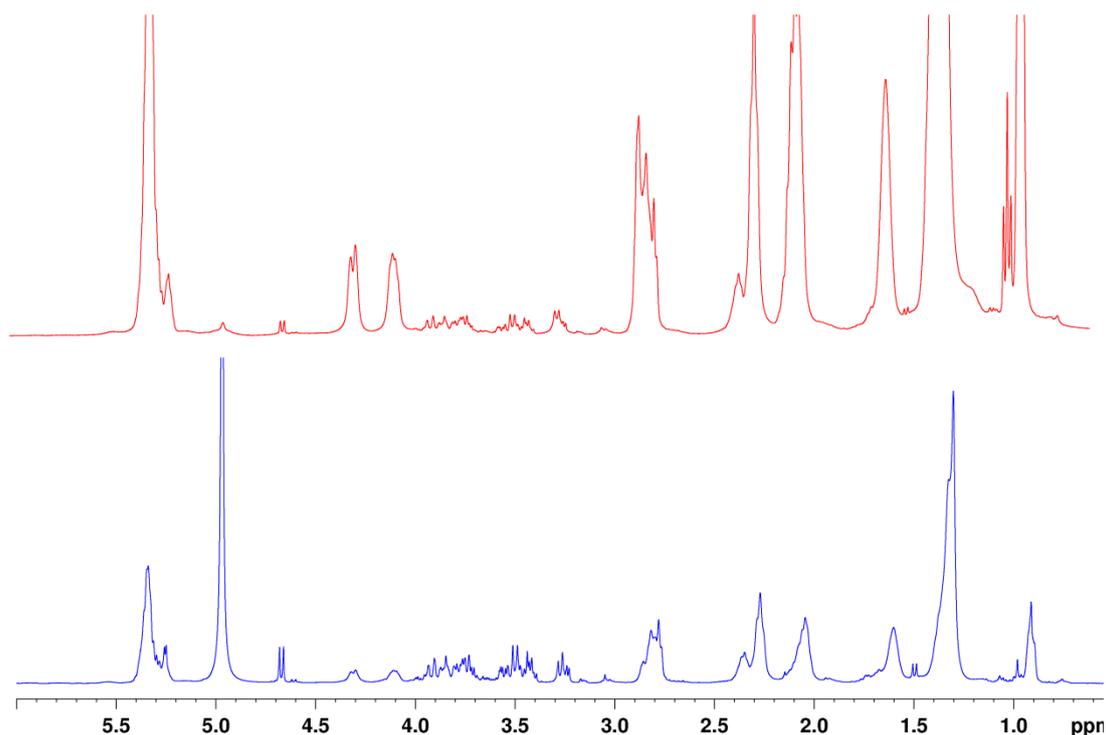


Figure 47 : Comparaison de spectres ^1H HR-MAS de caviar (L'Esturgeonnière) obtenus à l'aide (a) d'une acquisition proton classique (spectre rouge) et (b) d'une séquence CPMG avec une multiprésaturation (spectre bleu)

Comme nous pouvons le voir sur cette figure, les signaux des lipides sont nettement atténués par l'application de la multi présaturation. De plus, les signaux de certains métabolites qui étaient recouverts par les signaux lipidiques sont désormais observables comme par exemple, le doublet de l'alanine situé à 1,49 ppm.

Un autre problème rencontré lors de cette étude, a été la répétabilité entre les analyses distinctes conduites. En effet, des variations notables ont été observées entre les différents œufs d'une même boîte. Ainsi, nous avons décidé de modifier légèrement notre protocole de préparation des échantillons en prélevant cette fois-ci quelques dizaines d'œufs que nous avons placés dans un tube eppendorff pour les centrifuger. Nous avons récupéré alors entre 10 et 15 mg de cette solution que nous avons mise dans un insert avec les 10 μl de solvant deutéré (D_2O). Cette procédure nous a alors permis de réduire sensiblement la variabilité entre nos expériences.

Nous avons alors décidé de mener une première étude de faisabilité sur plusieurs boîtes commandées directement auprès de certains producteurs et provenant de différentes origines (françaises et étrangères). Pour chaque fournisseur, nous avons commandé deux boîtes correspondant à deux lots différents et réalisé trois analyses par RMN pour chaque boîte de caviar soit 6 mesures par producteur. Un tableau regroupant les différents échantillons analysés est donné ci-dessous (Tableau 15).

Producteur	Marque	Espèce	Origine
Salmo	Adamas Black	<i>Baerii</i>	Italie
	Adamas Green	<i>Transmontanus</i>	
Dom Petroff	Esturgeon Blanc	<i>Transmontanus</i>	Etats-Unis
Cassadotte	Ebène	<i>Baerii</i>	France
L'esturgeonnière	Perlita	<i>Baerii</i>	France
Strugeon	Sturia	<i>Baerii</i>	France
	Siberia	<i>Baerii</i>	

Tableau 15 : Plan d'échantillonnage mis en place dans le cadre de l'étude de faisabilité sur la détermination de l'origine du caviar par RMN HR-MAS.

Cependant, la présence, bien que faible de lipides résiduels ainsi que l'application de la multi présaturation nous oblige à exclure de nos études statistiques une large zone spectrale et donc, conduit à une perte significative d'informations. Nous avons alors pris la décision de développer une méthode d'extraction totale des lipides, basée sur la méthode Bligh & Dyer et présentée au sein de la **Partie III** et donc d'étudier le caviar par RMN liquide et non plus par RMN HR-MAS, la matrice obtenue à l'issue de l'extraction étant parfaitement adaptée à une étude par RMN liquide. En effet, le contenu métabolique des spectres obtenus par RMN liquide haute résolution (Figure 48) est beaucoup plus riche et le fait de travailler sur des extraits liquides (spectre rouge) permet un gain en résolution, et en sensibilité réduisant ainsi le temps d'acquisition.

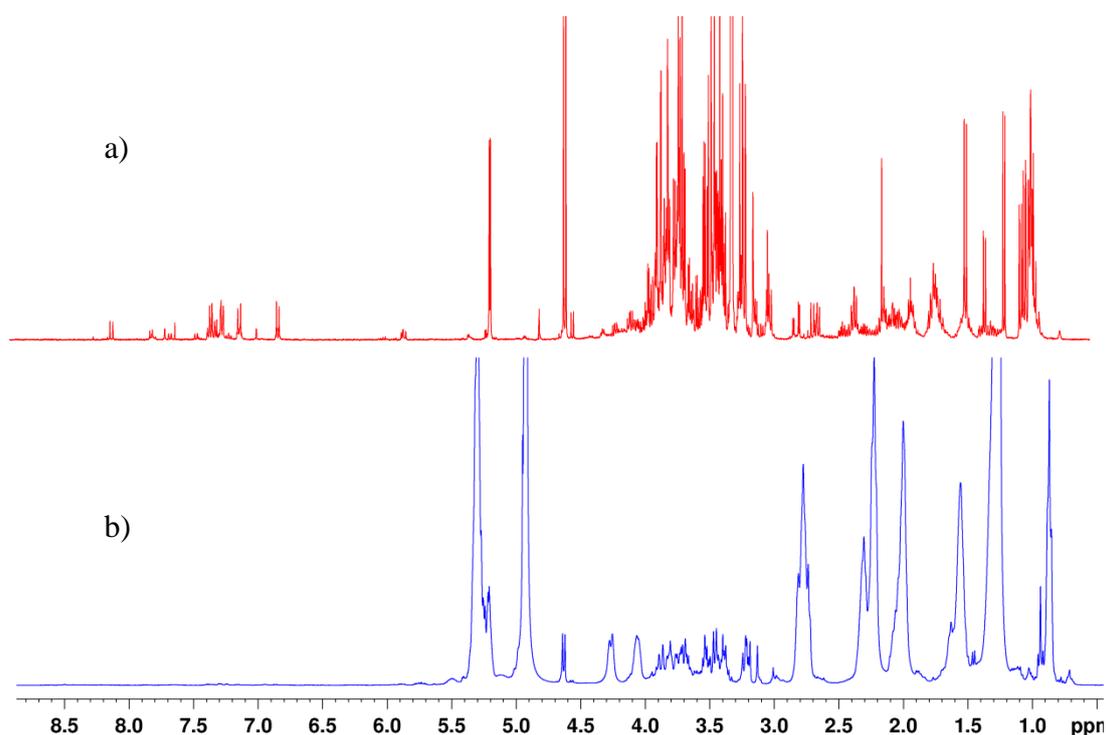


Figure 48 : Comparaison des spectres ^1H enregistrés à 400 MHz (a) en RMN liquide sur la phase aqueuse du caviar à l'aide de la séquence *noesygppr1d* et (b) en RMN HR-MAS sur des œufs de caviar à l'aide de la séquence *cpmaggpps* (multipresaturation) pour des durées d'acquisition équivalentes. Les différences en fréquence pour certains pics sont liées aux protocoles de préparation, celui utilisé pour la préparation d'échantillon en RMN HR-MAS n'utilisant pas de solution de tampon.

L'ensemble des analyses RMN effectuées sur le caviar que ce soit pour l'étude sur la détermination de l'origine du caviar ou celle portant sur la dégradation du caviar au cours de sa conservation, a été, par la suite, menée sur la phase aqueuse par RMN liquide.

II.2.2) Mise en place du plan d'échantillonnage

Dès lors que les conditions expérimentales étaient fixées, nous avons alors mis en place un plan d'échantillonnage avec la société Sturgeon. La société Sturgeon diffère des autres producteurs du bordelais de par sa structure. En effet, la plupart des producteurs français possèdent des petites structures composées d'un site d'élevage et d'un site de rinçage et n'élève qu'une seule espèce d'esturgeons. Le site d'élevage est le lieu où l'esturgeon grandit pendant 6 à 7 ans tandis que le site de rinçage, également appelé site de finition, est le l'endroit final où les esturgeons passent quelques semaines avant la collection des œufs. Les bassins des sites de rinçage sont composés d'eau très pure permettant l'élimination de

certaines molécules responsables du « goût de terre » que peut prendre le caviar, telles que la géosmine ou l'iso-bornéol.

La société Sturgeon dispose de 5 sites d'élevages différents et de 3 sites de rinçage. De plus, deux espèces différentes d'esturgeon sont élevées au sein de ce groupe, à savoir l'*Acipenser Baerii* et l'*Acipenser Guldenstaedtii*. Enfin, il arrive que certains poissons soient déplacés pour diverses raisons d'un site d'élevage à un autre au cours de son existence. Les différentes combinaisons classiques utilisées par la société Sturgeon sont données dans le Tableau 16. Il est important de signaler que le site d'élevage de Lacour est un site particulier puisque celui-ci consiste en un partenariat avec un petit producteur et ne représente qu'une très faible quantité de la production de la société Sturgeon.

Elevage	Elevage	Rinçage	Espèce
St Fort		Le Douhet	<i>Baerii</i>
Balizac		Le Douhet	<i>Baerii</i>
Villandraut		Le Douhet	<i>Baerii</i>
Lacour		Le Douhet	<i>Baerii</i>
St Fort		St Fort	<i>Baerii</i>
Ciré		Surgères	<i>Baerii</i>
Ciré	Souméras	Surgères	<i>Baerii</i>
Ciré		Surgères	<i>Guldenstaedtii</i>

Tableau 16 : Tableau récapitulatif des différentes combinaisons possibles de site d'élevage et de site de rinçage couramment employés au sein de la société Sturgeon au début de l'étude sur la détermination de l'origine géographique du caviar.

Dans un premier temps, après discussions avec le directeur d'exploitation de la société Sturgeon, nous avons choisi d'étudier, en premier lieu, certains paramètres les plus susceptibles d'influencer le profil métabolique déterminé par RMN. La détermination des différents paramètres pouvant introduire de la variabilité sur l'empreinte métabolique du caviar par RMN, est extrêmement importante pour mesurer la variabilité interne de la production de cette société.

Ainsi les paramètres suivants ont été pris en compte dans la mise en place du plan d'échantillonnage :

- le site d'élevage
- le site de rinçage
- l'espèce étudiée (*Guldenstaedtii* et *Baerii*)
- la période de production

Afin d'étudier l'influence de chaque paramètre, il a été capital de maintenir les autres à l'identique autant que possible. Pour chaque cas, nous avons étudié deux boîtes différentes correspondant à deux lots distincts dans le but de mesurer la variabilité pouvant exister entre différents lots mais également la répétabilité de nos analyses sur des boîtes relativement identiques. Pour cette entreprise, un numéro de lot coïncide à une journée de production et non pas à un poisson particulier. Ainsi, il peut arriver qu'un numéro de lot corresponde à plusieurs poissons du même bassin tués le même jour.

De plus, cette étude ayant commencé en 2013, quatre années de production, à travers quatre envois différents, ont ainsi pu être échantillonnées correspondant à des productions de 2012, 2013, 2014 et 2015. L'étude de la période de production étant très importante, celle-ci tenant compte à la fois des changements d'alimentation et de l'évolution du profil métabolique du caviar au cours de sa conservation. L'analyse de boîtes de caviar de 2012 a été rendue possible du fait que la société Sturgeon conserve, lorsque celles-ci n'ont pu être écoulées, certaines unités au sein de grosses boîtes au réfrigérateur entre -4 et -2°C avant d'être rempotées dans des boîtes commerciales en fonction des demandes clients. En effet, la dégradation du caviar est beaucoup plus lente lorsque celui-ci est conservé dans des boîtes de plusieurs kilogrammes et à des températures inférieures à celles des réfrigérateurs de particuliers.

II.2.3) Analyses statistiques

II.2.3.a) Prétraitement des données

Préalablement à cette analyse statistique, l'intégralité des spectres a été phasée manuellement et une correction de la ligne de base a été appliquée. Une fois traités, les spectres RMN ont été référencés sur le pic de gauche du lactate à 1,3376 ppm, ou ont été réalignés à l'aide de la fonction Matlab (MathWorks) que nous avons développée à partir de la fonction *icoshift*. Une fois réajusté, chaque enregistrement est alors référencé selon le même processus. Les spectres ont alors été découpés en 743 « buckets » de 0,01255 ppm soit à 4 points RMN, en excluant les zones spectrales correspondant aux solvants utilisés (méthanol, chloroforme et eau). Pour s'affranchir des variations d'intensité entre les différents échantillons analysés, la valeur de l'intégrale de chaque « bucket » est alors divisée par l'intensité totale mesurée. De plus, les variables possédant des valeurs nulles ont été supprimées avant l'application de méthodes statistiques. Enfin, les données ont été centrées et mise à l'échelle (Pareto), pour l'ACP, SIMCA et l'OPLS-DA). Pour les techniques de classification (SIMCA et OPLS-DA) le

nombre de composantes utilisées pour la construction des modèles statistiques a été déterminé par validation-croisée.

II.2.3.b) Attribution des composés présents dans la phase aqueuse du caviar

Parallèlement à nos travaux portant sur la détermination de l'origine géographique du caviar par RMN, une étude a été menée afin d'attribuer les différents signaux de la phase aqueuse de notre protocole de préparation. En effet, dans le but de comprendre les distinctions observées entre les différents échantillons sur les analyses statistiques, il est absolument nécessaire de pouvoir attribuer les différents pics du spectre RMN à leurs composés chimiques. Le caviar n'ayant jamais été étudié au préalable par RMN d'un point de vue métabolomique, nous avons, dans un premier temps, enregistré des spectres RMN 2D (^1H - ^1H TOCSY et ^1H - ^{13}C HSQC) et interrogé des bases de données (HMDB et BBIOREFCODE) pour attribuer les différents métabolites.

Par la suite, nous avons réalisé une analyse compositionnelle de la phase aqueuse de l'extraction Bligh & Dyer par CPG-SM pour confirmer les résultats de l'attribution et élargir la liste des composés détectés par RMN.

Tout d'abord, la méthode d'extraction Bligh & Dyer est appliquée sur 150 mg de caviar et 50 μl sur les 500 μl de la phase aqueuse sont collectés dans une fiole de 1,5 ml. Les solvants sont alors évaporés sous vide pendant une nuit sans chauffage. Une fois cette étape réalisée, 50 μl d'une solution de pyridine contenant 20mg/l de chlorhydrate de méthoxyamine sont alors ajoutés pour la dérivation (méthoxylation) des composés. Le flacon est alors agité au vortex pendant 10 min afin de dissoudre l'ensemble des molécules puis placé au sein d'un thermomixer pendant 90 minutes à 40 °C sous agitation (900 tours par minute). Ensuite, 5 μl d'une solution de n-alcanes et 80 μl de MSTFA sont ajoutés puis placés de nouveau au thermomixer pendant 30 minutes à 40°C. A ce stade, la solution peut alors être analysée par CPG-SM mais il est possible d'attendre une nuit afin que la silylation soit complète. La solution est alors conservée au réfrigérateur à 4°C jusqu'à son analyse.

Les analyses ont alors été effectuées sur un chromatographe Agilent 6890 couplé à un spectromètre de masse Agilent 5973 équipé d'un détecteur par impact électronique. La température de l'injecteur a été fixée à 230 °C en mode splitless et le débit du gaz vecteur (Hélium) à 0.6 ml par minute. Un gradient de température de 9°C par minute est appliqué de 70 °C à 320°C avec un plateau final de 10 minutes.

Au total 270 composés ont été détectés à l'aide de cette technique et une centaine de molécules ont pu être identifiées, 66 par comparaison avec des standards de référence et 34 à l'aide de leur temps de rétention et de leur spectre de masse par comparaison avec des bases de données (Golm et NIST) : tous les autres restant non identifiés. Il est important d'avoir à l'esprit que les principaux composés détectés par spectrométrie de masse ne sont pas nécessairement les espèces les plus abondantes sur le spectre RMN. En effet, en spectrométrie de masse, les métabolites détectés ainsi que leur teneur dépendent directement de la technique d'ionisation (impact électronique, ionisation chimique, electrospray, désorption-ionisation laser assistée par matrice) mais également du mode d'ionisation utilisé (positif ou négatif). Cependant, la liste des composés identifiés par MS est indispensable pour confirmer l'attribution des différents signaux du spectre RMN et élargir le nombre de signaux attribués.

La combinaison de toutes ces techniques et informations nous a permis d'attribuer 28 composés sur le spectre 1D RMN (Figure 49).

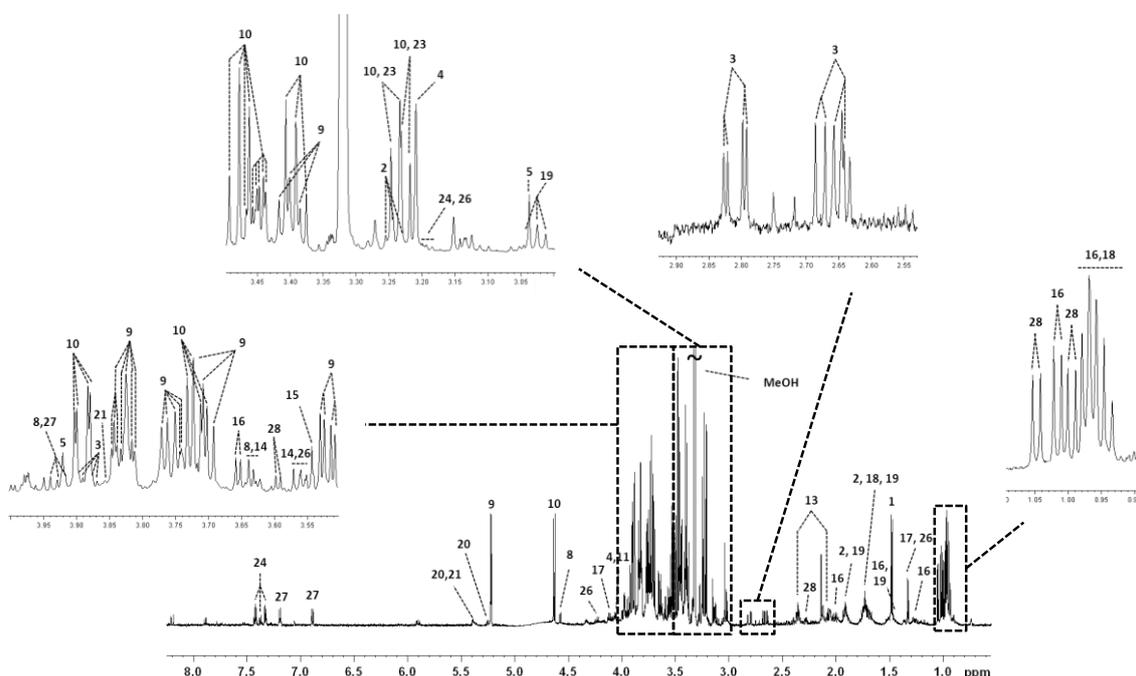


Figure 49 : Spectre 1D ^1H NOESY représentatif des composés détectés au sein de la phase aqueuse du caviar. L'attribution des différents composés a été réalisée à l'aide des analyses réalisées par Spectrométrie de Masse et des expériences RMN 2D (TOCSY et HSQC). Les numéros des différents métabolites réfèrent au Tableau 17.

La plupart des pics majoritaires observés sur le spectre ^1H 1D NOESY ont été attribués. Cependant beaucoup de signaux de composés moins abondants, ainsi que certains pics non

présents dans les bases de données RMN (notamment le signal présent à 3,15 ppm), sur le spectre n'ont pu être assignés.

Afin d'identifier les structures de ces molécules, plusieurs solutions sont envisageables. Tout d'abord, nous pourrions notamment utiliser des techniques séparatives telles que la CLHP ou des cartouches d'extraction en phase solide (SPE) pour collecter de plus grandes quantités de ces composés et enregistrés à nouveau des spectres 2D RMN (^1H - ^{13}C HSQC) sur ces fractions enrichies.

Pour les composés non présents dans les bases de données, une autre alternative serait de réaliser des analyses complémentaires par SM/SM haute résolution. En effet, la détermination des masses moléculaires précises des composés ainsi que les différents chemins de fragmentation établis, permettraient d'identifier la plupart des métabolites non identifiées lors de la première analyse par CPG-SM.

N°	Métabolite	Groupe	Déplacement chimique			N°	Métabolite	Groupe	Déplacement chimique		
			2D hetero		2D homo				2D hetero		2D homo
			¹ H (ppm)	¹³ C (ppm)	¹ H (ppm)				¹ H (ppm)	¹³ C (ppm)	¹ H (ppm)
1	Alanine	βCH ₃ αCH	1,49 3,77	19,09 53,38		17	Acide lactique	CH ₃ CH	1,33 4,09		4,09
2	Arginine	γ-CH ₂ β-CH ₂ δ-CH ₂ CH	1,70 1,92 3,25 3,74		1,92;3,23;3,73	18	Leucine	δCH ₃ δCH ₃ ' γCH βCH ₂ αCH	0,96 0,98 1,73 1,73 3,72	23,84 24,89 27,18 42,65 56,37	
3	Acide aspartique	CH ₂ CH ₂ ' CH	2,67 2,79 3,87	39,31 39,49 55,14		19	Lysine	γCH ₂ δ-CH ₂ β-CH ₂ ε-CH ₂ αCH	1,45 1,73 1,90 3,03 3,74		1,73;1,90;3,03;3,74
4	Choline/Phosphocholine	CH ₃ β-CH ₂ α-CH ₂	3,21 3,52 4,06	56,72 70,44 58,48		20	α-Maltose	CH-4 (2) CH-2 (2) CH-5 (2) CH ₂ (2) CH-3 (1) CH-1 (1) CH-1 (2)	3,39 3,53 3,70 3,82 3,96 5,23 5,39	72,55 74,48 75,72 63,58 76,24 94,36 102,62	
5	Creatine/P-Creatine	N-CH ₃ α-CH ₂	3,04 3,92	39,84 56,72		21	β-Maltose	CH-2 (2) CH-5 (2) CH-3 (1) CH ₂ (1) CH ₂ (2) CH-1 (1) CH-1 (2)	3,53 3,70 3,73 3,77 3,85 4,63 5,39	74,48 75,72 78,88 63,58 63,58 98,93 102,62	
6	Ethanolamine	α-CH ₂ β-CH ₂	3,13 3,82	44,24 60,59		22	α-Melibiose	CH-2 (1) CH-3 (1) CH ₂ (2) CH-2 (2) CH-4 (2) CH-1 (2) CH-1 (1)	3,53 3,70 3,73 3,81 3,98 4,95 5,23	74,48 75,72 63,76 71,32 72,20 100,86 95,59	
7	α-Galactose	CH ₂ CH-4	3,73 3,98	63,76 72,20		23	β-Melibiose	CH-5 (1) CH ₂ (2) CH-2 (2) CH-4 (2) CH-1 (2)	3,23 3,73 3,81 3,98 4,95	77,12 72,55 71,32 72,20 100,86	
8	β-Galactose	CH-2 CH-3 CH-5 CH ₂ CH-4 CH-1	3,48 3,64 3,69 3,73 3,92 4,58	74,84 75,72 78,00 63,76 71,67 99,46		24	Phenylalanine	βCH βCH' αCH δCH ζCH εCH	3,12 3,29 3,98 7,33 7,37 7,42		3,97 3,97 59,01 132,34 130,58 131,99
9	α-Glucose	CH-4 CH-2 CH-3 CH-6;CH'-6 CH-5 CH-1	3,39 3,53 3,70 3,82 3,82 5,22	72,55 74,48 75,72 63,58 74,31 95,06		25	Serine	αCH βCH ₂	3,82 3,98	59,36 63,23	
10	β-Glucose	CH-2 CH-4 CH-3;CH-5 CH-6 CH'-6 CH-1	3,23 3,39 3,46 3,70 3,89 4,63	77,12 72,55 78,88 63,76 63,76 98,93		26	Threonine	γCH ₃ αCH βCH	1,33 3,56 4,24	22,43 63,41 68,86	
11	Acide α-glucuronique	CH-2 CH-3 CH-5 CH-1	3,53 3,70 4,06 5,22	74,48 75,72 75,19 95,06		27	Tyrosine	βCH βCH' αCH εCH δCH	3,04 3,21 3,92 6,90 7,20		3,92 3,92 59,01 118,80 133,75
12	Acide β-glucuronique	CH-3 CH-1	3,48 4,63	74,84 98,93		28	Valine	γCH ₃ γCH ₃ βCH αCH	0,99 1,04 2,27 3,60	19,44 20,85 32,10 63,23	
13	Acide glutamique	βCH ₂ γCH ₂ CH	2,09 2,36 3,74		3,73;2,35						
14	Glycerol	CH ₂ CH ₂ ' CH	3,55 3,64 3,77	65,34 65,34 75,01							
15	Glycine	αCH	3,55	44,24							
16	Isoleucine	δCH ₃ γCH ₃ γCH γCH' βCH αCH	0,95 1,02 1,26 1,48 1,98 3,65	13,99 17,51	0,94;1,02;1,48;1,97 0,94;1,02;1,26;1,97						

Tableau 17 : Liste des métabolites élucidés au sein de la phase aqueuse du caviar grâce aux corrélations homonucléaires ¹H-¹H (TOCSY) et hétéronucléaires ¹H-¹³C (HSQC). (1) et (2) désignent respectivement le monosaccharide possédant l'extrémité réductrice et non réductrice des disaccharides. Les numéros des composés réfèrent au spectre représentatif présenté en Figure 4

II.2.3.c) ACP sur les données Sturgeon

Dans le but d'étudier l'influence des différents paramètres, une ACP a tout d'abord été effectuée sur l'ensemble des données acquises pour la société Sturgeon. Tout d'abord, l'ACP permettant de détecter les points aberrants, un total de 9 points ont ainsi été exclus de nos analyses, ceux-ci correspondant à des échantillons loin de leur duplicat/triplicat ou à des analyses anormalement différentes. Le graphe des scores de l'ACP obtenu en fonction de CP_1 et CP_2 sont donnés en Figure 50.

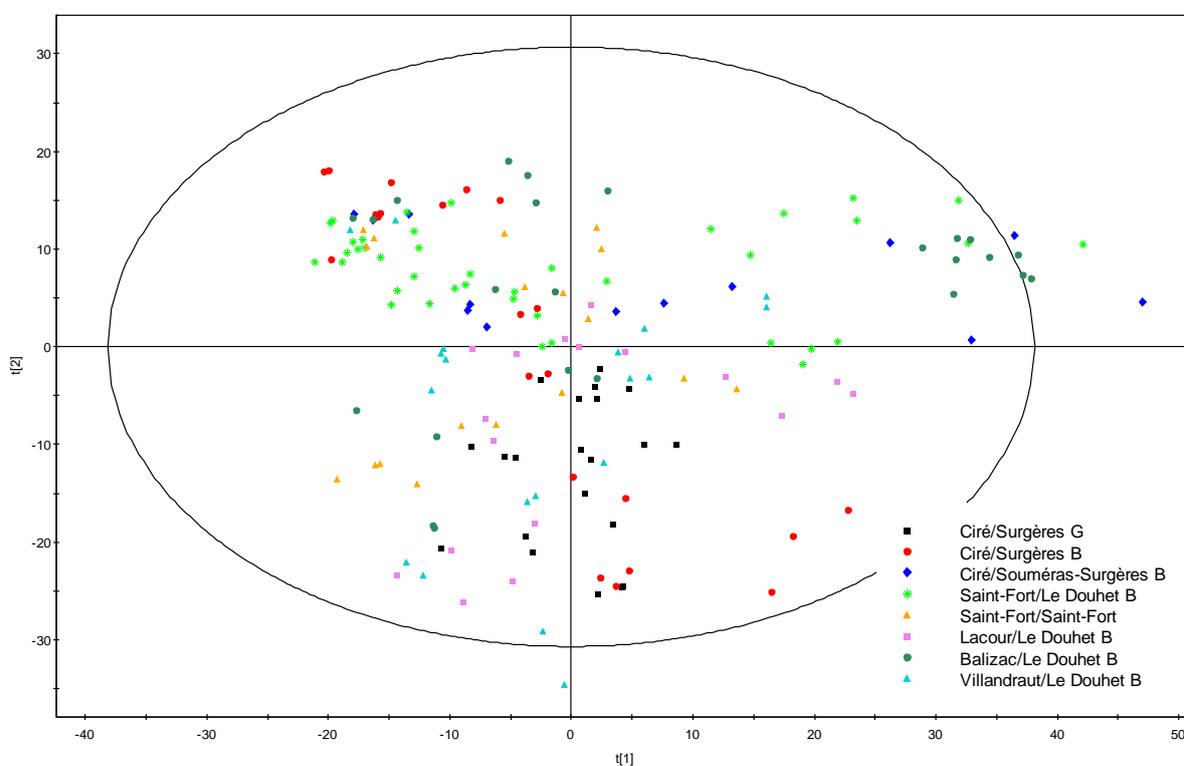


Figure 50 : Graphe des scores (CP_1 vs CP_2) de l'ACP appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur l'ensemble des échantillons de la société Sturgeon. La variance totale expliquée par ces deux composantes est de 52,74% (31,94% pour CP_1 et 20,7% pour CP_2). Chaque couleur représente une combinaison de site d'élevage et de site de rinçage de la société Sturgeon. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

Sur cette Figure, chaque combinaison de site d'élevage et de site de rinçage, est représentée par une couleur différente. Les analyses allant de 1 à 180 ont été effectuées en duplicat tandis que les analyses 181 à 249 ont été réalisées en triplicat, une analyse complémentaire étant réalisée par un autre opérateur dans le but d'étudier la répétabilité de la méthode entre deux manipulateurs différents.

Sur cette ACP, bien que certains paramètres n'aient pu être maintenus totalement identiques tels que la période de production (boîtes pouvant correspondre à plusieurs jours ou semaines différentes) et le régime alimentaire (celui-ci pouvant varier légèrement), certaines tendances intéressantes peuvent être observées.

Tout d'abord, le paramètre semblant introduire le plus de variation est la période de production (année). En effet, que ce soit pour les sites de Saint Fort/Saint Fort (triangles oranges), Balizac/Le Douhet (ronds vert foncé) et Ciré/Surgères *baerii* (points rouges) pour lesquels la production de 3 années différentes (fin 2012, fin 2013 et fin 2014) a pu être échantillonnée, nous observons que trois groupes peuvent être distingués correspondant aux trois années de production. Pour les échantillons couvrant deux années de production (fin 2013 et fin 2014), à savoir Saint Fort/ Le Douhet (étoiles vert clair), Villandraut/Le Douhet (triangles bleu ciel) et Lacour/ Le Douhet (carrés roses), une nette distinction est obtenue entre les deux années. Cependant, les deux ou trois périodes de production ne sont pas placées au même endroit sur le graphe des scores indiquant que d'autres paramètres influencent le profil métabolique du caviar déterminé par RMN (âge du caviar et alimentation). Le seul site pour lequel la période de production semble moins affecter nos résultats, est le site Ciré/Surgères *guldendaedtii*, l'ensemble des points étant relativement proches. Malgré tout, une légère distinction existe entre les 3 périodes couvertes.

De plus, les analyses portant sur les œufs de poisson partageant le même site de rinçage (Le Douhet) mais ayant été élevés au sein d'un site d'élevage différent (Saint Fort, Balizac, Lacour et Villandraut), présentent des profils différents du point de vue de nos analyses ceux-ci étant relativement éloignés sur le graphe des scores de l'ACP (CP_1 vs CP_2) bien que certains échantillons soient relativement proches. L'influence de ce paramètre est également visible sur le graphe des scores en fonction de CP_2 et CP_3 (Figure 51). En outre, les analyses réalisées sur des esturgeons élevés à Ciré mais ayant effectué un passage lors de leur élevage de plusieurs semaines sur le site particulier de Souméras (losanges bleu foncé), apparaissent différents sur l'ACP (CP_1 vs CP_2 mais surtout sur CP_2 vs CP_3) de ceux élevés seulement sur le site de Ciré (points rouges). Ce paramètre semble donc prépondérant comme pour la période de production ; celui-ci influençant directement l'empreinte métabolique du caviar.

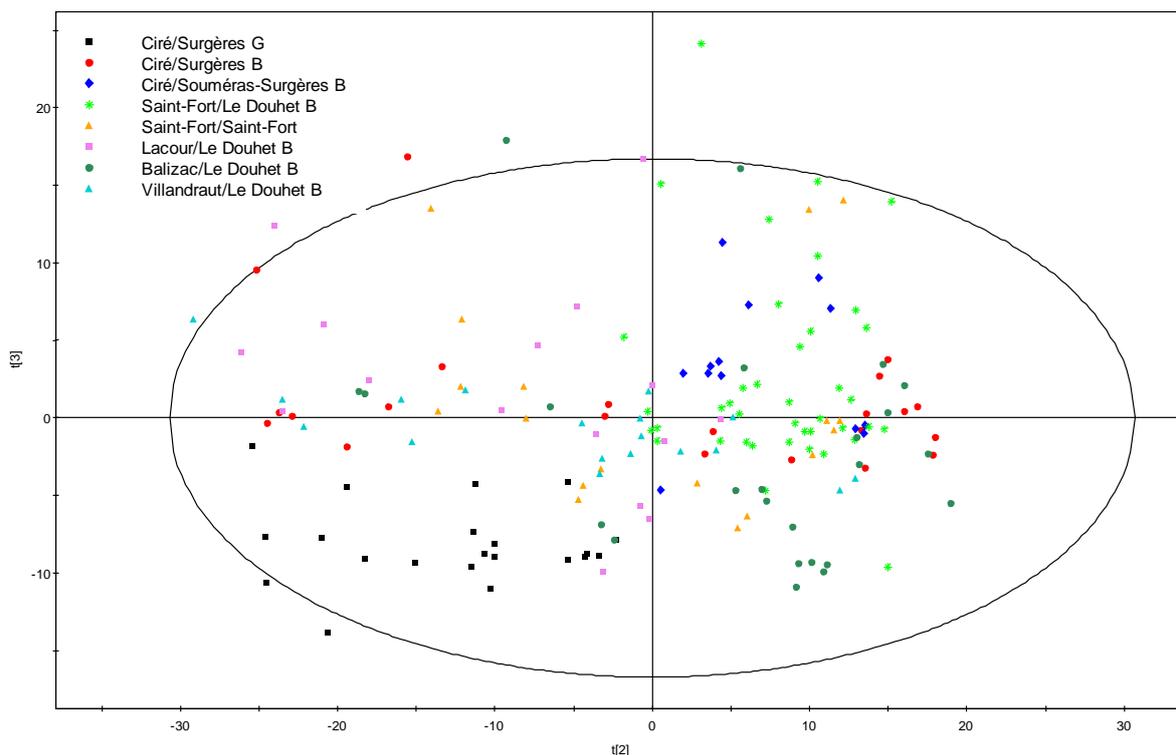


Figure 51 : Graphe des scores (CP_2 vs CP_3) de l'ACP appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur l'ensemble des échantillons de la société Sturgeon. La variance totale expliquée par ces deux composantes est de 26,79% (20,7% pour CP_2 et 6,09% pour CP_3). Chaque couleur représente une combinaison de site d'élevage et de site de rinçage de la société Sturgeon. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

Par ailleurs, les points orange et vert clair correspondent à des esturgeons élevés sur le même site d'élevage (Saint-Fort) mais pour lesquels le site de rinçage diffère (Saint-Fort et Le Douhet). Les analyses de ces échantillons apparaissent assez proches sur les 2 premières composantes pour la plupart d'entre eux. En revanche, pour les points éloignés, il est difficile de déterminer si cela provient réellement de la procédure de rinçage ou à d'autres paramètres. Malheureusement, seuls ces deux sites peuvent nous permettre d'étudier l'influence de ce paramètre; les autres combinaisons ne permettant pas de mettre en avant cette variable, il est donc difficile de conclure quant à son importance sur le profil métabolique déterminé par RMN.

Au vu de ces résultats, l'influence du site d'élevage semble beaucoup plus significative que celle du site de rinçage. Ceci n'est pas une surprise, du fait des 6 à 7 années passées par l'esturgeon sur le site d'élevage, tandis que seulement quelques semaines, le sont sur le site de rinçage. En effet, les conditions et structures d'élevage étant différentes entre les sites

(température de l'eau, qualité de l'eau, conditions climatiques extérieures, alimentation), le profil métabolique des œufs va être directement conditionné par ces paramètres ; celui-ci passant la majorité de sa vie à cet endroit.

D'autre part, sur le graphe des scores de l'ACP représentant la CP₂ en fonction de la CP₃ nous observons une bonne discrimination des carrés noirs, qui correspondent à l'autre espèce d'esturgeons élevée par la société Sturgeon (*guldenstaedtii*) sur le site de Ciré/Surgères, vis-à-vis des points rouges qui, quant à eux, se rapportent à l'espèce *baerii*. En regardant les coefficients de corrélation à l'origine de la distinction entre ces deux types de poissons (Figure 52), il s'avère que les variables les plus significatives dans cette différenciation sont principalement celles liées aux signaux de l' α - et du β -Glucose (Tableau 18). Ainsi, il semblerait que les œufs de l'espèce *guldenstaedtii* soient caractérisés par des teneurs plus élevées en glucose que l'espèce *baerii*.

Enfin, la plupart des duplicats ou triplicats sont très proches sur le graphe des scores, que ce soit sur la Figure 50 ou 51, attestant de la bonne répétabilité du protocole de préparation des échantillons mise au point pour permettre l'étude du caviar par RMN liquide.

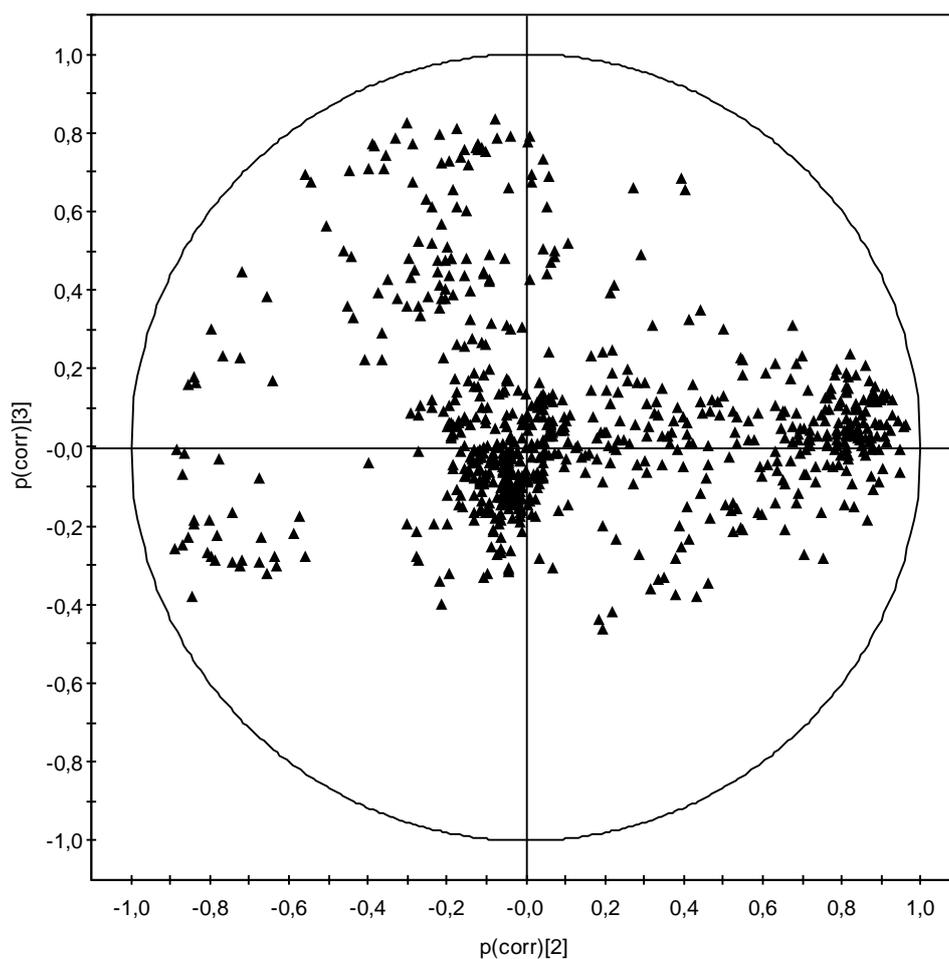


Figure 52 : Graphe des coefficients de corrélation entre les composantes principales (CP₂ et CP₃) et les variables initiales de l'ACP appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur l'ensemble des échantillons de la société Sturgeon

Variable	Coefficient corrélation CP ₂	Coefficient corrélation CP ₃
3,232 ppm	-0,851	-0,234
3,397 ppm	-0,885	-0,258
3,462 ppm	-0,726	-0,213
3,52 ppm	-0,629	-0,305
3,698 ppm	-0,779	-0,22
3,821 ppm	-0,846	-0,3816
3,897 ppm	-0,72	-0,301
4,622 ppm	-0,742	-0,312
4,641 ppm	-0,652	-0,32
5,225 ppm	-0,833	-0,29
5,235 ppm	-0,8665	-0,25

Tableau 18 : Tableau regroupant les coefficients de corrélation entre les composantes principales (CP₂ et CP₃) et les variables initiales à l'origine de la distinction entre les deux espèces d'esturgeon (*guldenstaedtii* et *baerii*) élevées au sein de la société Sturgeon.

II.2.3.d) ACP sur l'ensemble des données (caviars Sturgeon et étrangers)

Une ACP a ensuite été effectuée sur un nouveau jeu de données contenant cette fois-ci l'ensemble de celles enregistrées sur les échantillons Sturgeon ainsi que sur les spécimens de productions étrangères. Ce jeu de données, qui regroupe ainsi 172 spectres de référence (Tableau 19) de la société Sturgeon et 54 spectres étrangers (Tableau 20), est celui que nous utiliserons par la suite pour la construction des modèles SIMCA ainsi que pour les analyses OPLS-DA.

Elevage	Elevage	Rinçage	Espèce	Nombre de boîtes de caviar	Nombre de spectres RMN
Ciré		Surgères	<i>Baerii</i>	8	22
Ciré	Souméras	Surgères	<i>Baerii</i>	9	20
Ciré		Surgères	<i>Guldenstaedtii</i>	5	15
St Fort		Le Douhet	<i>Baerii</i>	18	37
St Fort		St Fort	<i>Baerii</i>	7	18
Lacour		Le Douhet	<i>Baerii</i>	7	13
Balizac		Le Douhet	<i>Baerii</i>	9	24
Villandraut		Le Douhet	<i>Baerii</i>	10	23

Tableau 19 : Tableau récapitulatif du nombre de boîtes de caviar utilisées pour la construction des modèles statistiques et du nombre de spectres RMN correspondant (duplicat et triplicat) de la société Sturgeon. Au total 172 spectres RMN sont utilisés pour les analyses statistiques.

Origine	Espèce	Nombre de boîtes de caviar	Nombre de spectres RMN
Chine	<i>Guldenstaedtii</i>	2	4
Chine	<i>Baerii</i>	3	6
Chine	<i>Skrenkii</i>	2	6
	<i>Dauricus/</i>		
Chine	<i>Skrenkii</i>	2	4
Chine	<i>Huso Dauricus</i>	1	2
Allemagne	<i>Baerii</i>	1	3
Etats-Unis	<i>Transmontanus</i>	2	4
Italie	<i>Baerii</i>	2	6
Italie	<i>Guldenstaedtii</i>	2	6
Italie	<i>Transmontanus</i>	2	4
Italie	<i>Naccarii/Baerii</i>	1	2
Uruguay	<i>Baerii</i>	2	4
Uruguay	<i>Baerii</i>	2	4

Tableau 20 : Tableau récapitulatif du nombre de boîtes de caviar utilisées pour la construction des modèles statistiques et du nombre de spectres RMN correspondant (duplicat et triplicat) des productions étrangères. Au total 54 spectres RMN sont utilisés pour les analyses statistiques.

L'intérêt de réaliser une ACP sur cette base de données est multiple. Tout d'abord, cette méthode constitue une excellente analyse exploratoire permettant de détecter des points aberrants et de les exclure avant l'application d'autres méthodes statistiques. Certaines analyses peuvent apparaître extrêmement différentes sur ce type d'analyse statistique du fait d'un souci rencontré lors de la préparation de l'échantillon. Une autre raison pouvant rendre une analyse aberrante est un mauvais traitement des données RMN. En effet, un spectre incorrectement référencé ou dont la phase a mal été ajustée sera très différent sur l'ACP des autres données.

De plus, comme nous l'avons vu dans la **Partie II**, l'ACP met en évidence des similitudes entre les différents échantillons. Ces informations sont très importantes pour une des techniques de classification que nous allons appliquer par la suite, à savoir les modèles SIMCA. Effectivement, la société Sturgeon présentant une grande variabilité interne, de par la diversité des conditions d'élevage et de rinçage existante au sein de sa production, la construction d'un seul modèle SIMCA pour ce producteur, est impossible. Il est alors nécessaire de regrouper les échantillons similaires au sein d'un même modèle statistique pour réduire la variabilité interne et ainsi améliorer la discrimination avec les productions étrangères.

Par ailleurs, la classification SIMCA est basée sur les résultats de l'ACP. Ainsi, si des distinctions sont observées sur l'ACP sur certaines composantes utilisées pour la construction des modèles SIMCA, alors ces derniers seront capables de différencier les échantillons. Comme nous pouvons le voir sur le graphe des scores donné en Figure 53, pour la plupart des échantillons, une discrimination entre les caviars étrangers (points rouges) et ceux de la société Sturgeon (points noirs) est obtenue principalement sur CP₂ mais également sur CP₄. En regardant les coefficients de corrélations sur le graphe des *loadings*, les différences sur CP₂ sont principalement liées pour les spécimens provenant de productions étrangères à des teneurs plus faibles (coefficients positifs le long de CP₂) en choline et phosphocholine, en tyrosine, en créatine et phosphocréatine, ainsi qu'en un composé non identifié apparaissant sous la forme d'un singulet à 3,15 ppm.

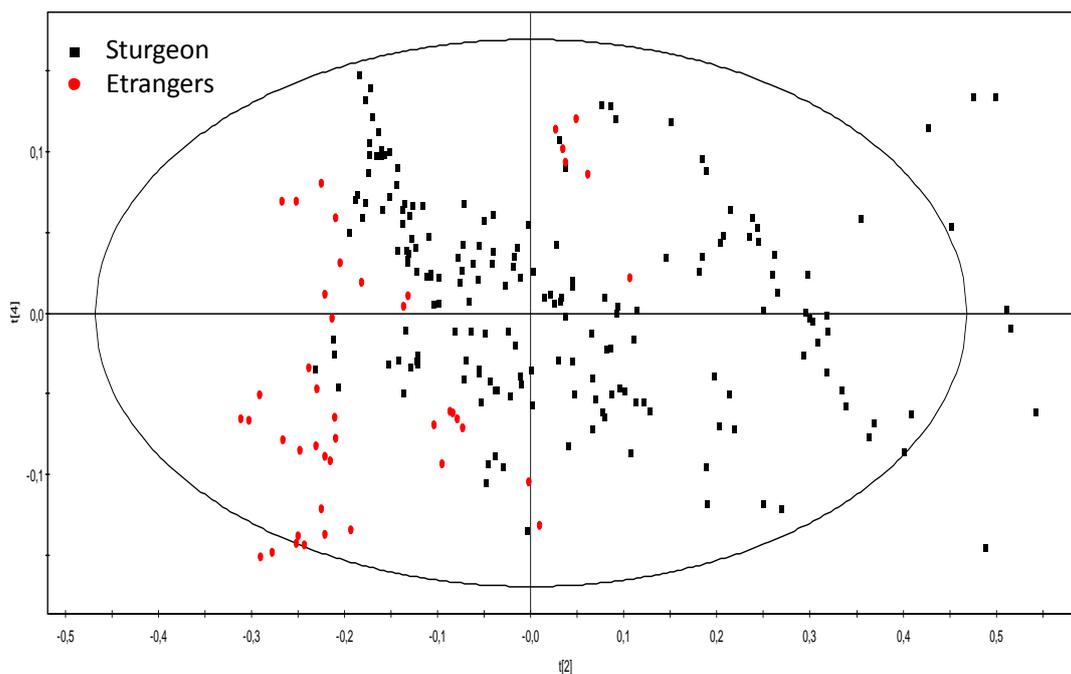


Figure 53 : Graphe des scores (CP_2 vs CP_4) de l'ACP appliquée à l'ensemble des données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur les échantillons de la société Sturgeon (points noirs) ainsi que ceux provenant de productions étrangères (points rouges). La variance totale expliquée par ces deux composantes est de 25,64% (22,68% pour CP_2 et 2,95% pour CP_3). L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

Les caviars étrangers sont également caractérisés par une teneur plus élevée pour certains acides aminés tels que la leucine et l'arginine, en α - et β -Glucose et en galactose pour certains cas. Cependant les coefficients de corrélation positifs sont beaucoup moins prépondérants pour la construction de CP_2 (valeurs maximum comprises entre 0,4 et 0,6 contre 0,7 et 0,9 pour les coefficients négatifs), l'importance de ces métabolites dans la discrimination est donc moins significative.

Enfin, comme pour l'ACP réalisée, précédemment, sur seulement les données de la société Sturgeon, les différents duplicats et triplicats apparaissent pour la plupart extrêmement proches, confirmant la bonne fiabilité et répétabilité du protocole d'extraction développé.

Cette étude exploratoire des données effectuée, nous allons, à présent, appliquer des techniques statistiques supervisées de classification (SIMCA et OPLS-DA) dans le but de mettre au point des modèles statistiques permettant de différencier les différents échantillons de caviar provenant des producteurs de l'Aquitaine de ceux issus de productions étrangères

II.2.3.e) SIMCA

Au début de l'étude sur la détermination de l'origine du caviar par RMN nous avons utilisé la technique de classification SIMCA. Cette méthode d'analyse statistique supervisée de

classification est totalement appropriée aux études portant sur la mise en place d'AOP (Appellation d'Origine Protégée) ou d'IGP, lorsqu'il est difficile de caractériser l'ensemble des productions différentes de celles sujettes à la protection. En effet, la classification SIMCA utilise les résultats de l'ACP pour construire un modèle ACP à l'aide des similarités entre les échantillons dont le nombre de composantes principales est déterminé par validation croisée. Comme nous l'avons vu précédemment, les spécimens inconnus sont alors testés en mesurant la distance de ces derniers par rapport au modèle. Il n'est donc pas nécessaire d'échantillonner toutes les productions, seule celles faisant l'objet d'une démarche IGP ou AOP doivent être caractérisées.

Les résultats obtenus à l'aide des premiers modèles SIMCA construits sur la base des deux premiers envois de boîtes de caviar de la société Sturgeon ainsi que sur les analyses des productions étrangères étaient excellents. Une bonne discrimination était alors obtenue entre les deux classes d'échantillon et la prédiction de boîtes testées en aveugle était correcte (Figures 54 et 55).

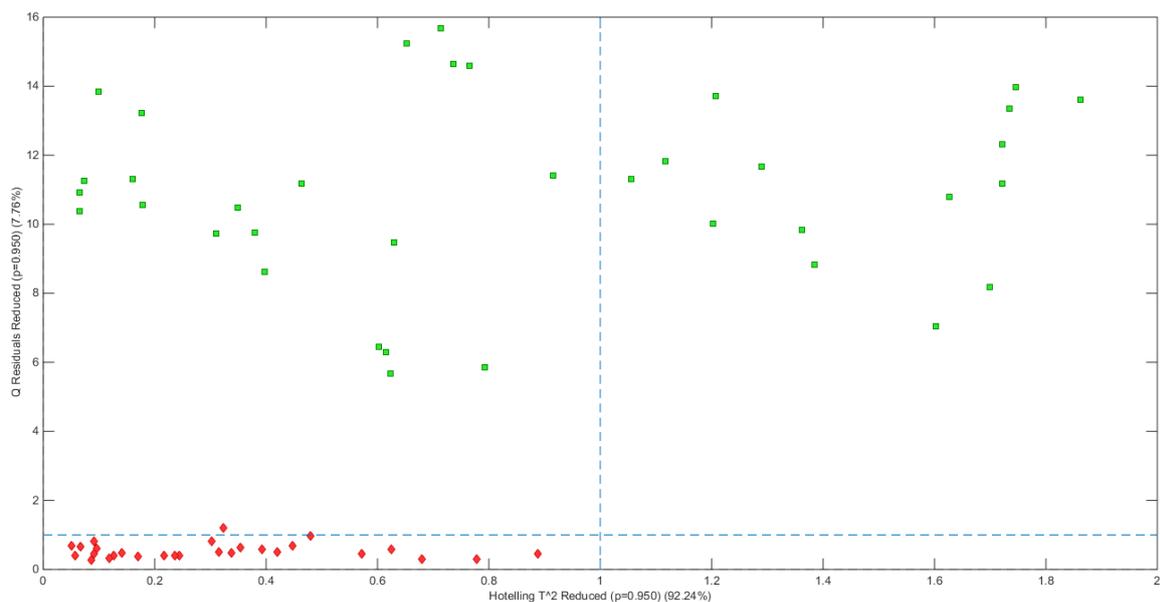


Figure 54 : Modèle SIMCA à 3 composantes, obtenu à l'aide des données acquises sur le premier envoi de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette Figure, la distinction entre les deux groupes est excellente.

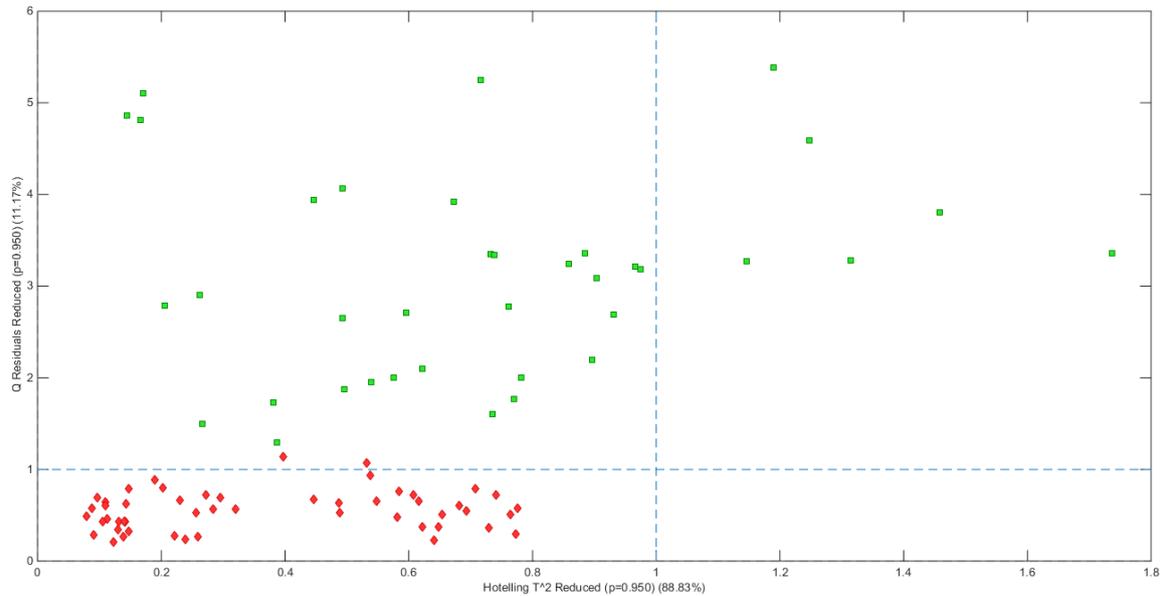


Figure 55 : Modèle SIMCA à 5 composantes, obtenu à l'aide des données acquises sur les deux premiers envois de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleu (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette Figure, certains points de la société Sturgeon commencent à sortir du modèle, cependant la distinction avec les étrangers est encore excellente.

Cependant, au fur et à mesure des envois de Sturgeon la variabilité interne entre les différents échantillons de cette société, liée au grand nombre de sites de rinçage et d'élevage, et à l'influence des différents paramètres influençant le profil métabolique du caviar déterminé par RMN (période de production, alimentation différentes et espèces d'esturgeon élevées), n'a cessé de croître et la qualité des modèles à différencier les deux classes a donc baissée (Figure 56).

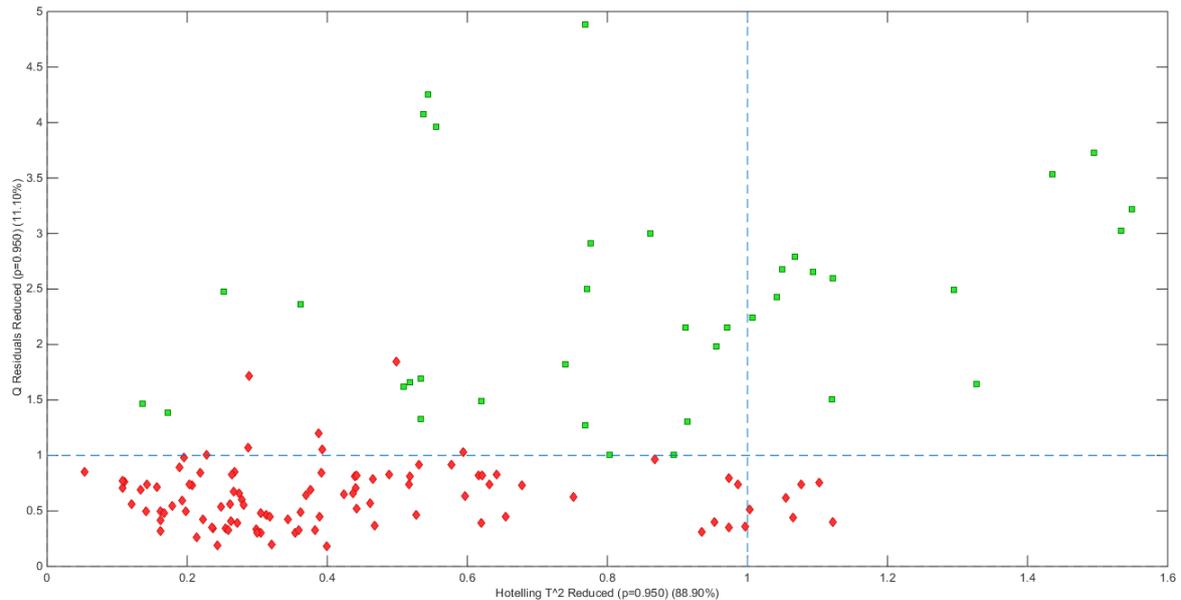


Figure 56 : Modèle SIMCA à 5 composantes, obtenu à l'aide des données acquises sur les trois premiers envois de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette Figure, certains points de la société Sturgeon commencent à sortir du modèle, et la distinction avec les étrangers est moins bonne, certaines boîtes étrangères apparaissent proches ou à l'intérieur du modèle.

Le modèle final à 4 composantes principales contenant l'ensemble des données acquises sur la société Sturgeon vis-à-vis des productions étrangères est donné en Figure 57. Comme nous pouvons le voir, il est impossible de construire un modèle SIMCA contenant l'ensemble des points de la société Sturgeon et discriminant toutes les productions étrangères. Des essais ont été effectués en construisant plusieurs modèles en regroupant les échantillons par similarité (sites spécifiques notamment) cependant la capacité prédictive sur des échantillons inconnus n'était pas satisfaisante. Enfin, des essais de différents traitements statistiques (PQN plutôt que normalisation par rapport à l'intensité totale, mise à l'échelle par unité de variance plutôt que pareto), de sélection de variables (seulement les plus significatives) ont été réalisés conduisant, cependant, à des résultats insuffisants.

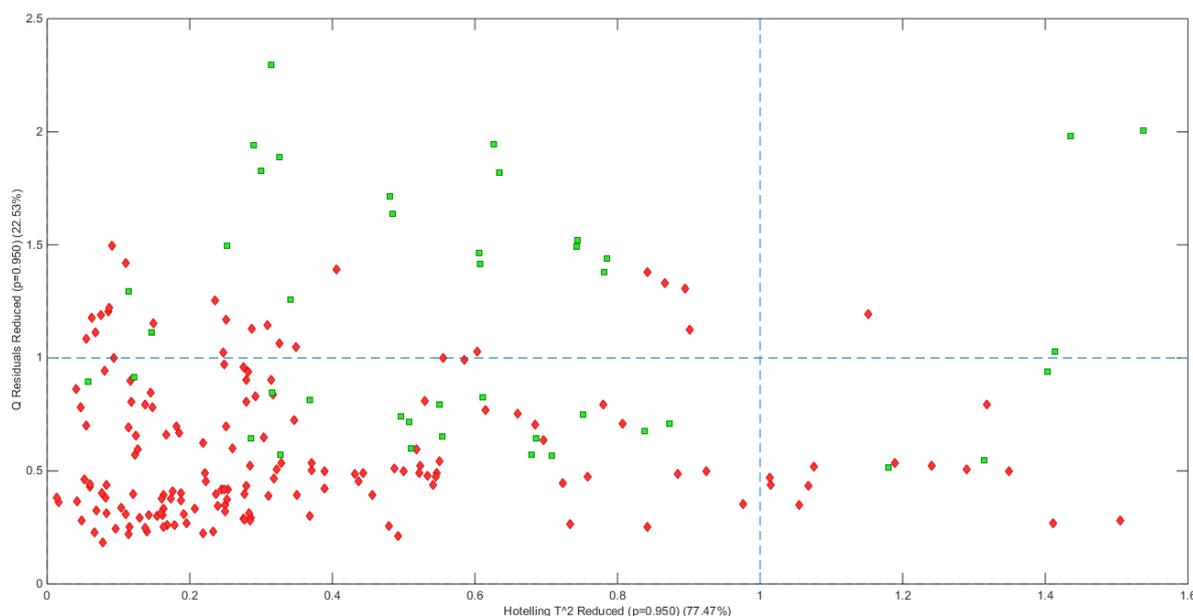


Figure 57: Modèle SIMCA à 5 composantes, obtenu à l'aide des données acquises sur l'ensemble des données de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette figure, la variabilité interne est trop importante et le modèle SIMCA ne fait plus la différence entre les productions.

II.2.3.f) OPLS-DA

Les modèles SIMCA ne permettant pas d'obtenir des résultats satisfaisants en vue de la détermination de l'origine géographique du caviar, nous avons alors décidé d'appliquer une autre méthode statistique l'OPLS-DA afin de pouvoir minimiser la variabilité interclasse observée sur les échantillons Sturgeon et d'améliorer la discrimination avec les productions étrangères. En effet, en regardant les modèles SIMCA construits précédemment, et bien que nous ayons essayé de regrouper les échantillons en fonction de leurs similitudes à l'aide des résultats de l'ACP, il nous a été impossible d'élaborer des modèles incluant l'ensemble des échantillons authentiques français sans réaliser de sur-apprentissage ou sans inclure certains échantillons étrangers à nos modèles.

Comme nous l'avons vu dans la **Partie III**, certains pics étant sensibles à de très légères variations du pH, nous avons mis au point une fonction Matlab basée sur la fonction « icoshift » pour réaligner l'ensemble des spectres afin d'éviter que les analyses statistiques ne soient potentiellement biaisées par le décalage de certains signaux sur le spectre RMN. Cependant, les variations entre les différentes solutions de tampon utilisées n'entraînaient que

des décalages sur les pics extrêmement sensibles, les autres pics présentaient, quant à eux, un excellent alignement.

Des modèles OPLS-DA ont ainsi été construits à l'aide des jeux de données alignés et non réalignés dans le but de comparer les résultats avant et après réalignement, mais également, de s'assurer qu'aucun biais n'était introduit par le processus de traitement utilisé sous Matlab. Les modèles calculés OPLS-DA ne se basant pas sur les quelques pics sensibles au pH, et ceux-ci étant relativement proches nous avons décidé d'utiliser ceux construits à l'aide des données non réalignés. En effet, les modèles élaborés à partir des données seulement référencées présentaient des résultats légèrement supérieurs en termes de classification et de prédiction.

Ainsi, un premier modèle de classification OPLS-DA a été construit, sous le logiciel SIMCA-P 13 (Umetrics), à l'aide de l'ensemble des données acquises sur les différentes boîtes de caviar (françaises et étrangères). Ce jeu de données est composé d'une première classe (classe 1) regroupant les données des producteurs de l'Aquitaine membre de la démarche IGP : Sturgeon (n=172 spectres), la Truite Argentière (n=13 spectres), l'Esturgeonnière (n=10 spectres) et Aquadem (n=24 spectres) et d'une seconde classe contenant celles enregistrées sur les échantillons provenant de productions étrangères (classe 2, n=54 spectres).

A la différence de la société Sturgeon, les exploitations des trois autres producteurs de l'Aquitaine, ne sont composées que d'un seul site d'élevage et de rinçage et n'élèvent qu'une seule espèce d'esturgeon (*baerii*). L'échantillonnage à réaliser pour ces trois producteurs est donc beaucoup moins exhaustif ; le seul paramètre important à considérer étant la période de production. Cependant, la société Aquadem a été échantillonnée de manière un peu plus importante, cette dernière se procurant des alevins à la fois au près de la société Sturgeon mais également au près de l'Esturgeonnière. Ce paramètre a donc été intégré au plan d'échantillonnage de cette exploitation au cas où celui-ci influencerait le profil métabolique du caviar par RMN.

Le graphique des scores obtenus, à l'aide de cette technique statistique, est donné en Figure 58. Comme nous pouvons le voir sur celle-ci, une excellente séparation entre les deux groupes d'échantillons est obtenue à l'aide de cette technique. Pour évaluer la qualité et la robustesse du modèle construit à l'aide de l'OPLS-DA à 10 composantes (1 prédictives et 9 orthogonales), deux paramètres sont couramment utilisés : le R^2Y qui exprime la variabilité expliquée et le Q^2 qui traduit la capacité de prédiction de celui-ci. Dans notre cas les valeurs

mesurées de 0,888 pour le R^2Y et de 0,761 pour le Q^2 attestent de la bonne qualité du modèle statistique construit.

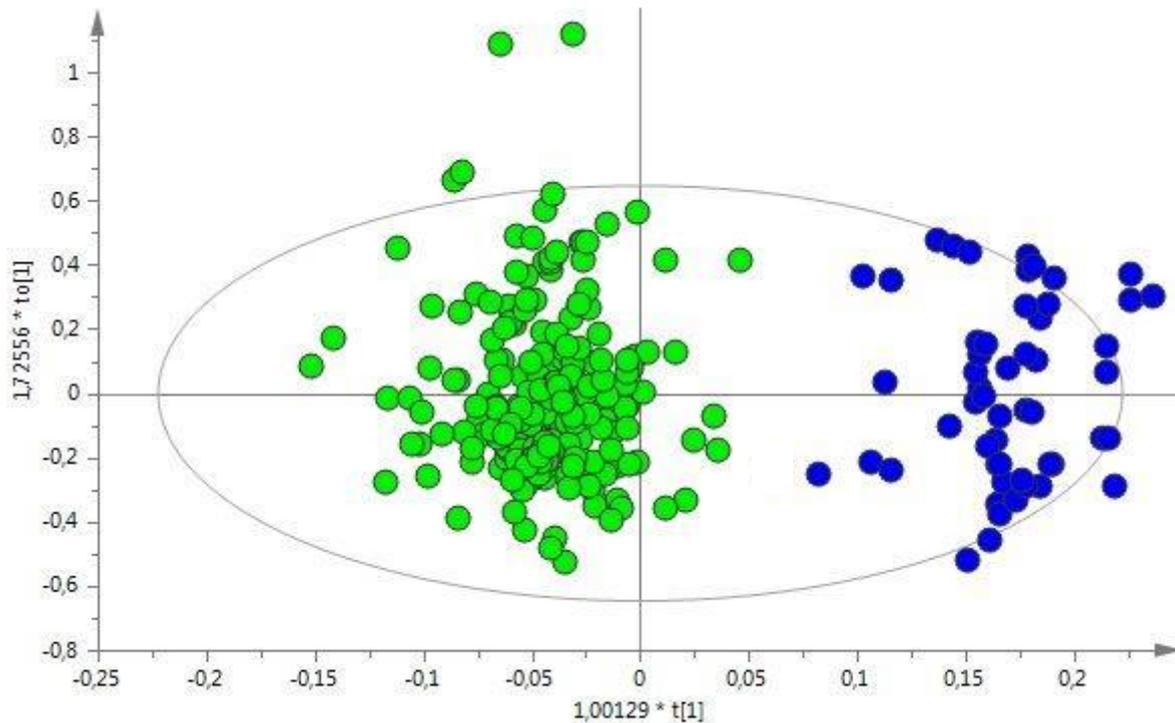


Figure 58 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur les échantillons des 4 producteurs membre de la démarche IGP (points verts, classe 1) ainsi que ceux provenant de productions étrangères (points bleus, classe 2). L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%. Paramètres du modèle : $R^2Y=0,888$ et $Q^2=0,761$.

Afin de tester le modèle statistique construit, 7 boîtes de la société Sturgeon ($n=16$ spectres), correspondant à 1 boîte représentative de chaque combinaison de site d'élevage et de site de rinçage de ce producteur, et 5 boîtes de producteurs étrangers (*baerii* chinois, *baerii* uruguayen, *guldenstaedtii* italien et *schrenckii* chinois, $n=13$ spectres) n'ayant pas été utilisées pour la construction de celui-ci, ont été testées en aveugle et soumises à une prédiction. La classification de ces échantillons est basée sur les valeurs Y prédites calculées pour chaque observation, et chaque spécimen est attribué à la classe pour laquelle sa valeur est la plus élevée.

Comme nous pouvons le voir sur la Figure 59 et le Tableau 21, l'ensemble des échantillons est correctement classifié, cependant, trois échantillons se trouvent à la limite entre les deux

populations. Ceux-ci correspondent respectivement à des caviars *baerii* chinois (n=2) et *guldenstaedtii* italiens (n=1). Le fait que les échantillons proches sur le modèle statistique correspondent à des esturgeons d'espèces identiques à celles que l'on retrouve au sein de la de l'Aquitaine paraît logique: ce paramètre influençant directement le profil métabolique du caviar. De plus, les conditions climatiques et environnementales entre la société Agroitica (Italie) et ceux que l'on retrouve dans la région du bordelais sont très similaires et il n'est pas étonnant que ces échantillons soient proches sur les analyses statistiques.

De manière générale, les duplicats ou triplicats apparaissent relativement proches sur le graphe des scores attestant de la répétabilité des analyses. Cependant, quelques répliquats sont un peu plus éloignés sur le graphique.

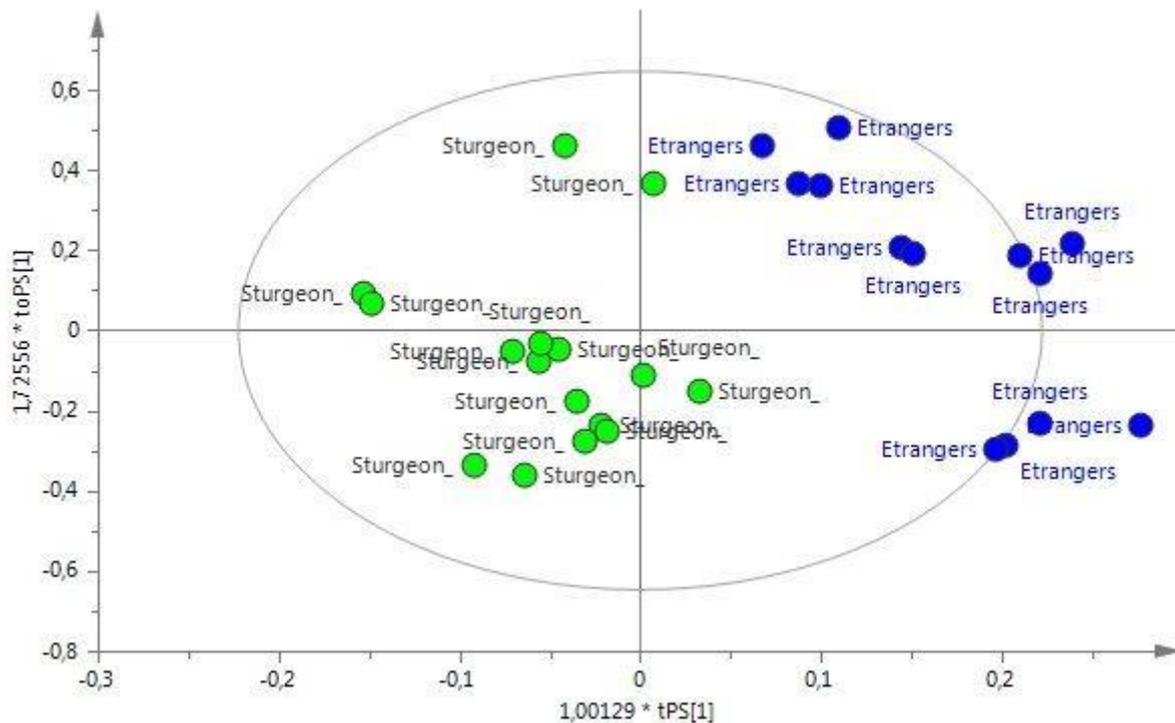


Figure 59 : Projection des échantillons testés en aveugle sur le graphe des scores du modèle OPLS-DA construit à l'aide des données collectées pour les quatre producteurs de l'Aquitaine (Sturgeon, L'Esturgeonnaire, La Truite Argentière et Aquadem), ainsi que sur les caviars étrangers. Les spectres correspondant à des productions de la société Sturgeon sont représentés par des points verts (7 boîtes soit 16 spectres) tandis que ceux provenant de productions étrangères sont représentés par des points bleus (5 boîtes soit 13 spectres). Tout comme le précédent modèle OPLS-DA, l'ensemble des spécimens testés sont correctement classifiés. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

ID	Valeurs prédites (IGP)	Valeurs prédites (Etrangers)
Sturgeon 1	0,7903	0,2096
Sturgeon 2	0,6543	0,3456
Sturgeon 3	1,0942	-0,0942
Sturgeon 4	1,0324	-0,0324
Sturgeon 5	1,4411	-0,4411
Sturgeon 6	1,4246	-0,4246
Sturgeon 7	0,8892	0,1107
Sturgeon 8	0,9428	0,0571
Sturgeon 9	0,9864	0,0135
Sturgeon 10	1,0299	-0,0299
Sturgeon 11	0,8738	-0,1262
Sturgeon 12	0,9264	-0,0735
Sturgeon 13	0,7654	-0,2345
Sturgeon 14	0,9713	0,0287
Sturgeon 15	1,0672	-0,0672
Sturgeon 16	1,1816	0,1816
Etranger 1	0,1880	0,8120
Etranger 2	0,1570	0,8430
Etranger 3	0,4277	0,5723
Etranger 4	0,3754	0,6246
Etranger 5	-0,0594	1,0594
Etranger 6	-0,0369	1,0369
Etranger 7	0,3320	0,6680
Etranger 8	0,4882	0,5118
Etranger 9	-0,1393	1,1393
Etranger 10	-0,3741	1,3741
Etranger 11	-0,0935	1,0935
Etranger 12	-0,2147	1,2147
Etranger 13	-0,1362	1,1362

Tableau 21 : Valeurs Y prédites pour les échantillons de caviar (Sturgeon et étrangers) testés en aveugle à l'aide du modèle OPLS-DA construit à l'aide des données des 4 producteurs de l'Aquitaine (Sturgeon, L'Esturgeonnière, la Truite Argentière et Aquadem) vis-à-vis des productions étrangères. Un échantillon est attribué à la classe pour laquelle sa valeur Y est la plus élevée. L'ensemble des spécimens testés sont correctement classifiés. Les résultats de classification pour ce modèle sont similaires à ceux obtenus précédemment.

Comme nous l'avons vu dans la **Partie II**, il existe plusieurs techniques pour identifier les variables à l'origine de la discrimination sur l'OPLS-DA à travers l'estimation de la contribution des variables individuelles ; les deux principales étant l'inspection des scores de l'importance des variables par projection (« VIP ») et celle des coefficients de régression. C'est la première que nous avons sélectionnée pour déterminer les variables les plus significatives pour la différenciation. De même que précédemment (ACP et donc également SIMCA), les niveaux de choline/phosphocholine, de créatine/phosphocréatine et de β -glucose apparaissent comme les plus significatifs pour la discrimination. La teneur en certains métabolites non identifiés, correspondant à des variables à 3,1288 ppm, 3,15 ppm et 3,26 ppm, font également partie des variables significatives pour la discrimination de ces producteurs. Ces composés sont ceux possédant les scores VIP les plus élevés (supérieurs à

3), cependant, entre 40 et 50 variables (dont les signaux correspondent notamment à l' α -glucose, la tyrosine et la leucine) présentent des scores supérieurs à 1. C'est donc le profil métabolique dans son ensemble (nombreuses variables) qui permet de différencier les différentes origines de caviar.

La choline est une vitamine essentielle et est un donneur de groupes méthyles fondamental en biologie : celle-ci étant impliquée dans plusieurs processus métaboliques ainsi que dans le métabolisme des lipides. Ce métabolite est présent en fortes quantités dans les œufs, le lait et le foie et il n'est donc pas étonnant d'en retrouver en quantité significative sur le spectre RMN de la phase aqueuse du caviar. La phosphocholine est quant à elle synthétisée par phosphorylation de la choline par réaction enzymatique (choline kinase). Les vitamines étant des substances nécessaires au métabolisme d'un organisme vivant, celles-ci ne pouvant être synthétisées naturellement en quantité suffisante, les différences observées sur le profil métabolique du caviar déterminé par RMN proviennent vraisemblablement de régimes alimentaires différents, ou des conditions d'élevage différentes influençant le métabolisme de ces composés, entre les producteurs de l'Aquitaine et les productions étrangères.

La créatine est formée à partir de certains acides aminés principalement la glycine, l'arginine et la méthionine mais peut provenir également de l'alimentation (riche en créatine ou en acides aminés utilisés pour sa formation). La phosphocréatine est comme pour la phosphocholine créée par phosphorylation (catalysée par la créatine kinase) le plus souvent par de l'ATP.

Des différences ayant déjà été observées sur l'ACP entre les deux espèces d'esturgeon de la société Sturgeon (*baerii* et *guldensiaedtii*) et au vu des résultats de l'OPLS-DA, il semblerait que les teneurs absolues soient légèrement différentes pour l' α et le β -glucose en fonction des espèces.

La tyrosine est également un des acides aminés de référence synthétisée à partir de la phénylalanine. Chez l'homme et certains mammifères une teneur élevée indique manque de vitamine C. De plus, la tyrosine participe notamment à la synthèse de l'adrénaline et de la dopamine. Ainsi, sous stress la teneur en tyrosine diminue. Au vu de ces informations, il semblerait donc que la différence de teneur en tyrosine entre les échantillons des producteurs de l'Aquitaine et les productions étrangères puissent être liée à un stress moins important des esturgeons élevés au sein des producteurs du bordelais. En effet, ces échantillons possèdent pour la plupart des teneurs en tyrosine supérieures à celles des étrangers (observées sur

l'ACP). Il est également possible que le régime alimentaire du poisson influence sa teneur, celle-ci étant influencée par la vitamine C.

L'inspection des VIP a également permis de s'assurer que des variables correspondant à du bruit n'était pas pris en considération dans la construction du modèle ce qui peut être le cas lorsqu'une mise à l'échelle est appliquée, et de confirmer que les pics à l'origine des distinctions n'étaient pas liées à des signaux mal alignés.

Un deuxième modèle OPLS-DA a, par la suite, été construit à l'aide du jeu de données précédemment utilisé auquel ont été ajoutées les données acquises sur les boîtes de caviar testés en aveugle et auquel ont été retirées une boîte de chaque combinaison de sites de la société Sturgeon (soit 7 boîtes) ainsi que de 5 origines étrangères différentes (identiques au premier jeu de données) afin de pouvoir réaliser une nouvelle prédiction. Le modèle obtenu et les paramètres d'évaluation de ce dernier ($R^2Y=0.854$ et $Q^2=0.734$) sont très similaires à celui calculé précédemment sur le premier jeu de données.

Les valeurs prédites ainsi que la projection des différents échantillons sur le graphe des scores sont données respectivement en Figure 60 et Tableau 22. L'ensemble des boîtes est correctement classifié sur l'OPLS-DA cependant deux échantillons étrangers correspondant à du caviar *baerii* chinois apparaissent proches de la frontière entre les deux classes comme sur le précédent modèle. Les spécimens de la société Sturgeon possédant les valeurs Y prédites les plus faibles, dont une valeur litigieuse, coïncident à des œufs provenant d'esturgeons *guldensstaedtii* élevés sur le site Ciré/Surgères. Cette espèce étant peu représentée au sein du modèle et différente de la majorité des autres échantillons (*baerii*), il n'est pas forcément étonnant que ce soit pour ces échantillons que le modèle est le plus de difficultés à prédire leurs appartenances. Les résultats de cette nouvelle classification semblent confirmer la tendance déjà observée sur l'ACP selon laquelle l'espèce d'esturgeon est un des principaux paramètres influençant le profil métabolique du caviar par RMN.

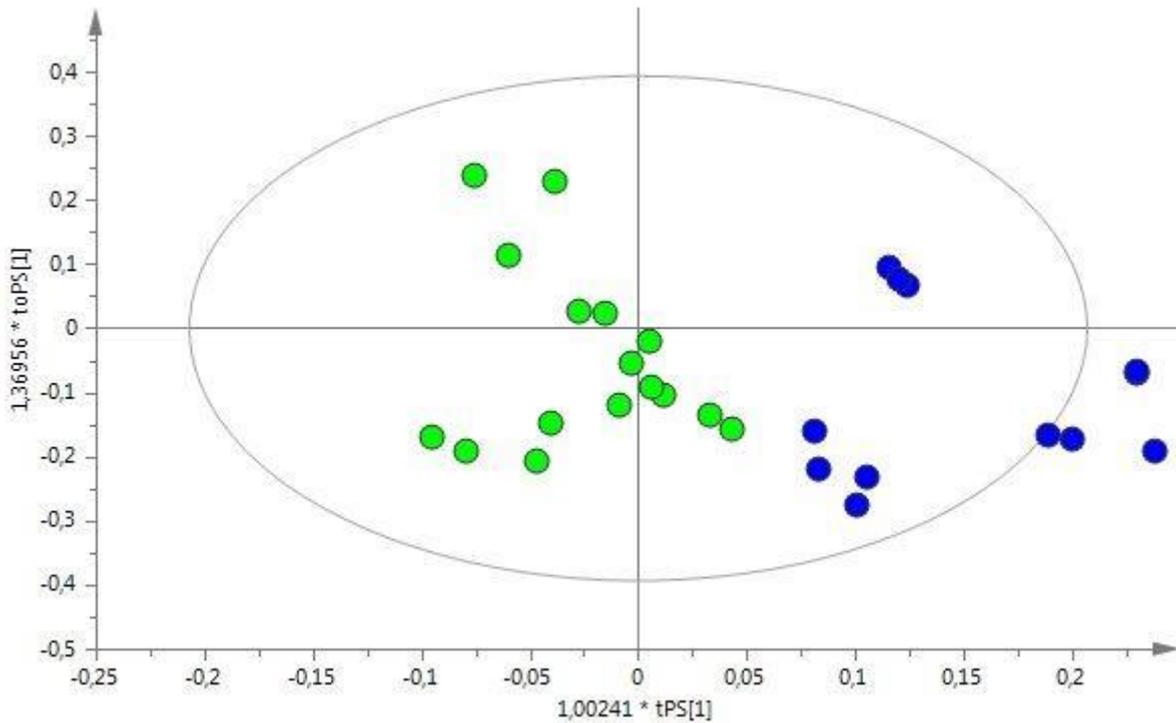


Figure 60 : Projection des échantillons testés en aveugle (second jeu de données) sur le graphe des scores du modèle OPLS-DA construit à l'aide des échantillons des 4 producteurs membre de la démarche IGP ainsi que ceux provenant de productions étrangères. Les spectres correspondant à des productions de la société Sturgeon sont représentés par des points verts (7 boîtes soit 16 spectres) tandis que ceux provenant de productions étrangères sont représentés par des points bleus (5 boîtes soit 12 spectres). L'ensemble des spécimens testés sont correctement classifiés. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.

ID	Valeurs prédites (IGP)	Valeurs prédites (Etrangers)
Sturgeon 1	1,1469	-0,1469
Sturgeon 2	1,2160	-0,2160
Sturgeon 3	0,5943	0,4057
Sturgeon 4	0,6409	0,3591
Sturgeon 5	0,9691	0,0309
Sturgeon 6	1,0008	-0,0008
Sturgeon 7	0,8583	0,1417
Sturgeon 8	0,9109	0,0891
Sturgeon 9	0,8020	0,1980
Sturgeon 10	0,7670	0,2330
Sturgeon 11	0,9638	0,0362
Sturgeon 12	1,1299	-0,1299
Sturgeon 13	1,0581	-0,0581
Sturgeon 14	0,7387	0,2613
Sturgeon 15	0,8295	0,1705
Sturgeon 16	0,7622	0,2378
Etranger 1	0,4256	0,5744
Etranger 2	0,4176	0,5824
Etranger 3	0,3148	0,6852
Etranger 4	0,3369	0,6631
Etranger 5	-0,0603	1,0603
Etranger 6	-0,1084	1,1084
Etranger 7	0,2717	0,7283
Etranger 8	0,2313	0,7687
Etranger 9	0,2507	0,7493
Etranger 10	-0,2800	1,2800
Etranger 11	-0,2426	1,2426
Etranger 12	-0,2427	1,2427

Tableau 22 : Valeurs Y prédites pour les échantillons de caviar testés en aveugle (second jeu de données) à l'aide du modèle OPLS-DA construit à l'aide de deux classes. Un échantillon est attribué à la classe pour laquelle sa valeur Y est la plus élevée. L'ensemble des spécimens testés sont correctement classifiés.

II.2.4) Conclusion

L'étude portant sur la détermination de l'origine géographique du caviar par RMN, menée en partenariat avec les différents producteurs de l'Aquitaine associés dans le cadre de l'obtention d'une IGP afin de protéger leur savoir-faire dans l'élevage d'esturgeon et la production de caviar, a démontré que le profil métabolique obtenu sur la phase aqueuse du caviar par RMN combiné avec une analyse statistique supervisée de classification (OPLS-DA) permettait de différencier les caviars français de ceux issus de productions étrangères. En effet, la méthode statistique initialement utilisée pour la classification (SIMCA) ne présentant pas des résultats satisfaisants à l'aide de l'ensemble des données acquises, nous avons dû appliquer une technique statistique réduisant la variabilité intra-groupe et maximisant celle intergroupe. Les modèles statistiques construits à l'aide de l'OPLS-DA, présentent des paramètres d'ajustement et de prédiction élevés attestant de leur qualité. De plus, les échantillons testés en aveugle ont tous été correctement classifiés.

Les principaux composés majoritaires détectés par RMN ne constituent pas des marqueurs biologiques propres aux productions françaises ou étrangères, cependant, la proportion relative de certains de ces métabolites permet de les discriminer (choline/phosphocholine, glucose, créatine/phosphocréatine et tyrosine). Néanmoins, certains composés minoritaires détectés par RMN mais non identifiés pour des raisons de sensibilité (mis à part pour le composé à 3,15 ppm) sur le spectre ^1H - ^{13}C HSQC, pourraient être des bio-marqueurs, propres à certaines origines ou productions.

Enfin, au vu des résultats obtenus sur les œufs d'esturgeon, les principaux paramètres influençant le profil métabolique du caviar semblent être l'alimentation, l'espèce d'esturgeon, la période de production, les conditions d'élevage et l'âge du caviar. Ce dernier paramètre a rendu plus difficile la mise au point de modèles statistiques, le profil métabolique du caviar évoluant au cours de sa conservation. Il est cependant difficile de conclure de manière certaine sur l'importance relative de ces différents facteurs, tant le nombre de paramètres pouvant entrer en compte ou varier est important.

Les futurs travaux à réaliser dans le cadre de cette étude seront déterminés avec l'ensemble des producteurs de l'Aquitaine lors d'une prochaine réunion. Il pourrait s'agir notamment de l'analyse d'autres boîtes en aveugles (des autres producteurs du bordelais non testés en aveugle ou de boîtes étrangères potentiellement frauduleuses).

II.3) Etude de la dégradation du caviar au cours de sa conservation par RMN

Parallèlement à la démarche d'obtention de l'IGP, les producteurs de caviar sont également à la recherche d'une méthode analytique pouvant évaluer la fraîcheur et la qualité des œufs d'esturgeon. En effet, comme nous l'avons vu dans la partie précédente, à la différence de certains produits agroalimentaires finis, le profil métabolique du caviar évolue au cours de sa conservation au réfrigérateur. Durant celle-ci, si les conditions de stockage ne sont pas respectées (température de 2°C), la qualité des œufs va se dégrader beaucoup plus rapidement en raison de réactions enzymatiques et de dégradations microbiologiques, affectant les propriétés nutritionnelles et donc le goût du caviar. Dans ce contexte, en plus des boîtes fournies par la société Sturgeon pour la construction des modèles statistiques, 3 boîtes issues de réclamations client ont également été analysées en duplicat afin de déterminer si le profil métabolique obtenu par RMN pouvait également être utilisé comme contrôle qualité. Sur l'ensemble de ces boîtes, les analyses ont révélées la présence sur les spectres de deux multiplets non présents dans les autres analyses : un triplet à 1,18 ppm ainsi qu'un quadruplet à 3,65 ppm. Afin d'identifier ce composé de façon certaine, des expériences RMN 2D ont été acquises (^1H - ^1H TOCSY et ^1H - ^{13}C HSQC) sur un des tubes RMN contenant la plus grande quantité de ce composé. Ces deux multiplets étaient corrélés sur le spectre TOCSY et ce composé a facilement été identifié comme étant de l'éthanol. Nous avons alors supposé que la présence d'éthanol pouvait provenir d'une fermentation alcoolique (transformation anaérobie de certains sucres en alcool) par des levures ou des bactéries, ou par des processus enzymatiques. Par ailleurs, les analyses ont également indiqué une quantité inhabituellement élevée d'acide acétique et dans un cadre moindre d'acide lactique. Ainsi nous avons décidé de mener deux études sur la dégradation du caviar dans le but de confirmer cette hypothèse. Une première investigation a alors été conduite pendant 3 mois sur des eppendorfs entièrement remplis et une seconde étude menée cette fois-ci sur 6 boîtes de caviar provenant du même lot pour une durée totale de 5 mois et demi.

Très peu d'études ont été menées sur la fraîcheur du caviar d'un point de vue chimique la plupart portant sur des analyses microbiologiques. Gussoni et al ont étudié l'effet de la conservation du caviar à l'aide de la Microscopie Electronique à Balayage (MEB), de mesures de temps de relaxation transversale et certaines méthodes chimiques (acides gras et ABVT)¹⁹⁰. Une étude plus récente s'est intéressée à l'oxydation du caviar et a différencié des

échantillons de caviar frais par spectroscopie de fluorescence frontale de caviar oxydés à l'aide de méthodes statistiques (PLS-DA, ACP et ADL)¹⁹¹.

II.3.1) Etude eppendorf

Ainsi une première étude a été conduite sur des eppendorfs remplis d'œufs d'esturgeon et conservés au réfrigérateur à 2°C pendant 3 mois, ce laps de temps correspondant à l'ancienne DLUO utilisée par la société Sturgeon. Immédiatement après l'ouverture de la boîte de caviar utilisée pour étudier ce phénomène, une analyse par RMN a été effectuée pour confirmer l'absence d'éthanol dans celle-ci. Une fois cette étape réalisée, 24 eppendorfs ont été remplis autant que possible pour minimiser la quantité d'air présente dans ces tubes et sans casser les œufs. Les eppendorfs ont alors été conservés au réfrigérateur à 2°C et maintenus à l'abri de la lumière afin d'essayer de reproduire les conditions de stockage classiques du caviar jusqu'à leur analyse par RMN.

En ce qui concerne cette investigation, nous avons préparé et analysé une fois par semaine un eppendorf ; deux analyses étant réalisées pour chaque tube correspondant, à une analyse de la partie supérieure, et une de la phase inférieure afin d'étudier si la dégradation était uniforme. Les résultats obtenus pour deux zones des tubes eppendorf étant similaires nous ne présenterons que ceux de la partie supérieure. Les évolutions des différents composés observées pour cette étude sur la partie supérieure des tubes pour différentes périodes de stockage sont données en Figure 61.

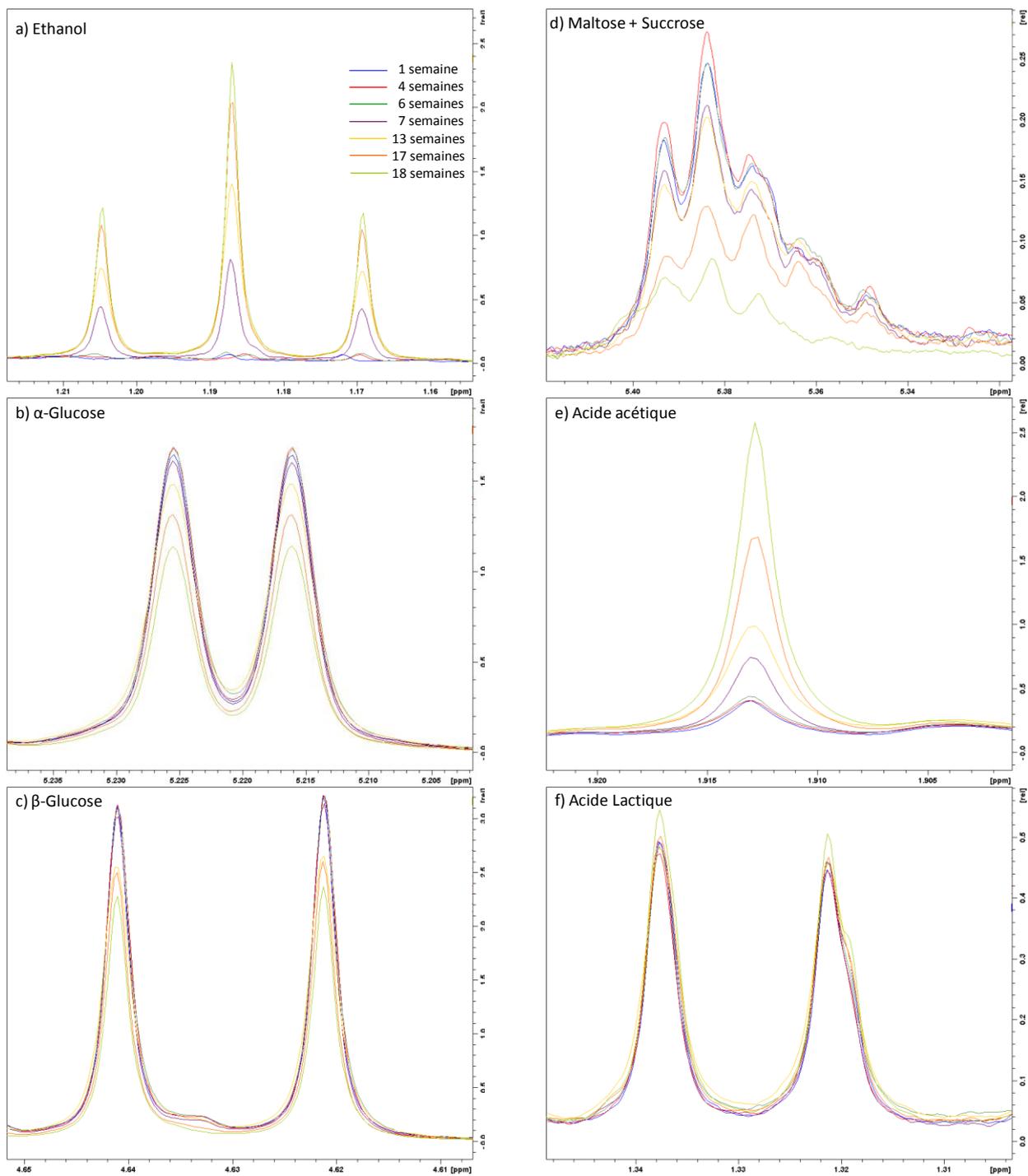


Figure 61 : Evolution des différents composés impliqués dans les processus de fermentation sur les spectres 1D ^1H NOESY d'échantillons de caviar conservés au réfrigérateur dans des tubes eppendorf pour différents temps de conservation jusqu'à 18 semaines a) Ethanol, b) α -Glucose, c) β -Glucose, d) Maltose et saccharose e) Acide acétique et f) Acide lactique

Tout d'abord, en ce qui concerne l'éthanol, celui-ci commence à apparaître sur le spectre à partir de 3 semaines mais en quantité extrêmement faible. Sa teneur devient réellement significative au bout de 5 à 7 semaines et augmente de manière continue tout au long de l'expérience comme nous pouvons le voir sur la Figure 61a).

Dans le même temps les quantités d' α -glucose et de β -glucose augmentent jusqu'à l'apparition de l'éthanol et décroissent ensuite progressivement jusqu'à la fin de l'étude (Figure 61b) et 61c)). De plus, il est important de signaler que non seulement les teneurs en glucose (α et β) diminuent à partir de la production d'éthanol, mais également les quantités de maltose, qui est un disaccharide composé de deux unités de glucose reliées par une liaison α (1-4) et de saccharose qui est également un disaccharide formé d'une unité de glucose et une unité fructose liés par leurs protons anomériques. Sur le spectre RMN ^1H , ces deux molécules peuvent être distinguées des autres sucres par le signal de leurs protons anomériques qui apparaissent sous la forme de doublets qui se chevauchent vers 5,38 ppm (Figure 61d)). Leur présence et leur distinction ont été confirmées par les expériences 2D enregistrées (^1H - ^1H TOCSY et ^1H - ^{13}C HSQC) pour l'attribution des composés présents dans la phase aqueuse du caviar.

De plus, un autre point important à souligner est que, non seulement la teneur en sucres augmente avant l'apparition de l'éthanol, mais aussi la quantité générale de l'ensemble des métabolites détectés par RMN. Au vu de ces résultats, nous supposons qu'au début de la dégradation, l'augmentation générale de l'intensité du spectre est due à la fois à la protéolyse (dégradation des protéines) et l'hydrolyse des polysaccharides, mais également à la dégradation des tissus qui facilite le processus d'extraction. L'hydrolyse des sucres a été confirmée par l'apparition de monosaccharides tels que le galactose sur le spectre RMN caractérisée par un doublet à 5.26 ppm.

En parallèle, la quantité d'acide acétique augmente doucement au début de l'étude jusqu'à l'apparition d'éthanol puis nettement jusqu'à la fin de l'expérimentation (Figure 61e)). La formation d'acide acétique semble indiquer qu'une fermentation aérobie ait lieu lors de la conservation du caviar, l'acide acétique provenant de l'oxydation de l'éthanol. Bien que nous ayons tâché de remplir au maximum les eppendorfs et de minimiser la quantité d'oxygène présent au sein de ces tubes, ce processus de dégradation a pu être accéléré par la préparation des contenants. L'acide lactique suit la même évolution que l'acide acétique, à savoir une lente augmentation au début de la conservation (jusqu'à la production d'éthanol) suivi d'une forte accumulation jusqu'à la fin de l'étude (Figure 61f)). Ces observations

semblent attester qu'une fermentation lactique existe lors du stockage, par des bactéries lactiques, ou par des processus enzymatiques. La fermentation lactique commence par la glycolyse du glucose en acide pyruvique, tout comme la fermentation alcoolique, cependant l'acide pyruvique est directement réduit par la lactate déshydrogénase et le coenzyme NADH (Nicotineamide Adénine Dinucléotide réduit) en acide lactique.

D'après ces résultats, la dégradation du caviar au cours de sa conservation au réfrigérateur, au sein des eppendorfs, peut être décrite en trois phases (Figure 62).

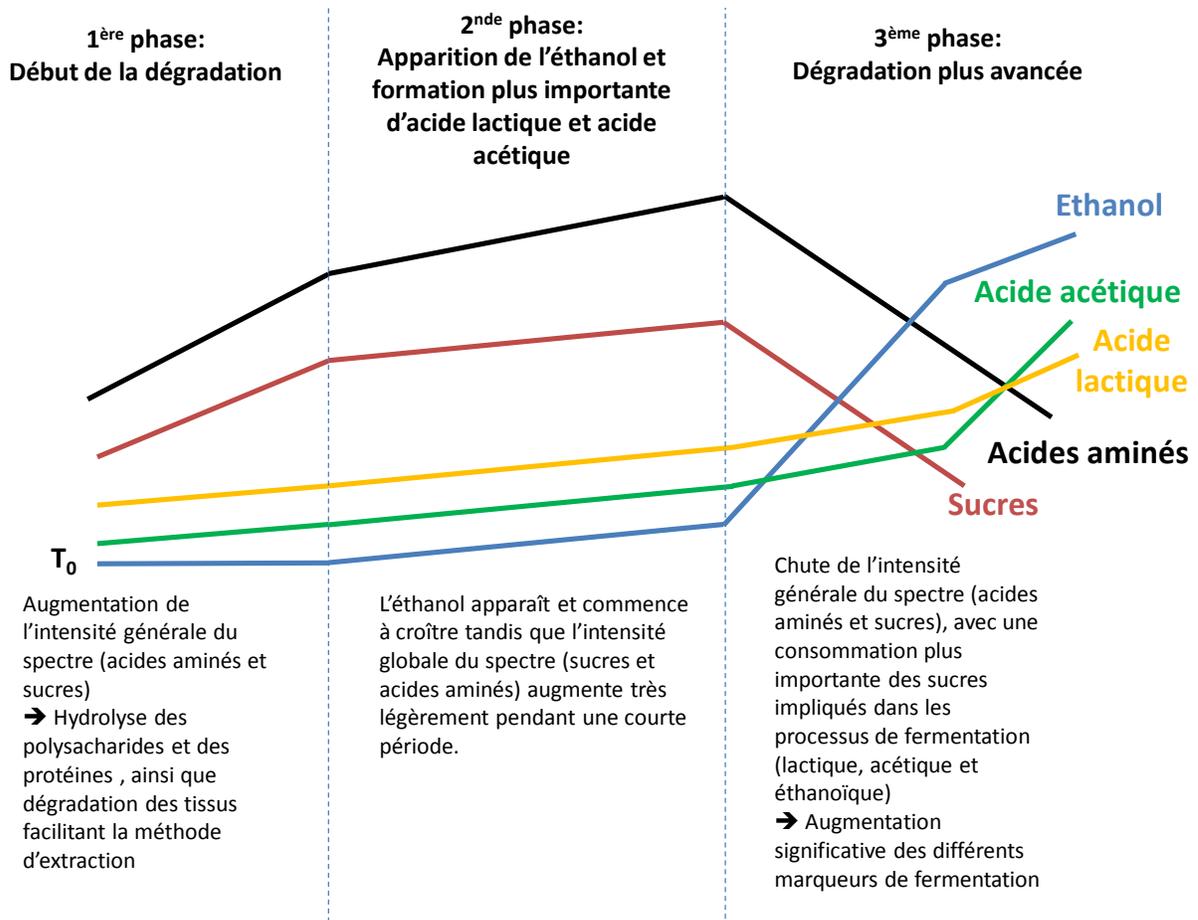


Figure 62 : Schéma de dégradation du caviar au cours de sa conservation au réfrigérateur, caractérisé par trois phases distinctes. Les traits de couleurs différentes représentent l'évolution de l'intensité sur le spectre RMN des différents métabolites (sucres, acides aminés, acide lactique, acide acétique et éthanol) au cours du temps. La durée de chaque phase n'est pas proportionnelle à la longueur sur le schéma.

Tout d'abord une première phase, correspondant au début de la dégradation, durant laquelle l'intensité générale du spectre augmente du fait de l'hydrolyse des protéines et des polysaccharides, ainsi que de l'altération des tissus favorisant le processus d'extraction. Cependant, pendant cette phase, l'intensité (et donc la quantité) des sucres (glucose, maltose et saccharose) augmente plus lentement que les acides

aminés, ceux-ci étant consommés par les différents processus de fermentation. Cette étape prend fin avec l'apparition de l'éthanol.

En effet, la présence d'éthanol semble indiquer le début d'une seconde étape de dégradation puisque, à partir de ce moment, les processus de fermentation commencent à s'accélérer (augmentation plus importante de la teneur en acide lactique et acétique). Pendant cette période relativement courte (quelques semaines), l'intensité générale du spectre augmente doucement.

Enfin, une troisième et dernière phase de dégradation, semble se dégager caractérisée par une nette accélération des processus de fermentation et une chute significative de l'ensemble des composés présents dans la phase aqueuse (acide aminés et sucres) avec une consommation plus importante des sucres impliqués dans les processus de fermentation. Lors de cette dernière étape, l'éthanol qui croît de manière significative au début, augmente par la suite de façon plus modérée, celui-ci étant visiblement oxydé en acide acétique, la formation de ce dernier s'accroissant dans le même temps. De plus, le processus de fermentation lactique s'accélère également ; les conditions devenant probablement plus favorables à sa réalisation (présence d'alcool, pH, teneur en coenzyme NADH).

Cette évolution avec une augmentation de l'intensité du spectre (et donc des métabolites) au début de la conservation du caviar suivie d'une baisse générale aux environs de l'apparition de l'éthanol a également été observée en comparant certains spectres utilisés pour l'étude sur la détermination de l'origine géographique du caviar. En effet, lors de l'envoi transmis par la société Sturgeon dans le but d'équilibrer les modèles statistiques et que ces derniers soient représentatifs de leur production, certains lots transmis correspondaient à des œufs récoltés à des périodes très proches et ayant suivi le même chemin d'élevage et de rinçage que certains échantillons précédemment envoyés, mais conservés pendant plus longtemps au sein de cette entreprise entre -2 et -4 °C. Pour les spécimens conservés pendant une période légèrement plus longue (quelques mois) la teneur des différents métabolites était plus élevée que sur les analyses effectuées précédemment. En revanche, pour les œufs d'esturgeon stockés pendant un intervalle plus long (entre 6 mois et un an), certaines analyses ont révélées la présence d'éthanol en faible quantité et celles-ci présentaient une intensité globale plus faible.

Au vu de l'ensemble des résultats de cette étude, il semblerait que plusieurs processus de fermentation (alcoolique, lactique et acétique) aient lieu lors de la conservation du caviar au réfrigérateur et que ces mécanismes puissent être reliés au niveau de fraîcheur ou d'altération des œufs d'esturgeon. De plus, les conditions de stockage du caviar employées pour cette étude sont tout de même très différentes des

conditions classiques (grammage, matière du contenant, quantité d'air). Ainsi, afin de confirmer que ces mécanismes de fermentation ont bien lieu dans les boîtes de caviar et que ceux-ci n'ont pas été influencés par la préparation des eppendorfs, une seconde étude a alors été menée, cette fois-ci, sur des boîtes commerciales conservées au réfrigérateur jusqu'à 5 mois et demi. L'intérêt de cette expérimentation est également de pouvoir définir les durées des trois phases de dégradation. Nous allons, à présent, nous intéresser à cette seconde étude et présenter les résultats obtenus sur ces boîtes.

II.3.2) Etude longue sur des boîtes commerciales

Pour cette expérimentation, chaque mois, une boîte de caviar conservée dans les mêmes conditions que les eppendorfs (à 2°C et à l'abri de la lumière) a été analysée. Au total 12 analyses sont conduites pour chaque boîte correspondant à 6 zones de prélèvement (centre de la boîte, deux parois différentes, une zone intermédiaire, centre inférieur et paroi inférieure) bien définies étudiées en duplicat (Figure 63).

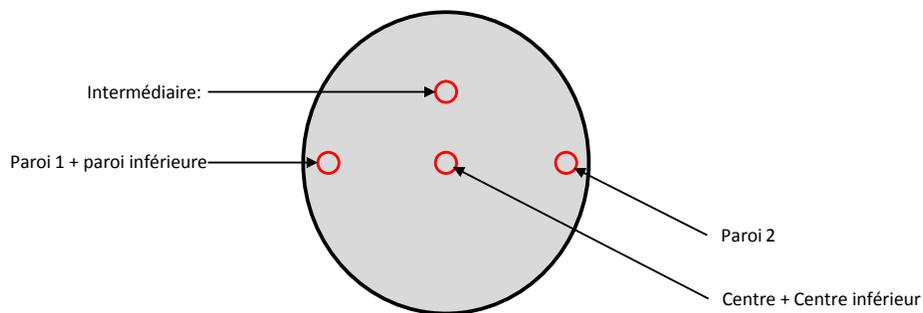


Figure 63. Différentes zones de prélèvement de l'étude longue pour chaque boîte étudiée. Chaque zone a été analysée en duplicat.

Lors des 4 premiers mois de cette étude, l'intensité globale du spectre RMN (acides aminés et sucres) augmente de manière continue tandis qu'en parallèle, les teneurs en acide acétique et en acide lactique croissent de manière légèrement plus significative. En revanche, aucune trace d'éthanol n'a été observée lors des de l'analyse des 5 premières boîtes (4 mois). A ce stade de l'étude l'évolution des différents marqueurs de dégradation ainsi que de l'ensemble des métabolites confirment le schéma de dégradation proposé lors de la précédente étude (tubes eppendorfs), cependant, la dégradation est beaucoup plus lente. Afin que notre étude couvre au maximum la période de DLUO utilisée par la société Sturgeon (6

mois), l'analyse de la dernière boîte a alors été reportée de deux semaines soit 5 mois et demi de conservation au réfrigérateur à 2°C.

Le spectre RMN de celle-ci a alors révélé la présence d'éthanol dans l'ensemble des 12 analyses et une forte augmentation de la teneur en acide acétique. Cette observation n'est pas étonnante puisque celui-ci se forme par oxydation de l'éthanol et confirme que les fermentations aérobie et anaérobie ont bien lieu lors de la conservation du caviar au sein de boîtes commerciales. En revanche, l'intensité de l'acide lactique augmente de manière nettement moins significative. En comparant l'ensemble des données (des deux études) et particulièrement celles correspondant aux deux dernières boîtes étudiées, les teneurs en acides aminés et en sucres étant légèrement supérieures pour les dernière analyse, et les marqueurs de dégradation étant présent en quantités significatives, celle-ci semble donc correspondre au début de la troisième phase de dégradation sur le schéma proposé précédemment. La forte consommation du saccharose et du maltose étant comparable à une dégradation avancée au sein des tubes eppendorfs.

L'ensemble des spectres obtenus au cours de cette étude confirme l'allure générale du schéma de dégradation de la précédente étude, cependant l'évolution est beaucoup plus lente. Malheureusement, la seconde phase de dégradation n'a pu être observée, celle-ci étant très courte et ayant probablement eut lieu entre les deux analyses. Au vue des différents résultats il semblerait que la première phase de dégradation pour des boîtes de caviar de 10 g conservées à 2°C dure environ 4 mois. La seconde phase est alors relativement courte (quelques semaines), le dernier stade de dégradation (avancé) commençant aux environs du cinquième mois de conservation (Figure 64).

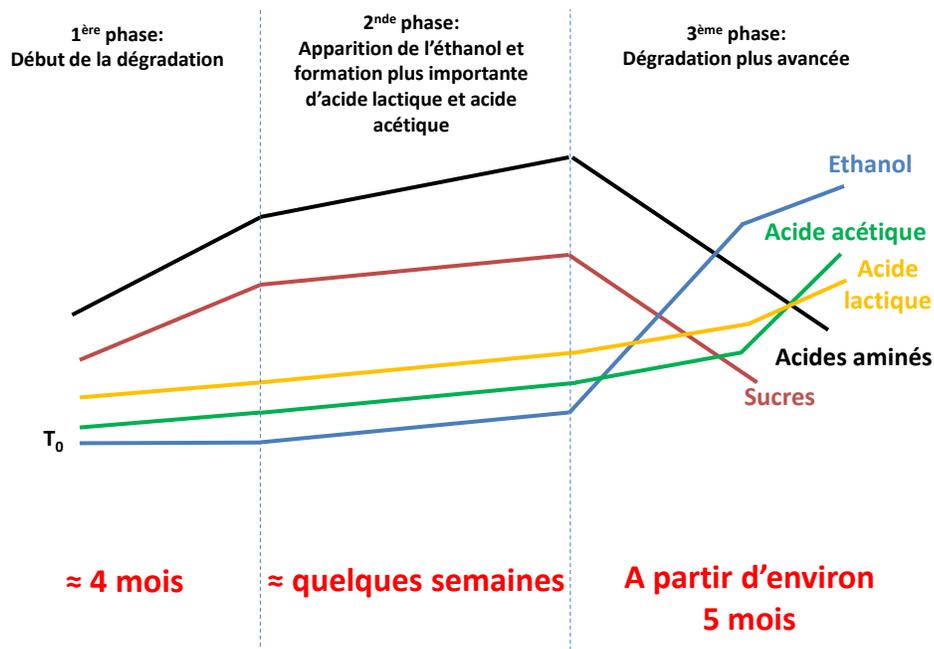


Figure 64 : Schéma de dégradation du caviar, au cours de sa conservation au réfrigérateur au sein de boîtes de 10 g, caractérisé par trois phases distinctes. Les traits de couleurs différentes représentent l'évolution de l'intensité sur le spectre RMN des différents métabolites (sucres, acides aminés, acide lactique, acide acétique et éthanol) au cours du temps. La durée de chaque phase est indiquée en rouge en bas de chaque phase.

II.3.3) Analyses microbiologiques

En parallèle de cette étude, nous avons également réalisé des analyses microbiennes dans le but de déterminer si ces processus étaient bien d'origine microbiologique (bactérienne ou levurienne) ou liés à des enzymes endogènes.

Tout d'abord, plusieurs observations par microscopie ont été effectuées. Dans un premier temps, du caviar a été placé entre la lame et lamelle et observer directement au microscope optique (x40). Ce type d'analyse permet de voir la morphologie, le regroupement et la mobilité des bactéries vivantes. Les mêmes observations ont ensuite été réalisées après avoir appliqué une technique de coloration (Gram) qui rend possible la détection d'espèces microbiologiques vivantes ou mortes et la différenciation des bactéries à paroi simple de celles à paroi double.

L'ensemble de ces tests n'a pas permis de mettre en évidence la présence de micro-organismes au sein des boîtes de caviar. Cependant, l'observation a été rendue difficile du fait de la substance visqueuse qui entoure les œufs d'esturgeon.

Plusieurs boîtes de milieu de culture ont alors été préparées (LB Miller agar et YPD agar) en prélevant la substance se trouvant à la surface des œufs. Trois boîtes LB Miller agar (contenant 10 g NaCl /l, 10 g/l de peptones et 5 g/l d'extrait de levure), qui sont des milieux riches et complexes favorisant la croissance de nombreux genres bactériens (notamment des entérobactéries et trois boîtes YPD agar (contenant 20 g/l de peptones, 10 g/l d'extrait de levure et 20g/l de glucose), qui sont, quant à elles, idéales pour la croissance des levures (plus exigeantes que les bactéries), ont alors été mis à incuber à 25, 30 et 37 ° C pendant 5 jours.

Une seule colonie bactérienne sur un des milieux LB Miller a été observée mais en très faible quantité. Les résultats des tests de croissance indiquent que les boîtes de caviar ne contiennent pas ou très peu de micro-organismes capables de se développer sur les deux milieux testés. Il faudrait, compléter ses analyses par d'autres tests de croissance sur des milieux plus spécifiques pour en avoir la certitude absolue. Cependant, l'observation au microscope n'ayant pas révélée la présence de micro-organismes et après discussion avec des microbiologistes et un des producteurs de caviar, l'intérêt de réaliser de telles analyses est faible.

Au vu des résultats des analyses microbiologiques réalisées la teneur en bactéries et/ou levures est extrêmement faible au sein des boîtes de caviar et les différents processus de fermentation résultent, ainsi, principalement de réactions enzymatiques endogènes suivant le schéma de fermentation présenté en Figure 65.

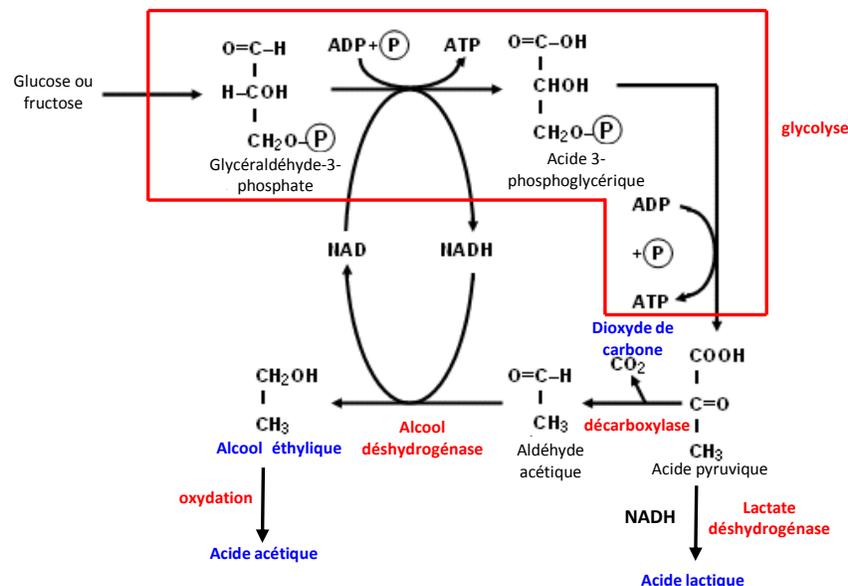


Figure 65. Schéma des différents processus de fermentations enzymatiques (alcoolique, lactique et acétique) à partir du glucose et du fructose. L'acide lactique et l'éthanol résultent de voies anaérobies tandis que l'acide acétique provient d'une réaction aérobie (oxydation de l'éthanol).

II.3.4) Conclusion

Parallèlement aux travaux portant sur l'origine géographique du caviar par RMN, les études conduites sur la dégradation du caviar au cours de sa conservation, dans le but de fournir une méthode de détermination de la qualité et de la fraîcheur des œufs d'esturgeon aux producteurs de caviar, ont démontré que le profil métabolique du caviar pouvait également être utilisé pour le contrôle qualité de boîtes de caviar.

Les deux expérimentations (au sein de tubes eppendorf ou de boîtes commerciales) ont révélées que des phénomènes de fermentation anaérobie (lactique et alcoolique) et aérobie (acétique) avaient lieu et que les produits de ces réactions enzymatiques pouvaient être utilisés comme des marqueurs de dégradation, ceux-ci étant observés à la fin de la DLUO (5 mois et demis) et étant présents en quantités significatives dans les boîtes sujettes à des retours clients. En effet, les niveaux d'acide acétique et d'éthanol déterminés pour ces dernières sont nettement supérieurs à ceux mesurés au sein de la boîte de 10 g analysée après 5 mois et demi de conservation au sein d'un réfrigérateur à 2°C.

Par ailleurs, les analyses microbiologiques (observations microscopiques et milieux de cultures) réalisées au sein de l'Institut de Biologie Moléculaire des Plantes de Strasbourg ont indiquées que la teneur en bactéries et levures était extrêmement faible au sein des boîtes de caviar et que les processus de fermentation étaient donc principalement d'origine enzymatique endogène.

La prochaine étape en vue de la mise au point d'une méthode de contrôle qualité est la réalisation d'analyses organoleptiques afin de déterminer le seuil de détection du palais humain pour l'acide acétique et pour l'éthanol, et ainsi fixer des quantités maximales sur le spectre RMN pour les différents marqueurs de fermentation.

Conclusion et Perspectives

L'objectif de cette thèse était de développer des nouvelles méthodes analytiques, principalement non ciblées, dans l'agroalimentaire par RMN afin de garantir la sécurité des produits étudiés que ce soit en termes de contaminants chimiques, de fraudes ou d'authenticité.

La premier grand axe de recherche de cette thèse a, tout d'abord, été l'étude de composés agroalimentaires par RMN HR-MAS. Ainsi, une méthode de détermination de la qualité et de la fraîcheur du poisson a ainsi été mise au point par RMN HR-MAS : cette technique étant totalement appropriée à l'analyse de ce type d'échantillon. A partir du profil métabolique déterminé à l'aide d'une séquence CPMG 1D ^1H , deux indicateurs chimiques sont estimés simultanément, à savoir, la valeur K qui est basée sur la mesure de la concentration des produits de dégradation de l'ATP, et la teneur du TMA-N qui est issue de la réduction de l'OTMA-N par des bactéries durant la conservation du poisson. La durée totale pour la préparation et l'analyse par RMN HR-MAS est d'environ 40 minutes. Bien que cette méthode présente plusieurs avantages vis-à-vis de celles traditionnellement employées (protocole de préparation extrêmement simple sans étape d'extraction, faibles quantité de poisson nécessaire, échantillons maintenus à 4°C pendant toute la durée de l'expérience), celle-ci se heurte pour le moment au faible intérêt des industriels.

Par ailleurs, la RMN HR-MAS s'est montré être une technique de choix pour l'étude de la texture du poisson à travers l'analyse du temps de relaxation transversale de l'eau à basse vitesse. La méthodologie mise au point, a permis de corréler les résultats obtenus lors des analyses organoleptiques au sein de la société Findus France à ceux acquis par RMN HR-MAS : ces échantillons présentant des profils particuliers que ce soit au niveau de la valeur des constantes de temps de relaxation transversale ou de la proportion des deux types d'eau (eau liée et eau libre). La nouvelle étude, menée en partenariat avec deux fournisseurs de la société Findus France, qui tâchera de relier l'influence du temps passé entre la pêche et la transformation du poisson (valeur K) et la texture (mesures du T_2 de l'eau) ouvre de nouvelles perspectives à ces méthodes.

Les travaux qui ont porté sur la détermination de la quantité absolue des métabolites majoritaires au sein de biopsies (de poisson ou de cerveau) lors d'une acquisition CPMG 1D ^1H ont conduit au développement d'un outil rapide, tenant compte des pertes d'aimantation liés au différents processus de relaxation ayant lieu lors de l'acquisition pour chaque signal, qui permettra d'affiner les mesures de quantification absolue par RMN HR-MAS lors d'études métaboliques. De plus, cette technique permet

de quantifier un composé dans le cas d'un chevauchement entre deux pics, à l'aide des rapports des aires relatives déterminées par déconvolution.

Le second grand axe de recherche de cette thèse, qui représente d'ailleurs la majeure partie des analyses, a été l'étude du caviar par RMN liquide initié avec la société Sturgeon, puis menée, par la suite, en partenariat avec les différents producteurs de l'Aquitaine membres de la démarche IGP, afin de protéger leur savoir-faire dans l'élevage d'esturgeons et la production de caviar. Dans le but de faciliter son obtention, une expérimentation a été conduite sur la détermination de l'origine géographique du caviar par RMN. Pour cela, un protocole de d'extraction et de préparation basé sur la méthode Bligh & Dyer a été mis au point afin de s'affranchir des signaux des lipides, ces derniers masquant une grande partie du profil métabolique du caviar sur le spectre RMN. Après avoir étudié les paramètres pouvant influencer l'empreinte métabolique du caviar par RMN (la période de production, l'âge du caviar, l'alimentation, l'espèce d'esturgeon, les conditions d'élevage et de rinçage) une base de données a été constituée celle-ci regroupant les données des 4 producteurs de l'Aquitaine ainsi qu'un échantillonnage des différents caviars étrangers (Chine, Uruguay, Italie, Allemagne et Etats-Unis). Le profil métabolique dans son ensemble déterminé à l'aide la phase aqueuse du caviar par RMN combiné à une analyse statistiques multivariée de classification (OPLS-DA) a permis de discriminer les œufs d'esturgeons provenant de l'Aquitaine de ceux provenant de productions étrangères. Les valeurs élevées des paramètres d'ajustement (R^2Y) et de prédiction (Q^2) ainsi que la classification correcte des boîtes analysées en aveugle au sein des modèles statistiques attestent de la bonne qualité de ces derniers. Cependant, le développement de modèles statistiques capables de différencier le caviar suivant son origine a été rendu difficile du fait des faibles différences entre les différentes productions et du fait que le profil métabolique du caviar évolue au cours de sa conservation.

Parallèlement aux travaux portant sur l'origine géographique du caviar par RMN, une étude sur la dégradation du caviar au cours de sa conservation a également été conduite : les producteurs de caviar étant également à la recherche d'une méthode de détermination de la qualité et de la fraîcheur des œufs. Celle-ci a permis de démontrer que le profil métabolique enregistré pouvait également être utilisé pour estimer ces paramètres, à travers certains marqueurs de fermentation. En effet, au cours de sa conservation au réfrigérateur, trois phénomènes de fermentation ont lieu au sein des boîtes de caviar : deux processus de fermentation anaérobie (lactique et alcoolique) et un aérobie (acétique). Les expérimentations menées ont permis de montrer que l'éthanol ainsi que la teneur en acide acétique pouvait être des marqueurs d'altération de la qualité des œufs ceux-ci étant présent dans les boîtes

sujettes à des retours clients, et puisqu'à partir du moment où l'éthanol apparaît, la teneur en acides aminés et sucres (et donc la qualité gustative) décroît de manière significative.

Ces deux méthodes pourraient conduire à une utilisation industrielle que ce soit pour vérifier l'authenticité de boîtes frauduleusement estampillées caviar d'Aquitaine pour la première, ou pour le contrôle qualité de boîtes de caviar pour la seconde dans le but d'éviter des retours clients. Il est encore trop tôt pour conclure de manière formelle sur la méthode de détermination de l'origine géographique du caviar, la base de données des producteurs français ainsi que celle des caviars étrangers devant être étoffées afin que le modèle OPLS-DA soit le plus complet possible. Cependant, l'étude de la détermination de l'origine du caviar par RMN est prometteuse et sera intégrée par les producteurs d'Aquitaine à leur dossier dans le but de faciliter l'obtention de l'IGP.

Evidemment, comme nous le savions avant le commencement des différentes études réalisées dans le cadre du projet Agrifood GPS, la RMN (liquide ou HR-MAS) n'est pas applicable au décèlement de contaminations ou de fraudes infimes puisque seuls les composés majoritaires peuvent être détectés à l'aide de cette technique du fait de sa sensibilité. La RMN peut, en revanche, être un outil analytique puissant pour d'autres applications notamment la détection d'adultérations importantes, la caractérisation de l'authenticité de certains produits ou encore l'évaluation de la qualité d'un produit.

Liste des Tableaux

Tableau 1 : Principaux « shims »(trois premiers ordres) exprimés dans le repère orienté à l'angle magique jusqu'au troisième ordre en tant que combinaisons linéaires des « shims » dans le repère du laboratoire.....	70
Tableau 2 : Paramètres d'acquisition de la séquence 1D ¹ H CPMG utilisés en RMN HR-MAS. La durée de l'impulsion 90°est la valeur moyenne observée sur l'ensemble des échantillons.....	76
Tableau 3 : Paramètres d'acquisition de la séquence RMN 1D ¹ H NOESY (bruker <i>noesygppr1d</i>) utilisés pour l'étude la phase aqueuse du caviar. La durée de l'impulsion 90°est la valeur moyenne observée sur l'ensemble des échantillons.....	78
Tableau 4 : Paramètres d'acquisition TOCSY et ATOCSY ¹ H- ¹ H utilisés pour l'identification des composés présents dans les échantillons de poisson (HR-MAS) et de la phase aqueuse du caviar (Liquide).....	83
Tableau 5 : Paramètres d'acquisition HSQC ¹ H- ¹³ C utilisés pour l'identification des composés présents dans les échantillons de poisson (HR-MAS) et de la phase aqueuse du caviar (Liquide).....	86
Tableau 6 : Attribution des différents métabolites détectés dans une biopsie de poisson. Un blanc dans une case de la colonne des déplacements chimiques du ¹³ C indiquent que le composé n'a pu être détecté sur le spectre ¹ H- ¹³ C HSQC pour des raisons de sensibilité.....	100
Tableau 7 : Plan d'échantillonnage conçu par la société Findus France pour la collecte d'échantillon de blocs de farce de colin d'Alaska dans le cadre du projet Agrifood GPS.....	103
Tableau 8 : Moyennes de la valeur K (en%) déterminées pour chaque échantillon de farce de colin d'Alaska de la société Findus. La lettre entre parenthèses représente le code utilisé pour chaque fournisseur.....	105
Tableau 9 : Mesures des constantes de temps de relaxation transversale de l'eau et des proportions relatives en eau libre et eau liée déterminées sur l'échantillon 13947-5 à différentes vitesses de rotation. Les mesures marquées d'un astérisque représentent les analyses réalisées sur des échantillons ayant passé une heure en rotation au sein du spectromètre.....	115
Tableau 10 : Résultats obtenus pour la quantification et la mesure des constantes de temps de relaxation (longitudinale et transversale) des deux signaux de l'Arginine sur une biopsie de poisson à l'aide de la méthode de quantification PULCON/ERETIC. La valeur en gras représente la quantité absolue d'Arginine déterminée pour chaque signal. Celle-ci tient compte de la perte de signal liée aux processus de relaxation.....	129
Tableau 11 : Résultats obtenus pour la quantification et la mesure des constantes de temps de relaxation (longitudinale et transversale) des deux signaux de l'Arginine sur une biopsie de cerveau (oligodendrogliome) à l'aide de la méthode de quantification PULCON/ERETIC. La valeur en gras représente la quantité absolue d'Arginine déterminée pour chaque signal. Celle-ci tient compte de la perte de signal liée aux processus de relaxation.....	131
Tableau 12 : Valeurs de T ₂ mesurées sur les différents dans les biopsies de cerveau ainsi que la perte de signal associée à la relaxation transversale. OLIGO : Oligodendrogliome GBM : Glioblastome.....	134
Tableau 13 : Valeurs de T ₁ mesurées sur les différents dans les biopsies de cerveau ainsi que la perte de signal associée à la relaxation transversale. OLIGO : Oligodendrogliome GBM : Glioblastome.....	135
Tableau 14 : Tableau récapitulatif de la perte totale d'aimantation au cours de la séquence CPMG utilisée pour l'étude de biopsies de cerveau par RMN HR-MAS au sein de l'hôpital d'Haute-pierre pour chaque métabolite sur les différents types de tumeurs. OLIGO : Oligodendrogliome GBM : Glioblastome.....	136

Tableau 15 : Plan d'échantillonnage mis en place dans le cadre de l'étude de faisabilité sur la détermination de l'origine du caviar par RMN HR-MAS.....	143
Tableau 16 : Tableau récapitulatif des différentes combinaisons possibles de site d'élevage et de site de rinçage couramment employés au sein de la société Sturgeon au début de l'étude sur la détermination de l'origine géographique du caviar.....	145
Tableau 17 : Liste des métabolites élucidés au sein de la phase aqueuse du caviar grâce aux corrélations homonucléaires ¹ H- ¹ H (TOCSY) et hétéronucléaires ¹ H- ¹³ C (HSQC). (1) et (2) désignent respectivement le monosaccharide possédant l'extrémité réductrice et non réductrice des disaccharides. Les numéros des composés réfèrent au spectre représentatif présenté en Figure 48.....	150
Tableau 18 : Tableau regroupant les coefficients de corrélation entre les composantes principales (CP ₂ et CP ₃) et les variables initiales à l'origine de la distinction entre les deux espèces d'esturgeon (<i>guldenstaedtii</i> et <i>baerii</i>) élevées au sein de la société Sturgeon.....	155
Tableau 19 : Tableau récapitulatif du nombre de boîtes de caviar utilisées pour la construction des modèles statistiques et du nombre de spectres RMN correspondant (duplicat et triplicat) de la société Sturgeon. Au total 172 spectres RMN sont utilisés pour les analyses statistiques.....	156
Tableau 20 : Tableau récapitulatif du nombre de boîtes de caviar utilisées pour la construction des modèles statistiques et du nombre de spectres RMN correspondant (duplicat et triplicat) des productions étrangères. Au total 54 spectres RMN sont utilisés pour les analyses statistiques.....	156
Tableau 21 : : Valeurs Y prédites pour les échantillons de caviar (Sturgeon et étrangers) testés en aveugle à l'aide du modèle OPLS-DA construit à l'aide des données des 4 producteurs de l'Aquitaine (Sturgeon, L'Esturgeonnière, la Truite Argentière et Aquadem) vis-à-vis des productions étrangères. Un échantillon est attribué à la classe pour laquelle sa valeur Y est la plus élevée. L'ensemble des spécimens testés sont correctement classifiés. Les résultats de classification pour ce modèle sont similaires à ceux obtenus précédemment.....	166
Tableau 22 : Tableau 22 : Valeurs Y prédites pour les échantillons de caviar testés en aveugle (second jeu de données) à l'aide du modèle OPLS-DA construit à l'aide de deux classes. Un échantillon est attribué à la classe pour laquelle sa valeur Y est la plus élevée. L'ensemble des spécimens testés sont correctement classifiés.....	170

Liste des Figures

- Figure 1. Niveaux d'énergie des spins 1/2 en l'absence et en présence de champ magnétique B_0 . La présence d'un champ magnétique conduit à l'apparition de deux niveaux d'énergie pour les spins 1/2 dont la différence d'énergie vaut ΔE . Ces niveaux d'énergie sont peuplés en accord avec la statistique de Boltzmann..... 6
- Figure 2. Représentation de l'orientation de l'axe z du laboratoire dans le système d'axes principaux (SAP)..... 16
- Figure 3. Schéma d'un échantillon hétérogène. Chaque partie de l'échantillon peut être assimilée à une somme de sphères de différents volumes V_n au sein desquelles la susceptibilité magnétique est homogène (χ_n)..... 19
- Figure 4 : Schéma illustrant la mise en rotation d'un rotor contenant l'échantillon à la vitesse ω_r autour d'un axe formant un angle de $54,74^\circ$ par rapport au champ magnétique externe, appelé « angle magique »..... 20
- Figure 5 : Principe de la démarche en métabolomique par RMN. Les flèches en pointillés indiquent les étapes menées en parallèle des études statistiques..... 26
- Figure 6 : Spectres ^1H enregistrés sur un spectromètre 400 MHz sur une biopsie de poisson par RMN HR-MAS à l'aide d'une séquence CPMG (spectre rouge) et NOESY 1D (spectre bleu). Pour les deux séquences une présaturation du signal de l'eau est utilisée..... 28
- Figure 7 : Spectre ATOCSY ^1H - ^1H enregistré sur un spectromètre 400MHz par RMN HR-MAS sur une biopsie de poisson au cours de cette thèse..... 30
- Figure 8 : Spectre HSQC ^1H - ^{13}C obtenu sur un échantillon d'urine tiré de la publication de Nicholson et al 31
- Figure 9 : Construction de la première composante principale par rapport au nuage de points des données initiales. Les scores (t_{i1}) sur la première composante principale (P_1) pour chaque observation correspondent aux projections de ceux-ci sur ce nouvel axe..... 44
- Figure 10 : Construction de la seconde composante principale (P_2) par rapport au nuage de points des données initiales. Celle-ci est perpendiculaire à la première composante et les scores pour chaque observation sont obtenus de la même manière que pour P_1 (projection sur P_2)..... 45
- Figure 11 : Exemple de (a) graphe des scores et (b) graphe des loadings ; en fonction de CP_1 et CP_2 obtenus à l'aide de l'ACP sur des échantillons de caviar. Sur le graphe des scores, une distinction est faite entre les deux classes correspondant à deux espèces d'esturgeon différentes. Le graphe des loadings permet de déterminer quelles sont les variables initiales (origine chimique) responsables de la discrimination entre les deux groupes. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%..... 46
- Figure 12 : Exemple de modèle SIMCA obtenu à l'aide des données acquises sur des échantillons de caviar. En rouge les points utilisés pour la construction du modèle et en noir ceux utilisés pour le classement des échantillons inconnus. L'axe des abscisses représente les valeurs de $T_{\text{red},kA}^2$ tandis que celui des ordonnées représente les valeurs de $Q_{\text{red},kA}$. Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent)..... 53
- Figure 13 : Schéma de l'apport de l'OPLS vis-à-vis de la PLS inspiré de la publication de Trygg et Wold¹¹⁵. La variation systématique non corrélée à Y est séparée du jeu de données facilitant l'interprétation et réduisant la complexité de modèles PLS. Celle-ci peut alors être étudiée et analysée séparément et les sources de variation orthogonale peuvent être identifiées..... 57
- Figure 14 : Spectres ^1H NOESY 1D de la phase aqueuse du caviar enregistrés à 400 MHz a) avant réalignement et b) après réalignement à l'aide la fonction Matlab « icoshift ». L'utilisation de cette fonction permet de corriger les légères variations de pH des différentes solutions de tampon employées pour la réalisation des travaux de cette thèse et de permettre..... 60
- Figure 15 : Exemple de diagramme d'éboullis, représentant les valeurs propres (λ_a) en fonction de chaque composante principale. Dans le cas présent, le coude est obtenu au niveau de la 4^{ème} composante principale..... 63

Figure 16 : Vue du spectromètre Agrifood (AVANCE III 400 MHz) utilisé pour l'acquisition de l'ensemble des données du projet. Celui-ci est équipé d'une unité pneumatique permettant la mise en rotation de l'échantillon, d'un passeur d'échantillon ainsi que d'une unité de refroidissement garantissant le maintien des échantillons à la température souhaitée.....	66
Figure 17 : Schéma du dispositif HR-MAS avec insert jetable (représenté en bleu) de 33 µl en Kel-F situé à l'intérieur d'un rotor. Ce dispositif est composé d'un capuchon et d'une vis permettant de rendre l'insert hermétique. Le rotor est alors fermé à l'aide d'un autre capuchon (ailettes) permettant la mise en rotation rapide du rotor.....	67
Figure 18 : Photographie du passeur d'échantillons HR-MAS Sample Pro. Celui-ci est composé d'un compartiment réfrigéré au sein duquel les rotors sont stockés jusqu'à leur analyse. Une unité pneumatique assure le passage via un tube des rotors jusqu'au spectromètre au sein de la sonde.....	72
Figure 19 : Photographie du dispositif HR-MAS utilisé pour la préparation des différents échantillons étudiés à l'aide de cette technique. De gauche à droite sur la Figure : le rotor, l'insert jetable de 33 µl et le capuchon (ailettes).....	72
Figure 20 : Séquence d'impulsions RMN bruker zgpr. Celle-ci est composée d'une présaturation pendant le temps de répétition puis d'une impulsion 90°.....	74
Figure 21 : Séquence d'impulsions CPMG (bruker cpmgpr1d), avec une présaturation du signal de l'eau pendant le temps de répétition, synchronisée sur la période du rotor. Le temps d'écho total employé pour l'ensemble des acquisitions a été fixé à 75.6 ms ($\tau = 200 \mu s$ et $L4 = 188$).....	75
Figure 22 : Séquence d'impulsions 1D 1H NOESY (bruker noesygppr1d), avec une présaturation du signal de l'eau et l'utilisation de gradients, employée pour l'ensemble des analyses réalisées sur la phase aqueuse du caviar.....	77
Figure 23 : Séquence d'impulsions Inversion-Récupération utilisée pour les mesures de temps de relaxation longitudinale. La durée de τ varie pour chaque incrément de la liste contenant les différents délais (vdlist).....	79
Figure 24 : Séquence d'impulsions TOCSY utilisée pour l'acquisition des spectres 2D homonucléaires 1H - 1H permettant la mise en évidence des couplages d'un système de spins facilitant l'identification des composés. En RMN HR-MAS le mélange isotrope (« spin-lock ») est composée d'impulsions adiabatiques (ATOCSY).....	81
Figure 25 : Séquence d'impulsions HSQC utilisée, avec présaturation du signal de l'eau et l'utilisation de gradients, pour mettre en évidence les corrélations hétéronucléaires 1H - ^{13}C permettant l'identification des composés. L'abréviation ad en indice de certaines impulsions indique que celle-ci est une adiabatique.....	84
Figure 26 : Spectre 400 MHz 1H CPMG HR-MAS enregistré sur une biopsie de bar. Le spectre correspond à l'analyse d'un poisson datant de 6 jours, la teneur en ATP, ADP et AMP est donc faible. L'attribution des différents métabolites réfère aux Tableau 6. Les signaux utilisés pour la détermination de la valeur K sont indiqués par un carré noir. L'abréviation NA est utilisée pour signaler les composés non attribués.....	92
Figure 27 : Spectre 400 MHz HSQC 1H - ^{13}C enregistré sur un échantillon de bar. La correspondance des différentes tâches avec son métabolite est indiquée.....	93
Figure 28 : Evolution de la valeur K (en %) au cours de sa conservation au réfrigérateur dans de la glace à 0°C pour les différentes espèces de poissons étudiées. Les barres représentent la déviation standard observée sur les mesures de la valeur K sur les deux poissons analysés en duplicat pour chaque espèce.....	95
Figure 29 : Evolution de la quantité de TMA-N (en mg/100g) pour le bar et le rouget au cours de sa conservation au réfrigérateur dans de la glace à 0°C. Les barres représentent la déviation standard observée sur les deux poissons analysés pour chaque espèce.....	96
Figure 30 : Evolution des différents produits de dégradation de l'ATP pour la truite marron au cours de sa conservation au réfrigérateur dans de la glace à 0°C. Pour cette espèce, la dégradation de l'ATP, ADP, AMP et IMP est très rapide tandis que	

les quantités d'inosine (HxR) et hypoxanthine (Hx) croissent tout au long de l'étude.....	97
Figure 31 : Graphe des scores (CP_1 vs CP_2) de l'ACP appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. La variance totale expliquée par ces deux composantes est de 64,37% (52,98% pour CP_1 et 11,39% pour CP_2). Chaque couleur représente un Fournisseur différent. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%.....	107
Figure 32 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. Chaque classe représente un fournisseur différent. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95% Paramètres du modèle : $R^2Y=0,807$ et $Q^2=0,554$	109
Figure 33 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. Chaque classe représente un mois de pêche différent. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%. Paramètres du modèle : $R^2Y=0,803$ et $Q^2=0,48$	110
Figure 34 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN HR-MAS sur les blocs de farce de colin d'Alaska de la société Findus. Chaque classe représente une saison de pêche différente. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%. Paramètres du modèle : $R^2Y=0,903$ et $Q^2=0,831$	111
Figure 35 : Exemple de courbe d'ajustement de la décroissance du signal de l'eau dans le plan transversal obtenue avec une fonction bi-exponentielle sur l'échantillon 13878-1 sous le logiciel Dynamics Center. La corrélation entre les mesures et la courbe d'ajustement est excellente.....	117
Figure 36 : Graphe regroupant les mesures des constantes de temps de relaxation transversale (T_{21}) de l'eau liée sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.....	118
Figure 37 : Graphe regroupant les mesures des constantes de temps de relaxation transversale de l'eau libre (T_{22}) sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.....	119
Figure 38 : Proportion en eau liée (pT_{21}) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.....	120
Figure 39 : Proportion en eau libre (pT_{22}) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des six analyses réalisées pour chaque échantillon.....	120
Figure 40 : Graphe regroupant les mesures des constantes de temps de relaxation transversale de l'eau liée (T_{21}) sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.....	123
Figure 41 : Graphe regroupant les mesures des constantes de temps de relaxation transversale de l'eau libre (T_{22}) sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.....	123
Figure 42 : Proportions en eau liée (pT_{21}) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.....	124
Figure 43 : Proportions en eau libre (pT_{22}) déterminées sur les différents blocs de farce de colin d'Alaska. Chaque symbole représente une des quatre analyses réalisées pour chaque échantillon. L'échantillon non conforme est indiqué par l'abréviation NC entre parenthèses après le numéro de lot.....	124

Figure 44 : Fenêtre spectrale d'un spectre ^1H 1D CPMG enregistré sur biopsie de poisson. Le triplet de l'arginine situé à 3.77 ppm présente un chevauchement entre deux composés.....	128
Figure 45 : Mesures des constantes de temps de relaxation longitudinale (T1) des principaux métabolites présents dans les différents types de biopsies de cerveau (\blacklozenge) oligodendrogliome, (\blacksquare) cortex et (\blacktriangle) glioblastome.....	132
Figure 46 : Mesures des constantes de temps de relaxation transversale des principaux métabolites présents dans les trois types de biopsies de cerveau étudiés sur deux échantillons différents représentés par deux symboles distincts (\times et \blacklozenge) oligodendrogliome, (\blacksquare et \blacktimes) cortex et (\blacktriangle et \bullet) glioblastome. Les valeurs mesurées sur les mêmes types de biopsies sont extrêmement proches.....	132
Figure 47 : Comparaison de spectres ^1H HR-MAS de caviar (L'Esturgeonnière) obtenus à l'aide (a) d'une acquisition proton classique (spectre rouge) et (b) d'une séquence CPMG avec une multipresaturation (spectre bleu).....	142
Figure 48 : Comparaison des spectres ^1H enregistrés à 400 MHz (a) en RMN liquide sur la phase aqueuse du caviar à l'aide de la séquence <i>noesygppr1d</i> et (b) en RMN HR-MAS sur des œufs de caviar à l'aide de la séquence <i>cpmggpps</i> (multipresaturation) pour des durées d'acquisition équivalentes. Les différences en fréquence pour certains pics sont liées aux protocoles de préparation, celui utilisé pour la préparation d'échantillon en RMN HR-MAS n'utilisant pas de solution de tampon.....	144
Figure 49 : Spectre 1D ^1H NOESY représentatif des composés détectés au sein de la phase aqueuse du caviar. L'attribution des différents composés a été réalisée à l'aide des analyses réalisées par Spectrométrie de Masse et des expériences RMN 2D (TOCSY et HSQC). Les numéros des différents métabolites réfèrent au Tableau 17.....	148
Figure 50 : Graphe des scores (CP_1 vs CP_2) de l'ACP appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur l'ensemble des échantillons de la société Sturgeon. La variance totale expliquée par ces deux composantes est de 52,74% (31,94% pour CP_1 et 20,7% pour CP_2). Chaque couleur représente une combinaison de site d'élevage et de site de rinçage de la société Sturgeon. L'ellipse du T2 d'Hotelling correspond à un niveau de confiance de 95%.....	151
Figure 51 : Graphe des scores (CP_2 vs CP_3) de l'ACP appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur l'ensemble des échantillons de la société Sturgeon. La variance totale expliquée par ces deux composantes est de 26,79% (20,7% pour CP_2 et 6,09% pour CP_3). Chaque couleur représente une combinaison de site d'élevage et de site de rinçage de la société Sturgeon. L'ellipse du T2 d'Hotelling correspond à un niveau de confiance de 95%.....	153
Figure 52 : Graphe des coefficients de corrélation entre les composantes principales (CP_2 et CP_3) et les variables initiales de l'ACP appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur l'ensemble des échantillons de la société Sturgeon.....	155
Figure 53 : Graphe des scores (CP_2 vs CP_4) de l'ACP appliquée à l'ensemble des données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur les échantillons de la société Sturgeon (points noirs) ainsi que ceux provenant de productions étrangères (points rouges). La variance totale expliquée par ces deux composantes est de 25,64% (22,68% pour CP_2 et 2,95% pour CP_3). Sturgeon. L'ellipse du T ² d'Hotelling correspond à un niveau de confiance de 95%.....	158
Figure 54 : Modèle SIMCA à 3 composantes, obtenu à l'aide des données acquises sur le premier envoi de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleu (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette Figure, la distinction entre les deux groupes est excellente.....	159
Figure 55: Modèle SIMCA à 5 composantes, obtenu à l'aide des données acquises sur les deux premiers envois de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses	

représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette Figure, certains points de la société Sturgeon commencent à sortir du modèle, cependant la distinction avec les étrangers est encore excellente

..... 160

Figure 56 : Modèle SIMCA à 5 composantes, obtenu à l'aide des données acquises sur les trois premiers envois de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette Figure, certains points de la société Sturgeon commencent à sortir du modèle, et la distinction avec les étrangers est moins bonne, certaines boîtes étrangères apparaissant proches ou à l'intérieur du modèle..... 161

Figure 57: Modèle SIMCA à 5 composantes, obtenu à l'aide des données acquises sur l'ensemble des données de la société Sturgeon (points rouges). Les points en vert représentent les caviars étrangers testés vis-à-vis du modèle. L'axe des abscisses représente les valeurs de T_{red}^2 tandis que celui des ordonnées représente les valeurs de Q_{red} . Les limites du modèle sont délimitées par les axes en pointsillés bleus (valeur limite de 1 dans le cas présent). Comme nous pouvons le voir sur cette figure, la variabilité interne est trop importante et le modèle SIMCA ne fait plus la différence entre les productions..... 162

Figure 58 : Graphe des scores obtenus à l'aide de l'OPLS-DA, sous le logiciel SIMCA-P, appliquée aux données enregistrées par RMN liquide haute résolution sur la phase aqueuse du caviar sur les échantillons des 4 producteurs membre de la démarche IGP (points verts, classe 1) ainsi que ceux provenant de productions étrangères (points bleus, classe 2). L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%. Paramètres du modèle : $R^2Y=0,888$ et $Q^2=0,761$ 164

Figure 59 : Projection des échantillons testés en aveugle sur le graphe des scores du modèle OPLS-DA construit à l'aide des données collectées pour les quatre producteurs de l'Aquitaine (Sturgeon, L'Esturgeonnière, La Truite Argentière et Aquadem), ainsi que sur les caviars étrangers. Les spectres correspondant à des productions de la société Sturgeon sont représentés par des points verts (7 boîtes soit 16 spectres) tandis que ceux provenant de productions étrangères sont représentés par des points bleus (5 boîtes soit 13 spectres). Tout comme le précédent modèle OPLS-DA, l'ensemble des spécimens testés sont correctement classifiés. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%..... 165

Figure 60 : Projection des échantillons testés en aveugle (second jeu de données) sur le graphe des scores du modèle OPLS-DA construit à l'aide des échantillons des 4 producteurs membre de la démarche IGP ainsi que ceux provenant de productions étrangères. Les spectres correspondant à des productions de la société Sturgeon sont représentés par des points verts (7 boîtes soit 16 spectres) tandis que ceux provenant de productions étrangères sont représentés par des points bleus (5 boîtes soit 12 spectres). L'ensemble des spécimens testés sont correctement classifiés. L'ellipse du T^2 d'Hotelling correspond à un niveau de confiance de 95%..... 169

Figure 61 : Evolution des différents composés impliqués dans les processus de fermentation sur les spectres 1D 1H NOESY d'échantillons de caviar conservés au réfrigérateur dans des tubes eppendorf pour différents temps de conservation jusqu'à 18 semaines a) Ethanol, b) α -Glucose, c) β -Glucose, d)Maltose et saccharose e)Acide acétique et f)Acide lactique..... 174

Figure 62 : Schéma de dégradation du caviar au cours de sa conservation au réfrigérateur, caractérisé par trois phases distinctes. Les traits de couleurs différentes représentent l'évolution de l'intensité sur le spectre RMN des différents métabolites (sucres, acides aminés, acide lactique, acide acétique et éthanol) au cours du temps. La durée de chaque phase n'est pas proportionnelle à la longueur sur le schéma..... 176

Figure 63. Différentes zones de prélèvement de l'étude longue pour chaque boîte étudiée. Chaque zone a été analysée en duplicat..... 178

Figure 64 : Schéma de dégradation du caviar, au cours de sa conservation au réfrigérateur au sein de boîtes de 10 g, caractérisé par trois phases distinctes. Les traits de couleurs différentes représentent l'évolution de l'intensité sur le spectre RMN des différents métabolites (sucres, acides aminés, acide lactique, acide acétique et éthanol) au cours du temps. La durée de chaque phase est indiquée en rouge en bas de chaque phase..... 180

Figure 65 : Schéma des différents processus de fermentations enzymatiques (alcoolique, lactique et acétique) à partir du glucose et du fructose. L'acide lactique et l'éthanol résultent de voies anaérobies tandis que l'acide acétique provient d'une réaction aérobie (oxydation de l'éthanol)..... 181

Liste des publications et communications orales

Publications

- 1) Heude, C., Lemasson, E., Elbayed, K., and Piotto, M. (2014). Rapid Assessment of Fish Freshness and Quality by ¹H HR-MAS NMR Spectroscopy. *Food Analytical Methods*, 8(4), 907-915.
- 2) Heude, C., Elbayed, K., Jezequel, T., Fanuel, M., Lugan, R., Heintz, D., Benoit, P. and Piotto, M. (2015) Metabolic characterization of caviar specimens by ¹H NMR spectroscopy: Towards caviar authenticity and integrity, *Analytical Chemistry* (soumis)

Communication orale

Heude, C. Metabolic characterization of caviar specimens by ¹H NMR spectroscopy: Towards caviar authenticity and integrity, **Euroanalysis 2015** (Bordeaux, France)

Posters

Heude, C., Lemasson, E., Elbayed, K., and Piotto, M. Rapid Assessment of Fish Freshness and Quality by ¹H HR-MAS NMR Spectroscopy, **RFMF 2014** (Lyon, France)

Heude, C., Elbayed, K. and Piotto, M. Développement de nouvelles méthodes analytiques dans l'agroalimentaire par RMN HR-MAS, **Journée Poster ED269 2013** (Strasbourg, France)

References

1. Rinke, P., Moitrier, S., Humpfer, E., Keller, S., Moertter, M., Godejohann, M., Hoffmann, G., Schaefer, H., & Spraul, M. A ^1H -NMR-technique for high throughput screening in quality and authenticity control of fruit juice and fruit juice raw materials -SGF- profiling. *Fruit Processing* **10**, 10-18 (2007).
2. Alonso-Salces, R. M., Moreno-Rojas, J. M., Holland, M. V., Reniero, F., Guillou, C., & Héberger, K., Virgin olive oil authentication by multivariate analyses of ^1H NMR fingerprints and $\delta^{13}\text{C}$ and $\delta^2\text{H}$ data. *J. Agric. Food Chem.* **58(9)**, 5586-5596 (2010).
3. Bloch, F., Nuclear Induction. *Physical review* **70(7-8)**, 460 (1946).
4. Fitch, W. L., Detre, G., Holmes, C. P., Shoolery, J. N., & Keifer, P. A., High-resolution ^1H NMR in solid-phase organic synthesis. *J. Org. Chem.* **59(26)**, 7955-7956 (1994).
5. Cheng, L. L., Ma, M. J., Becerra, L., Ptak, T., Tracey, I., Lackner, A., & Gonzalez, R. G., Quantitative neuropathology by high resolution magic angle spinning proton magnetic resonance spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6408-13 (1997).
6. Griffin, J. L., Walker, L. A., Garrod, S., Holmes, E., Shore, R. F., & Nicholson, J. K., NMR spectroscopy based metabonomic studies on the comparative biochemistry of the kidney and urine of the bank vole (*Clethrionomys glareolus*), wood mouse (*Apodemus sylvaticus*), white toothed shrew (*Crocidura suaveolens*) and the laboratory rat. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **127(3)**, 357-367 (2000).
7. Li, W., Multidimensional HRMAS NMR: a platform for in vivo studies using intact bacterial cells. *Analyst* **131(7)**, 777-781 (2006).
8. Simpson, A. J., Kingery, W. L., Shaw, D. R., Spraul, M., Humpfer, E., & Dvortsak, P., The application of ^1H HR-MAS NMR spectroscopy for the study of structures and associations of organic components at the solid-aqueous interface of a whole soil. *Environmental science & technology* **35(16)**, 3321-3325 (2001).
9. Lippens, G., Bourdonneau, M., Dhalluin, C., Warras, R., Richert, T., Seetharaman, C., ... & Piotto, M., Study of compounds attached to solid supports using high resolution magic angle spinning NMR. *Current Organic Chemistry* **3**, 147-169 (1999).
10. Riedl, R., Tappe, R., & Berkessel, A., Probing the scope of the asymmetric dihydroxylation of polymer-bound olefins. Monitoring by HRMAS NMR allows for reaction control and on-bead measurement of enantiomeric excess. *Journal of the American Chemical Society* **120(35)**, 8994-9000 (1998).
11. Plante, O. J., Palmacci, E. R., & Seeberger, P. H., Automated solid-phase synthesis of oligosaccharides. *Science* **291(5508)**, 1523-1527 (2001).
12. Martínez-Bisbal, M. C., Martí-Bonmatí, L., Piquer, J., Revert, A., Ferrer, P., Llacer, J. L., ... & Celda, B., ^1H and ^{13}C HR-MAS spectroscopy of intact biopsy samples ex vivo and in vivo ^1H MRS study of human high grade gliomas. *NMR in Biomedicine* **17(4)**, 191-205 (2004).

13. Chan, E. C. Y., Koh, P. K., Mal, M., Cheah, P. Y., Eu, K. W., Backshall, A., ... & Keun, H. C., Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). *Journal of proteome research* **8(1)**, 352-361 (2008).
14. Abragam, A. *The Principles of Nuclear Magnetism* (ed. ed. Oxford University Press) (, 1961).
15. Pauling, L., Robinson, A. B., Teranishi, R., & Cary, P., Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography. *Proc Natl Acad Sci USA* **68(10)**, 2374-2376 (1971).
16. Horning, E., & Horning, M. G., Metabolic profiles : gas-phase methods for analysis of metabolites. *AM Assoc Clin Chem* **17(8)**, 802-809 (1971).
17. Mamer, O. A., & Crawhall, J. C., The identification of urinary acids by coupled gas chromatography-mass spectrometry. *Clinica Chimica Acta* **32(2)**, 171-184 (1971).
18. Oliver, S. G., Winson, M. K., Kell, D. B., & Baganz, F., Systematic functional analysis of the yeast genome. *Trends in Biotechnology* **16(9)**, 373-378 (1998).
19. Moco, S., Forshed, J., De Vos, R. C., Bino, R. J., & Vervoort, J., Intra-and inter-metabolite correlation spectroscopy of tomato metabolomics data obtained by liquid chromatography-mass spectrometry and nuclear magnetic resonance. *Metabolomics* **4(3)**, 202-215 (2008).
20. Pan, Z., & Raftery, D., Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Analytical and Bioanalytical Chemistry* **387(2)**, 525-527 (2007).
21. Kovacs, H., Moskau, D., & Spraul, M., Cryogenically cooled probes-a leap in NMR technology. *Progress in Nuclear Magnetic Resonance* **46(2)**, 131-155 (2005).
22. Bunescu, A., Garric, J., Vollat, B., Canet-Soulas, E., Graveron-Demilly, D., & Fauvelle, F., In vivo proton HR-MAS NMR metabolic profile of the freshwater cladoceran *Daphnia magna*. *Molecular BioSystems* **6(1)**, 121-125 (2010).
23. Bon, D., Gilard, V., Massou, S., Pérès, G., Malet-Martino, M., Martino, R., & Desmoulin, F., In vivo ³¹P and ¹H HR-MAS NMR spectroscopy analysis of the unstarved *Aporrectodea caliginosa* (Lumbricidae). *Biology and Fertility of Soils* **43(2)**, 191-198 (2006).
24. Robertson, D.G., Metabonomics in toxicology: a review. *Toxicological Sciences* **85(2)**, 809-822 (2005).
25. Madsen, R., Lundstedt, T., & Trygg, J., Chemometrics in metabolomics-a review in human disease diagnosis. *Analytica Chimica Acta* **659(1)**, 23-33 (2010).
26. Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R. N., & Willmitzer, L., Metabolite profiling for plant functional genomics. *Nature biotechnology* **18(11)**, 1157-1161 (2000).
27. Fontaine, J. X., Molinié, R., Tercé-Laforgue, T., Cailleu, D., Hirel, B., Dubois, F., & Mesnard, F., Use of ¹H-NMR metabolomics to precise the function of the third glutamate dehydrogenase gene in *Arabidopsis thaliana*. *Comptes Rendus Chimie* **13(4)**, 453-458 (2010).

28. Afman, L., & Müller, M., Nutrigenomics: from molecular nutrition to prevention of disease. *Journal of the American Dietetic Association* **106(4)**, 569-576 (2006).
29. Reaves, M. L., & Rabinowitz, J. D., Metabolomics in systems microbiology. *Current opinion in biotechnology* **22(1)**, 17-25 (2011).
30. Bundy, J. G., Davey, M. P., & Viant, M. R., Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **5(1)**, 3-21 (2009).
31. Mesnard, F., & Ratcliffe, R. G., NMR analysis of plant nitrogen metabolism. *Photosynthesis Research* **83(2)**, 163-180 (2005).
32. Beckonert, O., Keun, H. C., Ebbels, T. M., Bundy, J., Holmes, E., Lindon, J. C., & Nicholson, J. K., Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature protocols* **2(11)**, 2692-2703 (2007).
33. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., ... & Sykes, B. D., ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *Journal of Biomolecular NMR* **6(2)**, 135-140 (1995).
34. Freeman, R. Spin Choreography (ed. ed. Oxford University Press) (Oxford, 1998).
35. Morris, G. A., & Freeman, R., Selective excitation in Fourier transform nuclear magnetic resonance. *Journal of Magnetic Resonance* **29**, 433-462 (1978).
36. Patt, S.L., & Sykes, B.D., Water Eliminated Fourier Transform NMR Spectroscopy. *Journal of Magnetic Resonance* **56**, 3182-3184 (1972).
37. Piotto, M., Saudek, V., & Sklenář, V., Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR* **2(6)**, 661-665 (1992).
38. McKay, R.T., How the 1D-NOESY suppresses solvent signal in metabolomics NMR spectroscopy: An examination of the pulse sequence components and evolution. *Concepts in Magnetic Resonance Part A* **38(5)**, 197-220 (2011).
39. Carr, H. Y., & Purcell, E. M., Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical review* **94(3)**, 630-638 (1954).
40. Meiboom, S., & Gill, D., Modified spin-echo method for measuring nuclear relaxation times. *Review of scientific instruments* **29(8)**, 688-691 (1958).
41. Nicholson, J., & Lindon, J., Analysis of drug-protein binding using nuclear magnetic resonance based molecular diffusion measurements. *Analytical Communications* **34(8)**, 225-228 (1997).
42. Hu, F., Furihata, K., Kato, Y., & Tanokura, M., Nondestructive quantification of organic compounds in whole milk without pretreatment by two-dimensional NMR spectroscopy. *Journal of Agricultural and Food Chemistry* **55(11)**, 4307-4311 (2007).

43. Ludwig, C., Ward, D. G., Martin, A., Viant, M. R., Ismail, T., Johnson, P. J., ... & Günther, U. L., Fast targeted multidimensional NMR metabolomics of colorectal cancer. *Magnetic Resonance in Chemistry* **47(S1)**, S68-S73 (2009).
44. Martineau, E., Tea, I., Akoka, S., & Giraudeau, P., Absolute quantification of metabolites in breast cancer cell extracts by quantitative 2D 1H INADEQUATE NMR. *NMR in Biomedicine* **25(8)**, 985-992 (2012).
45. Aue, W. P., Bartholdi, E., & Ernst, R. R., Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *The Journal of Chemical Physics* **64(5)**, 2229-2246 (1976).
46. Forvefle, L., Vercauteren, J., & Rutledge, D. N., Multivariate statistical analysis of two-dimensional NMR data to differentiate grapevine cultivars and clones. *Food Chemistry* **57(3)**, 441-450 (1996).
47. Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., ... & Querengesser, L., HMDB: the human metabolome database.. *Nucleic acids research* **35**, 521-526 (2007).
48. Cui, Q., Lewis, I. A., Hegeman, A. D., Anderson, M. E., Li, J., Schulte, C. F., ... & Markley, J. L., Metabolite identification via the madison metabolomics consortium database. *Nature Biotechnology* **26(2)**, 162-164 (2008).
49. Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., ... & Markley, J. L., BioMagResBank. *Nucleic acids research* **36**, 402-408 (2008).
50. Braunschweiler, L., & Ernst, R. R., Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *Journal of Magnetic Resonance* **53(3)**, 521-528 (1983).
51. Bodenhausen, G., & Ruben, D. J., Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* **69(1)**, 185-189 (1980).
52. Merrifield, C. A., Lewis, M., Claus, S. P., Beckonert, O. P., Dumas, M. E., Duncker, S., ... & Nicholson, J. K., A metabolic system-wide characterisation of the pig: a model for human physiology. *Molecular BioSystems* **7(9)**, 2577-2588 (2011).
53. Aue, W.P., Karhan, J., & Ernst, R.R., Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy. *Journal of Chemical Physics* **64**, 4226-4227 (1976).
54. Bharti, S. K., & Roy, R., Quantitative 1 H NMR spectroscopy. *TrAC Trends in Analytical Chemistry* **35**, 5-26 (2012).
55. Weljie, A. M., Newton, J., Mercier, P., Carlson, E., & Slupsky, C. M., Targeted profiling: quantitative analysis of 1H NMR metabolomics data. *Analytical Chemistry* **78(13)**, 4430-4442 (2006).
56. Akoka, S., Barantin, L., & Trierweiler, M., Concentration measurement by proton NMR using the ERETIC method. *Analytical Chemistry* **71(13)**, 2554-2557 (1999).
57. Mehr, K., John, B., Russell, D., & Avizonis, D., Electronic referencing techniques for quantitative NMR: pitfalls and how to avoid them using amplitude-corrected referencing through signal injection. *Analytical Chemistry* **80(21)**, 8320-8323 (2008).

58. Wider, G., & Dreier, L., Measuring protein concentrations by NMR spectroscopy. *Journal of the American Chemical Society* **128(8)**, 2571-2576 (2006).
59. Frydman, L., Scherf, T., & Lupulescu, A., The acquisition of multidimensional NMR spectra within a single scan. *Proceedings of the National Academy of Sciences* **99(25)**, 15858-15862 (2002).
60. Guennec, A. L., Giraudeau, P., & Caldarelli, S., Evaluation of fast 2D NMR for metabolomics. *Analytical Chemistry* **86(12)**, 5946-5954 (2014).
61. Sousa, S. A. A., Magalhaes, A., & Ferreira, M. M. C., Optimized bucketing for NMR spectra: Three case studies. *Chemometrics and Intelligent Laboratory Systems* **122**, 93-102 (2013).
62. De Meyer, T., Sinnaeve, D., Van Gasse, B., Tsiporkova, E., Rietzschel, E. R., De Buyzere, M. L., ... & Van Crielinge, W. Crielinge, NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. *Analytical Chemistry* **80(10)**, 3783-3790 (2008).
63. Piotto, M. M. F. M., Moussallieh, F. M., Dillmann, B., Imperiale, A., Neuville, A., Brigand, C., ... & Namer, I. J., Metabolic characterization of primary human colorectal cancers using high resolution magic angle spinning ¹H magnetic resonance spectroscopy. *Metabolomics* **5(3)**, 292-301 (2009).
64. Gil, A. M., Duarte, I. F., Delgado, I., Colquhoun, I. J., Casascelli, F., Humpfer, E., & Spraul, M., Study of the compositional changes of mango during ripening by use of nuclear magnetic resonance spectroscopy. *Journal of Agricultural and Food Chemistry* **48(5)**, 1524-1536 (2000).
65. Sobolev, A. P., Segre, A., & Lamanna, R., Proton high-field NMR study of tomato juice. *Magnetic Resonance in Chemistry* **41(4)**, 237-245 (2003).
66. Pérez, E. M. S., Iglesias, M. J., Ortiz, F. L., Pérez, I. S., & Galera, M. M., Study of the suitability of HRMAS NMR for metabolic profiling of tomatoes: Application to tissue differentiation and fruit ripening. *Food Chemistry* **122(3)**, 877-887 (2010).
67. Iglesias, M. J., García-López, J., Collados-Luján, J. F., López-Ortiz, F., Díaz, M., Toresano, F., & Camacho, F., Differential response to environmental and nutritional factors of high-quality tomato varieties. *Food Chemistry* **176**, 278-287 (2015).
68. Mallamace, D., Corsaro, C., Salvo, A., Cicero, N., Macaluso, A., Giangrosso, G., ... & Dugo, G., A multivariate statistical analysis coming from the NMR metabolic profile of cherry tomatoes (The Sicilian Pachino case). *Physica A: Statistical Mechanics and its Applications* **401**, 112-117 (2014).
69. Ritota, M., Marini, F., Sequi, P., & Valentini, M., Metabolomic characterization of Italian sweet pepper (*Capsicum annum* L.) by means of HRMAS-NMR spectroscopy and multivariate analysis. *Journal of Agricultural and Food Chemistry* **58(17)**, 9675-9684 (2010).
70. Ritota, M., Casciani, L., Han, B. Z., Cozzolino, S., Leita, L., Sequi, P., & Valentini, M., Traceability of Italian garlic (*Allium sativum* L.) by means of HRMAS-NMR spectroscopy and multivariate data analysis. *Food Chemistry* **135(2)**, 684-693 (2012).

71. Vermathen, M., Marzorati, M., & Vermathen, P., Exploring high-resolution magic angle spinning (HR-MAS) NMR Spectroscopy for metabonomic analysis of apples. *CHIMIA International Journal for Chemistry* **66(10)**, 747-751 (2012).
72. Vermathen, M., Marzorati, M., Baumgartner, D., Good, C., & Vermathen, P., Investigation of different apple cultivars by high resolution magic angle spinning NMR. A feasibility study. *Journal of agricultural and food chemistry* **59(24)**, 12784-12793 (2011).
73. Shintu, L., Ziarelli, F., & Caldarelli, S., Is high-resolution magic angle spinning NMR a practical speciation tool for cheese samples? Parmigiano Reggiano as a case study. *Magnetic Resonance in Chemistry* **42(4)**, 396-401 (2004).
74. Shintu, L., & Caldarelli, S., High-resolution MAS NMR and chemometrics: characterization of the ripening of Parmigiano Reggiano cheese. *Journal of Agricultural and Food Chemistry* **53(10)**, 4026-4031 (2005).
75. Shintu, L., & Caldarelli, S., Toward the determination of the geographical origin of Emmental (er) cheese via high resolution MAS NMR: a preliminary investigation. *Journal of Agricultural and Food Chemistry* **54(12)**, 4148-4154 (2006).
76. Mazzei, P., & Piccolo, A., ¹H HRMAS-NMR metabolomic to assess quality and traceability of mozzarella cheese from Campania buffalo milk. *Food Chemistry* **132(3)**, 1620-1627 (2012).
77. Sacco, A., Neri Bolsi, I., Massini, R., Spraul, M., Humpfer, E., & Ghelli, S., Preliminary investigation on the characterization of durum wheat flours coming from some areas of south Italy by means of ¹H high-resolution magic angle spinning nuclear magnetic resonance. *Journal of Agricultural and Food Chemistry* **46(10)**, 4242-4249 (1998).
78. Brescia, M. A., Di Martino, G., Fares, C., Di Fonzo, N., Platani, C., Ghelli, S., ... & Sacco, A., Characterization of Italian durum wheat semolina by means of chemical analytical and spectroscopic determinations. *Cereal Chemistry* **79(2)**, 238-242 (2002).
79. Brescia, M. A., Sgaramella, A., Ghelli, S., & Sacco, A., ¹H HR-MAS NMR and isotopic investigation of bread and flour samples produced in southern Italy. *Journal of the Science of Food and Agriculture* **83(14)**, 1463-1468 (2003).
80. Larsen, F. H., Blennow, A., & Engelsen, S. B., Starch granule hydration-A MAS NMR investigation. *Food Biophysics* **3(1)**, 25-32 (2008).
81. Sacco, D., Brescia, M. A., Buccolieri, A., & Jambrenghi, A. C., Geographical origin and breed discrimination of Apulian lamb meat samples by means of analytical and spectroscopic determinations. *Meat Science* **71(3)**, 542-548 (2005).
82. Shintu, L., Caldarelli, S., & Franke, B. M., Pre-selection of potential molecular markers for the geographic origin of dried beef by HR-MAS NMR spectroscopy. *Meat Science* **76(4)**, 700-707 (2007).
83. Ritota, M., Casciani, L., Failla, S., & Valentini, M., HRMAS-NMR spectroscopy and multivariate analysis meat characterisation. *Meat Science* **92(4)**, 754-761 (2012).

84. Aursand, M., Gribbestad, I. S., & Martinez, I., Omega-3 Fatty Acid Content of Intact Muscle of Farmed Atlantic Salmon (*Salmo salar*) Examined by ¹H MAS NMR Spectroscopy. *Modern Magnetic Resonance*, **Part I**, 941-945 (2006).
85. Nestor, G., Bankefors, J., Schleichriem, C., Brännäs, E., Pickova, J., & Sandström, C., High-resolution ¹H magic angle spinning NMR spectroscopy of intact Arctic char (*Salvelinus alpinus*) muscle. Quantitative analysis of n-3 fatty acids, EPA and DHA. *Journal of Agricultural and Food Chemistry* **58(20)**, 10799-10803 (2010).
86. Castejón, D., Villa, P., Calvo, M. M., Santa-María, G., Herraiz, M., & Herrera, A., ¹H-HRMAS NMR study of smoked Atlantic salmon (*Salmo salar*). *Magnetic Resonance in Chemistry* **48(9)**, 693-703 (2010).
87. Bankefors, J., Kaszowska, M., Schleichriem, C., Pickova, J., Brännäs, E., Edebo, L., ... & Sandström, C., A comparison of the metabolic profile on intact tissue and extracts of muscle and liver of juvenile Atlantic salmon (*Salmo salar* L.)—Application to a short feeding study. *Food Chemistry* **129(4)**, 1397-1405 (2011).
88. Martin, G. J., & Martin, M. L., Deuterium labelling at the natural abundance level as studied by high field quantitative ²H NMR. *Tetrahedron letters* **22(36)**, 3525-3528 (1981).
89. Martin, G. J., & Martin, M. L., Determination of specific isotopic fractionation by natural abundance deuterium NMR—Application to the detection of wine chaptalization. *J. Chim. Phys. Phys.-Chim. Biol.* **80**, 293-297 (1983).
90. Martin, G. J., Benbernou, M., & Lantier, F., Application of site specific natural isotope fractionation (SNIF-NMR) of hydrogen to the characterization of European beers. *Journal of the Institute of Brewing* **91(4)**, 242-249 (1985).
91. Remaud, G. S., Martin, Y. L., Martin, G. G., & Martin, G. J., Detection of sophisticated adulterations of natural vanilla flavors and extracts: application of the SNIF-NMR method to vanillin and p-hydroxybenzaldehyde. *Journal of Agricultural and Food Chemistry* **45(3)**, 859-866 (1997).
92. Martin, G. G., Wood, R., & Martin, G. J., Detection of added beet sugar in concentrated and single strength fruit juices by deuterium nuclear magnetic resonance (SNIF-NMR method): collaborative study. *Journal of AOAC International* **79(4)**, 917-928 (1995).
93. Remaud, G. S., Martin, Y. L., Martin, G. G., Naulet, N., & Martin, G. J., Authentication of mustard oils by combined stable isotope analysis (SNIF-NMR and IRMS). *Journal of Agricultural and Food Chemistry* **45(5)**, 1844-1848 (1997).
94. Godelmann, R., Fang, F., Humpfer, E., Schütz, B., Bansbach, M., Schäfer, H., & Spraul, M., Targeted and nontargeted wine analysis by ¹H NMR spectroscopy combined with multivariate statistical analysis. Differentiation of important parameters: grape variety, geographical origin, year of vintage. *Journal of Agricultural and Food Chemistry* **61(23)**, 5610-5619 (2013).
95. Son, H. S., Kim, K. M., Van Den Berg, F., Hwang, G. S., Park, W. M., Lee, C. H., & Hong, Y. S., ¹H nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas. *Journal of Agricultural and Food Chemistry* **56(17)**, 8007-8016 (2008).

96. Monakhova, Y. B., Godelmann, R., Hermann, A., Kuballa, T., Cannet, C., Schäfer, H., ... & Rutledge, D. N., Synergistic effect of the simultaneous chemometric analysis of ^1H NMR spectroscopic and stable isotope (SNIF-NMR, ^{18}O , ^{13}C) data: Application to wine analysis. *Analytica Chimica Acta* **833**, 29-39 (2014).
97. J.E. Lee, B.J. Lee, J.O. Chung, J.A. Hwang, S.J. Lee, C.H. Lee, Y.S. Hong, Geographical and climatic dependencies of green tea (*Camellia sinensis*) metabolites: a ^1H NMR-based metabolomics study. *Journal of Agricultural and Food Chemistry* **58(19)**, 10582-10589 (2010).
98. Tarachiwin, L., Ute, K., Kobayashi, A., & Fukusaki, E., ^1H NMR based metabolic profiling in the evaluation of Japanese green tea quality. *Journal of Agricultural and Food Chemistry* **55(23)**, 9330-9336 (2007).
99. Charlton, A. J., Farrington, W. H., & Brereton, P., Application of ^1H NMR and multivariate statistics for screening complex mixtures: quality control and authenticity of instant coffee. *Journal of Agricultural and Food Chemistry* **50(11)**, 3098-3103 (2002).
100. Kang, J., Lee, S., Kang, S., Kwon, H. N., Park, J. H., Kwon, S. W., & Park, S., NMR-based metabolomics approach for the differentiation of ginseng (*Panax ginseng*) roots from different origins. *Archives of pharmacal research* **31(3)**, 330-336 (2008).
101. Aursand, M., & Grasdalen, H., Interpretation of the ^{13}C -NMR spectra of omega-3 fatty acids and lipid extracted from the white muscle of Atlantic salmon (*Salmo salar*). *Chemistry and Physics of Lipids* **62(3)**, 239-251 (1992).
102. Aursand, M., Standal, I. B., Prael, A., McEvoy, L., Irvine, J., & Axelson, D. E., ^{13}C NMR pattern recognition techniques for the classification of Atlantic salmon (*Salmo salar* L.) according to their wild, farmed, and geographical origin. *Journal of Agricultural and Food Chemistry* **57(9)**, 3444-3451 (2009).
103. Jung, Y., Lee, J., Kwon, J., Lee, K. S., Ryu, D. H., & Hwang, G. S., Discrimination of the geographical origin of beef by ^1H NMR-based metabolomics. *Journal of Agricultural and Food Chemistry* **58(19)**, 10458-10466 (2010).
104. Kowalski, B.R. Chemometrics: mathematics and statistics in chemistry (ed. ed. Kowalski, B.R.) (Springer Science & Business Media, 2013).
105. Marini, F. Chemometrics in food chemistry (ed. ed. Marini, F.) (Elsevier, Amsterdam, 2013).
106. Pearson, K., Principal components analysis. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **6(2)**, 559-572 (1901).
107. Hotelling, H., Analysis of a complex of statistical variables into principal components. *Journal of educational psychology* **24(6)**, 417-441 (1933).
108. Wold, H., Soft modeling by latent variables: the nonlinear iterative partial least squares approach. *Perspectives in Probability and Statistics, papers in honour of MS Bartlett*, 520-540 (1975).
109. Golub, G. H., & Reinsch, C., Singular value decomposition and least squares solutions. *Numerische Mathematik* **14(5)**, 403-420 (1970).

110. Wold, S., Pattern recognition by means of disjoint principal components models. *Pattern recognition* **8(3)**, 127-139 (1976).
111. Wold, S., & Sjöström, M., SIMCA: A method for analyzing Chemical Data in Terms of Similarity and Analogy. *Chemometrics: Theory and Application* **52**, 243-282 (1977).
112. Nomikos, P., & MacGregor, J. F., Multivariate SPC charts for monitoring batch processes. *Technometrics* **37(1)**, 41-59 (1995).
113. Mahalanobis, P.C., On the generalized distance in statistics. *Proceedings of the National Institute of Sciences (Calcutta)* **2**, 49-55 (1936).
114. Jackson, J. E., & Mudholkar, G. S., Control procedures for residuals associated with principal component analysis. *Technometrics* **21(3)**, 341-349 (1979).
115. Trygg, J., & Wold, S., Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics* **16(3)**, 119-128 (2002).
116. Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J., OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics* **20(8-10)**, 341-351 (2006).
117. Tang, C.G., An analysis of baseline distortion and offset in NMR spectra. *Journal of Magnetic Resonance* **109(2)**, 232-240 (1994).
118. Savorani, F., Tomasi, G., & Engelsen, S. B., icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* **202**, 190-202 (2010).
119. Dieterle, F., Ross, A., Schlotterbeck, G., & Senn, H., Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Analytical Chemistry* **78(13)**, 4281-4290 (2006).
120. Jolliffe, I. Principal Component Analysis (ed.) (Springer Verlag, 1986).
121. Kaiser, H.F., The application of electronic computers to factor analysis.. *Educational and psychological measurement* **20**, 141-151 (1960).
122. Cattell, R.B., The scree test for the number of factors. *Multivariate behavioral research* **1(2)**, 245-276 (1966).
123. Maas, W. E., Laukien, F. H., & Cory, D. G., Gradient, high resolution, magic angle sample spinning NMR. *Journal of the American Chemical Society* **118(51)**, 13085-13086 (1996).
124. Sodickson, A., & Cory, D. G., Shimming a high-resolution MAS probe. *Journal of Magnetic Resonance* **128(1)**, 87-91 (1997).
125. Piotto, M., Elbayed, K., Wieruszkeski, J. M., & Lippens, G., Practical aspects of shimming a high resolution magic angle spinning probe. *Journal of Magnetic Resonance* **173(1)**, 84-89 (2005).

126. Bligh, E. G., & Dyer, W. J., A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **37(8)**, 911-917 (1959).
127. Bodenhausen, G., Freeman, R., & Turner, D. L., Suppression of artifacts in two-dimensional J spectroscopy. *Journal of Magnetic Resonance* **27(3)**, 511-514 (1977).
128. Hoult, D. I., & Richards, R. E., Critical factors in the design of sensitive high resolution nuclear magnetic resonance spectrometers. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **344**, 311-340 (1975).
129. Patt, S. L., & Sykes, B. D., Water eliminated Fourier transform NMR spectroscopy. *The Journal of Chemical Physics* **56(6)**, 3182-3184 (1972).
130. Davis, D. G., & Bax, A., Assignment of complex proton NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. *Journal of the American Chemical Society* **107(9)**, 2820-2821 (1985).
131. Shaka, A. J., Lee, C. J., & Pines, A., Iterative schemes for bilinear operators; application to spin decoupling. *Journal of Magnetic Resonance* **77(2)**, 274-293 (1988).
132. Bax, A. D., & Davis, D. G., MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *Journal of Magnetic Resonance* **65(2)**, 355-360 (1985).
133. Kupče, Ě., Keifer, P. A., & Delepierre, M., Adiabatic TOCSY MAS in liquids. *Journal of Magnetic Resonance* **148(1)**, 115-120 (2001).
134. Peti, W., Griesinger, C., & Bermel, W., Adiabatic TOCSY for C, C and H, H J-transfer. *Journal of Biomolecular NMR* **18(3)**, 199-205 (2000).
135. Bendall, M. R., & Pegg, D. T., Uniform sample excitation with surface coils for in vivo spectroscopy by adiabatic rapid half passage. *Journal of Magnetic Resonance* **67(2)**, 376-381 (1986).
136. Hardy, C. J., Edelstein, W. A., & Vatis, D., Efficient adiabatic fast passage for NMR population inversion in the presence of radiofrequency field inhomogeneity and frequency offsets. *Journal of Magnetic Resonance* **66(3)**, 470-782 (1986).
137. Tenailleau, E., & Akoka, S., Adiabatic 1 H decoupling scheme for very accurate intensity measurements in 13 C NMR. *Journal of Magnetic Resonance* **185(1)**, 50-58 (2007).
138. Basus, V. J., Ellis, P. D., Hill, H. D., & Waugh, J. S., Utilization of chirp frequency modulation with 180-phase modulation for heteronuclear spin decoupling. *Journal of Magnetic Resonance* **35(1)**, 19-37 (1979).
139. States, D.J., Haberkorn, R.A., & Ruben, D.J., A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *Journal of Magnetic Resonance* **48(2)**, 286-292 (1982).
140. Marion, D., & Wüthrich, K., Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1 H-1 H spin-spin coupling constants in proteins. *Biochemical and biophysical research communications* **113**, 967-974 (1983).

141. Marion, D., Ikura, M., Tschudin, R., & Bax, A. D., Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *Journal of Magnetic Resonance* **85(2)**, 393-399 (1989).
142. Morris, G. A., & Freeman, R., Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* **101(3)**, 760-762 (1979).
143. Davis, A. L., Keeler, J., Laue, E. D., & Moskau, D., Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *Journal of Magnetic Resonance* **98(1)**, 207-216 (1992).
144. Maudsley, A.A., Wokaun, A., & Ernst, R.R., Coherence transfer echoes. *Chemical Physics Letters* **55(1)**, 9-14 (1978).
145. Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletsky, O., Glaser, S. J., ... & Griesinger, C., A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *Journal of Biomolecular NMR* **4(2)**, 301-306 (1994).
146. Woyewoda, A. D., Shaw, S. J., & Ke, P. J., Recommended laboratory methods for assesment of fish quality . *Canadian Technical Report of Fisheries and Aquatic Sciences* **1448**, 143 (1986).
147. Ocaño-Higuera, V. M., Maeda-Martínez, A. N., Marquez-Ríos, E., Canizales-Rodríguez, D. F., Castillo-Yáñez, F. J., Ruíz-Bustos, E., ... & Plascencia-Jatomea, M., Freshness assesment of ray fish stored in ice by biochemical, chemical and physical methods. *Food Chemistry* **125**, 49-54 (2011).
148. Civera, T., Turi, R. M., Parisi, E., & Fazio, G., Further investigations on total volatile basic nitrogen and trimethylamine in some Mediterranean teleosts during cold storage. *Sciences des aliments* **15(2)**, 179-186 (1995).
149. Karube, I., Matsuoka, H., Suzuki, S., Watanabe, E., & Toyama, K., Determination of Fish Freshness with an Enzyme Sensor System. *Journal of Agricultural and Food Chemistry* **32**, 314-319 (1984).
150. Saito, T., Arai, K. I., & Matsuyoshi, M., A new method for estimating the freshness of fish. *Bulletin of the Japanese Society of Scientific Fisheries* **24(9)**, 749-750 (1959).
151. Conway, E. J., & Byrne, A., An absorption apparatus for the micro-determination of certain volatile substances: The micro-determination of ammonia. *Biochemical Journal* **27(2)**, 419-429 (1933).
152. Antonacopoulos, N., & Vyncke, W., Determination of volatile basic nitrogen in fish: a third collaborative study by the West European Fish Technologists' Association (WEFTA). *Zeitschrift für Lebensmittel-Untersuchung und Forschung* **189(4)**, 309-316 (1989).
153. Wong, K., & Gill, T. A., Enzymatic determination of trimethylamine and its relationship to fish quality . *Journal of Food Science* **52(1)**, 1-3 (1987).
154. Kawai, M., Okiyama, A., & Ueda, Y., Taste enhancements between various amino acids and IMP. *Chemical Senses* **27(8)**, 739-745 (2002).

155. Ehira, S., & Uchiyama, H. Determination of fish freshness using the K value and comments on some other biochemical changes in relation to freshness. In: *Seafood Quality Determination* (ed. ed. D.E. Kramer, J.L.) (Amsterdam:Elsevier, 1987).
156. Okuma, H., & Watanabe, E., Flow system for fish freshness determination based on double multi-enzyme reactor electrodes. *Biosensors and Bioelectronics* **17(5)**, 367-372 (2002).
157. Itoh, D., Koyachi, E., Yokokawa, M., Murata, Y., Murata, M., & Suzuki, H., Microdevice for on-site fish freshness checking based on K-value measurement. *Analytical Chemistry* **85(22)**, 10962-10968 (2013).
158. Sitter, B., Krane, J., Gribbestad, I. S., Jørgensen, L., & Aursand, M., Quality Evaluation of Atlantic Halibut (*Hippoglossus hippoglossus* L) during Ice Storage Using ¹H NMR Spectroscopy. *Special Publication-Royal Society of Chemistry* **231**, 226-240 (1999).
159. Howell, N., Shavila, Y., Grootveld, M., & Williams, S., High-Resolution NMR and Magnetic Resonance Imaging (MRI) Studies on Fresh and Frozen Cod (*Gadus morhua*) and Haddock (*Melanogrammus aeglefinus*). *Journal of the Science of Food and Agriculture* **72(1)**, 49-56 (1996).
160. Erikson, U., Standal, I. B., Aursand, I. G., Veliyulin, E., & Aursand, M., Use of NMR in fish processing optimization: a review of recent progress. *Magnetic Resonance in Chemistry* **50(7)**, 471-480 (2012).
161. Chang, K. L. B., Chang, J., Shiau, C. Y., & Pan, B. S., Biochemical, microbiological, and sensory changes of sea bass (*Lateolabrax japonicus*) under partial freezing and refrigerated storage. *Journal of Agricultural and Food Science* **46(2)**, 682-686 (1998).
162. Cakli, S., Kilinc, B., Dincer, T. , & Tolasa S., Effects of gutting and ungutting on microbiological, chemical, and sensory properties of aquacultured sea bream (*Sparus aurata*) and sea bass (*Dicentrarchus labrax*) stored in ice. *Critical Reviews in Food Science and Nutrition* **46(7)**, 519-527 (2006).
163. Lougovois, V.P., Kyranas, E.R., & Kyrana, V.R., Comparison of selected methods of assessing freshness quality and remaining storage life of iced gilthead sea bream (*Sparus aurata*). *Food research International* **36(6)**, 551-560 (2003).
164. Grikorakis, K., Taylor, K.D.A., Alexis, M.N. Seasonal patterns of spoilage of ice-stored cultured gilthead sea bream (*Sparus aurata*). *Food Chemistry* **81(2)**, 263-268 (2003).
165. C. Alasalvar, K.D.A. Taylor, A. Oksüz, T. Garthwaite, M.N. Alexis, K. Grigorakis, Freshness assessment of cultured sea bream (*Sparus aurata*) by chemical, physical and sensory methods. *Food Chemistry* **72(1)**, 33-40 (2001).
166. Sallam, K.I., Antimicrobial and antioxidant effects of sodium acetate, sodium lactate, and sodium citrate in refrigerated sliced salmon. *Food Control* **18(5)**, 566-575 (2007).
167. Spinelli, J., & Kemp, B., Nucleotide Measurement, Rapid Measurement of Inosine Monophosphate and Total Adenosine Nucleotides in Fish Tissue. *Journal of Agricultural and Food Chemistry* **14(2)**, 176-178 (1966).
168. Marseno, D.W., Hori, K., & Miyazawa, K., Distribution of 5'-nucleotidase in muscle of some marine fishes. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **102(2)**, 247-253 (1992).

169. Kassemarn, B.O., Perez, B., Murray, J. , & Jones, N.R. Nucleotide degradation in the muscle of iced haddock (*Gadus aeglefinus*), lemon sole (*Pleuronectes microcephalus*), and plaice (*Pleuronectes platessa*). *Journal of Food Science* **28(1)**, 28-37 (1963).
170. Iwamoto, M., Yamanaka, H., Watabe, S., & Hashimoto, K., Effect of Storage Temperature on Rigor-Mortis and ATP Degradation in Plaice *Paralichthys olivaceus* Muscle. *Journal of Food Science* **52(6)**, 1514-1517 (1987).
171. Hattula, T., Kiesvaara, M. & Moran, M., Freshness evaluation in European whitefish (*Coregonus wartmanni*) during chill storage. *Journal of Food Science* **58(6)**, 1212-1215 (1993).
172. Ortiz, J., Vivanco, J.P., Quitral, V., Larrain, M.A., & Aubourg, S.P., Changes in freshness during frozen storage of farmed coho salmon: effect of replacement of synthetic antioxidants by natural ones in fish feeds. *North American Journal of Aquaculture* **74(2)**, 224-229 (2012).
173. Bertram, H. C., Karlsson, A. H., Rasmussen, M., Pedersen, O. D., Dønstrup, S., & Andersen, H. J., Origin of multiexponential T₂ relaxation in muscle myowater. *Journal of Agricultural and Food Chemistry* **49(6)**, 3092-3100 (2001).
174. Cole, W. C., Leblanc, A. D., & Jhingran, S. G., The origin of biexponential T₂ relaxation in muscle water. *Magnetic Resonance in medicine* **29(1)**, 19-24 (1993).
175. Yamada, T., ¹H-NMR spectroscopy of the intracellular water of resting and rigor frog skeletal muscle. *Mechanisms of Work Production and Work Absorption in Muscle II*, 145-155 (1998).
176. Lillford, P. J., Clark, A. H., & Jones, D. V., Distribution of water in heterogeneous food and model systems. *Water in Polymers* **127**, 177-195 (1980).
177. Site de la FAO <http://www.fao.org/docrep/003/v7180f/v7180f06.htm>.
178. Abe, H., Okuma, E., Rigor mortis progress of carp acclimated to different water temperatures. *Nippon Suisan Gakkaishi* **57**, 2095-2100 (1991).
179. Azam, K., Mackie, I. M., & Smith, J., Effect of stunning method on the time of onset, duration and resolution of rigor in rainbow trout (*Salmo gairdneri*) as measured by visual observation and analysis for lactic acid, nucleotide-degradation products and glycogen. , (1990).
180. Skjervold, P. O., Rørå, A. M. B., Fjæra, S. O., Vegusdal, A., Vorre, A., & Einen, O, Effects of pre-, in-, or post-rigor filleting of live chilled Atlantic salmon. *Aquaculture* **194(3)**, 315-326 (2001).
181. Martinez-Bisbal, M. C., Monleon, D., Assemat, O., Piotto, M., Piquer, J., Llacer, J. L., & Celda, B., Determination of metabolite concentrations in human brain tumour biopsy samples using HR-MAS and ERETIC measurements. *NMR in Biomedicine* **22(2)**, 199-206 (2009).
182. Billard, R., Sturgeons and Caviar. , (2002).
183. Williot, P., Sabeau, L., Gessner, J., Arlati, G., Bronzi, P., Gulyas, T., & Berni, P., Sturgeon farming in Western Europe: recent developments and perspectives. *Aquatic Living Resources* **14(6)**, 367-374 (2001).

184. Rezvani Gilkolaie, S., DNA PCR amplification for species diagnosis of caviar from Caspian Sea Sturgeon. *Journal of Agricultural Science and Technology* **4**, 51-61 (2010).
185. Birstein, V. J., Doukakis, P., Sorkin, B., & DeSalle, R., Population Aggregation Analysis of Three Caviar-Producing Species of Sturgeons and Implications for the Species Identification of Black Caviar. *Conservation Biology* **12(4)**, 766-775 (1998).
186. Rehbein, H., Gonzales-Sotelo, C., Perez-Martin, R., Quinteiro, J., Rey-Mendez, M., Pryde, S., ... & Santos, A. T., Differentiation of sturgeon caviar by single strand conformation polymorphism (PCR-SSCP) analysis. *Archiv Fur Lebensmittelhygiene* **50**, 13-16 (1999).
187. Czesny, S., Dabrowski, K., Christensen, J. E., Van Eenennaam, J., & Doroshov, S., Discrimination of wild and domestic origin of sturgeon ova based on lipids and fatty acid analysis. *Aquaculture* **189(1)**, 145-153 (2000).
188. Gong, Y., Huang, Y., Gao, L., Lu, J., Hu, Y., Xia, L., & Huang, H., Nutritional Composition of Caviar from Three Commercially Farmed Sturgeon Species in China. *Journal of Food and Nutrition Research* **1(5)**, 108-112 (2013).
189. Park, K. S., Kang, K. H., Bae, E. Y., Baek, K. A., Shin, M. H., Kim, D. U., ... & Im, J. S., General and biochemical composition of caviar from Sturgeon (*Acipenser ruthenus*) farmed in Korea. *International Food Research Journal* **22(2)**, 777-781 (2015).
190. Gussoni, M., Greco, F., Vezzoli, A., Paleari, M. A., Moretti, V. M., Beretta, G., ... & Zetta, L., Monitoring the effects of storage in caviar from farmed *Acipenser transmontanus* using chemical, SEM, and NMR methods. *Journal of Agricultural and Food Chemistry* **54(18)**, 6725-6732 (2006).
191. Mohamadi Monavar, H., Alimardani, R., Omid, M., & Wold, J. P., Prediction of species and freshness of caspian caviar during storage by front-face fluorescence spectroscopy. *Journal of Agricultural Technology* **8(5)**, 1555-1569 (2012).

Annexes

Annexe 1 :

A) Mode opératoire pour la préparation des échantillons :

Les boîtes de caviar doivent être conservées au réfrigérateur à 2°C et à l'abri de la lumière jusqu'à leur analyse.

1. Préparation du caviar

- 1) Peser environ 150 mg de caviar dans un eppendorf de 1,5 mL
- 2) Centrifuger à 13000 tours.min⁻¹ pendant 1 min
- 3) Broyer manuellement le caviar à l'aide d'une spatule puis placer le tube eppendorf dans de la glace pilée. Le tube y est maintenu pendant toute la durée de la préparation. Une fois les masses prélevées mettre la boîte de caviar au congélateur à -80°C afin de limiter sa dégradation.

2. Extraction des lipides

- 1) Ajouter 187,5 µl de CDCl₃
- 2) Ajouter 375 µl de MeOD
- 3) Agitation vigoureuse au vortex pendant 1 min
- 4) Ajouter 187,5 µl de CDCl₃
- 5) Ajouter 187,5 µl de D₂O
- 6) Agitation vigoureuse au vortex pendant 1 min
- 7) Laisser décanter au réfrigérateur pendant 15 min
- 8) Oter la majeure partie des résidus de caviar
- 9) Laisser décanter au réfrigérateur pendant 5 min
- 10) Centrifuger à 13000 tours.min⁻¹ pendant 1 min

3. Préparation du tube RMN

- 1) Prélever 150 L de phase aqueuse (phase supérieure) et les insérer dans un Eppendorf de 1,5 mL
- 2) Ajouter 450 L de Buffer D₂O tamponné à pH 7.30
- 3) Centrifuger à 13000 tours.min⁻¹ pendant 1 min

Prélever la quasi-totalité du contenu de l'Eppendorf en laissant 20 à 30 L du fond de celui-ci, pouvant contenir du chloroforme et les insérer dans un tube RMN)

B) Paramètres d'acquisition RMN

Séquence bruker *noesygppr1d* :

Paramètres d'acquisition ¹ H NOESYGPPR1D	
Caviar	
Echantillonnage	65536
Fenêtre spectrale (ppm)	20,5643
Temps d'acquisition (s)	4
Gain du receveur	90,5
Délai de relaxation (s)	4
Temps d'écho fixe (μs)	200
Temps de mélange (ms)	10
Nombre de scans fictifs	8
Nombre de scans	512
Durée impulsion ¹ H 90° (μs)	10,5
Forme gradient	Sinusoïdale 100 points
Force des gradients	GP1 : 50% GP2:-10%
Champ RF présaturation (Hz)	50

Les points C) et D) ne sont à réaliser que dans le cas d'un modèle OPLS-DA construit à l'aide des données réalignées.

C) Traitement des données en vue de leur réalignement

- 1) Ajuster la phase de chaque spectre RMN (seulement à l'aide de l'ordre 0)
- 2) Appliquer manuellement une correction de la ligne de base
- 3) Appliquer alors un SR de 0 Hz

D) Réalignement des données sous Matlab

- 1) Placer les spectres à réaligner dans le dossier contenant les spectres de référence de la société Sturgeon
- 2) Modifier dans la fonction Matlab (MathWorks) le chemin jusqu'au répertoire si celui-ci est différent de celui spécifié dans la fonction
- 3) Choisir les différentes options (alignement, affichage des graphes de vérification avant et après alignement, création d'un nouveau fichier contenant les données réalignées) dans la fonction Matlab

(*true* ou *false*). Il est également possible de définir de nouveaux intervalles si certains pics sont mal réalignés. Par défaut les spectres réalignés sont stockés en PROCNO 2000 (modifiable).

E) Création des « buckets »

- 1) Une fois réalignés les spectres (réalignés ou non) sont alors référencés sur le pic de gauche du doublet de l'acide lactique à 1,3376 ppm
- 2) Les « buckets » de 0,01255 ppm peuvent être générés soit par Matlab (MathWorks) soit par Amix (Bruker) en prenant soin d'exclure les zones du spectre correspondant au signal des différents solvants eau (5,08 à 4,7325 ppm), chloroforme (7,72 à 7,69 ppm) et méthanol (3,3525 à 3,298 ppm) ainsi que du TSP (0,6 à 0 ppm). Les données doivent être normalisées par rapport à l'intensité totale et centrées. La réduction des variables est réalisée sous le logiciel SIMCA-P (Umetrix).
- 2) Les données sont alors exportées sous la forme d'un fichier texte (.txt), réduites sous SIMCA-P et sont alors utilisables en vue d'une analyse OPLS-DA.

Annexe 2 :

Rapid assessment of fish freshness and quality by ^1H HR-MAS NMR Spectroscopy

Annexe 3:

Metabolic characterization of caviar specimens by ^1H NMR Spectroscopy: Towards caviar authenticity and integrity

Résumé :

La majorité des méthodes d'analyse et de contrôle actuelles dans l'agroalimentaire sont basées sur une approche ciblée, c'est-à-dire avec une définition en amont des contaminants recherchés, et présentent ainsi le risque de ne pas détecter certaines fraudes ou sources potentielles de contaminations de produits authentiques. C'est autour de cette problématique que le projet Agrifood GPS (Global Protection System), dont fait partie cette thèse, a été initié. Celui-ci a pour objectif principal la mise en place de nouvelles méthodes analytiques de criblage (non-ciblées) afin de garantir l'intégrité des produits analysés.

Cette thèse regroupe ainsi les différents résultats obtenus sur des matrices semi-solides (poisson principalement), par RMN Haute Résolution en Rotation à l'Angle Magique (HR-MAS), et sur des extraits de caviar par RMN liquide haute résolution (HR). Ce manuscrit présente, tout d'abord, une méthode de détermination rapide de la fraîcheur et la qualité du poisson basée sur la mesure de deux indicateurs chimiques (le TMA-N et la valeur K) ainsi que les résultats portant sur l'évaluation de la texture du poisson à travers l'étude du temps de relaxation transversale (T_2) de l'eau contenue dans les tissus musculaires, ces deux études étant réalisées par RMN HR-MAS.

Puis, les résultats des travaux réalisés sur la détermination de l'origine géographique du caviar à l'aide du profil métabolique enregistré par spectroscopie RMN liquide haute résolution et d'analyses statistiques multivariées, dans le cadre de la démarche d'obtention d'une IGP (Indication Géographique Protégée) des producteurs de l'Aquitaine, et sur l'étude de la dégradation du caviar au cours de sa conservation au réfrigérateur sont alors présentés.

Mots clés : RMN HR-MAS, RMN HR, criblage, authenticité, fraîcheur, texture, origine, poisson, caviar

Abstract :

Most of the current analytical and quality control methods in the food industry are based on a targeted approach, with an upstream definition of the intended contaminants, and may fail to detect some frauds or contaminations of genuine products. It is around this issue that the Agrifood GPS (Global Protection System), of which this thesis is part of, has been initiated. This project aims at developing new holistic analytical methods (untargeted) in order to ensure the integrity of the foodstuff analyzed.

This thesis manuscript gathers the results obtained on semi-solid foodstuffs (mainly fish), by High Resolution Magic Angle Spinning NMR, and on caviar extracts by high resolution liquid-state NMR (HR NMR). First of all, it presents a rapid method to evaluate fish freshness and quality based on the determination of two chemical indicators (the TMA-N and the K-value) and the results of a fish texture study through the measurement of the transverse relaxation time (T_2) of water in muscle tissues, both by HR-MAS NMR spectroscopy.

Thereafter, are presented the results of the work carried out on the determination of the geographical origin of caviar using the metabolic profile acquired by liquid-state NMR spectroscopy and multivariate statistical analysis in the context of the PGI (Protected Geographical Indication) status for the Aquitaine producers, and on the degradation study of caviar during its storage in a fridge.

Keywords : HR-MAS NMR, HR NMR, profiling, authenticity, freshness, texture, origin, fish, caviar

