Université Pierre et Marie Curie

École doctorale de Chimie Physique et de Chimie Analytique de Paris-Centre

CNRS USR 3149 / Spectrométrie de Masse Biologique et Protéomique

Analyse protéomique ciblée à haute résolution : un outil puissant pour le diagnostic clinique à partir de prélèvements de tissus amyloïdes bruts

Présentée par

Sophie LIUU

Thèse de doctorat en Chimie Analytique

Dirigée par Dr. Joëlle VINH

Présentée et soutenue publiquement le 24/11/2014

Devant un jury composé de :

Dr. Jean ARMENGAUD, CEA Marcoule	Rapporteur
Dr. Virginie REDEKER, Inserm	Rapporteur
Dr. Gérard BOLBACH, CNRS, UPMC	Examinateur
Pr. Jean-Michel CAMADRO, CNRS, Institut Jacques Monod	Examinateur
Pr. Gilles GRATEAU, Inserm, UPMC, Hôpital Tenon	Examinateur

Except where otherwise noted, this work is licensed under
http://creativecommons.org/licenses/by-nc-nd/3.0/

À mes parents,

À ma famille.

Choisissez un travail que vous aimez et vous n'aurez pas à travailler un seul jour de votre vie.

(Proverbe chinois)

Remerciements

Je tiens tout d'abord à remercier Joëlle Vinh pour son accueil au sein du laboratoire de Spectrométrie de Masse Biologique et Protéomique (SMBP) de l'École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI ParisTech). Merci pour ta confiance, ton écoute et ta présence pendant ces trois années de thèse.

Mes remerciements vont également aux membres du jury qui m'ont fait le plaisir d'accepter d'évaluer ce travail de thèse : Jean Armengaud, Virginie Redeker, Gérard Bolbach, Jean-Michel Camadro et Gilles Grateau.

Je suis très reconnaissante envers le personnel du service d'anatomie pathologique de l'hôpital Tenon à Paris qui a initié ce projet et fourni tous les échantillons. Je remercie également l'Association Française contre l'Amylose d'avoir participé au financement d'une partie de ce projet.

Mes sincères remerciements à toute l'équipe SMBP. À commencer par Emmanuelle Demey, pour m'avoir tant aidé dans cette aventure du début à la fin et avoir accepté que je participe à ce projet. Merci mille fois pour tous ces « Proxeons », confitures et brioches vendéennes. Merci pour ta patience. À Yann Verdier pour ses bonnes idées et sans qui toute l'équipe mourrait de faim. Toujours les bons mots, les bonnes remarques et les bonnes blagues, sacré Yann ! Un grand merci à Iman Haddad pour toutes les réponses à mes questions, les macros, les dépannages informatiques, les ordres de missions, les commandes et pour ses superbes confitures aux mirabelles. Merci à Giovanni Chiappetta pour les discussions, la découverte de Naples et ses mets culinaires. Merci à Chiara Giangrande pour ces belles discussions et les sfogliatella ! Je retournerai à Naples, rien que pour en manger ! Merci à toi, Sega Ndiaye, de m'avoir tant aidé au début de cette thèse. Tu vois, j'ai réussi à tenir le cap jusqu'à là fin ! La nano ne m'a pas tuER ! On est devenu amies, même. Je laisse une trace de mon passage sur ma chaîne préférée : la RSLC Sophie ! Prenez soin d'elle, elle vous le rendra bien ! Merci aussi à Anne-Marie Hesse pour son aide avec les logiciels. Enfin, merci à mon « petit » thésard adoré, Shakir Shakir qui m'a beaucoup aidé, et supporté, surtout ces derniers temps ! Je te souhaite bon courage pour ta dernière année de thèse !

Je tiens également à remercier les personnes qui m'ont soutenue, accordé leur confiance et suivi durant de ce long parcours : mes professeurs d'IUT et de la faculté d'Orsay, mes professeurs de l'ECPM de Strasbourg et mes anciens encadrants et collègues de stages

français et anglais. Merci d'avoir contribué à mon intérêt pour ce domaine magnifique qui est la protéomique.

Les remerciements ne pourraient pas être considérés comme tels sans citer les personnes grâce à qui je suis arrivée là et, ai pu tenir ces longues années d'étude : mes parents, mes sœurs et mon frère et tous mes amis.

Merci aussi à toi, Etienne, qui m'a soutenue de près et de loin dans cette longue aventure. Je te promets, j'essayerai d'être plus disponible à présent !

Résumé

Les amyloses sont un ensemble de maladies rares. Elles se manifestent par des agrégats extracellulaires de protéines insolubles mal repliées. Trente types d'amyloses sont répertoriés, systémiques ou localisés, et qui se différencient par la protéine amyloïdogène associée [1]. Leur identification repose essentiellement sur l'analyse immunohistochimique après un prélèvement sur le tissu malade. Cette technique, nécessitant des anticorps spécifiques, permet de mettre en évidence jusqu'à une dizaine de protéines, mais son interprétation reste fortement dépendante du pathologiste et du matériel employé, et peut donc être ambigüe. Récemment de nouvelles techniques ont fait leur apparition, combinant la microdissection par capture laser (LCM) et la spectrométrie de masse [2].

Notre étude propose d'évoluer vers une approche protéomique ciblée à partir de tissus bruts traités par ultrasons, en minimisant le recours au LCM. Après la phase d'optimisation du système nano-chromatographique (nanoLC), nous montrons que les biopsies prélevées chez des patients atteints de différents types d'amyloses peuvent être protéolysées par la trypsine sous ultrasons [3]. Cette technique permet d'augmenter l'efficacité et réduire le temps de la digestion (90s au lieu de 15h conventionnellement), tout en minimisant la quantité de matériel biologique nécessaire.

Le contenu peptidique de ces échantillons protéolytiques est ensuite caractérisé au niveau moléculaire par nanoLC couplée à la spectrométrie de masse tandem (MS/MS) à haute résolution de type nanoESI-LIT/FTICR [4]. Cette première étape protéomique pour le diagnostic clinique des patients présentant une amylose dans différents organes/tissus (reins, poumons, glandes salivaires, testicule, humeur vitrée, rate), nous a permis de mettre au point une méthode de criblage dédiée afin d'implémenter notre approche dans les départements cliniques français. Pour cela, deux méthodes qui montrent des résultats très encourageants, ont été évaluées : la SRM (*selected reaction monitoring*) sur un spectromètre de masse à basse résolution de type triple-quadripolaire, disponible dans certains laboratoires cliniques, et la PRM (*parallel reaction monitoring*) accessible sur les instruments de dernière génération à haute résolution de type quadripôle-Orbitrap, sur lesquels la mise en œuvre de méthodes ciblées est simplifiée [5].

En conclusion, nous avons développé un protocole spécifique qui permet de discriminer rapidement les cas amyloïdes des autres pathologies sans a priori puis de typer les amyloses détectées.

Références:
[1]. JD. Sipe et al., Amyloid. 2012;19(4):167-70.
[2]. S. Sethi et al., Clin. J. Am. Soc. Nephrol., 2010; 5(12): 2180-7.
[3]. L. Fernandez, et al JIOMICS, 2011, 1(1): 144-150.
[4]. S.Liuu et al. Proceedings of the XIIIth International Symposium on Amyloidosis, 2012 May 6-10; Groningen, The Netherlands, GUARD, 201p. ISBN : 978-90-821593-0-1.
[5]. S. Liuu et al., soumis.

Table générale des matières

PUBLICATIONS **17**

LISTE DES FIGURES **19**

LISTE DES TABLEAUX **25**

LISTE DES ABRÉVIATIONS **27**

INTRODUCTION GÉNÉRALE **29**

TRAVAIL REALISE DANS CETTE THESE **30**

PARTIE I : CONTEXTE SCIENTIFIQUE ET TECHNOLOGIQUE **33**

CHAPITRE 1 : LES AMYLOSES **35**

I. GENERALITES **35**

I.1 Historique 35

I.2 Pathogenèse 36

II. CLASSIFICATION **37**

II.1 Composants communs 38

II.1.1 Les glycosaminoglycanes 38

II.1.2 Composant P amyloïde (Serum Amyloid P component) 38

II.1.3 Apolipoprotéine E (ApoE) 39

II.2 Principales protéines amyloïdogènes 39

II.2.1 Immunoglobulines (Ig) 39

II.2.2 Sérum amyloïde A (SAA) 40

II.2.3 Transthyrétine (TTR) 40

II.2.4 Apolipoprotéine A 1 (ApoA1) 41

III. MOYENS DE DIAGNOSTIC DES AMYLOSES **41**

III.1 Histologie 42

III.1.1 Coloration au rouge Congo ... *42*
III.1.2 Immunohistochimie ... *43*
III.2 Spectrométrie de masse ... 44
III.2.1 Digestion sous ultrasons ... *44*

CHAPITRE 2 : IDENTIFICATION ET QUANTIFICATION DES PROTÉINES 47

I. STRATEGIE D'ANALYSE DES PROTEINES ... 47
I.1 Apport de la spectrométrie de masse ... 47
I.2 Approche descendante (Top-Down) ... 48
I.3 Approche ascendante (Bottom-up) ... 49

II. STRATEGIE ANALYTIQUE EN MODE BOTTOM-UP ... 50
II.1 Identification des protéines par couplage chromatographique ... 50
II.2 Recherche dans des banques de données ... 52
II.3 Validation des données ... 53
II.3.1 Validation par le score et le nombre de peptides ... *53*
II.3.2 Validation par le taux de faux positifs (FDR) ... *54*

III. QUANTIFICATION DES PROTEINES PAR SPECTROMETRIE DE MASSE ... 55
III.1 Quantification relative par marquage isotopique ... 56
III.1.1 Marquage métabolique ... *56*
III.1.2 Marquage chimique ou enzymatique ... *57*
III.2 Quantification absolue par marquage isotopique ... 59
III.2.1 Absolute quantification of proteins (AQUA) ... *61*
III.2.2 Quantification concatamer (QconCAT) ... *61*
III.2.3 Protein standard absolute quantification(PSAQ) ... *61*
III.3 Quantification relative sans marquage (label-free) ... 62
III.3.1 Décompte des spectres MS/MS (spectral count) ... *63*
III.3.2 Mesure d'aire sous le pic (AUC) ... *64*
III.4 Conclusions ... 65

CHAPITRE 3 : LA SPECTROMÉTRIE DE MASSE ... 69

I. HISTORIQUE ... 69

II. PRINCIPE 70
II.1 Source d'ionisation: NanoElectrospray (NanoESI) 70
II.1.1 Des capillaires NanoESI aux émetteurs NanoESI sur puce 71
II.2 Les analyseurs 73
II.2.1 Résolution 73
II.2.2 Précision de masse 74
II.2.3 Gamme de masse 76
II.2.4 Vitesse d'acquisition 76
II.2.5 Gamme dynamique 76

III. LA SPECTROMETRIE DE MASSE TANDEM (MS/MS) 77
III.1 La fragmentation générale des peptides 77
III.2 Dissociation induite par collision à basse énergie (low-energy CID) 79
III.3 Dissociation induite par collision à plus haute énergie (HCD) 80

IV. LE COUPLAGE AVEC LA CHROMATOGRAPHIE 82
IV.1 Généralités 82
IV.2 La nano-chromatographie (nanoLC) 83

V. CONFIGURATIONS INSTRUMENTALES 84
V.1 LIT-FTICR 84
V.1.1 Généralités 84
V.1.2 Data-dependent acquisition (DDA) 85
V.2 Triple-Quadripôle (QqQ) 85
V.2.1 Généralités 85
V.2.2 Mode de balayage 86
V.2.2.1 Selected reaction monitoring (SRM) 86
V.2.2.2 Time-scheduled SRM (tSRM) 88
V.2.2.3 Intelligent SRM (iSRM) 89
V.3 QqOrbitrap 89
V.3.1 Généralités 89
V.3.2 Modes d'acquisition 90
V.3.2.1 Data-dependent acquisition (DDA) 90
V.3.2.2 Parallel reaction monitoring (PRM) 91
V.3.2.3 Multiplexage PRM 92
V.4 QqOrbitrap HF (Q Exactive HF) 93

VI. MEGA DONNEES (BIG DATA) 95

PARTIE II : RÉSULTATS ET DISCUSSIONS 97

CHAPITRE 4 : ANALYSE PROTÉOMIQUE GLOBALE D'AGRÉGATS DE PROTÉINES INSOLUBLES : APPLICATION AUX AMYLOSES 99

I. INTRODUCTION 99

II. PASSER DE L'AGREGAT INSOLUBLE A LA SOLUTION DE PEPTIDES PROTEOLYTIQUES 100

II.1 Les échantillons 100

II.2 Traitement des échantillons 102

III. DU MELANGE PROTEOLYTIQUE A LA SEQUENCE DES PEPTIDES 103

III.1 Optimisations chromatographiques 103

III.1.1 Optimisation du gradient d'élution 104

III.1.2 Longueur de la colonne 105

III.1.3 Conclusion 107

III.2 Conditions nanoLC-MS/MS finale 107

IV. TRAITEMENT DES DONNEES 108

IV.1 Recherche dans les banques de données 108

IV.1.1 Critères de recherche 108

IV.1.2 Critères de validation 108

IV.2 Quantification des protéines 109

V. RESULTATS 110

V.1 Cas des reins 111

V.1.1 Témoins 111

V.1.2 Patients amyloses 112

V.1.3 Intérêt de la microdissection 113

V.2 Cas de la rate 114

V.3 Cas des bronches/poumons 114

V.3.1 Témoins 114

V.3.2 Patient amylose 115

V.3.3 Patients non amyloses *115*
V.4 Conclusion 116
VI. DATA MINING: VISUALISATION GRAPHIQUE GLOBALE 116
VI.1 Cartographie heatmap sous un tableur 117
VI.1.1 Traitement des données *117*
VI.1.2 Résultats *117*
VI.2 Visualisation plus automatisée avec MaxQuant et Perseus 121
VI.2.1 Recherche dans les banques de données *121*
VI.2.1.1 Paramètres expérimentaux 121
VI.2.2 Résultats de la classification par regroupement hiérarchique *122*
VII. CONCLUSION 125

CHAPITRE 5 : ANALYSE PROTÉOMIQUE CIBLÉE EN MODE SRM : APPLICATION AUX AMYLOSES 127
I. L'INTERET DES APPROCHES CIBLEES 127
II. ANALYSE CIBLEE EN MODE SINGLE REACTION MONITORING (SRM) 128
II.1 Stratégie expérimentale pour l'analyse tSRM 129
II.1.1 Conditions nanoLC-MS/MS *129*
II.1.2 Choix des peptides cibles *129*
II.1.2.1 Acquisition de nouvelles données 129
II.1.2.2 Critères de sélection des peptides 130
II.1.2.3 Validation des peptides expérimentaux 134
II.1.2.4 Ajout des peptides des banques de données 136
II.1.3 Choix des transitions *136*
II.1.4 Vérification du nombre de transitions sélectionnées *137*
II.1.5 Choix des énergies de collision *138*
II.2 Traitements des données 138
II.2.1 Quantification des peptides *138*
II.2.2 Quantification des protéines *138*
II.3 Interprétation des données 139
III. ANALYSES CIBLEES EN MODE TSRM ET TPRM, ANALYSE SHOTGUN 140
III.1 Méthode tPRM 140

III.1.1 Vérification du nombre de peptides sélectionnés ... *140*
III.1.2 Paramètres d'acquisition ... *141*
III.1.3 Traitement des données ... *141*
III.1.4 Résultats ... *142*
III.2 Comparaison sous forme de cartographie heatmap ... 143
III.2.1 Traitement des données ... *143*
III.2.1.1 Quantification des peptides ... 143
III.2.1.2 Quantification des protéines ... 143
III.2.2 Résultats ... *144*

IV. OPTIMISATION DE LA METHODE TSRM ... 147
IV.1 Optimisation des transitions ... 147
IV.1.1 Méthode ... *147*
IV.1.2 Résultats ... *148*
IV.2 Optimisation des énergies de collision ... 149
IV.2.1 Méthode ... *149*
IV.2.2 Résultats ... *150*
IV.3 Comparaison tSRM /tPRM à leur optimum ... 151
IV.3.1 Méthode tSRM ... *151*
IV.3.2 Méthode PRM ... *151*
IV.3.3 Comparaison des résultats ... *152*

V. CONCLUSION ... 154

CHAPITRE 6 : ANALYSE PROTÉOMIQUE CIBLÉE EN MODE PRM : APPLICATION AUX BIOPSIES AMYLOÏDES ... 157

I. ANALYSE CIBLEE A HAUTE RESOLUTION : VERS UNE APPLICATION CLINIQUE ? 157

II. STRATEGIE ANALYTIQUE POUR UN DIAGNOSTIQUE ROBUSTE ... 158
II.1 Choix des peptides ... 158
II.1.1 Caractéristiques de rétention chromatographiques ... *158*
II.1.2 Critères de choix ... *160*
II.1.3 Validation des peptides expérimentaux ... *161*
II.1.4 Ajout des peptides des banques de données ... *161*
II.2 Optimisation des paramètres d'acquisitions ... 162

II.2.1 Vérification du nombre de peptides sélectionnés *162*

II.2.2 Choix des transitions *163*

III. RESULTATS **164**

III.1 Quantification des transitions par Pinpoint 164

III.1.1 Traitement des données *164*

III.1.2 Data mining *164*

III.1.2.1 Etude qualitative et cartographie heatmap 164

III.1.2.2 Analyse en composantes principales 168

III.1.2.3 Analyse statistique et cartographie heatmap sous R 171

III.1.3 Interprétation générale des données *172*

III.2 Quantification des transitions par Skyline 174

III.2.1 Traitement des données *174*

III.2.2 Data mining *174*

III.2.2.1 Cartographie heatmap sous tableur 174

III.2.2.2 Analyse en composantes principales 178

III.2.2.3 Cartographie heatmap sous R 180

IV. CONCLUSION : PROPOSITION D'UN PROTOCOLE **183**

CHAPITRE 7 : COMPARAISON DES ANALYSES SUR DEUX APPAREILS DE DIFFÉRENTE GÉNÉRATION **185**

I. INTRODUCTION **185**

II. ANALYSE SHOTGUN **186**

II.1 Méthodologie 186

II.2 Résultats 187

III. ANALYSE CIBLEE EN MODE PRM **190**

III.1 Méthodologie 191

III.1.1 Choix des peptides *191*

III.1.2 Méthodes tPRM *191*

III.2 Traitement des données 192

III.3 Résultats 193

IV. CONCLUSION **197**

CONCLUSION GÉNÉRALE 199

ANNEXES 205

ANNEXE 1: PRINCIPE DE L'ELECTROSPRAY 207

ANNEXE 2 : FRAGMENTATIONS DES PEPTIDES 215

ANNEXE 3 : MONTAGE DU SYSTEME NANO-CHROMATOGRAPHIQUE 217

ANNEXE 4 : PRINCIPE DES ANALYSEURS QUADRIPOLAIRES 221

ANNEXE 5 : PRINCIPE DES ANALYSEURS FTMS 229

ANNEXE 6 : TABLEAU DES TRANSITIONS CHOISIES POUR LES EXPERIENCES TSRM. 239

ANNEXE 7 : TABLEAU DES TRANSITIONS CHOISIES POUR LES EXPERIENCES TPRM. 251

ANNEXE 8 : TABLEAU DES TRANSITIONS CHOISIES POUR LES EXPERIENCES TPRM POUR LA COMPARAISON DES INSTRUMENTS Q EXACTIVE ET Q EXACTIVE HF 256

RÉFÉRENCES 257

FICHE ENVOI DES ÉCHANTILLONS POUR UNE ANALYSE PAR SPECTROMÉTRIE DE MASSE 271

Publications

I. Articles scientifiques

1. **Liuu S.**, Demey E., Durighello E., Grateau G., Vinh J. *Subtyping of amyloidosis by direct proteomic analysis of biopsy samples* (soumis)
2. **Liuu S.**, Demey E., Durighello E., Colombat M., Grateau G., Vinh J. *Subtyping of amyloidosis by direct proteomic analysis of fixed biopsy samples.* Proceedings of the XIIIth International Symposium on Amyloidosis, 2012 May 6-10; Groningen, the Netherlands, GUARD, 201p. . ISBN 978-90-821593-0-1.
3. Fukuyama H., Ndiaye S., Hoffmann J., Rossier J., **Liuu S.**, Vinh J., Verdier Y. *On-bead tryptic proteolysis: an attractive procedure for LC-MS/MS analysis of the Drosophila caspase 8 protein complex during immune response against bacteria.* Journal of Proteomics. 2012 Aug 3;75 (15):4610-9.

II. Conférences orales

1. **Liuu S.**, Demey E., Durighello E., Grateau G., Vinh J. *Targeted analysis in PRM mode towards clinical application: Optimization of experimental setup and data processing for accurate amyloid diagnosis*. Congrès français de Spectrométrie de Masse et d'Analyse Protéomique, Lyon, France. 30 juin - 2 juillet 2014.
2. **Liuu S**., Demey E., Durighello E., Grateau G., Vinh J. *Typage direct de biopsies amyloïdes par analyse protéomique ciblée.* Journée de l'école doctorale Interdisciplinaire pour le vivant, Paris, France. 6 décembre 2013.
3. **Liuu S**., Demey E., Durighello E., Grateau G., Vinh J. *Typage direct de biopsies amyloïdes par analyse protéomique.* 7th annual conference oh the European Proteomics Association et Congrès français de Spectrométrie de Masse et d'Analyse Protéomique, Club jeune, Saint-Malo, France. 14 - 17 octobre 2013.
4. **Liuu S**., Demey E., Durighello E., Grateau G., Vinh J. *Typage direct de biopsies amyloïdes par analyse protéomique.* Congrès biennal de l'Association Francophone des Sciences Séparatives, Paris, France. 4-7 juin 2013.
5. **Liuu S**., Demey E., Durighello E., Colombat M., Grateau G., Vinh J. *Typage direct de biopsies amyloïdes par analyse protéomique.* Congrès français de Spectrométrie de Masse et d'Analyse Protéomique, Club jeune, Avignon, France. 18-22 septembre 2011.

III. Posters

1. **Liuu S.**, Demey E., Durighello E., Grateau G., Vinh J. *Targeted analysis in PRM mode towards clinical application: Optimization of experimental setup and data processing for accurate amyloid diagnosis*. 13th Annual World Congress of the Human Proteome Organization, Madrid, Spain, 5-8 October 2014.
2. **Liuu S.**, Demey E., Durighello E., Grateau G., Vinh J. *Subtyping of amyloidosis by direct targeted proteomics analysis of biopsy samples*. 7th annual conference oh the European Proteomics Association et Congrès français de Spectrométrie de Masse et d'Analyse Protéomique, Saint-Malo, France. 14 - 17 October 2013.
3. **Liuu S.**, De Simone V., Califano D. Chiappetta G., Chiappetta G., Vinh J. Nouvelles cibles thérapeutiques et diagnostiques du Cancer Anaplasique de la Thyroïde (CAT). 29èmes Journées Françaises de Spectrométrie de Masse, Orléans, France. 17-20 septembre 2012.
4. Demey E., **Liuu S.**, Durighello E., Colombat M., Grateau G., Vinh J. *Typage direct de biopsies amyloïdes par analyse protéomique.* Journées Françaises de Spectrométrie de Masse, Orléans, France. 17-20 septembre 2012.
5. **Liuu S.**, Demey E., Durighello E., Colombat M., Grateau G., Vinh J. *Subtyping of amyloidosis by direct proteomic analysis of fixed biopsy samples.* XIIIth International Symposium on Amyloidosis, Groningen, the Netherlands. 6-10 Mai 2012. **Sélectionné pour une conférence orale.**
6. Demey E., **Liuu S.,** Durighello E., Colombat M., Grateau G., Vinh J. *Subtyping of amyloidosis by direct proteomic analysis of fixed biopsy samples.* 10th Annual World Congress of the Human Proteome Organization, Geneva, Switzerland, 4–7 September 2011.

Liste des figures

Figure 1 : Pathogenèse des amyloses 36
Figure 2 : Stratégie histochimique 42
Figure 3 : Molécule rouge Congo 43
Figure 4 : Coloration au rouge Congo d'un dépôt amyloïde artériolaire 43
Figure 5 : schéma de principe de l'immunohistochimie 44
Figure 6 : Stratégie *top-down* et *bottom-up* 47
Figure 7 : L'approche *top-down* fournit des informations sur les phénotypes 48
Figure 8 : Schéma de la stratégie *bottom-up*. 50
Figure 9 : Interprétation des données MS/MS 52
Figure 10 : Procédures classiques de quantification en spectrométrie de masse. 55
Figure 11 : Protocole SILAC. 56
Figure 12 : Stratégie de quantification par dilution isotopique. 60
Figure 13 : Quantification absolue des protéines. 60
Figure 14 : Schéma des deux principales approches en quantification sans marquage de deux échantillons présentant un changement d'expression de 2 unités. 62
Figure 15 : Illustration des deux principales approches de la quantification sans marquage. .. 64
Figure 16 : Schéma récapitulatif des principales méthodes de quantification actuellement utilisées dans les laboratoires de protéomique. 65
Figure 17 : Représentation schématique des fractions du protéome (en bleu) qui peuvent être identifiées (en rouge) et quantifiées (en orange) par spectrométrie de masse 70
Figure 18 : Photographie d'un capillaire nanoESI 71
Figure 19 : Puce à émetteurs nanospray 72
Figure 20 : Schéma de principe du nanospray conventionnel (A) et du nanospray sur puce (B) 72
Figure 21 : Schéma des lignes de champ produit en nanospray conventionnel (A) et en nanospray sur puce (B) 72
Figure 22 : Résolution à 50% 73
Figure 23 : Illustration de la fidélité et la justesse 74
Figure 24 : Définition de la précision et justesse en français (A) et en anglais (B). 74
Figure 25 : Concentration de 70 protéines dans le plasma 77
Figure 26 : Répartition des principales protéines plasmatiques 77
Figure 27 : Nomenclature de la fragmentation de la chaine peptidique pour un tétrapeptide. 78
Figure 28 : Exemple de fragmentation d'un peptide à quatre acides aminés en CID à basse énergie 79
Figure 29 : Fragmentation HCD des ions précurseurs dans un QqOrbitrap. 81
Figure 30 : Potentiel appliqué pour la fragmentation en HCD. 81
Figure 31 : Spectromètre de masse LTQ FT Ultra 84
Figure 32 : Schéma temporel de l'acquisition en parallèle dans le LIT/FT-ICR 85
Figure 33 : Schéma détaillé du TSQ Vantage 86
Figure 34 : Balayage des transitions père-fils sur un triple quadripôle 86
Figure 35 : Schéma détaillé du Q Exactive. 90
Figure 36 : Acquisition data-dépendant (DDA) sur le QqOrbitrap. 91
Figure 37 : Mode d'acquisition PRM sur le Q Exactive. 92
Figure 38 : Mode d'acquisition multiplexe PRM sur le Q Exactive 93
Figure 39 : Schéma du Q Exactive HF 94

Figure 40 : Superposition des chromatogrammes de courant d'ions total obtenus avec différents gradients d'élution de 2 à 40% de tampon B 104
Figure 41 : Comparaison de la largeur de pic à mi-hauteur (FWHM) et le nombre de protéines identifiées avec 2 peptides pour différent temps d'élution chromatographique. 105
Figure 42 : Comparaison de A) la largeur de pic à mi-hauteur (FWHM) et B) le nombre de protéines identifiées avec 2 peptides entre une colonne de 15cm et 50 cm 106
Figure 43 : Biopsies de rein de contrôles négatifs. (A, B, C) aucune protéine amyloïdogène n'a été quantifiée. (D) Ce témoin présente une proportion non négligeable de protéine Ig kappa mais ne présente par les protéines SAMP et ApoE, marqueurs des amyloses 111
Figure 44 : Biopsies de rein de patients atteints d'une amylose. (A) Amylose de type AA avec 80% de dépôt fibrillaire (B) Amylose de type AL avec 10% de dépôt fibrillaire 112
Figure 45 : Biopsies de rein de patients atteints d'une amylose. (A) Tissu non fixé (frais) avec 50% de dépôt fibrillaire, présentant une amylose de type AL (B) Tissu fixé découpé par laser (LCM) avec 50% de dépôt fibrillaire présentant une amylose de type AL 113
Figure 46 : Biopsie de rate. (A) Échantillon présentant une amylose de type AA avec 90% de dépôt fibrillaire. (B) Témoin rate. 114
Figure 47 : Biopsies contrôles de poumon. 115
Figure 48 : Biopsie de bronches présentant une amylose de type AH 115
Figure 49 : Biopsie de poumon malade présentant des maladies de dépôt de chaines légères Ig kappa non amyloïdes (LCDD) 116
Figure 50 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans les expériences protéomiques *shotgun* de différents tissus avec Proteome Discoverer. 120
Figure 51 : Visualisation sous forme de cartographie *heatmap* des moyennes des intensités des peptides quantifiés dans les expériences protéomiques *shotgun* de différents tissus avec MaxQuant et Perseus 124
Figure 52 : Comparaison entre les performances des analyses shogun (A) et les analyses ciblées (B) 128
Figure 53 : Temps de rétention en fonction du temps de rétention de la librairie. 135
Figure 54 : Temps de rétention en fonction du facteur d'hydrophobicité prédit par l'algorithme SSRCalc 135
Figure 55 : Nombre de transitions en fonction du temps d'élution chromatographique 137
Figure 56 : Biopsie de rate atteint d'une amylose de type AA analysée en *time-scheduled* SRM. A) Quantification avec tous les peptides. B) Quantification avec les trois peptides les plus intenses 139
Figure 57 : Biopsie de rate atteint d'une amylose de type AA analysée en *time-scheduled* PRM. La quantification est réalisée avec tous les peptides sélectionnés pour la protéine 142
Figure 58 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans un échantillon de biopsie de rate. 146
Figure 59 : Nombre de transitions en fonction du temps d'élution chromatographique. 147
Figure 60 : Optimisation des transitions 148
Figure 61 : Comparaison des aires de pics pour les 10 peptides de la protéine SAMP sélectionnés pour une expérience tSRM. 148
Figure 62 : Nombre de transitions en fonction du temps d'élution chromatographique. 149
Figure 63 : Courbe d'efficacité de fragmentation des transitions sélectionnées pour le peptide IVLGQEQDSYGGK. 150
Figure 64 : Comparaison des aires de pics (en log10) avant et après optimisation des énergies de collisions pour les 3 transitions sélectionnées par peptide de la protéine SAMP 150
Figure 65 : Nombre de transitions en fonction du temps d'élution chromatographique. 151
Figure 66 : Comparaison des aires de pics (en log10) des peptides sélectionnées pour la protéine SAMP obtenues sur le Q Exactive et le TSQ Vantage à leur optimum 153

Figure 67 : Comparaison du signal sur bruit (S/B) de pics des peptides sélectionnées pour la protéine SAMP obtenues sur le Q Exactive et le TSQ Vantage à leur optimum. 153
Figure 68 : Profil chromatographique d'un échantillon de biopsie de rate analysé à différents temps. 158
Figure 69 : Profil chromatographique d'un échantillon peptidique de protéines standards sur deux colonnes de 15 cm de même nature. 159
Figure 70 : Temps de rétention prédit en fonction du facteur d'hydrophobicité prédit par l'algorithme SSRCal pour les peptides choisis pour le *time-scheduled PRM*. 161
Figure 71 : Nombre de transitions en fonction du temps d'élution chromatographique 162
Figure 72 : Spectre en acquisition DDA (TOP7) sur le LTQ FT Ultra de l'échantillon bronche 10H12063 165
Figure 73 : Visualisation sous forme de cartographie *heatmap* des moyennes d'aires normalisées des peptides quantifiés dans les expériences protéomiques ciblées *time-scheduled PRM* de différents tissus avec Pinpoint. 167
Figure 74 : Analyse en composantes principales des données *shotgun* après traitement via Proteome Discoverer. 168
Figure 75 : Analyse en composantes principales des données *time-scheduled* PRM après traitement via Pinpoint. 169
Figure 76 : Analyse en composantes principales des données *time-scheduled* PRM des patients présentant une amylose sans les peptides appartenant aux deux marqueurs fibrillaire (SAMP et ApoE) après traitement via Pinpoint. 170
Figure 77 : Visualisation sous forme de cartographie *heatmap* des données *time-scheduled* PRM normalisées verticalement, des patients présentant une amylose sans les peptides appartenant aux deux marqueurs fibrillaires (SAMP et ApoE) après traitement via Pinpoint. 173
Figure 78 : Visualisation sous forme de cartographie *heatmap* des moyennes d'aires normalisées des peptides quantifiés dans les expériences protéomiques ciblées de différents tissus avec Skyline. 177
Figure 79 : Analyse en composantes principales des données *time-scheduled* PRM après traitement via Skyline. 178
Figure 80 : Analyse en composantes principales des données *time-scheduled* PRM des patients présentant une amylose, sans les peptides appartenant aux deux marqueurs fibrillaires (SAMP et ApoE) après traitement via Skyline. 179
Figure 81 : Visualisation sous forme de cartographie *heatmap* des données *time-scheduled* PRM normalisées verticalement des patients présentant une amylose sans les peptides appartenant aux deux marqueurs fibrillaires (SAMP et ApoE) après traitement via Skyline. 182
Figure 82 : Profil chromatographique des analyses *shotgun* DDA TOP 20 de l'échantillon contrôle (numéro RT1.01.1) sur le Q Exactive (en haut en noir) et sur le Q Exactive HF (en bas en rouge). 188
Figure 83 : Profil chromatographique des analyses *shotgun* DDA TOP 20 de l'échantillon patient (numéro 11H06267) sur le Q Exactive (en haut en noir) et sur le Q Exactive HF (en bas en rouge). 188
Figure 84 : Comparaison du nombre de peptide uniques identifiés et le nombre de groupes de protéines avec au minimum 2 peptides uniques entre le Q Exactive (en bleu) et le Q Exactive HF (en rouge). 189
Figure 85 : Biopsie de rate. Témoin rate analysé sur le Q Exactive (A) et le Q Exactive HF (C). Échantillon présentant une amylose de type AA avec 90% de dépôt fibrillaire analysé sur le Q Exactive (B) et le Q Exactive HF (D) 190
Figure 86 : Nombre de transitions en fonction du temps d'élution chromatographique. 191
Figure 87 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans un échantillon de biopsie de rate témoin et patient. 193

Figure 88 : Extrait du chromatogramme d'ions MS correspondant au peptide LQAEAFQAR (ApoE) sur le Q Exactive HF (A) et le Q Exactive (B) 194
Figure 89 : Extrait du chromatogramme d'ions MS/MS correspondant au peptide LQAEAFQAR (ApoE) sur le Q Exactive HF (A) et le Q Exactive (B). 195
Figure 90 : Distribution du temps d'injection des ions dans le Q Exactive HF au cours de l'élution chromatographique. 195
Figure 91 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans un échantillon de biopsie de rate patient. 196
Figure 92 : Extrait du chromatogramme d'ions MS correspondant au peptide LQAEAFQAR (ApoE) acquis sur le Q Exactive HF à 30 000 de résolution en MS/MS (à gauche) et à 15 000 (à droite). 196
Figure 93: Processus d'ionisation en électrospray. 207
Figure 94: Shadowgraphes flash illustrant la fission coulombienne des gouttelettes d'heptane. La direction du champ électrique est du haut vers le bas. 209
Figure 95 : Modèle d'évaporation de l'ion proposé par le groupe de Dole pour les macromolécules globulaires 210
Figure 96 : Etat de charge maximum et moyen en fonction de la masse de différentes protéines qui suivent le modèle du résidu chargé comparé avec la limite de Rayleigh avec ϱ=1g.cm^{-3} et γ=0.072 N.m^{-1} 210
Figure 97 : Modèle d'évaporation de l'ion proposé par le groupe de Iribane et Thomson pour les petites molécules 211
Figure 98: Modèle de l'éjection de chaine proposé par le groupe de Konermann pour les protéines dépliées 212
Figure 99: Simulation d'un spectre de masse d'une protéine multichargée en ESI. 213
Figure 100: Exemple d'un spectre de masse d'une protéine multichargée 214
Figure 101: Nomenclature de la fragmentation de la chaine peptidique pour un tétrapeptide. 215
Figure 102: Ions fragments a_2 et x_2 issus de la rupture C-CO. 215
Figure 103: Ions fragments b_2 et y_2 issus de la rupture de la liaison peptidique CO-NH. 215
Figure 104: Ions fragments c_2 et z_2 issus de la rupture NH-CH. 216
Figure 105: Ions fragments d_2 et w_2 issus de la rupture des chaines latérales en C_β - C_γ. 216
Figure 106: Ion fragment v_2 issus de la rupture des chaines latérales en C_α - C_β. 216
Figure 107: Fragments internes 216
Figure 108 : Schéma du montage nanoLC 217
Figure 109: Schéma d'un quadripôle.. 221
Figure 110: Diagramme de stabilité 222
Figure 111: photomultiplicateur à dynodes discrètes.. 223
Figure 112 : Piège ionique 224
Figure 113: Courbe de Lissajous 224
Figure 114: Diagramme de stabilité. 225
Figure 115: Piège à ions linaire 225
Figure 116: Simulation du champ électrique dans un piège à ions à une section (A) et trois sections (B) 226
Figure 117 : Piège à ions linaire à double pression 226
Figure 118 : Cellule ICR 229
Figure 119: Signaux avant et après transformée de Fourier. 230
Figure 120: Effet de l'apodisation des spectres FT-ICR 232
Figure 121: Effet du zéro-filling sur les spectres FT-ICR 232
Figure 122: Electrodes idéales. A) Piège 3D de type piège de Penning pour le piégeage des ions B) un quadrupole 2D pour l'axialisation des ions, C) des électrodes planes infinies pour l'excitation et la détection des ions 233

Figure 123: cellules ICR harmonique dynamique 233
Figure 124: Orbitrap..... 234
Figure 125: Comparaison de l'évolution de la résolution en fonction du rapport m/z entre l'analyseur Orbitrap et FT-ICR. 235
Figure 126: Ejection de la C-Trap vers l'Orbitrap. 236
Figure 127: Spectre de la protéine ubiquitine 10 fois chargée 237
Figure 128: Augmentation des performances de l'Orbitrap par diminution de la géométrie. 237

Liste des tableaux

Tableau 1 : Différentes classes d'amylose et leurs principales caractéristiques 37

Tableau 2 : Récapitulatif des réactifs pour le marquage chimique des protéines et peptides... 59

Tableau 3 : Récapitulatif des différentes performances des analyseurs. 73

Tableau 4 : Temps d'acquisition du transient. 90

Tableau 5 : Tableau comparatif entre le Q Exactive et le Q Exactive HF..... 95

Tableau 6 : Liste des échantillons analysés au cours du projet. 101

Tableau 7 : Listes des protéines amyloïdogènes pris en compte dans cette étude. 109

Tableau 8 : Paramètres de recherche pour les analyses shotgun avec les Proteome Discoverer et MaxQuant..... 122

Tableau 9 : Paramètres de recherche des expériences *shotgun* sur les instruments LTQ FT Ultra et Q Exactive..... 130

Tableau 10 : Liste des protéines amyloïdogènes sélectionnées pour la quantification en mode time-scheduled PRM. 132

Tableau 11 : Comparaison des peptides de la protéine SAMP obtenus sur les spectromètres de masse LTQ FT Ultra et Q Exactive et Peptide Atlas. 133

Tableau 12 : Exemple de sélection de peptides (protéines SAMP et ApoE)..... 134

Tableau 13 : Paramètres d'extractions des données pour la méthode tSRM et tSRM par Pinpoint..... 142

Tableau 14 : Listes des transitions avant optimisation sur le triple-quadripôle. 149

Tableau 15 : Listes des transitions les plus intenses après optimisation sur le triple-quadripôle.. 149

Tableau 16 : Paramètres d'extractions des données pour la méthode tSRM et tSRM à leur optimum par Pinpoint. 152

Tableau 17 : Liste des protéines amyloïdogènes sélectionnées pour la quantification en mode *time-scheduled PRM*..... 160

Tableau 18 : Comparaison des conditions chromatographiques entre les deux laboratoires. 186

Tableau 19 : Comparaison des conditions de détection entre les deux laboratoires 187

Tableau 20 : Comparaison des conditions de détection entre les deux laboratoires pour l'analyse en mode *tPRM*. 192

Liste des abréviations

ACN	Acétonitrile, cyanure de méthyle (CH_3CN)
ACP	Analyse en composantes principales
AGC	Contrôle automatique du gain (*automatic gain control*)
ApoA1	Apolipoprotéine A1
ApoE	Apolipoprotéine E
AQUA	Peptide marqué par des isotopes stables, rajouté après digestion (*absolute protein quantification*)
AUC	Aire sous la courbe (*area under the curve*)
BGSA	Biopsie de glandes salivaires accessoires
CV	coefficient de variation (rapport de l'écart-type sur la moyenne)
DDA	Analyse data dépendante, se base sur l'intensité des précurseurs pour déclencher une fragmentation MS/MS (*data dependent analysis)*
D.E.	Diamètre externe
D.I.	Diamètre interne
DIA	Analyse data indépendante, fragmentation de tous les précurseurs par petites gammes de masse (*data independent analysis).*
ESI	Électrospray
FA	Acide formique (CH_3COOH)
FDR	Taux de faux positifs (*false discovery rate*)
FFPE	Tissu fixé au formol et inclus en paraffine (*formalin-fixed paraffin-embedded*)
FIBA	Fibrinogène alpha
FT	Transformée de Fourier (*Fourier transform*)
	FT-ICR : cellule de résonance cyclotronique à transformée de Fourier
FWHM	Largeur à mi-hauteur (*full width at half maximum*)
GAGs	Glycosaminoglycanes
Go	Giga octets
HIFU	Ultrasons intenses et focalisés (*high intensity focused ultrasound*)
HR/AM	Haute résolution et grande exactitude de mesure de la masse (*high-resolution*, *accurate-mass*)
Ig	Immunoglobuline
L.	Longueur
LC	Chromatographie liquide
LCDD	Maladie des dépôts de chaines légères non amyloïde (*light-chain deposition disease*)
LCM	Micro-dissection par capture laser (*laser capture microdissection*)
LIT	Piège à ions linéaire
LMCO	Seuil de basse masse (*low mass cut-off*)
MALDI	Désorption-ionisation laser assistée par matrice *(matrix-assisted laser desorption/ionisation)*

mmu	un millième d'une unité de masse (*milli mass unit*)
MS	Spectrométrie de masse (*mass spectrometry*)
	MS/MS (ou MS^2) : Spectrométrie de masse tandem, fragmentation des peptides
m/z	rapport masse sur charge
NL	Intensité normalisée correspondant au TIC sur le temps d'accumulation des ions (*normalized level*).
Orbitrap HF	
	Orbitrap à haut champ électrostatique, la tension appliquée sur l'électrode centrale est plus élevée que sur les autres Orbitrap (*high field*).
PPM	Partie par million
PRM	Parallel reaction monitoring
	tPRM : time-scheduled Parallel reaction monitoring
PSAQ	Protéine recombinante marquée par des isotopes stables, rajoutée avant traitement de l'échantillon (*protein standard absolute quantification*)
PSM	Spectre assimilé à un peptide (*peptide spectrum match*)
Q	Quadripôle
	QqQ : triple-quadripôle
	QqOrbitrap : quadripôle couplé à un Orbitrap
QconCAT	Protéine chimérique, marquée par des isotopes stables, rajoutée avant la digestion enzymatique (*Quantification concatamer*)
SAA	Sérum amyloïde A
	SAA1 : Sérum amyloïde A 1
	SAA2 : Sérum amyloïde A 2
SAMP	Composant P du sérum amyloïde ou composant P (*Serum Amyloid P component*)
S/B	Signal sur bruit
SIM	Selected ion monitoring
SRM	Selected reaction monitoring
	iSRM : intelligent Selected reaction monitoring
	tSRM : time-scheduled Selected reaction monitoring
	MRM : multiple reaction monitoring ou multiplexe SRM
T3PQ	Quantification de protéine par les trois pics les plus intenses (*Top 3 Protein quantification*)
TFA	Acide trifluoroacétique (CF_3COOH)
TIC	Courant total d'ions (*total ion current*)
To	Téra octets
TRFL	Lactotransferrine
TTHY	Transthyrétine (ou TTR)
UMA	Unité de masse atomique
XIC	Chromatogramme d'ion extrait (*Extracted ion chromatogram*)

Introduction générale

La protéomique désigne la science qui étudie les protéomes, qu'il s'agisse d'études à grande échelle ou d'analyses ciblées très particulières. **Un protéome, entité dynamique et complexe, est défini comme l'ensemble des protéines exprimées par le génome d'une cellule, d'un tissu ou d'un organe dans un environnement donné et à un moment donné** [1]. Les protéines sont directement corrélées à l'information fonctionnelle alors que les gènes sont le support de l'information génétique. Chaque protéine a une structure et une fonction qui lui est propre. Grâce au projet « Human Genome Project », le génome humain est désormais connu et les séquences des protéines peuvent être déterminées. Chez l'humain, il existe environ 100 000 protéines pour environ 20 000 gènes [2]. Cependant, ces séquences théoriques ne suffisent pas. En effet, un gène peut donner divers produits d'épissage et les protéines sont souvent modifiées après leur expression (par diverses modifications post-traductionnelles). Un gène peut alors être à l'origine de diverses formes de protéines, appelée protéoformes [3]. Un programme similaire au génome sur le protéome a alors été initié par l'organisation du protéome humain (HUPO: *Human Proteome Organisation*) en 2002 [4,5] afin d'identifier et de caractériser en détail toutes les protéines séquencées par traduction du génome. Deux groupes de travail se sont constitués. Le premier groupe, réparti sur 24 pays, travaille sur les chromosomes (CPP: *Chromosome-Centric Human Proteome Project*) [6]. Le second vise à étudier le rôle biologique des protéines dans les maladies (B/D-HPP: *Biology/Disease-driven Human Proteome Project*) [7].

L'analyse protéomique est dorénavant envisagée à deux échelles. Tout d'abord **une analyse qualitative** qui permet de mettre en évidence les protéines, dont certaines étaient jusqu' alors prédites mais non observées, dans différents échantillons complexes tels que des lysats cellulaires ou encore des tissus. La deuxième approche qui se veut quantitative, est définie par Anderson et Anderson en 1998 comme une **analyse dynamique et quantitative de la régulation d'expression du produit des gènes** qui caractérise un processus biologique donné dans le but de déchiffrer les mécanismes d'interactions cellulaires [8]. Par exemple, une comparaison d'extraits protéiques issus de différentes conditions (analyse différentielle) est réalisée pour comprendre les processus biologiques intervenants dans des pathologies et trouver des marqueurs de ces processus, aussi appelés biomarqueurs. Il s'agit ici de trouver de nouvelles voies de thérapies et/ou d'établir un diagnostic fiable plus précoce de la maladie pour augmenter les chances de survie du patient. Du fait de la diversité des

protéines et de la très grande gamme dynamique de concentration de cette famille de molécules au sein d'un même échantillon, l'analyse protéomique est un réel défi analytique [9].

Travail réalisé dans cette thèse

C'est dans ce contexte d'analyse protéomique que s'intègre ce projet de thèse. Les amyloses sont un ensemble de maladies, peu connues du grand public, qui sont caractérisées par un dépôt extracellulaire de protéines insolubles diffus ou localisées sur un organe. Ces dépôts peuvent conduire à une défaillance de l'organe. Aujourd'hui, il existe une trentaine de types d'amyloses qui se différencient par la nature biochimique de la protéine précurseur. Le typage est alors important pour connaitre la protéine amyloïdogène mise en jeu pour un traitement efficace.

Le but de ce projet est d'évaluer un outil diagnostique complémentaire aux analyses histochimiques. Cette technique est aujourd'hui la méthode de choix pour le diagnostic des amyloses. Cependant, l'interprétation des données histologiques reste fortement dépendante du pathologiste, du matériel employé (coloration, anticorps, révélateur) et des tissus (diffus ou non) [10–13]. Le diagnostic par spectrométrie de masse est une alternative qui a émergé ces dernières années. Cette application est évaluée en routine dans la fédération hospitalo-universitaire et de recherché américaine Mayo Clinic, sur des prélèvements de tissus ne contenant que la partie atteinte grâce à une micro-dissection par capture laser (LCM) [14]. Nous avons développé au sein de notre laboratoire une approche similaire pour la détection et la quantification des protéines amyloïdogènes dans les tissus bruts par spectrométrie de masse, pour un diagnostic en routine en collaboration avec le service d'anatomie pathologique de l'hôpital Tenon. Afin de définir un vocabulaire commun, nous considèrerons dans la suite que **les tissus bruts sont les tissus sans préparation préalable** (telle que la LCM) aux étapes de digestion enzymatique. Les **tissus bruts peuvent être frais ou fixés**, généralement au formol et inclus en paraffine (FFPE*, formalin-fixed paraffine-embedded*), c'est-à-dire sans fixation. Les tissus prélevés peuvent également être coupés en fines coupes, souvent de 10µm d'épaisseur, à l'aide d'un microtome.

Après une introduction (Chapitre 1) sur la maladie des amyloses pour mettre en évidence l'importance de l'analyse protéomique pour cette maladie particulière, nous reviendrons (Chapitre 2) plus en détail sur les stratégies utiles en approche protéomique et (Chapitre 3) les techniques de spectrométrie de masse appropriées et mises en œuvre dans le cadre de cette

étude. Les résultats propres de notre étude sont rassemblés dans la deuxième partie (Chapitres 4, 5, 6 et 7).

Les amyloses étant associées à des protéines insolubles, nous avons évalué l'extraction et la digestion de ces protéines amyloïdogènes à l'aide d'ultrasons. Le service d'anatomie pathologique ne disposant pas d'appareil LCM, nous avons également évalué dans ce projet la possibilité de faire du typage sans cette étape de préparation de l'échantillon par spectrométrie de masse. L'extrait biologique restant complexe, une étape de séparation chromatographique était nécessaire. Pour cela, nous avons d'abord optimisé le système de chromatographie en phase liquide (durée du gradient d'élution des peptides et longueurs des colonnes). Les peptides ont ensuite été analysés en spectrométrie de masse selon deux méthodes : l'analyse globale appelée couramment *shotgun proteomics* et l'analyse ciblée appelée *targeted proteomics* pour une quantification plus robuste des protéines amyloïdogènes. Pour cette dernière approche, deux modes d'acquisition ont été testés : la SRM (*selected reaction monitoring*) sur un spectromètre de masse à basse résolution de type triple-quadripolaire [15,16], et la PRM (*parallel reaction monitoring*) possible sur un instrument à haute résolution de type quadripôle-Orbitrap [17,18] commercialisé au cours de la thèse. Enfin, un comparatif d'analyse en mode d'acquisition PRM sur le même type d'instrument mais de dernière génération (commercialisé à l'été 2014) sera présenté (Q Exactive et Q Exactive HF, Thermo Scientific).

Partie I : Contexte scientifique et technologique

Chapitre 1 : Les amyloses

I. Généralités

L'amylose, voie finale d'un désordre du métabolisme protéique, est un ensemble de maladies rares caractérisées par un dépôt extracellulaire (agrégat) de protéines mal repliées, qui parfois engagent le pronostic vital. Ce dépôt amyloïde peut être localisé dans un seul organe (dit amylose localisée) ou dans plusieurs organes (dit amylose systémique, généralisée ou diffuse) et se dépose généralement dans les parois des vaisseaux sanguins et sur les membranes basales. La raison de la résistance de cette agrégation de protéines aux processus de dégradation normaux est vraisemblablement due à sa structure spatiale en feuillets ß-plissés caractéristique et son interaction avec la matrice extracellulaire. Du fait de la diversité de ces dépôts, le diagnostic d'une amylose découle alors d'un long processus clinique. Disposer d'un moyen de diagnostic robuste, universel et rapide par spectrométrie de masse est le but de ce projet.

I.1 Historique

Une des premières descriptions modernes de la substance (dépôt) amyloïde comme une infiltration des organes par une substance d'un aspect « lardacée » et « cireux » remonte en 1842 par Rokitansky, constatée lors d'une autopsie macroscopique d'une rate. En 1854, Virchow [19] montra que cette substance se colorait en bleu puis en violet respectivement avec l'iode et l'acide sulfurique. Pensant que ces réactions étaient dues à la présence de cellulose, il utilisa le nom d'amylose, *amylum* en latin (contraction d'amidon et de cellulose), pour qualifier cette substance. En 1859, Friedreich et Kekule démontrèrent l'absence de composant glucidique dans l'amylose et la décrivent comme une association de plusieurs protéines. Le concept de la nature cellulosique de l'amylose fut donc discrédité, et l'amylose est désormais considérée comme une substance essentiellement protéique. En 1922, Bennhold introduit la coloration au rouge Congo comme test diagnostique clinique puis comme marqueur histologique [20]. Une biopsie des organes atteints permettait alors de diagnostiquer une amylose avant le stade de l'autopsie. Les travaux de recherche de Cohen en 1956 [21] en microscopie électronique confirment la structure fibrillaire commune des amyloses et mettent en évidence la présence d'un composant commun : le composant P [22–25]. En 1968, une étude de cristallographie aux rayons X des dépôts amyloïdes

révèle une conformation commune spatiale en feuillets β-plissés par Eanes et Glenner [26].

I.2 Pathogenèse

Les précurseurs protéiques à l'origine des dépôts amyloïdes sont pour la plupart du temps, des protéines circulantes ou produites localement, et parfaitement solubles. Les protéines amyloïdogènes, responsables de la formation des fibrilles, proviennent dans la majorité des cas, d'un clivage protéolytique de leur précurseur. Par un mécanisme encore inconnu, elles subissent une modification de leur repliement qui conduit à une structure majoritairement en feuillets β plissés. Elles s'incorporent et/ou interfèrent ensuite avec d'autres précurseurs de constituants fibrillaires et mènent à une polymérisation et à la formation de fibrilles amyloïdes insolubles (Figure 1). La pathologie peut résulter d'une modification structurale, d'une surproduction de la protéine précurseur, d'une réduction de sa dégradation dans l'organisme ou encore d'une anomalie génétique associée aux protéines variantes [25,27].

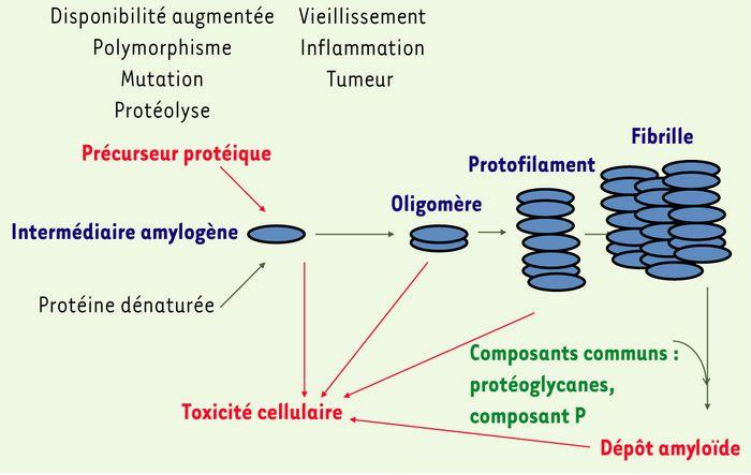


Figure 1 : Pathogenèse des amyloses. Les protéines subissent une protéolyse ou transformation par un mécanisme méconnu, formant une structure en feuillets β-plissés insolubles. Un empilement des protéines amyloïdes a ensuite lieu donnant naissance à des dépôts amyloïdes insolubles qui résistent aux protéolyses [28].

Il existe actuellement une controverse quant à la toxicité réelle des polymères amyloïdes dans la maladie d'Alzheimer. Certains groupes démontrent que dans le cadre de cette maladie, associée à la présence du peptide Aβ dans des dépôts amyloïdes, ce sont les oligomères qui sont les plus toxiques [29–31]. Cette question a été récemment posée pour certaines atteintes organiques des amyloses systémiques. Cependant dans ces maladies la toxicité du dépôt est considérée comme le principal mécanisme pathogénique.

II. Classification

La principale caractéristique structurale des dépôts amyloïdes extracellulaires est donc son aspect fibrillaire, de 7 à 10 nm de diamètre et de longueurs variables. **Bien que cette structure fibrillaire soit commune à tous les dépôts amyloïdes, les différents types d'amyloses se différencient par leur protéine précurseur**. La classification des amyloses, proposée à l'origine par Glenner en 1970 [32], se fait ainsi suivant leur composition biochimique. En 1993 [33], un comité de nomenclature des amyloses dirigé par Westermark a été créé pour uniformiser les noms et mettre à jour les différentes classes d'amyloses résumées dans le Tableau 1.

Tableau 1 : Différentes classes d'amylose et leurs principales caractéristiques [23,25,33].

Forme Amyloïde	Protéines Précurseurs	Systémiques (S) ou localisées (L)	Héréditaire (H) ou Acquis (A)	Syndrômes ou tissus impliqués
AL	Chaînes légères d'immunoglobuline (λ, κ)	S et L	A	(Primaire) isolée ou associée à un myélome multiple (tous types d'organes sauf SNC)
AH	Chaînes lourdes d'immunoglobuline (γ)	S et L	A	isolée (tous types d'organes sauf SNC)
AA	(Apo) Sérum Amyloïde A	S et L	A	(Secondaire) infection, inflammation chronique, tumeur, fièvre méditerranéenne, familiale, syndrome de Muckle-Wells (tous types d'organes sauf SNC)
ATTR	Transthyrétine normale	S	A	Cœur principalement
	Transthyrétine mutée	S	H	SNP, SNA, cœur, œil
Aβ2M	β2-microglobuline normale	L	A	Système musculo-squelettique, associée à l'insuffisance rénale chronique terminale
	β2-microglobuline mutée	S	H	SNA
AApoAI	Apolipoprotéine AI normale	L	A	Aorte (intima)
	Apolipoprotéine AI mutée	S	H	Cœur, foie, rein, SNP, testicules, larynx, peau
AApoAII	Apolipoprotéine AII	S	H	Rein
AApoAIV	Apolipoprotéine AIV	S	A	Rein
AGel	Gelsoline	S	H	SNP, cornée
ALys	Lysozyme	S	H	Rein
AFib	Fibrinogène α	S	H	Rein principalement
ACys	Cystatine C	S	H	Hémorragie cérébrale de type islandais (SNP), peau
Aβ	Précurseur de la protéine Aβ (AβPP)	L	H / A	Maladie d'Alzheimer, Trisomie 21, Angiopathie amyloïde cérébrale héréditaire ou sporadique (SNC)
APrP	Protéine prion normale	L	A	Maladie de Creutzfelt-Jackob, insomnie fatale familiale
	Protéine prion mutée	L	H	Maladie de Creutzfelt-Jackob, insomnie fatale familiale, syndrome de Gerstmann–Sträussler–Scheinker
ACal	(Pro)calcitonine	L	A	Cancer médullaire de la thyroïde
AANF	Facteur atrial natriurétique	L	A	Amylose auriculaire isolée (cœur)
AIAPP	Polypeptide islet amyloïde	L	A	Ilôts de Langerhans du diabète de type 2, insulinome
AIns	Insuline	L	A	Iatrogènique
APro	Prolactine	L	A	Prolactinome, hypophyse sénile
AKer	Kératoépithéline	L	A	Cornée

ALac	Lactoferrine	L	A	Dystrophie coréenne sous épithéliale
ABri *	ABriPP	S	H	Démence héréditaire britannique (SNC)
ADan *	ADanPP	L	H	Démence héréditaire Danoise (SNC)
AMed	Lactadhérine	L	A	Aorte (média)
ALect2	Leucocyte chemotactic factor-2	S	A	Rein principalement
ASPC	Proteine surfactant pulmonaire	L	A	Poumon
AGal7	Galectine-7	L	A	Peau
Acor	Corneodesmosine	L	A	Épithélium cornifié, follicule pileux
AOaap	Odontogenic ameloblast associated protein	L	A	Tumeur odontogénique
ASem1	Semenogelin I	L	A	vésicule séminale

* ABriPP et ADanPP sont dérivées du même gène * ABriPP et ADanPP sont dérivées du même gène ; SNC = système nerveux central ; ANS = système nerveux autonome ; SNP = système nerveux periphérique

II.1 Composants communs

Dans tous les cas d'amylose, il existe des protéines communes souvent glycosylées qui participent à la maladie des amyloses.

II.1.1 Les glycosaminoglycanes

Les glycosaminoglycanes (GAGs) sont un des constituants majeurs de la matrice extracellulaire et jouent un rôle facilitateur dans la constitution des dépôts [34]. Ce sont des polysaccharides linéaires formés de structure répétitive d'un disaccharide non ramifié. Ils sont attachés au résidu sérine d'un noyau protéique par un trisaccharide. Tout cet ensemble est appelé protéoglycane.

II.1.2 Composant P amyloïde (Serum Amyloid P component)

Le composant P amyloïde (composant P) est une protéine glycosylée plasmatique de structure pentagonale formée de sous-unités globulaires [35]. Il se lie spécifiquement par une interaction calcium dépendante, à un ligand (GAGs) présent dans tout type d'amylose y compris la maladie d'Alzheimer [22,36]. L'interaction forte avec l'héparane sulfate et le dermatane sulfate en présence de calcium suggère que ces GAGs pourraient être les intermédiaires dans l'interaction du composant P avec les fibrilles amyloïdes [37,38]. Le composant P est reconnu par la communauté comme un marqueur des amyloses et est un constituant majeur (jusqu'à 20%) dans tous les dépôts amyloïdes [22,23,39]. Son rôle biologique n'est pas encore clairement défini mais il participerait à la formation des plaques amyloïdes et il inhiberait la protéolyse des fibrilles amyloïdes [40,41].

II.1.3 Apolipoprotéine E (ApoE)

L'apolipoprotéine E (ApoE) est une protéine chaperonne qui en présence des GAGs et du composant P induit une structure en feuillet ß-plissé des protéines amyloïdogènes dans les amyloses systémiques et localisées. Elle est surtout reconnue comme représentant un facteur risque de la maladie d'Alzheimer. Son rôle dans la formation des fibrilles reste encore inconnu mais elle a été répertoriée comme interagissant avec les lipoprotéines AA, AI et AII, elles-mêmes marqueurs d'amyloses, et retrouvée dans les amyloses de type AA et TTR [23].

II.2 Principales protéines amyloïdogènes

Dans les paragraphes suivants sont décrites certaines protéines amyloïdogènes. Seules les amyloses suivies au cours de ce projet, associées aux pathologies les plus fréquemment rencontrées, seront détaillées dans ce manuscrit.

II.2.1 Immunoglobulines (Ig)

L'amylose dite AL *(Amyloid Light chain)* ou amylose primaire, est associée principalement au myélome (cancer hématologique). Elle est caractérisée par la présence anormale de chaînes légères d'immunoglobuline (Ig) monoclonales complètes ou de certains de leurs fragments, qui résulte d'un dysfonctionnement du plasmocyte. La majorité des chaînes légères amyloïdogènes est de type lambda (Ig λ) mais dans certains cas les chaines légères d'Ig kappa (Ig ϰ) peuvent être identifiées. Il existe plusieurs isoformes de chaines légères d'immunoglobulines qui ont une région constante et variable mais elles ne sont pas tous amyloïdogènes [27].

L'accumulation de substance amyloïde de type AL peut être détectée dans plusieurs organes, cependant c'est le plus souvent dans les reins qu'elle est diagnostiquée [42]. On note 500 nouveaux cas par an en France, maladie rare donc, qui touche généralement les patients entre 60 et 70 ans.

Pour **l'amylose AL**, l'objectif du traitement est de faire diminuer la production de ces dépôts d'immunoglobuline. Un traitement par chimiothérapie peut être proposé si la maladie est prise à un stade précoce. Le type de chimiothérapie utilisé ressemble à celui que l'on utilise dans les maladies plasmocytaires.

Une autre maladie liée au dépôt de chaines légères d'immunoglobulines (LCDD, *light-chain deposition disease*) existe et ne doit pas être confondue avec l'amylose de type AL. Ces dépôts

sont négatifs au Rouge Congo et présentent un aspect granuleux et non fibrillaire en microscopie électronique [43].

II.2.2 ***Sérum amyloïde A (SAA)***

La protéine précurseur de l'amylose AA secondaire est la protéine sérique amyloïde A (SAA) synthétisée sous une stimulation de diverses cytokines pro-inflammatoires et dégradée dans les hépatocytes (foie) [23,27]. C'est une apolipoprotéine, c'est-à-dire qu'elle est constitutive des lipoprotéines qui ont pour but de transporter des molécules hydrophobes dans le sang (triglycérides, cholestérols). Dans l'espèce humaine, 3 gènes codant la SAA ont été identifiés *(SAA1, SAA2, SAA4)*. Le quatrième (*SAA3*) est un pseudo-gène [44]. Ces protéines sont retrouvées dans le plasma des malades, ce qui signifie qu'il y a surproduction de la SAA ou une mauvaise dégradation de cette protéine par le foie, ce qui aurait pour conséquence la formation de dépôts amyloïdes. Cette protéine est également un marqueur de la réaction d'inflammation. Dans ce cas, sa concentration augmente de trois ordres de grandeur au moins. L'amylose AA peut donc être une complication de n'importe quelle maladie inflammatoire chronique, par exemple la fièvre méditerranéenne familiale [23]. Dans 6% des cas, elle n'a pu être liée à aucune pathologie.

L'accumulation de ces dépôts amyloïdes de protéines SAA se rencontre dans 90% des cas dans le rein.

Pour **l'amylose AA**, le traitement consiste à faire diminuer l'inflammation et dépend de la maladie inflammatoire chronique sous-jacente.

II.2.3 ***Transthyrétine (TTR)***

La transthyrétine (TTR), protéine non glycosylée et tétramérique de 4 sous-unités de 127 acides aminés, est la protéine impliquée majoritairement dans les amyloses héréditaires. Elle fut appelée autrefois pré-albumine du fait de son élution en électrophorèse capillaire avant l'albumine. Principalement synthétisée par le foie, elle a pour rôle physiologique de transporter les hormones thyroïdiennes et le rétinol (vitamine A) dans le plasma. La TTR est également un marqueur de la malnutrition, de l'inflammation et d'insuffisance hépatocellulaire : sa concentration diminue dans ces trois cas. Au contraire, en cas d'insuffisance rénale chronique, la concentration en TTR augmente. [23]

L'amylose ATTR est une maladie héréditaire autosomique dominante très rare, de l'ordre de 500 cas au total en France. Cependant tout porteur du gène ne développera pas

obligatoirement la maladie (15% des patients porteurs n'auront aucun symptôme). Il existe plus d'une centaine de mutations génétiques. La forme protéique normale de la transthyrétine peut également être à l'origine de dépôt amyloïde (amylose sénile).

Les dépôts amyloïdes de type ATTR se retrouvent souvent dans le système nerveux périphérique, le cœur et le rein.

Pour **l'amylose ATTR**, le traitement de fond de la maladie recommandé est la transplantation hépatique. De petites molécules inhibant la formation de l'agrégat de transthyrétine sont prometteuses.

II.2.4 Apolipoprotéine A 1 (ApoA1)

L'apolipoprotéine A 1 (ApoA1) est le principal composant des lipoprotéines de haute densité HDL *(High Density Lipoprotein)* et intervient dans le transport du cholestérol des tissus périphériques vers le foie. Elle est synthétisée principalement dans le foie et dans l'intestin sous forme d'un précurseur à 267 acides aminés qui après protéolyse, donne la protéine ApoA1 constituée de 234 acides aminés [23].

Comme la protéine TTR, l'amylose de type AApoA1 est une maladie héréditaire qui peut être causée par des mutations et une maladie liée au vieillissement [45]. Les dépôts amyloïdes de type ApoAI peuvent être retrouvés dans les aortes (amylose localisée) en cas de mauvaise élimination du cholestérol dans les artères mais également dans d'autres organes ou tissus (rein, cœur, nerf, foie, peau, larynx) en cas de maladie héréditaire.

Pour **l'amylose AApoA1**, l'objectif du traitement est de diminuer (de 50% environ) la production de l'apolipoprotéine A1 par transplantation hépatique.

III. Moyens de diagnostic des amyloses

Aucun élément clinique ne permet le diagnostic d'une amylose. Seul un prélèvement invasif tel que prélèvement de tissus (par biopsie) direct de l'organe malade ou sur des zones plus facile d'accès en cas d'amyloses systémiques (biopsies des glandes salivaires accessoires ou rectales), permet de poser un diagnostic clair. La biopsie peut être réalisée à l'œil nu lors d'une opération à ciel ouvert ou sous endoscopie. Récemment, certaines équipes travaillent sur l'aspiration de la graisse sous cutanée abdominale, moins invasive, pour le diagnostic [46–48].

Dans le cadre de ce projet, les biopsies ont été effectuées par nos collaborateurs de l'hôpital Tenon à Paris (coordinateur de l'équipe : Dr. Gilles Grateau). Une fois ces tissus prélevés, deux techniques peuvent être employées : (1) l'histologie conventionnelle et l'immunohistochimie, très utilisées en milieu clinique ou (2) la spectrométrie de masse, technique de plus en plus utilisée au service de la biologie, qui est l'objet de notre étude.

III.1 Histologie

Le diagnostic en France se fait principalement par histologie. Le principe de cette étude des différentes réactions chimiques, du métabolisme et la constitution chimique des tissus et leurs cellules est résumé sur la Figure 2.

III.1.1 Coloration au rouge Congo

Les tissus prélevés sont d'abord congelés ou fixés puis une microtomie est effectuée avant dépose sur lamelle de verre. **Un test à la coloration au rouge Congo** est ensuite effectué. Si ce test est positif, **un test plus spécifique avec des anticorps** (immunohistochimie) est réalisé pour connaitre la nature de l'amylose. Si le test au Rouge Congo est négatif, plusieurs autres pathologies de type non amyloïde sont suspectées.

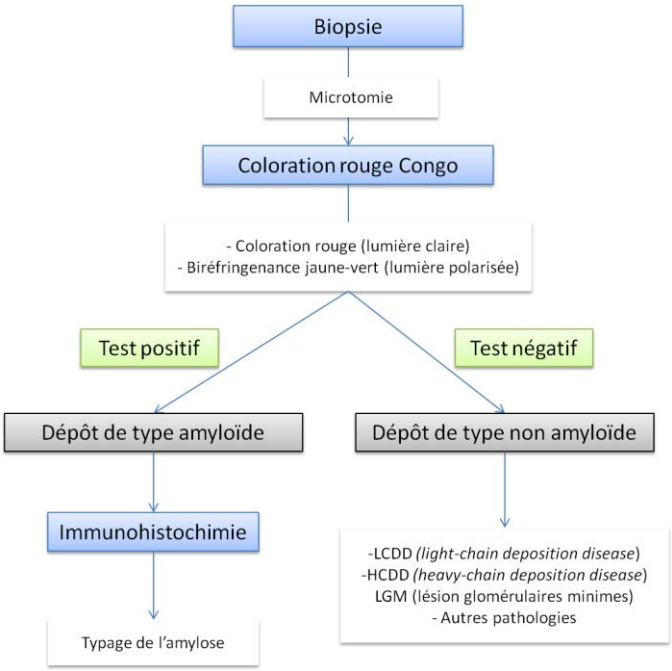


Figure 2 : Stratégie histochimique.

Les coupes sont analysées par microscope optique après coloration au rouge Congo. Ce dernier est un colorant azosulfonaté symétrique qui possède un centre hydrophobe, formé par un groupe biphényle situé entre deux extrémités qui sont chargées (Figure 3). Le rouge Congo

forme des interactions hydrophobes avec les parties hydrophobes des fibrilles amyloïdes. Des interactions électrostatiques ont également lieu entre les acides aminés de charge positive et les groupements sulfonates, chargés négativement. [23]. Cette affinité tinctoriale entre dépôt amyloïde et rouge Congo permet de distinguer les tissus affectés par leur coloration : en lumière claire les substances amyloïdes se colorent et le rouge est plus ou moins prononcé suivant le tissu, la coupe et le type d'amylose (Figure 4A). Cette fixation n'est pas suffisante pour identifier sans ambiguïté un dépôt amyloïde, car certains composants fixent également le rouge Congo, comme le collagène. En revanche, en lumière polarisée (Figure 4B), les substances amyloïdes ont une propriété biréfringente spécifique de couleur jaune-vert (dichroïsme). D'autres colorations sont utilisées en histologie, telles que la thioflavine T et le cristal violet, mais ils n'ont pas la spécificité de la coloration au rouge Congo.

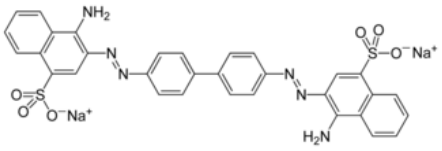


Figure 3 : Molécule rouge Congo

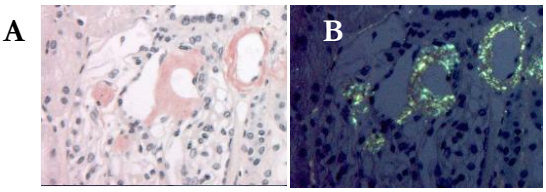
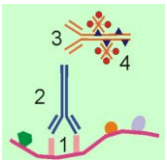


Figure 4 : Coloration au rouge Congo d'un dépôt amyloïde artériolaire

III.1.2 ***Immunohistochimie***

L'immunohistochimie permet de localiser une protéine (antigène) dans des tissus ou cellules à l'aide d'anticorps. Elle repose ainsi sur le concept de clé-serrure : anticorps-antigène.

Pour cela, un anticorps primaire dirigé contre la ou les protéine(s) fibrillaire(s) amyloïdogène(s) est utilisé. Il peut être poly ou monoclonal, sachant que l'anticorps monoclonal sera plus spécifique. Puis un anticorps secondaire couplé à un agent réactif, souvent une enzyme reporteur (tel que la peroxydase) ou un fluorophore, est utilisé pour se fixer à l'anticorps primaire (Figure 5). S'il y a reconnaissance, une coloration a lieu pour l'enzyme rapporteur ou une lumière fluorescente est émise par le fluorophore qui sera détectée en fluorimétrie. Cette méthode indirecte est plus sensible que la méthode directe de l'antigène avec son anticorps seul.



1- Antigène
2- Anticorps primaire
3- Anticorps secondaire
4- Complexe

Figure 5 : schéma de principe de l'immunohistochimie

La technique de reconnaissance par immunohistochimie, bien qu'assez spécifique, est limitée par la sensibilité incomplète de certains anticorps, notamment ceux dirigés contre les chaînes légères d'immunoglobuline, et par une spécificité faible avec des réactions croisées. C'est une technique délicate qui pose des difficultés d'interprétation et ne permet pas toujours de conclure à un typage certain. De plus, le pathologiste ne dispose bien souvent que d'un ensemble limité d'anticorps dirigés contre les protéines amyloïdogènes les plus fréquentes [49].

Une autre technique, sensible, spécifique et sans *a priori*, telle que la spectrométrie de masse est donc nécessaire pour aider/confirmer le typage des amyloses et trouver de nouvelles formes d'amylose.

III.2 Spectrométrie de masse

Récemment, plusieurs équipes ont évalué le diagnostic des amyloses par détection en spectrométrie de masse [46,50–54,14,55–58]. Cette technique permet de s'affranchir des anticorps qui nécessitent une connaissance particulière de la technique. Bien souvent les tissus issus de biopsies sont fixés au formaldéhyde, appelé aussi formol, et enrobés dans de la paraffine (FFPE) pour une conservation à l'état d'origine. C'est sous cette forme que les échantillons peuvent être coupés en fines tranches pour une observation au microscope. Cette fixation engendre la formation des ponts méthylène inter et intra moléculaires entre des acides aminés des chaines latérales (lysine, arginine, histidine et cystéine) et amides des chaines peptidiques, qui aboutissent à une modification des protéines et mènent souvent à la perte des épitopes. L'hydrophobicité des tissus est également accrue par la fixation au formol, augmentant les colorations non spécifiques. Pour solubiliser les protéines dans les tissus fixés, des étapes supplémentaires de récupération de ces protéines sont nécessaires avant la digestion enzymatique. Pour réduire le volume de traitement et récupérer seulement la partie du tissu présentant le dépôt amyloïdogène,, une étape de micro-dissection LCM est souvent rajoutée[50–54,14,55–60].

III.2.1 Digestion sous ultrasons

La stratégie générale d'identification des protéines par l'approche *bottom-up* requiert une digestion enzymatique des protéines avant l'analyse en spectrométrie de masse. Pour faciliter la protéolyse des protéines, après les avoir solubilisées, une réduction des ponts disulfure puis une alkylation des cystéines libres (pour empêcher leur réassociation) sont nécessaires. La

durée nécessaire pour réaliser une protéolyse enzymatique spécifique (avec la trypsine par exemple) avec un rendement correct oscille entre 2 et 16 heures. Ces étapes de préparation et digestion augmentent considérablement le temps de traitement des échantillons. Ces dernières années, beaucoup d'efforts ont été menés pour accélérer la réaction enzymatique en jouant sur **deux principaux paramètres qui tendent à accroitre l'activité enzymatique : la température et la pression**. Ainsi, l'utilisation des radiations micro-ondes permet une digestion en moins de 5 min due à une augmentation locale de la température et une agitation locale des molécules [61,62]. En appliquant en plus des ultrasons de haute intensité et focalisés (HIFU, *high intensity focused ultrasound*) pour augmenter la pression, la digestion en solution ou en gel s'effectue en quelques secondes [63–67]. Cette dernière technique est l'objet de notre étude.

Les ultrasons peuvent être transférés à l'échantillon de deux manières : (1) grâce à un bain à ultrasons où les ondes émises par des transducteurs se propagent à travers l'eau vers l'échantillon ou (2) grâce à une sonde à ultrasons directement immergée dans la solution d'échantillon. Ce système offre une puissance locale bien supérieure au cas du bain à ultrasons [65,66] et sera utilisé au cours de ce projet. L'augmentation de l'activité enzymatique n'est pas due à une interaction directe entre la sonde à ultrasons et les protéines mais au **phénomène de cavitation acoustique**. Lorsque l'onde se propage dans le milieu liquide, elle génère de petites bulles de gaz, appelées bulles de cavitation. Ces bulles grossissent jusqu'à atteindre une taille où elles implosent créant des jets ultra-rapides de liquide avec une onde de choc si intense que la surface des échantillons est endommagée. **Cette implosion produit des conditions de température et pression locales très élevées (environ 5000°C et 1000 atm.), fournissant l'énergie nécessaire pour accélérer les réactions chimiques** [68–70]. Cependant, des produits dérivés et une perte d'échantillon peuvent se produire par modification chimique, évaporation, nébulisation ou échauffement localisé. Il est donc nécessaire de maitriser l'environnement de la digestion sous ultrasons et d'utiliser entre autres un système de refroidissement, comme des blocs de froid, pour éviter la dégradation des protéines et peptides [70].

Les ultrasons peuvent également être utilisés en amont de la digestion, comme lors de l'étape de solubilisation des protéines amyloïdogènes agrégées dans les tissus bruts. Certaines équipes ont utilisé des bains à ultrasons pour récupérer ces protéines contenues dans des tissus FFPE. Cette technique est très efficace, cependant elle nécessite de nombreuses heures, voire des jours entiers, d'extraction [50,56]. Afin de réduire le temps de préparation des

échantillons pour une analyse à haut débit dans un contexte clinique, nous avons évalué le broyage des tissus et la digestion sous ultrasons à l'aide d'une sonde à ultrasons puissante.

L'objectif de ce projet de thèse est d'évaluer la récupération des protéines amyloïdogènes insolubles dans les tissus congelés et fixés au formaldéhyde, enrobés ou non de paraffine en utilisant les ultrasons. Les tissus testés ont été obtenu après biopsie avec ou sans micro-dissection LCM. Le traitement choisi utilise une sonde à ultrasons directement mise en contact avec l'échantillon. Le but est d'établir une méthode par détection en spectrométrie de masse, pour un diagnostic rapide des amyloses dans les départements cliniques français, en collaboration avec le service d'anatomie pathologique de l'hôpital Tenon, l'un des centres de référence des amyloses.

Chapitre 2 : Identification et quantification des protéines

I. Stratégie d'analyse des protéines

I.1 Apport de la spectrométrie de masse

Les techniques de spectrométrie de masse utilisées dans le cadre de cette étude et les concepts associés, sont détaillés dans le chapitre suivant. Brièvement ici notons que deux modes d'utilisation principaux d'analyse sont utilisés en routine : la spectrométrie de masse (MS) simple qui donne accès à la masse des molécules présentes dans le mélange et la spectrométrie de masse tandem (MS/MS) qui donne accès à des éléments de structure et notamment à des éléments de séquence pour les protéine et/ou peptides présents dans le mélange. Deux approches complémentaires en protéomique peuvent être utilisées pour analyser les protéines : l'approche *Top-down* qui consiste à analyser les protéines entières et les fragmenter pour avoir une information de séquence et l'approche *Bottom-up*, par laquelle les protéines sont identifiées par analyse des peptides issus de la digestion enzymatique (Figure 6). Cette dernière, qui reste de loin la plus utilisée dans les laboratoires de protéomique, a été choisie au cours de ce projet.

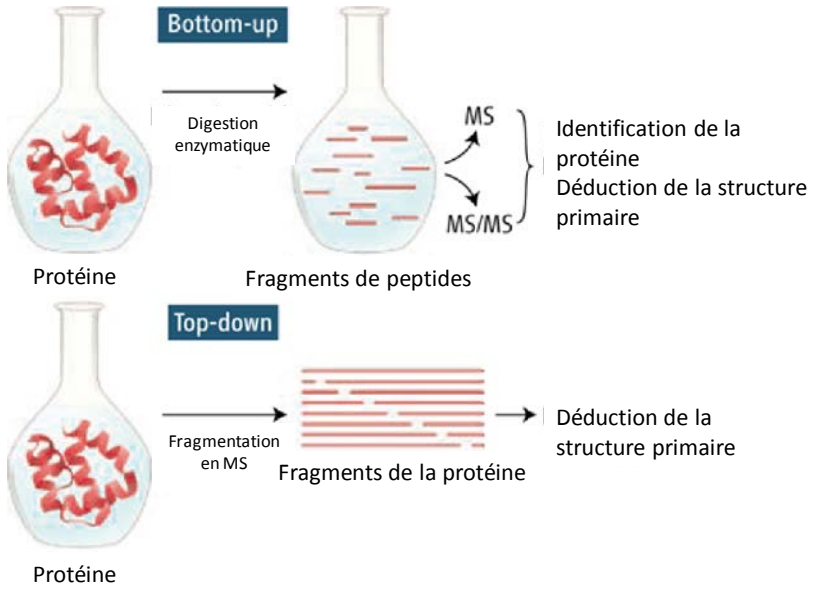


Figure 6 : Stratégie *top-down* et *bottom-up*. La stratégie *top-down* part de la protéine intacte pour arriver aux fragments peptidiques. La stratégie *bottom-up* part quant à elle aux peptides pour arriver aux protéines (figure adaptée de Chait [71]).

I.2 Approche descendante (Top-Down)

Dans cette approche appelée ***top-down***, littéralement du haut vers le bas, les protéines restent intactes et ne subissent pas de protéolyse chimique et enzymatique. Les protéines entières sont introduites en phase gazeuse par une source d'ionisation puis sont directement fragmentées dans le spectromètre de masse. Etant donné que ces protéines ont une grande masse moléculaire, il est nécessaire d'avoir des fragmentations spécifiques, telles que l'ECD (*electron captured dissociation*) [72], ETD (*electron transfert dissociation*) [73] ou encore l'IRMPD (*infrared multiphoton dissociation*) [74] (cf. § III.1). Les spectres de protéines ainsi produits, aussi bien sur des protéines entières que des fragments de protéine (*middle-down*), sont complexes. C'est pourquoi cette méthode nécessite des spectromètres de masse à haute résolution avec une grande gamme de masse. Un des intérêts majeurs de ce type d'approche est la mise en évidence de la coexistence de différentes formes moléculaires d'une protéine donnée, produit d'un même gène et présentant ou non des modifications post-traductionnelles, appelées protéoformes (Figure 7) [3] , puisqu'aucun traitement n'est réalisé sur les protéines.

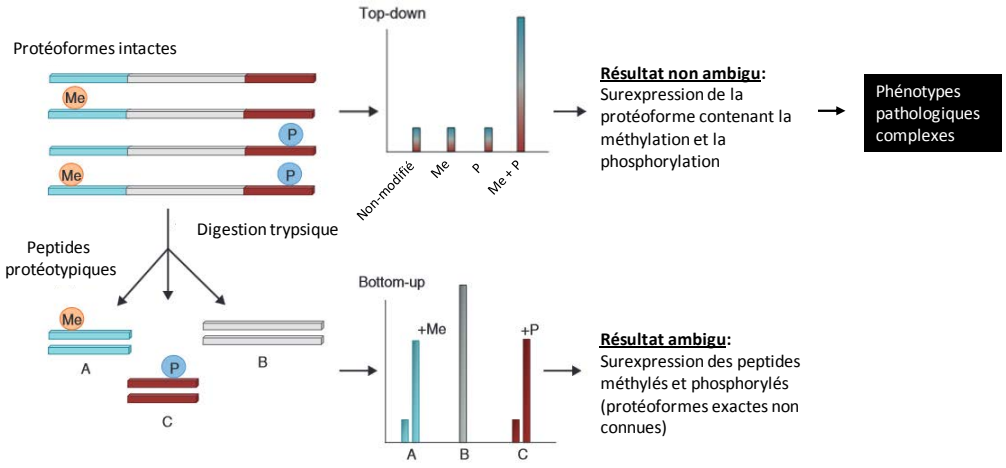


Figure 7 : L'approche *top-down* fournit des informations sur les phénotypes. Un gène peut donner plusieurs protéines modifiées par une(des) méthylation(s) (Me) et/ou une(des) phosphorylation(s) (P) par exemple. Ces formes, qui peuvent être simultanément exprimées dans le protéome, sont appelées protéoformes. L'approche *bottom-up* dans cette exemple permet d'observer une surexpression de peptides méthylés et phosphorylés sans identifier la forme protéique en jeu. L'approche *top-down* quant à elle, permet d'affirmer que c'est la forme méthylé et phosphorylé qui est surexprimée dans ce complexe (figure adaptée de Savaryn *et al.* [75]).

Cette approche, bien que très élégante, n'en n'est pas moins à ses prémices. De plus, il faut des appareils couteux, de très haute résolution et grande gamme de masse pour mesurer la masse précise des protéines. Ce dernier peut être contourné par l'utilisation de l'ionisation par électrospray (ESI) qui produit des ions multichargés et diminue donc le rapport masse sur charge (m/z) des protéines (cf. Annexe 1) mais le signal des grosses protéines est alors dilué dans les nombreux massifs isotopiques. Dans ce cas, il faut s'assurer que la protéine étudiée

soit isolée ou séparée des autres protéines par chromatographie, étape qui reste encore aujourd'hui délicate, pour éviter la superposition, la co-élution de différentes espèces et l'effet de suppression des signaux. En outre, les logiciels de retraitement et validation de ces données sont encore peu développés. C'est pour ces raisons que très peu de laboratoires de protéomique *top-down* existent à l'inverse du *bottom-up*, qui reste aujourd'hui la méthode de choix pour l'analyse du protéome.

I.3 Approche ascendante (Bottom-up)

L'approche dite ***bottom-up***, illustrée à la Figure 8, part du bas vers le haut : Ici les protéines subissent soit une étape de digestion enzymatique « in-gel » après séparation sur gel SDS PAGE 1D ou 2D (selon la complexité de l'échantillon), soit un fractionnement en solution « off-gel » ou encore une digestion liquide sans séparation préliminaire ou, pour analyser le mélange de peptides protéolytique obtenus dans une stratégie appelée *shotgun proteomics*. En général, l'enzyme utilisée est la trypsine car elle présente une bonne spécificité et génère des peptides de taille relativement homogène. Les peptides obtenus sont ensuite analysés par spectrométrie de masse en mode MS (mesure de la masse de peptides) ou en mode MS/MS (mesure de la masse des fragments de peptides). C'est cette approche *shotgun* que nous avons choisi de mettre en œuvre dans ce projet. Une fois les spectres de masse obtenus, ils sont comparés aux spectres de masse théoriques calculés à partir des séquences de protéines contenues dans les banques de données pour une identification des protéines réellement présentes dans l'échantillon.

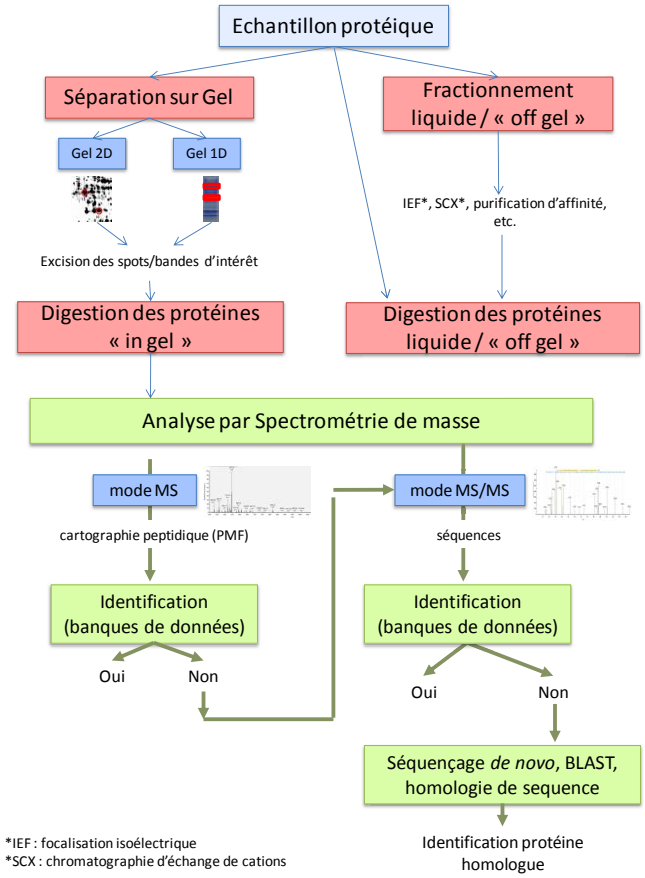


Figure 8 : Schéma de la stratégie *bottom-up*. Les protéines sont séparées ou non sur un gel SDS-PAGE 1D/ 2D ou liquide avant de subir une digestion enzymatique qui génèrent des peptides protéotypiques. Ces derniers sont ensuite analysés en mode MS ou en mode MS/MS. Les spectres obtenus sont ensuite confrontés au spectre théorique des protéines dont leur séquence est répertoriée dans des bases de données pour l'identification des protéines.

Avant l'analyse en spectrométrie de masse, les peptides générés peuvent être séparés par chromatographie liquide à polarité de phases inversée, technique robuste pour les peptides, afin de réduire la complexité des échantillons et ainsi éviter les effets de suppression de signal en masse (cf. §II.1)

II. Stratégie analytique en mode bottom-up

II.1 Identification des protéines par couplage chromatographique

Lorsque les mélanges sont très complexes, il est indispensable de simplifier le mélange analysé par spectrométrie de masse (MS) en fractionnant l'échantillon, en recourant à des techniques de séparation robustes. La stratégie reine dans le cas de l'analyse protéomique est le couplage de la chromatographie en phase liquide (LC) à polarité de phase inversée (RP, *reverse phase*) avec la MS en tandem, qui sera désigné sous l'acronyme LC-MS/MS par la suite.

Ce couplage a été revu dans de nombreuses publications [76]. Nous nous intéresserons ici au traitement des données obtenues.

L'identification des protéines en approche *shotgun* s'effectue par combinaison des informations obtenues en mode MS (masse précise des peptides) et en mode MS/MS (séquençage de ces mêmes peptides). De nombreux programmes bioinformatiques et méthodes mathématiques ont été mis en place pour attribuer les séquences peptidiques aux spectres de fragmentation MS/MS de façon automatique. Ces logiciels peuvent être classés en trois catégories [77] présentées dans la suite, mais **seule la recherche dans les banques de données été utilisée dans ce projet (détaillée au § II.2).**

1) **Recherche dans des banques de données**. Les spectres expérimentaux MS/MS sont confrontés aux spectres théoriques prédits pour chaque peptide, issu de la digestion *in silico* des protéines contenues dans une banque de données, selon différents paramètres de recherche : la taxonomie, l'enzyme de digestion, le nombre de sites de coupure manqués, les modifications chimiques ou post-traductionnelles autorisées, le type de fragmentation, l'erreur tolérée en masse pour les précurseurs et les fragments. Différents algorithmes de recherche Mascot [78], Sequest [79] et Andromeda [80] ont été utilisés au cours de ce projet. Mais de nombreux autres ont été développés (Protein prospector [81], ProbID [82], X!Tandem [83], OMSSA [84], ProSight [85] et Inspect [86]) La comparaison peut également se faire à l'aide d'une bibliothèque de spectres MS/MS de peptides acquis et identifiés au préalable. Cet outil est disponible sur le logiciel MyProMS [87].
2) **Séquençage *de novo*,** où la séquence en acides aminés, non répertoriée, est généralement retrouvée directement à partir des spectres de fragmentation MS/MS, sans recours à une base de données de séquences. Des logiciels ont été conçus pour faciliter la tâche des analystes tels que PEAKS [88].
3) **Approche hybrides**, qui combinent la recherche dans les banques de données et le séquençage *de novo*. La première étape consiste en une identification de courts *tags* peptidiques (3 à 5 acides aminés) d'après les spectres MS/MS, suivi d'une recherche de ces *tags* dans les banques de données en *error tolerant,* c'est-à-dire que la recherche accepte une ou plusieurs erreurs d'attribution. De telles recherches sont possible sur des sites accessibles au public telle que Matrix Science en utilisant l'algorithme Mascot en mode *sequence query* [78].

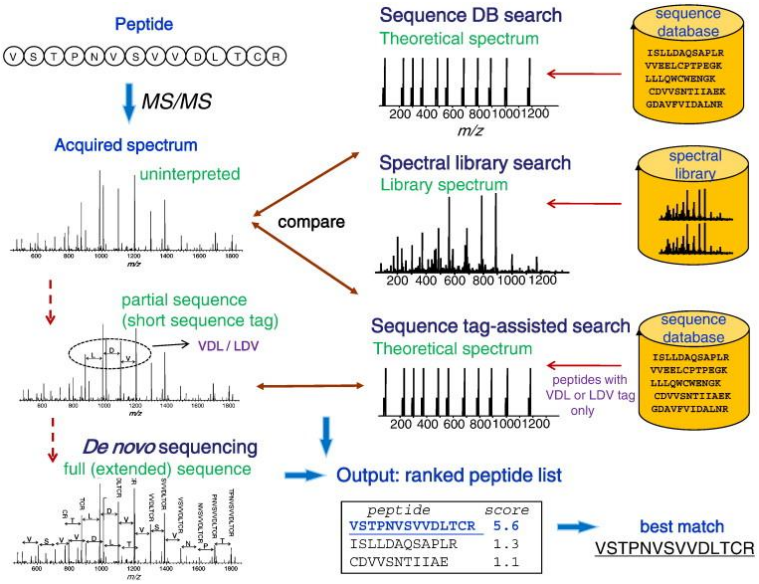


Figure 9 : Interprétation des données MS/MS. L'identification des peptides peut se faire par (1) corrélation des spectres expérimentaux avec des spectres théoriques des peptides prédits grâce à des banques de données ou à partir de spectres dans une librairie (recherche dans les banques de données), (2) en remontant la séquence manuellement à partir des spectres MS/MS (séquençage *de novo*) ou (3) en combinant le séquençage d'un petit enchainement d'acides aminés (*tag*) puis recherche dans les banques de données restreinte à ce tag (approche hybride) [77].

II.2 Recherche dans des banques de données

En pratique, les logiciels considèrent en entrée chaque spectre expérimental et lui attribuent un score en le comparant aux profils théoriques de fragmentation des peptides de la banque de données. Le nombre de peptides candidats est réduit en fonction des critères spécifiés par l'utilisateur, cités au paragraphe précédent. En sortie, les logiciels génèrent une liste de séquences peptidiques classées selon leur score de corrélation avec le spectre expérimental (Figure 9). **Ce score mesure le degré de similitude entre le spectre expérimental et le spectre théorique (Sequest) et la probabilité qu'une autre séquence de la banque puisse être utilisée pour l'interprétation de ce spectre (Mascot, Andromeda). Ceci explique pourquoi le score d'identification de Mascot pour une séquence dépend de la banque utilisée et contenant cette séquence. La tolérance de masse permet aussi de filtrer les identifications les meilleures de celles qui sont fausses** [77]. Nous reviendrons dans la suite sur les spécificités des moteurs de recherche utilisés dans le cadre de cette étude.

Bien que cette méthode soit la plus utilisée, il ne faut pas oublier qu'elle repose sur des banques de données qui contiennent parfois des erreurs ou des informations manquantes, telles que l'existence de modifications donnant lieu à différentes protéoformes de la protéine ou encore à des mutations de séquence. Cela contribue au fait qu'une partie seulement des spectres MS/MS expérimentaux soient finalement attribués à un peptide (PSM, *peptide spectrum match*).

Le choix de la banque de données est crucial. En général, les banques de protéines annotées manuellement et non redondantes, telles que Swiss-Prot [89], sont privilégiées. Lorsqu'il s'agit d'identifier des polymorphismes, de plus grandes banques de données avec des séquences redondantes peuvent être utilisées sachant que plus la banque est large plus le nombre de faux positifs peut être grand. De la même façon, avec l'amélioration des performances des instruments, le nombre d'entrées de spectres MS/MS s'est considérablement accru et le risque de faux positifs également. Des outils de statistiques sont alors nécessaires pour réduire au maximum ces faux positifs et valider une identification.

II.3 Validation des données

II.3.1 Validation par le score et le nombre de peptides

Certains groupes valident les peptides selon leur score de similitude avec les spectres théoriques. Plus ce score est élevé, plus la probabilité que l'attribution soit fausse est faible.

Pour l'algorithme de recherche Mascot [78], le score repose sur la probabilité P qu'un spectre expérimental soit attribué à un peptide dans la banque de référence (PSM) par chance. Plus cette probabilité est faible, plus l'attribution est alors significative. Pour simplifier l'interprétation, cette probabilité P est convertie en score par l'opération mathématique décrite ci-dessous (Equation1). Chaque score est également accompagné d'une valeur de probabilité d'occurrence (*expectation value ou E-value*). Cette valeur reflète le nombre PSMs attribués au hasard qui ont un score identique ou meilleur. Plus cette valeur est faible, plus l'attribution sera sure.

$$Ion\ score = -10\ log_{10}P \qquad \text{(Equation1)}$$

Concernant Sequest [79], l'algorithme sélectionne les séquences en acides aminés dans la banque sélectionnée qui coïncident avec la masse mesurée du peptide. Ensuite, il compare les ions fragments théoriques au spectre MS/MS et génère un score préliminaire pour chaque séquence. Enfin, pour les 500 meilleurs scores, il effectue une corrélation croisée en comparant les spectres théoriques aux spectres expérimentaux, donnant le score de corrélation Sequest appelé Xcorr. Un Xcorr supérieur à 2 est généralement signe d'une très bonne corrélation.

Le score d'Andromeda est proche de la logique de Mascot : il est calculé avec la même relation (Equation1) avec P la probabilité de trouver une correspondance par chance pour au

moins le même nombre de pics parmi les n masses théoriques calculées, en considérant le même nombre de pics par intervalles de masse que celui utilisé pour l'identification proposée.

En outre, **un nombre élevé de peptides identifiés par protéine** va accroître la confiance attestant de la présence de la protéine dans l'échantillon. En général, un minimum de deux peptides par protéine est requis. Dans le cadre de notre étude et afin de réaliser des diagnostiques les plus robustes possibles, nous avons considéré que trois peptides par protéines devaient être identifié de façon reproductible pour que la présence de la protéine soit retenue. L'intensité des peptides validés sera ensuite utilisée pour la classification.

II.3.2 ***Validation par le taux de faux positifs (FDR)***

L'approche *target-decoy* [90] est actuellement la plus utilisée pour valider les identifications de protéines. Les spectres expérimentaux ne sont pas seulement comparés aux spectres de la banque de protéines cible (*target database*) mais également à une banque de séquences leurre (*decoy database*). Cette banque, de même taille que la banque cible, contient des protéines fictives souvent composées de séquences inversées ou aléatoires de protéines existantes.

Cette approche suppose que les associations (PSMs) positives des spectres expérimentaux aux spectres de la banque leurre sont fausses. Le nombre de faux positifs dans la banque de données cible est ainsi estimé par le nombre de PSMs de la banque leurre. Grâce à cette approche, **le taux de faux positifs (FDR, *false discovery rate*) qui est la proportion de faux positifs PSMs parmi les vrais PSMs,** peut être calculé (Equation 2). En général, un taux de faux positifs de 1% est toléré.

$$FDR\ (\%) = \frac{False\ positive\ PSMs}{False\ positive\ PSMs + True\ positive\ PSMs} \qquad \text{(Equation 2)}$$

III. Quantification des protéines par spectrométrie de masse

L'efficacité d'ionisation et la réponse de détection dépend des caractéristiques physico-chimiques des molécules d'intérêt. Ainsi, la réponse en spectrométrie de masse diffère d'un peptide à un autre, rendant la technique non quantitative entre molécules différentes. Cependant plusieurs procédures ont été développées ces dernières années pour quantifier les variations d'expression de protéines entre deux ou plusieurs états physiologiques différents, en comparant les peptides spécifiques à une protéine donnée. Ces méthodes reposent en général sur la comparaison de l'intensité de molécule de formules développées identiques. La plupart des méthodes de quantification emploient des isotopes stables pour créer un marqueur de masse (*tag*) spécifique, qui est capable d'être reconnu et quantifié en spectrométrie de masse [91]. Ces marqueurs isotopiques peuvent être introduits à différents niveaux de l'expérience (Figure 10). D'autres techniques de quantification sans marquage (*label-free*) sont d'un intérêt particulier pour ce projet puisqu'elles ne nécessitent aucun réactif de marquage parfois couteux, et ne sont pas limitées en termes de nombres d'échantillons.

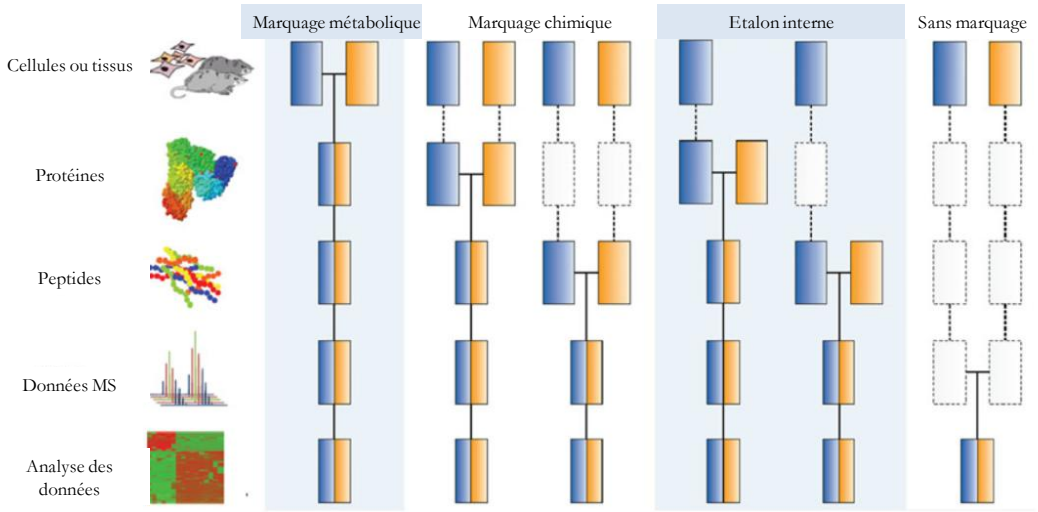


Figure 10 : Procédures classiques de quantification en spectrométrie de masse. Les différentes couleurs représentent deux conditions expérimentales différentes. Les lignes représentent les différents niveaux de l'expérience *bottom-up*. Les traits discontinus montrent les étapes auxquelles des variations expérimentales peuvent engendrer une erreur sur la quantification finale des protéines. Le risque d'erreur augmente de l'expérience de gauche (marquage métabolique) à l'expérience de droite (sans marquage) (figure adaptée de Bantscheff [92]).

III.1 Quantification relative par marquage isotopique

III.1.1 Marquage métabolique

En 2002, l'équipe de Ong [91] développe une stratégie de marquage isotopique des acides aminés dans des cellules, appelée SILAC (*stable isotope labelling by amino acids in cell culture*), qui consiste à mettre en culture des cellules d'intérêts et des cellules contrôles, dans deux conditions différentes, dans un environnement aérobique et à température contrôlés (en général en 5% CO_2 et 37°C) : l'une avec un milieu de culture contenant des acides aminés normaux (isotopes légers), l'autre contenant des acides aminés marqués par des isotopes stables : le carbone 13 et l'azote 15 (isotopes lourds). En protéomique classique, il s'agit souvent du couple d'acides aminés lysine et arginine, sites basiques de coupure tryptique. Pour incorporer ces acides aminés et disposer de cellules sous une condition unique [93], un nombre minimum de passages est nécessaire (généralement 6 passages). Chaque type de cellule présente donc une composition déterminée en acides aminés (lourd ou léger). Un même peptide issu de chaque échantillon aura la même structure (et donc les mêmes caractéristiques physico-chimiques en première approximation) et pourra être quantifié grâce à la différence d'abondance de 2 espèces comparables et différenciées par la différence de masse entre les isotopes lourds et légers qui seront co-élués en chromatographie (Figure 11). Le mélange des deux échantillons étant réalisé très tôt dans la procédure, toutes les erreurs dues à la manipulation, les conditions opératoires de la chromatographie ou la mesure de masse, se répercutent de la même manière sur les deux conditions. Le rapport d'intensité des signaux MS des deux espèces donne une information sur la variation d'expression des protéines d'une condition par rapport à une autre.

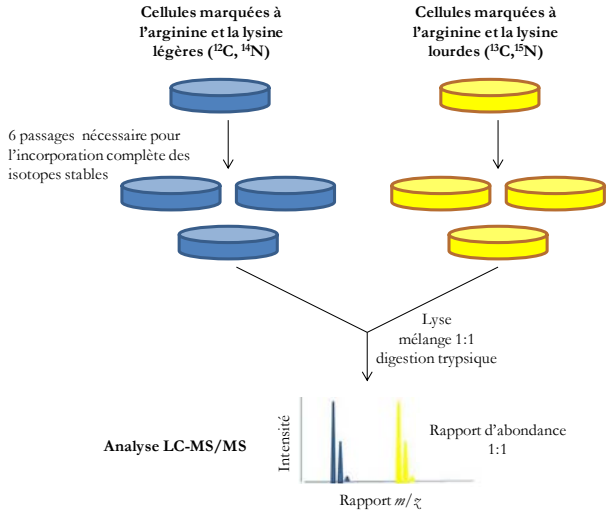


Figure 11 : Protocole SILAC. Chaque lignée cellulaire est cultivée dans un milieu contenant exclusivement les résidus arginine et lysine légers (^{12}C, ^{14}N) ou lourds (^{13}C, ^{15}N). Au bout de six passages, chaque cellule contient des acides aminés d'une condition donnée. Ces cellules sont ensuite mélangées en quantité équivalente avant l'analyse protéomique *bottom-up*.

En 2010, la même équipe étend cette application aux tissus, nommée super SILAC, qui permet d'éviter le marquage métabolique des animaux [94]. Dans ce cas, un mélange de plusieurs lignées cellulaires marquées SILAC du même organisme avec des acides aminés lourds est utilisé comme standard interne et, est rajouté au broyat tissulaire. La quantification se déroule alors en deux étapes. Tout d'abord, les protéines des tissus sont quantifiées par rapport au standard. Puis, la quantification entre tissus est calculée en comparant leurs rapports SILAC obtenus à la première étape. L'emploi d'un mélange de plusieurs et non d'une seule lignée cellulaire permet d'avoir une représentation plus complète du protéome d'un organisme engendrant une meilleure quantification.

Ce marquage SILAC suppose que les cellules soient sous une seule condition donnée, or dans la réalité, ce n'est pas toujours le cas. Cela mène alors à des quantifications erronées. De plus, elle n'est accessible qu'aux cellules capables de se multiplier dans un milieu modifié et de métaboliser les acides aminés fournis. En outre, seules 3 conditions peuvent être comparées.

*III.1.2 **Marquage chimique ou enzymatique***

Les protéines et peptides peuvent être marqués via différentes réactions enzymatiques ou chimiques, résumées au Tableau 2. La plus ancienne est l'incorporation de un atome ^{18}O dans le groupement carboxyle C-terminal durant la digestion enzymatique en eau lourde ($H_2{}^{18}O$) pour le séquençage *de novo* [95]. Cette technique a été employée plus tard pour comparer deux protéomes d'échantillons différents, l'un réalisé dans des conditions de digestion normales et l'autre dans de l'eau lourde créant une différence de masse de 4 daltons, aisément mesurable en spectrométrie de masse [96]. Cependant pour les molécules les plus lourdes, cet écart de masse peut être insuffisant car se superposent les isotopes M+2 de l'espèce ^{16}O et M de l'espèce ^{18}O.

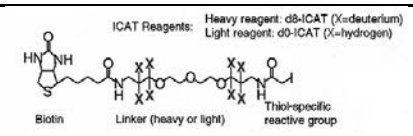
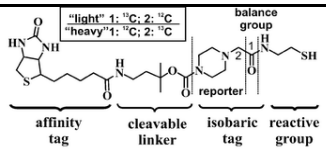
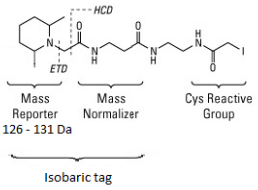
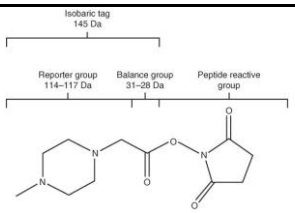
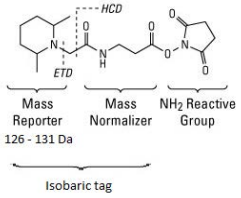
Des marqueurs de masses isotopiques peuvent également être introduits par réaction chimique sur différents acides aminés. Dans une première approche par l'équipe d'Aebersold en 1999, les groupements thiols des chaines latérales des cystéines ont été modifiés par un réactif deutéré appelé ICAT (*Isotope-Coded Affinity Tag*), pour comparer deux protéomes entre eux [97]. Le réactif était constitué d'un groupement réagissant sur les résidus cystéines, un bras espaceur contenant zéro ou huit deutériums et un groupement biotine permettant la purification de ces peptides par affinité sur une colonne de chromatographie d'affinité avidine-biotine. Cependant, le deutérium ayant une interaction avec la phase stationnaire de la chromatographie, il provoque un retard du temps de rétention des peptides par rapport à leurs homologues non deutérés. Une seconde génération avec un marquage au carbone 13,

appelée cICAT (*cleavable ICAT*), a donc été développée. Il comprend un bras espaceur clivable en milieu acide permettant de se débarrasser de la partie biotinylée et d'obtenir des fragmentations de meilleure qualité [98]. Plus récemment, un autre réactif réagissant également sur le thiol a été développé par l'équipe de Thomson en 2003, nommé l'iodoTMT (*iodoacetamide-Tandem Mass Tags*) permettant de comparer plus de deux protéomes à la fois [99]. Le réactif est lui aussi composé de trois parties, avec une partie réactive envers les résidus cystéines, un bras espaceur qui va servir de balance de masse et un groupement rapporteur marqué ou non. Contrairement à l'ICAT et cICAT où la quantification relative se fait par intégration du signal MS des isotopes lourds versus légers co-élués, les marqueurs de masse (*mass tags*) apportés par les réactifs IodoTMT ont tous la même masse et ne sont donc pas différentiables en MS. La différence de masse ne sera visible qu'après fragmentation MS/MS. La quantification relative s'effectue alors sur les ions rapporteurs spécifiques à chaque condition, de masse allant de m/z 126 *à* 131 suivant les réactifs utilisés.

Cependant, étant donné que le résidu cystéine est de fréquence assez modérée, il n'est pas présent dans tous les peptides issus de la digestion. Ainsi, d'autres marqueurs comme l'iTRAQ (*isotope Tags for Relative and Absolute Quantification*) [100] ou le TMT (*Tandem Mass Tag*) [101] ont été développés. Dans ces cas, les peptides de chaque échantillon sont marqués au niveau des amines primaires avec un réactif spécifique après la digestion enzymatique, puis les peptides des différents échantillons sont mélangés ensemble avant l'analyse en spectrométrie de masse. Chaque condition est différenciée grâce à l'ion rapporteur après fragmentation MS/MS.

Les méthodes iTRAQ, TMT et leurs variantes sont des méthodes sensibles. Cependant les ions rapporteurs ont des rapports de basse masse allant de m/z 113 à m/z 131, ce qui implique que ces réactifs ne sont pas utilisables avec des analyseurs de masse de type piège à ions 3D et 2D qui présentent un *cut-off* des basses masses (Annexe 4).

Tableau 2 : Récapitulatif des réactifs pour le marquage chimique des protéines et peptides

Nom	Site de marquage	Réactif
^{18}O	Groupement carboxyle terminal	$H_2{}^{18}O$
ICAT	Thiol	
cICAT	Thiol	
IodoTMT	Thiol	
ITRAQ	Amine primaire	
TMT	Amine primaire	

III.2 Quantification absolue par marquage isotopique

Les méthodes décrites dans les paragraphes précédents permettent de connaitre l'abondance relative des peptides/protéines d'une condition par rapport à une autre, mais ne permettent pas de réaliser une quantification absolue, c'est à dire de donner une concentration exacte de l'analyte dans l'échantillon.

Pour remédier à ce problème, plusieurs méthodes ont été développées ces dernières années. En rajoutant une quantité connue d'étalon interne au mélange étudié, il est possible de faire de la quantification absolue par spectrométrie de masse. Cet étalon peut être introduit à différents stades de l'analyse protéomique résumés sur la Figure 12 et détaillés dans les prochains paragraphes: S'il intervient au niveau de la protéine, on parlera de PSAQ (*Protein*

Standard Absolute Quantification) [102]; avant la digestion enzymatique, de QconCAT (*Quantification conCATamer*) [103] et enfin au niveau peptidique, du modèle AQUA (*Absolute QUAntification of proteins*) [104].

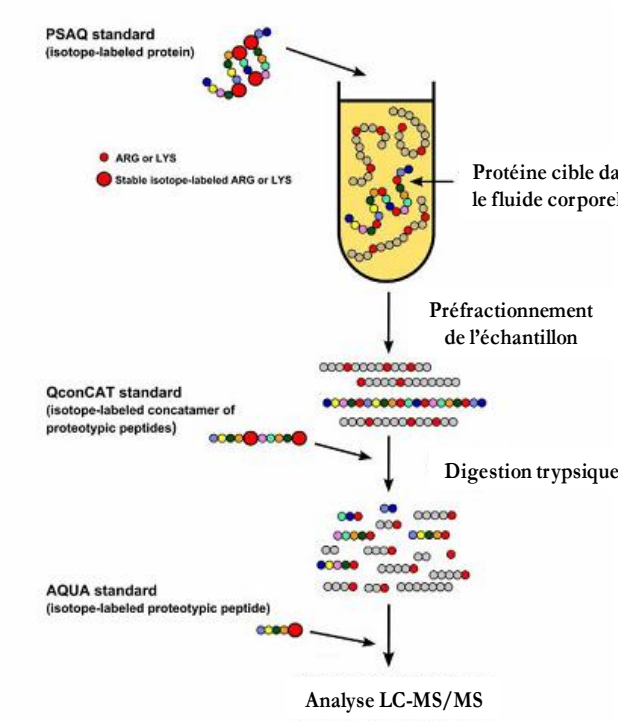


Figure 12 : Stratégie de quantification par dilution isotopique. Trois types de standards peuvent être utilisés pour la quantification absolue des protéines dans les matrices biologiques: (1) le standard PSAQ (Protein Standard Absolute Quantification) qui est une protéine recombinante marquée, est directement introduit dans l'échantillon en amont du protocole protéomique, (2) le standard QconCAT (Quantification conCATamer) qui est une protéine chimérique contenant un ou plusieurs peptides protéotypiques marqués, est introduit avant l'étape de digestion pour libérer le(s) peptide(s) protéotypique(s) d'intérêt, (3) les peptides AQUA (Absolute QUAntification of proteins) qui sont des peptides synthétiques marqués, analogues aux peptides d'intérêt, introduits avant l'analyse en LC-MS (figure adaptée de Brun *et al.* [102]).

L'étalon interne idéal doit avoir des caractéristiques physico-chimiques très proches de l'analyte à doser. Le meilleur étalon est l'analogue de l'analyte, marqué par des isotopes stables, tels que le carbone 13 ou l'azote 15. La différence de masse apportée par ces isotopes permet de les distinguer en spectrométrie de masse (Figure 13). Les propriétés physico-chimiques sont très proches. La concentration de l'analyte sera alors donnée par l'équation (Equation 3):

$$Concentration_{analyte\ engogène} = \frac{XIC_{analyte\ endogène}}{XIC_{analyte\ marqué}} \times Concentration_{analyte\ marqué} \qquad \text{(Equation 3)}$$

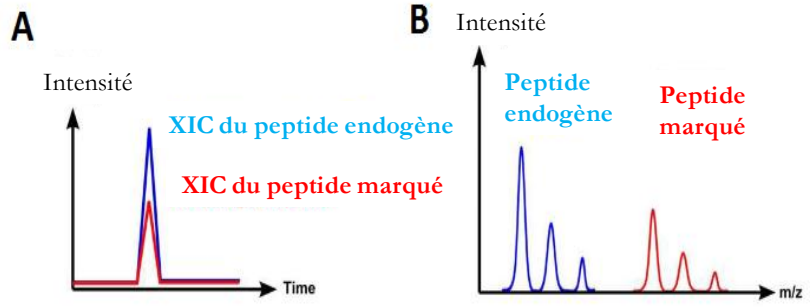


Figure 13 : Quantification absolue des protéines. (A) Le peptide endogène et le peptide marqué sont co-élués en chromatographie. (B) Les deux peptides sont différentiables par spectrométrie de masse. La quantification se fait par comparaison de l'extraction d'ion du chromatogramme (XIC) du peptide endogène versus peptide marqué (figure adaptée de Brun *et al.* [102]).

III.2.1 Absolute quantification of proteins (AQUA)

La méthode AQUA a été introduite par l'équipe de Gygi en 2003 [104]. Elle consiste à rajouter au mélange, après digestion enzymatique, une quantité connue de peptide protéotypique (spécifique à une protéine) synthétique marqué (^{13}C et/ou ^{15}N), analogue au peptide endogène. La quantification se fait par comparaison du signal du peptide endogène avec le même peptide marqué. Pour obtenir plus de précision sur la mesure ainsi qu'une bonne sensibilité et gamme dynamique de détection, la quantification absolue à l'aide de peptide AQUA est réalisée en mode SRM (*selected reaction monitoring*) (cf. chap.3 §V.2.2.1).

Contrairement au marquage métabolique où le marquage se fait sur la totalité du protéome, cette méthode ne s'intéresse qu'à un nombre particulier de protéines. Elle est très utilisée pour des approches de validation de biomarqueurs potentiels sur une grande cohorte d'échantillons biologiques. Cette méthode nécessite bien sûr une parfaite maitrise de la concentration des peptides de références qui ne sont pas exempts de dégradation ou de perte par collage non spécifiques sur les divers récipients intervenant dans la manipulation. Par ailleurs, cette méthode permet de s'affranchir des bais après la digestion mais pas de ceux précédant cette étape, notamment au niveau de la digestion même. Elle souffre alors d'un biais due aux variations de rendement de digestion qui font varier la concentration des peptides tryptiques à concentration de protéine égale. Il faut alors maîtriser la digestion et corriger la quantification en aval lorsqu'elle n'est pas totale, ce qui est bien trop souvent le cas.

III.2.2 Quantification concatamer (QconCAT)

En 2005, l'équipe de Prat [103] développe une méthode complémentaire à l'AQUA permettant de suivre plusieurs protéines. Plusieurs peptides marqués au sein d'une même protéine ou de protéines différentes, sont concaténés pour former le standard QconCAT, réduisant ainsi le nombre de peptides à introduire. Cette protéine chimérique est alors introduite avant la digestion pour libérer des peptides AQUA marqués tout en corrigeant les biais de la digestion, sans toutefois les corriger totalement. En effet, cette protéine chimérique étant différente de celle étudiée, elle peut se comporter différemment lors de la digestion enzymatique.

III.2.3 Protein standard absolute quantification(PSAQ)

En 2009, l'équipe de Brun [102], va encore plus loin en employant des protéines recombinantes entièrement marquées. Elles sont introduites en amont de la préparation

réduisant alors tout biais de la digestion ainsi que de la purification ou du fractionnement. Cependant l'obtention de ces protéines recombinantes purifiées est assez laborieuse et couteuse.

III.3 Quantification relative sans marquage (label-free)

Le marquage par dilution isotopique est la référence standard pour la quantification. Cependant, elle nécessite une ou plusieurs étapes de préparation additionnelles et peut s'avérer couteuse. Ainsi, la quantification sans marquage, dite *label-free*, a été exploitée ces dernières années. Du fait de l'absence de réactif de marquage, cette quantification permet de comparer un nombre infini d'échantillons quelle que soit leur nature (tissu, cellule). Cependant pour que cette quantification soit la plus représentative possible, il est impératif de contrôler chaque étape pour limiter les différents biais dus à la manipulation des échantillons. Il est alors important de limiter le nombre d'étapes de préparation des échantillons et si possible de les préparer en parallèle. Les échantillons sont ensuite analysés séparément ce qui réduit la complexité intrinsèque du mélange analysé par rapport aux échantillons mélangés après marquage différentiel.

Plusieurs méthodes pour la quantification sans marquage existent, et sont illustrées sur la Figure 14 : le décompte des spectres MS/MS nommée *spectral count*, et la mesure sous l'aire des pics du signal MS, nommée AUC (*Area Under the Curve*) ou XIC (*eXtracted-ion Chromatogram*).

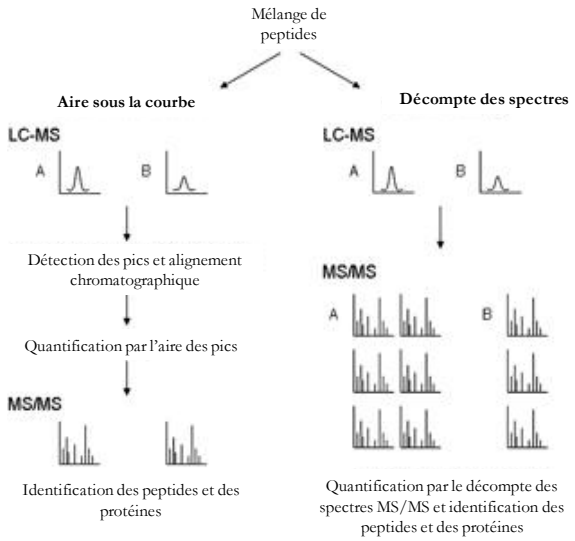


Figure 14 : Schéma des deux principales approches en quantification sans marquage de deux échantillons présentant un changement d'expression de 2 unités. A) La quantification par mesure d'aire sous le pic (AUC area under the curve) est réalisée en MS et les peptides sont identifiés par fragmentation MS/MS. B) La quantification par décompte des spectres (spectral count) et l'identification sont réalisées simultanément en fragmentation MS/MS (figure adaptée de Neilson et al. [105]).

III.3.1 ***Décompte des spectres MS/MS (spectral count)***

Cette méthode empirique repose sur un constat simple : plus la protéine est abondante dans l'échantillon, plus le nombre de peptides protéolytiques détectables de cette protéine est important et plus le nombre de spectres MS/MS assignés à cette protéine sera élevé [106].

Ainsi, pour évaluer la quantité de protéine dans l'échantillon, plusieurs calculs se basant sur ce décompte des spectres MS/MS ont ainsi été ensuite développés. Tout d'abord l'équipe de Mann [107] introduit l'index d'abondance de la protéine, PAI (*protein abundance index*) définit par l'équation suivante :

$$PAI = \frac{nombre\ peptides\ observés}{nombre\ peptides\ observables\ dans\ la\ gamme\ de\ masse\ du\ spectrometre} \quad \text{(Equation 4)}$$

Une version améliorée de cette dernière méthode, appelée emPAI (*exponentially modified Protein Abundance Index)* avec $emPAI=10^{PAI}-1$, a été introduite plus tard par ce même groupe [108]. D'autres équipes proposent d'améliorer cette méthode en rajoutant un facteur de corrélation, reposant sur la probabilité d'observer le peptide, nommée APEX (*Absolute Protein EXpression profiling*) [109] ou en prenant en compte le signal global et la longueur de la protéine, nommée SI_N (*normalized Spectral Index*) [110].

Etant donné que ces échantillons sont analysés à des temps différents et indépendamment les uns des autres, il n'est pas rare d'observer des variations entre les analyses, même dans le cas de réplicas techniques. Ces variations notables touchent la quantité exacte injectée, le gradient et les solvants chromatographique ou encore le caractère stochastique du déclenchement des événements MS et MS/MS. Une voie simple pour prendre en compte ces variations est de normaliser les données par le nombre total de spectres MS/MS, nommée NSAF (*normalized spectral abundance factor*) [111].

Toutes ces méthodes permettent de quantifier les protéines dans les mélanges biologiques à condition d'avoir suffisamment de données (Figure 15). Or, la méthode d'acquisition employée dans la plupart des laboratoires utilise une exclusion dynamique pendant un intervalle de temps précis, des ions déjà fragmentés (cf. chap3 §V.1.2), qui tend à sous-estimer la quantité des protéines les plus abondantes. Le nombre de points de données est également influencé par la largeur du pic chromatographique et la rapidité d'acquisition du spectromètre de masse.

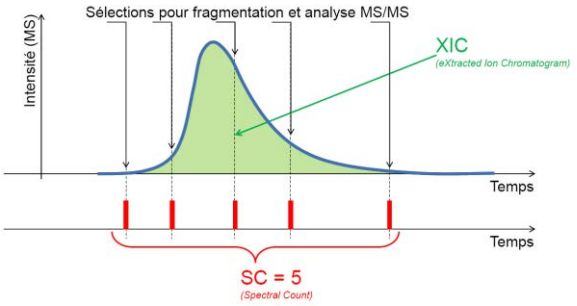


Figure 15 : Illustration des deux principales approches de la quantification sans marquage. La quantification par mesure d'aire sous le pic est plus représentative que le décompte des spectres MS/MS (Ecole thématique CNRS, FTMS 2014, Burger).

III.3.2 ***Mesure d'aire sous le pic (AUC)***

Cette méthode repose sur une linéarité de la réponse en spectrométrie de masse en fonction de la concentration du peptide donné. Ainsi l'intégration des pics chromatographiques pour des rapports m/z donnés en MS, correspondant à la masse des peptides ionisés (XIC), sera prise en compte pour la comparaison entre les échantillons. L'abondance des protéines sera ensuite calculée en prenant la moyenne ou la somme des aires des peptides. La plus utilisée est la méthode T3PQ (*Top 3 Protein quantification*) développée par l'équipe de Silva en 2006 [112], qui permet de corréler la moyenne des trois pics les plus intenses à l'abondance de la protéine dans l'échantillon. Puisque la quantification repose sur l'intégration d'un rapport m/z, à partir du signal MS, un spectromètre de masse de haute résolution et de bonne justesse de masse (HR/AM) est donc nécessaire pour éviter toute erreur de quantification due à des interférences. Pour que le pic chromatographique soit décrit le mieux possible, il faut enregistrer le maximum de spectres MS possible. Or, les méthodes LC-MS/MS actuelles alternent entre les spectres MS pour la mesure exacte de la masse des peptides et les spectres de fragmentation MS/MS de ces mêmes peptides pour leur séquençage (cf. chap3 § V.1.2). Il faut alors trouver un compromis entre ces deux expériences pour obtenir une bonne quantification en MS tout en maximisant la couverture du protéome, en accumulant les spectres MS/MS. Nous discuterons plus en détail de ces paramètres dans la partie décrivant les spectromètres de masse utilisés. D'autres alternatives au DDA existent pour augmenter la couverture du protéome sans sacrifier la quantification telles que l'AMRT (*Accurate Mass, Retention Time*) [113], qui alterne une analyse en mode MS/MS pour l'identification des peptides qui sera stockée dans une base de données et une analyse en mode MS pour la quantification. Un logiciel va ensuite corréler la masse et le temps de rétention à la base de données créée par les spectres MS/MS. Une autre alternative est la méthode MS^E (*elevated energy mass spectrometry*) qui consiste à enregistrer le signal MS à faible énergie de collision puis à plus haute énergie de collision [101]. Un traitement logiciel est ensuite

nécessaire pour associer les fragments aux précurseurs par corrélation avec les temps de rétention. Cependant, dans tous les cas de mesure d'aire sous le pic, un alignement des chromatogrammes est nécessaire pour comparer les peptides entre les analyses.

Depuis peu, une autre méthode fait son apparition dans les laboratoires de recherche. Il s'agit du DIA ou SWATH™ (ABSciex)[114]. L'intérêt de cette méthode réside dans le fait que tous les précurseurs sont fragmentés après sélection dans des séries fenêtres de rapport *m/z* assez larges (en général de 25 Da), et sans a priori, par rapport à une liste comme dans les méthodes ciblées classiques (SRM et PRM). Ainsi, plus de protéines peuvent être quantifiées. De plus, cette méthode est très simple à mettre en place. Enfin, avec cette méthode, une protéine peut être quantifiée *a posteriori* puisqu'il n'y a pas de sélection de précurseurs à fragmenter. La quantification repose alors sur l'intégration des pics MS/MS à l'aide d'une librairie de spectres.

Comme dans le cas du *spectral count*, ces échantillons sont analysés à différents temps, engendrant des variations d'une analyse à l'autre. Une manière simple pour prendre en compte ces variations consiste à normaliser les données par le courant ionique total (*TIC*).

III.4 Conclusions

Nous avons vu qu'il existe deux grandes catégories de méthodes de quantification résumées en Figure 16: la quantification avec marquage isotopique stable et la quantification sans marquage.

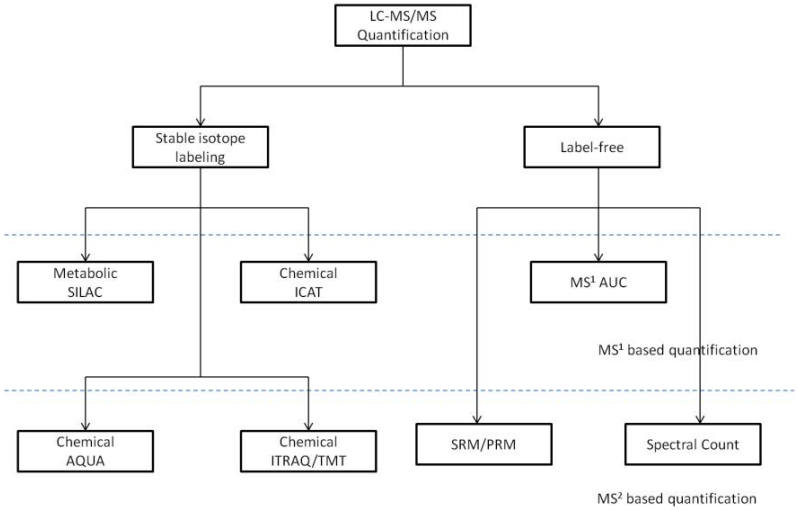


Figure 16 : Schéma récapitulatif des principales méthodes de quantification actuellement utilisées dans les laboratoires de protéomique.

Dans le cas d'une analyse protéomique *bottom-up* **d'un mélange complexe biologique, des milliers de peptides sont générés** après digestion. Pour simplifier ce mélange, une

séparation chromatographique en amont est nécessaire (cf. chap3 § IV) mais sa capacité de pic est en général insuffisante au regard du nombre de peptides présents. La deuxième dimension, la spectrométrie de masse, est alors nécessaire pour séparer ces peptides co-élués. Le temps que met l'instrument pour analyser ces peptides (temps de cycle) devient alors le facteur limitant de l'analyse.

Or, il s'avère que dans le cadre des échantillons biologiques complexes, la quantité de peptides co-élués dans le mélange est telle que le spectromètre de masse, malgré les avancées techniques, n'a pas le temps de tous les analyser, menant à un sous échantillonnage du mélange. De plus, la présence de peptides quasi-isobares, qui seront potentiellement sélectionnés simultanément pour l'analyse en MS/MS, conduit à une superposition des spectres de fragmentation de différents précurseurs peptidiques dans le même enregistrement. L'interprétation du spectre est alors plus délicate et la robustesse d'interprétation en est diminuée. Cet **effet de masse est d'autant plus important dans le cas d'analyse d'échantillons marqués** (SILAC, multiplexe-TMT/ITRAQ/ICAT) puisque les n échantillons, chacun sous une condition donnée, sont injectés ensemble pour l'analyse LC-MS/MS, ce qui augmente de n fois la complexité du mélange [115]. En outre, le marquage métabolique SILAC ne peut être réalisé sur nos biopsies amyloïdes même si l'alternative pour les tissus existe (super SILAC). En effet, dans ce cas-là, il faut savoir quel type de cellules de référence prendre, sachant que nous analysons des tissus de différents organes. Le marquage chimique après digestion serait envisageable si nous pouvions avoir des contrôles pour réaliser l'analyse différentielle. Et même si cela était possible, quel contrôle choisir pour la quantification relative des protéines? De plus, dans le cadre de mélanges complexes la sélection spécifique d'un peptide isolé est peu probable. Le bruit de fond de l'analyse est en partie le résultat d'un bruit chimique provenant notamment de multiples espèces peptidiques en faible abondance. Etant donné que tous les peptides protéolytiques sont marqués c'est-à-dire que leur fragmentation, indépendamment de leur séquence, génère les mêmes ions rapporteurs, ces quantités massives de peptides marqués risquent de contaminer les spectres MS/MS avec les marqueurs de masse tout au long de l'analyse et cet effet est additionnel sur l'intensité des rapporteurs. C'est ainsi que le signal des rapporteurs provenant des peptides mineurs et non identifiés se retrouve parfois non négligeable voire prépondérant devant le signal des rapporteurs provenant du signal du peptide d'intérêt même. Dans ce cas, la quantification est fausse. Ceci sera d'autant plus vrai que la fenêtre de sélection est large. Une fenêtre de sélection étroite fera à son tour chuter la sensibilité de l'analyse. La quantification absolue par marquage isotopique (AQUA, QconCAT, PSAQ) pourrait être utilisée mais le

nombre de protéines à suivre est tel que le cout des réactifs devient un obstacle à une analyse en routine. De plus, dans le cas des réactifs PSAQ, la meilleure représentante de la protéine d'intérêt, l'obtention de la protéine pure est laborieuse.

Malgré une variabilité inter-échantillon plus grande, nous nous sommes alors orientés vers une quantification sans marquage tout en gardant à l'esprit que le problème de normalisation est aussi présent. En effet, dans le cas des marquages, la normalisation se fait en prenant la moyenne ou médiane globale des valeurs logarithmiques des ratios d'abondance relative de l'ensemble des peptides et, dans le cas du *label-free*, une normalisation similaire peut être réalisée avec le TIC global. Ces deux normalisations supposent que les changements d'expression de protéines sont mineurs et n'affectent pas les logs des ratios ou TIC globaux, ce qui n'est pas toujours vrai [115] dans le cas de nos prélèvements.

Parmi les méthodes *label-free* existantes, nous avons choisi dans une première étape pour l'analyse *shotgun* de mettre en place la méthode reposant sur la mesure d'aire sous le pic (MS^1 AUC). Elle a l'avantage de mieux décrire le pic chromatographique et donne donc une meilleure quantification que le *spectral count* (Figure 15) avec nos setups expérimentaux où nous cherchons à minimiser le nombre d'analyse MS/MS redondantes.

Un autre problème commun à toutes les quantifications décrites ci-dessus est la **grande gamme dynamique des protéines dans les échantillons biologiques**. Or, la plupart des méthodes employées pour ces analyses à grande échelle (*shotgun*) se base sur la sélection des ions les plus intenses. Un sous-échantillonnage des espèces faiblement abondantes a donc également lieu.

Pour limiter l'effet de masse et gamme dynamique limitée, nous proposons ensuite une analyse ciblée (SRM/PRM) des peptides d'intérêt, moins discriminante en vers les protéines abondants et plus sensible et spécifique. Cependant, cela nécessite de connaître les cibles au préalable pour d'établir une liste avec leurs paramètres (m/z, état de charge, temps de rétention chromatographique).

Chapitre 3 : La spectrométrie de masse

I. Historique

Historiquement, l'analyse du protéome utilisait la technique de gel bidimensionnel (2DGel) suivi d'une détection de protéines par des moyens colorimétriques, marquage par fluorophores ou encore mesure de radioactivité. Bien que ces méthodes de détection permettent de visualiser des milliers de protéines, elles souffrent d'un certain nombre de limitations (sensibilité pour le bleu de Coomassie, défaut de linéarité pour le nitrate d'argent). De plus, sauf protocoles particuliers, seules les protéines solubles sont visualisables sur le gel, les protéines membranaires demandant une mise au point complexe pour être analysées [116]. En outre, il n'y a aucune identification de la (les) protéine(s) sous-jacente(s), et les gels 2D ont bien souvent une répétabilité limitée [117,118]. Les marquages par fluorescence et radioactivité sont extrêmement sensibles mais limitent l'analyse *a posteriori*.

C'est pourquoi des techniques complémentaires telles que la spectrométrie de masse, ont connu un essor fulgurant dans le domaine de la protéomique avec l'avènement dans les années 80 de deux nouvelles techniques d'ionisation douce capables de mesurer la masse des grosses molécules: l'ionisation électrospray (ESI) [119] et la désorption-ionisation laser assistée par matrice (MALDI) [120]. La spectrométrie de masse est aujourd'hui un outil incontournable pour caractériser les biomolécules de façon sensible et universelle. Ces deux modes d'ionisation ont valu le prix Nobel en 2002 à J.B. Fenn et K. Tanaka. Seul le processus d'ionisation ESI, utilisé au cours de ce projet sera détaillé dans ce manuscrit.

Désormais, les spectromètres de masse permettent d'analyser des échantillons biologiques complexes, pouvant contenir des milliers de protéines, dont certaines sont présentes en faible quantité. Cependant, l'étude des systèmes biologiques représente un défi technique d'actualité pour la spectrométrie de masse. En effet, d'après le groupe de Bantscheff [121], du fait de l'étendue de la gamme dynamique des protéines cellulaires, seulement une fraction des protéines présentes dans l'échantillon est identifiée, dont une partie minime est quantifiable à cause de la qualité des spectres, du faible rapport signal sur bruit (S/B), de la résolution ou encore de la suppression spectrale (Figure 17). Aujourd'hui, l'enjeu analytique est d'augmenter la couverture du protéome.

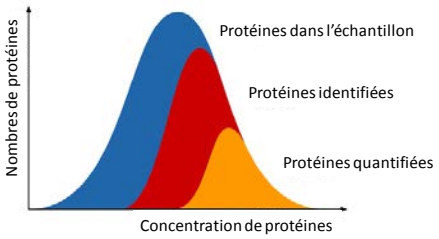


Figure 17 : Représentation schématique des fractions du protéome (en bleu) qui peuvent être identifiées (en rouge) et quantifiées (en orange) par spectrométrie de masse. Etant donné la grande gamme dynamique des protéines cellulaires, seule une partie est identifiée et quantifiée (figure adaptée de Bantscheff et al. [121]).

II. Principe

Tous les spectromètres de masse sont constitués de 4 éléments : (1) une **source d'ionisation** pour produire des ions en phase gazeuse, (2) un **analyseur de masse** pour séparer les ions produits selon leur rapport masse sur charge (m/z), (3) un (ou plusieurs) **détecteur** qui enregistre les ions séparés par l'analyseur de manière à produire un courant électrique dont l'intensité est corrélée au nombre d'ions détectés, et enfin(4) un **traitement de signal** pour obtenir un spectre de masse d'intensité en fonction du rapport m/z.

L'analyseur et le détecteur constituent un même élément dans le cas des analyseurs à haute résolution à piège d'ions et analyse par transformée de Fourier (FT, *Fourier transform*) de type résonance cyclotronique d'ions (ICR) et piège à ions orbitalaire (Orbitrap).

II.1 Source d'ionisation: NanoElectrospray (NanoESI)

Les cibles d'étude en protéomique ne sont souvent malheureusement présentes qu'à de très faibles concentrations (sub fentomoles) et de très faibles quantités (microlitres). Ainsi de nombreux efforts ont été menés par différents équipes de recherches afin d'augmenter la sensibilité de la méthode ESI (principe rappelé en Annexe 1). Tous les efforts ont été motivés par le fait **qu'une partie infime des ions analytes parvient au détecteur**. **La perte d'information a lieu dans l'otique ionique (transfert et analyse) mais aussi à cause de la désolvatation incomplète des gouttelettes et de la faible transmission de ces ions après production à pression atmosphérique vers une région à très basse pression.** Pour diminuer ces pertes, l'évaporation des solvants (chauffage, gaz) et le guidage des ions (optique, géométrie) sont continuellement optimisés dans les développements instrumentaux [122].

En plus des modifications à l'intérieur du spectromètre de masse, une miniaturisation de l'ESI même a été entreprise pour gagner en sensibilité dans les années 90 [123–126]. L'abaissement du débit améliore fortement l'efficacité d'ionisation. En effet, la formation d'un plus petit cône de Taylor et donc la production de gouttelettes chargées plus petites

permettent d'avoir une meilleure efficacité de désolvatation. Les gouttelettes étant plus petites, le processus d'évaporation est plus efficace, et le capillaire peut être placé à une distance plus proche de l'orifice d'entrée dans l'analyseur (1-2mm). Ceci a pour conséquence une plus grande transmission des ions vers l'analyseur par rapport à l'ESI conventionnel. Une meilleure efficacité d'ionisation est également observée grâce à une charge par unité de volume plus élevée. Enfin, la réduction du débit de travail engendre une diminution de la consommation de solvant et une diminution du flux gazeux dans les analyseurs. Ce dernier point est un avantage pour coupler ce type de source qui travaille à pression atmosphérique à un analyseur à très basse pression (10^{-10} mbar) tel que la cellule ICR (cf. Annexe 5).

Un petit volume d'échantillon (quelques µL) est introduite dans un capillaire en général de silice ou métal de faible diamètre interne (10µm en général), permettant d'atteindre de faibles débits, de l'ordre du nl.min^{-1}. Ce capillaire est recouvert ou non de matériau conducteur pour appliquer la tension de l'ordre de 1 à 2 kV.

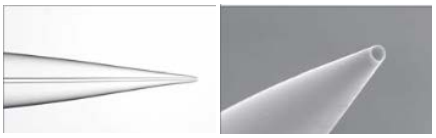


Figure 18 : Photographie d'un capillaire nanoESI (SilicapTip™ , New Objective).

II.1.1 ***Des capillaires NanoESI aux émetteurs NanoESI sur puce***

Certains paramètres, comme la position de l'aiguille ou encore le spray, varient d'une expérience à une autre. Depuis quelques années, un effort est mené pour développer des systèmes multi-émetteurs ESI, plus robustes et plus sensibles, permettant de changer d'émetteur rapidement quand l'un d'eux se bouche ou lorsque le spray devient instable [127,128].

Un des systèmes aujourd'hui mis sur le marché a été testé au cours de ma thèse et désormais installé sur le spectromètre de masse LTQ FT Ultra (TriVersa Nanomate, Advion). La puce est fabriquée à base de silice par gravure ionique réactive et se compose d'un panel de 400 émetteurs nanospray robustes et fiables. Différents diamètres internes (5µm à 2 µm) sont proposés pour des débits de 30 à 300nl/min (Figure 19). Ce système automatique permet de s'affranchir des variations engendrées par les capillaires classiques individuels. Des cônes conducteurs de courant pour l'infusion direct et un *LC–Chip coupler*, un coupler conducteur pour la chromatographique, sont employés pour introduire l'échantillon dans les émetteurs.

En ESI ou nanoESI conventionnel, le spray est obtenu par application d'une différence de potentiel entre le fluide et l'entrée du spectromètre de masse, la contre-électrode (Figure 20 A). Pour les puces nanospray, une différence de potentiel est également appliquée mais elle a lieu entre le fluide, via le cône conducteur, et les émetteurs nanoESI, la contre-électrode. Le spray peut donc être établi à distance du spectromètre de masse (Figure 20 B). Cette configuration particulière permet d'obtenir une meilleure sensibilité de détection [128]. En effet, la distance entre l'origine du potentiel et la contre électrode est de quelques microns, soit 1000 fois plus petite qu'en ESI conventionnel, créant un champ électrique plus intense et donc plus efficace (Figure 21).

De plus, la fonction "*spray sensing*" mesure en temps réel le courant d'ion intrinsèque aux bornes de la puce et permet d'avoir le contrôle sur le spray d'ionisation et de passer d'un émetteur nanoESI à un autre lorsque le courant d'ions est anormal, en moins de 5 secondes, ce qui réduit le risque de perte d'échantillon. En outre, ce système facilite le passage d'une analyse en mode d'injection direct à une analyse en mode indirect (couplage avec la nano-chromatographie), réalisable en moins de 5 min.

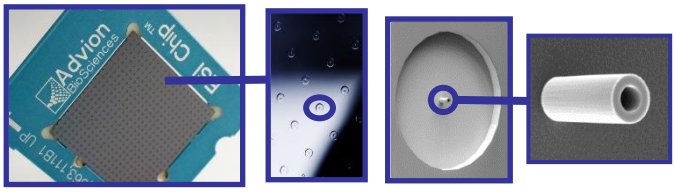


Figure 19 : Puce à émetteurs nanospray. Chaque puce continent 400 émetteurs nanoESI (images en microscopie électronique à balayage; TriVersa Nanomate, Advion)

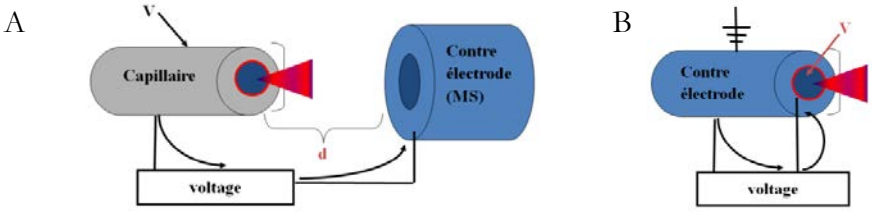
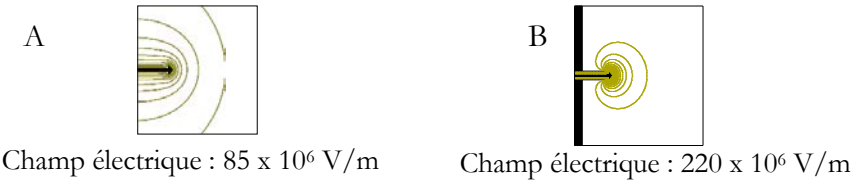


Figure 20 : Schéma de principe du nanospray conventionnel (A) et du nanospray sur puce (B). Contrairement au nanospray, la production du spray nanoESI ne nécessite pas d'être à une distance proche du spectromètre de masse. Aucune perte de sensibilité n'a été détectée (figure adaptée de la présentation d'Advion).



.

Figure 21 : Schéma des lignes de champ produit en nanospray conventionnel (A) et en nanospray sur puce (B). Le champ électrique est plus focalisé et intense avec le système multi-émetteurs nanoESI sur puce (figure adaptée de la présentation d'Advion).

II.2 Les analyseurs

L'analyseur reçoit les molécules ionisées, extraites puis focalisées, via une optique ionique de transmission constituée d'un jeu de lentilles et de multipôles. **Les analyseurs permettent de séparer les ions selon leur rapport masse sur charge (*m/z*)**. Presque tous les analyseurs peuvent être couplés à l'ESI : le quadrupole (Q), le temps de vol (TOF), le piège à ions quadripolaire (IT), le piège à ions orbitalaire (Orbitrap) ou encore la cellule à résonance cyclotronique d'ions (ICR). Tous ces analyseurs opèrent avec un champ électrique et/ou magnétique statique ou dynamique suivant deux lois physiques: la force de Lorentz et la seconde loi de Newton.

Le choix de l'analyseur conditionne la mesure de masse. Chaque type d'analyseur présente des caractéristiques différentes qui peuvent être résumées dans le Tableau 3. Tous ces analyseurs à l'exception du TOF sont utilisés au cours de ce projet et seront détaillés dans les Annexe 4 et Annexe 5.

Tableau 3 : Récapitulatif des différentes performances des analyseurs.

Analyseur	Pression	Résolution	Précision (étalonnage externe)	Gamme de masse	Vitesse d'acquisition	Gamme dynamique
Q	10^{-6} mbar	2 000 à 400*m/z*	100 ppm	8000	élevée	10^{6}
IT	10^{-5} mbar	5 000 à 400*m/z*	100 ppm	6000	élevée	10^{4}
TOF	10^{-8} mbar	20 000 à 400*m/z*	20-50 ppm	500000	moyen	10^{4}
Orbitrap	10^{-10} mbar	100 000 à 200*m/z*	10 ppm	10^{3} Da	moyen	10^{4}
FT-ICR	10^{-10} mbar	1 000 000 à 400*m/z*	5 ppm	4000	lente	10^{4}

II.2.1 Résolution

La résolution est le pouvoir de l'analyseur à séparer deux signaux de *m/z* proches. Plus la résolution est élevée, plus la distance entre deux signaux considérés comme distincts est petite. La résolution à mi-hauteur (FWHM, *Full Width at Half Maximum*) est définie comme le rapport $m/\Delta m$ où m est la masse de l'ion que l'on considère et Δm la largeur à mi-hauteur (Figure 22).

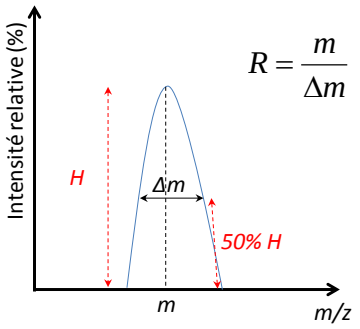


Figure 22 : Résolution à 50%

II.2.2 *Précision de masse*

La précision qualifie l'aptitude de l'analyseur à fournir des masses qui sont proches de la valeur vraie. En métrologie, **un analyseur précis est donc à la fois fidèle (erreur aléatoire) et juste (erreur systématique)**, c'est à dire que le pic doit être étroit et centré sur la valeur vraie (Figure 23 D).

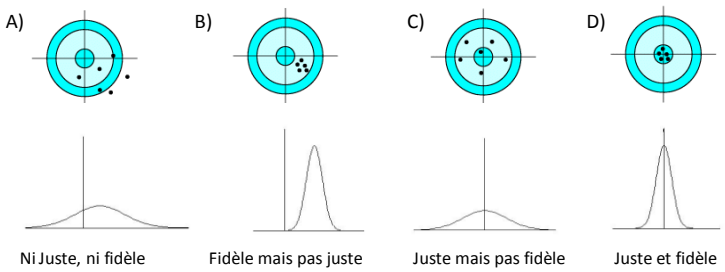


Figure 23 : Illustration de la fidélité et la justesse. A) la valeur finale est ni juste ni fidèle, on obtient un pic large éloigné de la valeur vraie, B) la valeur finale est fidèle mais pas juste, on obtient un pic étroit mais la valeur finale est éloignée de la valeur vraie, C) la valeur finale est juste mais pas fidèle, on obtient un pic large centré à la valeur vraie, D) la valeur finale est juste et fidèle, on obtient un pic étroit centré à la valeur vraie (figure adaptée de http://www.univ-brest.fr/lpo/instrumentation/02.htm et http://www.yorku.ca/psycho/fr/postscript.asp).

La fidélité (*precision* en anglais) de la mesure définit la dispersion des résultats. Des résultats sont dits fidèles quand leur coefficient de variation (CV) est faible (cas B et D sur la Figure 23).

La justesse ou exactitude de la mesure (*accuracy* en anglais) est la capacité de l'analyseur à donner une mesure proche de la valeur vraie de référence (cas C et D sur la Figure 23).

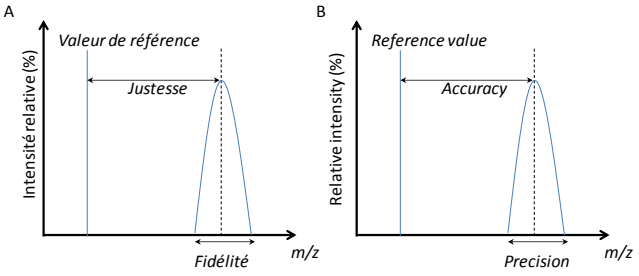


Figure 24 : Définition de la précision et justesse en français (A) et en anglais (B).

La justesse est exprimée par le rapport de l'écart absolu entre la valeur vraie et la valeur mesurée sur la valeur vraie (Equation 5), exprimée en ppm (partie par million).

$$Erreur\ (ppm) = \frac{\left| M_{thèorique} - M_{mesurée} \right|}{M_{theorique}} \text{ x } 10^6 \qquad \text{(Equation 5)}$$

En spectrométrie de masse, la justesse par abus de langage, est souvent confondue avec la précision de mesure. C'est souvent cette erreur de justesse qui est donnée dans les publications. Lors des interrogations dans les banques de données, chaque spectre est considéré individuellement et la masse correspond à la masse du centroïde du pic, la tolérance en masse renseignée correspond à l'erreur de justesse. La fidélité n'est donc pas prise en

compte. Pour prendre en compte un défaut de fidélité sur certains spectromètres, il est parfois nécessaire d'élargir la fenêtre d'interrogation pour intégrer ces erreurs aléatoires [129].

Dans le cas de mélanges complexes, le nombre de peptides générés est conséquent (plusieurs dizaines de milliers). De plus, plus la banque de données comporte un grand nombre de séquences, plus le nombre de faux positifs augmente. Ainsi, il est indispensable d'utiliser un spectromètre de masse capable de fournir des masses exactes pour limiter le nombre de candidats possibles. Plus l'erreur de justesse est faible plus le nombre de séquences candidates erronées décroit rapidement. Ainsi une justesse de l'ordre du ppm permet d'exclure 99% des peptides qui présentent la même masse nominale (masse obtenue à partir des masses isotopiques arrondies à l'unité) mais de composition en acide aminé différente. Une grande exactitude de mesure induit ainsi une grande robustesse dans l'interprétation des résultats [129,130].

Pour qu'un analyseur donne des résultats justes, un étalonnage est nécessaire. En effet, nous ne mesurons pas directement le rapport m/z mais pour le FT-ICR et l'Orbitrap, des courant induits ou pour le quadripôle et les pièges à ions quadripolaires, un signal électrique. Deux procédures d'étalonnage peuvent alors être utilisées :

- ✓ **l'étalonnage externe**, qui consiste à analyser en amont un mélange d'étalons standards dont la masse est connue. Les paramètres sont alors fixés pour toute la série d'analyses qui suit avant une nouvelle calibration de l'instrument. Si une variation d'un paramètre tel que l'effet de charge a lieu, l'erreur sur la justesse peut augmenter de façon significative. Pour les appareils de type FT-ICR à très faible dérive, ce type d'étalonnage est généralement suffisant.
- ✓ **l'étalonnage interne** qui consiste à rajouter un mélange d'étalons standards dont la masse exacte est connue. L'intérêt est de corriger extemporanément une dérive temporelle potentielle de l'exactitude de mesure. Sur des appareils de type Orbitrap, une calibration interne, appelée *lock mass* est possible, utilisant les contaminants ubiquitaires environnementaux (dérivés siloxanes) comme étalons internes [131]. La calibration interne est plus fine puisque les étalons internes vont subir les mêmes effets que les analytes, et en particulier les effets d'espace de charge ou de variations environnementales (température de la pièce par exemple).

La justesse de la mesure et la résolution de l'analyseur sont deux notions interdépendantes. Une bonne précision sur la mesure est obtenue avec un analyseur à pouvoir de résolution élevé et un S/B élevé.

II.2.3 Gamme de masse

La gamme de masse est l'entendue du domaine *m/z* que l'analyseur est capable de mesurer. En ESI, la gamme de masse requise est moins importante qu'en MALDI car la formation des multichargés diminue l'étendue des *m/z* des analytes.

II.2.4 Vitesse d'acquisition

La vitesse d'acquisition correspond au nombre de spectres par unité de temps que peut fournir l'analyseur. Elle est souvent exprimée en hertz (s^{-1}). Plus cette vitesse est grande, plus le nombre de protéines identifiées dans un mélange sera grand et donc plus grande sera la couverture du protéome. Ceci dit, il est important de rappeler que la très haute vitesse se fait souvent au détriment de la résolution.

Cette vitesse d'acquisition est liée au temps de cycle, c'est-à-dire le temps mis pour faire une acquisition MS et les MS/MS associées (voir §V page 84). Ce temps cycle doit être inférieur au temps d'élution des peptides dans le système chromatographique. En effet, plus un spectromètre est rapide, plus le nombre d'événements réalisables au cours du cycle est important. En analyse *shotgun*, un cycle court permet de fragmenter plus de peptides. En analyse ciblée, un temps de cycle réduit permet d'obtenir une quantification plus robuste avec plus de points de mesure.

II.2.5 Gamme dynamique

La gamme dynamique d'un analyseur est l'étendue des concentrations relatives détectables simultanément. C'est-à-dire, sa capacité à détecter un analyte minoritaire en présence d'un analyte très abondant. La détection d'analytes très peu abondants est critique, car généralement les biomarqueurs des maladies sont présents en de très faibles quantités.

Le sang [132] présente un bon exemple de très grande gamme dynamique avec une étendue sur plus de 10 décades (Figure 25). Les 10 protéines majoritaires représentent 90% de la totalité des protéines plasmatiques (Figure 26). Et parmi ces protéines, l'albumine représente à elle seule, plus de 50 % de la quantité totale de protéine [133]. Ainsi plusieurs stratégies d'analyse ont été mises en place pour augmenter la gamme dynamique de l'analyse telles que la déplétion des protéines majoritaires [134], un fractionnement ou encore une séparation chromatographique en amont de l'analyse de masse (cf. § IV.2page 83).

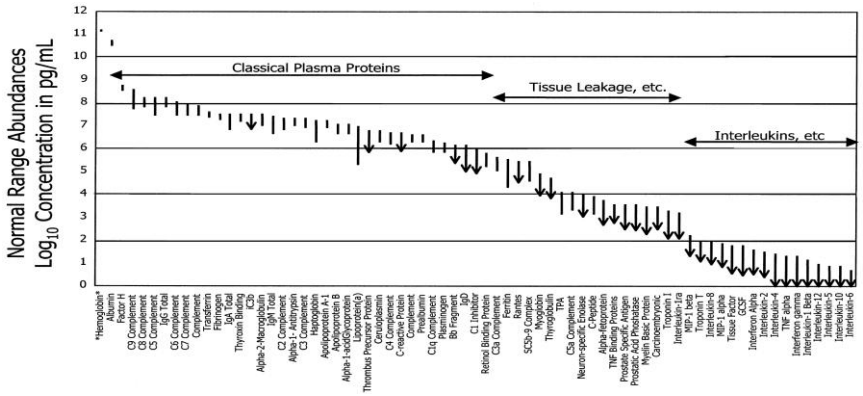


Figure 25 : Concentration de 70 protéines dans le plasma. Une différence de concentration de plus de 10 décades existe entre la protéine la plus abondante et la moins abondantes [132].

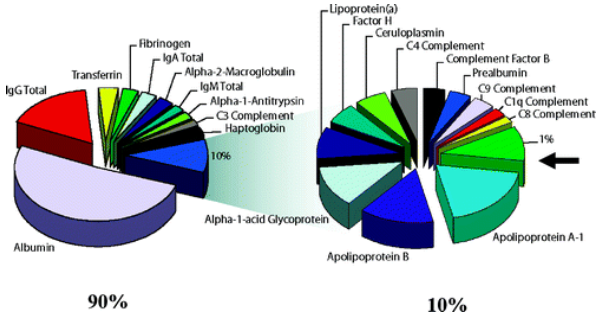


Figure 26 : Répartition des principales protéines plasmatiques [133].

III. La spectrométrie de masse tandem (MS/MS)

Lorsque plusieurs analyseurs sont couplés séquentiellement, on parle de spectrométrie tandem (MS^2 ou MS/MS). Dans ce cas, l'ion d'intérêt (dit ion parent ou ion précurseur) est isolé dans le premier analyseur, puis cet ion est fragmenté par collision d'un gaz inerte,électron ou photon. Enfin, les ions fragments produits sont séparés suivant leur rapport masse sur charge (m/z) dans le troisième analyseur. Selon les instruments utilisés, ces étapes se font soit dans l'espace (QqQ) soit dans le temps (IT) soit dans l'espace et dans le temps (QqOrbitrap, IT/FT-ICR).

III.1 La fragmentation générale des peptides

Pour l'identification des protéines, une fragmentation des peptides protéotypiques est souvent nécessaire afin de recueillir des informations de séquence. Pour cela, la connaissance du mécanisme de fragmentation des peptides en phase gazeuse est nécessaire pour corréler les spectres expérimentaux obtenus aux spectres théoriques découlant des règles de fragmentation. Les différents types de fragments générés sont dépendants de différents

facteurs tels que la séquence primaire, l'énergie apportée, de la pression ou encore l'état de charge de l'analyte.

La nomenclature des fragments adoptée aujourd'hui a été initiée par Roepstorff et Fohlman en 1984 [135] puis complétée par Biemann en 1990 [136].

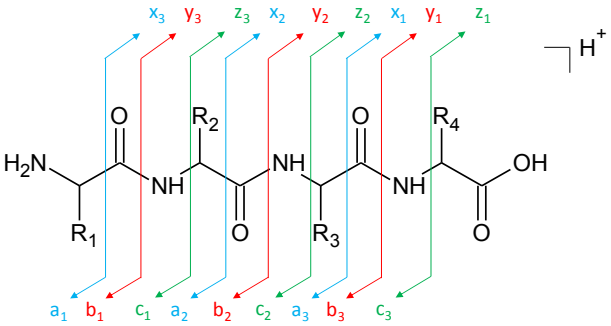


Figure 27 : Nomenclature de la fragmentation de la chaine peptidique pour un tétrapeptide.

La rupture d'une liaison dans la chaine peptidique peut avoir lieu à trois endroits différents décrits à la Figure 27. La numérotation correspond au nombre de résidus contenus dans le fragment.

Lorsque l'énergie de fragmentation fournie est élevée, **une rupture de la chaine latérale** peut avoir lieu, donnant plus d'information sur la structure des acides aminés mis en jeu [137,138]. D'autres fragments, appelés fragments secondaires, peuvent également se produire par perte d'eau, d'ammoniaque ou du groupement carbonyle CO (Annexe 2).

Il existe quatre grands types de fragmentation permettant d'apporter de l'énergie au système et menant chacun à des types d'ions fragments préférentiels :

- ✓ **CID à basse énergie** (*low-energy collision-induced dissociation*) qui nécessite des énergies faibles de l'ordre de l'électron volt (eV). C'est le cas des triple-quadripôles, pièges à ions quadripolaires et qToF. Pour les appareils équipés d'un Orbitrap, c'est la fragmentation HCD (*higher energy dissociation*) qui est employée. Elle est similaire au CID à basse énergie. **Les ions b et y sont principalement générés** par ces collisions. Ce sont les types de fragmentation qui ont été utilisés au cours de ce projet, et seront donc les seuls à être détaillés dans ce manuscrit.
- ✓ **CID à haute énergie** (*high-energy collision-induced dissociation*) qui emploie des énergies de collision de l'ordre du kilo-électron volt (keV). C'est le cas des instruments ToF/ToF. **Tous les types d'ions peuvent être générés** par ces collisions. Une rupture de la chaine latérale peut également avoir lieu avec cette fragmentation.

- ✓ **ECD** (*electron capture dissociation*) qui utilise des électrons de basse énergie pour induire la dissociation des grands peptides ou des protéines. Ce type de fragmentation est employé dans les appareils de type FT-ICR. **Les ions c et z sont principalement générés** par cette collision.
- ✓ **ETD** (*electron transfer dissociation*) où les peptides/protéines se dissocient par transfert d'électron d'un donneur anionique. Ce type de fragmentation est employé dans les appareils de type qTOF ou Orbitrap. **Les ions c et z sont principalement générés** par cette collision.
- ✓ **IRMPD** (*infrared mutiphoton dissociation*) où les peptides/protéines se dissocient par excitation vibrationnelle. Ce type de fragmentation est employé dans les appareils de type FT-ICR. **Les ions b et y sont principalement générés** par cette collision.

III.2 Dissociation induite par collision à basse énergie (low-energy CID)

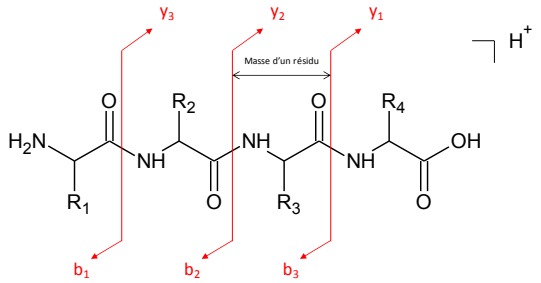


Figure 28 : Exemple de fragmentation d'un peptide à quatre acides aminés en CID à basse énergie. **Les ions b et y sont les ions majoritaires.**

La dissociation induite par collision à basse énergie (CID, *low-energy collision-induced dissociation*) est la méthode de fragmentation la plus couramment utilisée pour les peptides, et plus particulièrement dans ce projet. La fragmentation a lieu quand le peptide entre en collision avec des molécules de gaz (généralement argon ou hélium) dans la cellule de collision. Les ions sont accélérés sous tension (5-40eV) dans cette cellule afin d'accélérer la collision. **Il y'a fragmentation quand l'énergie de collision est assez élevée pour convertir une partie de l'énergie interne du peptide en énergie vibrationnelle, distribuée en fonction des affinités protoniques des groupements chimiques le long de la séquence primaire**. Ainsi, la dissociation du peptide a lieu quand l'énergie apportée est suffisamment élevée pour passer la barrière d'activation de rupture de la liaison. C'est pour

cela que le CID est parfois appelé dissociation activée par collision (CAD, *collision-activated dissociation*)

La fragmentation CID à basse énergie induit essentiellement une rupture des liaisons peptidiques (CO-NH). **La charge peut être retenue soit du côté N-terminal (ions fragments de type b) soit du côté C-terminal (ions fragments de type y).** La fragmentation des peptides permet de déduire la séquence en acide aminés. En effet, la différence de masse entre les ions fragments, consécutifs d'une même série d'ions (b_n-b_{n-1} ou y_n-y_{n-1}), correspond à la masse d'un résidu (Figure 28). Afin de lever l'ambigüité d'attribution d'une masse à un acide aminé, tel que pour la glutamine (128.06 Da) et la lysine (128.09 Da), un instrument d'une grande exactitude de mesure tel que les appareils à transformée de Fourier est nécessaire. Dans le cas d'acides aminées isobariques tels que la leucine et l'isoleucine (113.08 Da), d'autres voies de fragmentation à plus haute énergie sont nécessaires pour lever l'ambiguïté [137,138].

III.3 Dissociation induite par collision à plus haute énergie (HCD)

La fragmentation C-Trap à plus haute énergie (HCD, *higher-energy C-Trap dissociation*) a été introduite en 2007 par l'équipe de Olsen [139] pour une analyse des fragments à haute résolution avec l'Orbitrap sans limitation de seuil des basses masses (LMCO). À cette époque, la C-Trap servait de lieu de stockage des ions avant leur entrée dans l'analyseur et également de lieu de fragmentation. Les mots « plus haute énergie » se réfèrent à une tension de radiofréquence appliquée plus élevée en HCD pour retenir les fragments dans la C-Trap et non à une plus grande énergie utilisée pour fragmenter les peptides. Cette tension est de 2500V alors qu'elle est de 1500V dans les autres instruments. De même qu'en CID, l'énergie de collision appliquée est entre 20 et 50eV pour fragmenter les peptides.

Aujourd'hui les fragmentations de type HCD, renommée « dissociation à plus haute énergie » (*higher energy collisionna dissociation*), ont lieu dans une cellule de collision constituée d'octopoles, distincte de la C-Trap. Une radiofréquence (RF *only*) est appliquée dans cette cellule. Cette dernière est remplie d'un gaz identique à celui de la C-Trap pour ralentir les ions, l'azote, qui sert à fragmenter les peptides.

Les ions précurseurs, sélectionnés par le quadripôle, traversent les multipôles de transfert et la C-Trap pour être fragmentés dans la cellule de collision HCD. Ces ions sont ensuite accélérés vers cette cellule HCD par application d'une différence de potentiel entre la C-Trap et la cellule HCD. Etant donné que la pression dans cette cellule est assez élevée (effet dû à la

présence du gaz), un gradient de potentiel doit être appliqué pour assurer la bonne transmission des ions dans la cellule. Les ions entrent en collision avec des molécules de gaz neutre (azote par défaut) dans la cellule entrainant la rotation, puis l'étirement et finalement rupture des liaisons, aboutissant à la création d'ions fragments chargés et de neutres. Ces derniers sont éliminés du système par pompage. Les fragments sont ensuite renvoyés vers la C-Trap en appliquant un gradient de potentiel pour être focalisés et accélérés avant leur détection en Orbitrap (Figure 29 et Figure 30). Tout comme en **fragmentation CID, ce sont des ions fragments de type b et y qui sont principalement générés par la fragmentation HCD**. De plus, tous les ions fragments générés et transférés sont détectés, sur le modèle de fonctionnement du piège ionique (ce qui est favorable à une haute sensibilité).

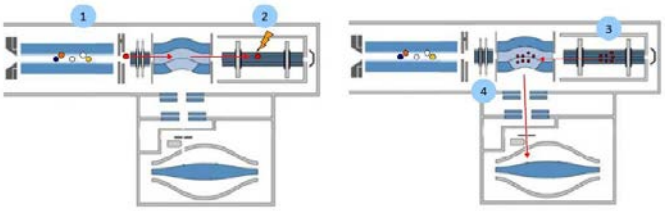


Figure 29 : Fragmentation HCD des ions précurseurs dans un QqOrbitrap. Les ions précurseurs sont d'abord sélectionnés par le premier quadripôle puis transférés vers la cellule de collision HCD pour être fragmentés. Ces derniers sont ensuite transférés vers la C-Trap avant leur détection sur l'Orbitrap [figure tirée du manuel de l'utilisateur du Q Exactive, Thermo Scientific].

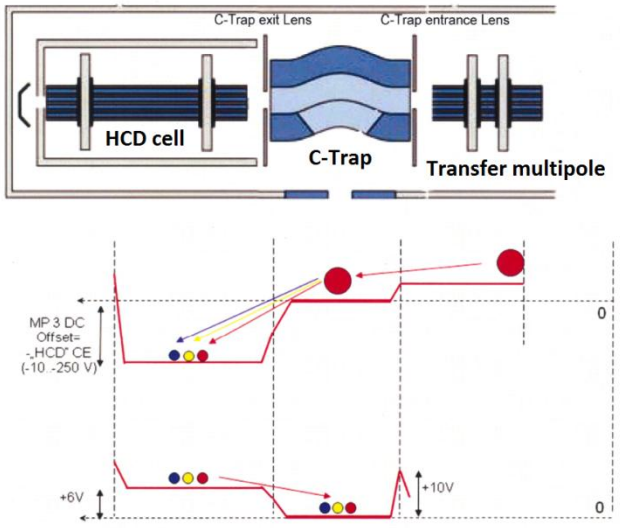


Figure 30 : Potentiel appliqué pour la fragmentation en HCD. Les ions sélectionnés dans le quadripôle sont transférés vers la cellule de collision HCD via le multipôle de transfert et la C-Trap avec des différences de potentiel. Un gradient de potentiel est également appliqué dans la cellule de collision pour augmenter la transmission des ions dans cette cellule où ils vont être fragmentés par collision dans un gaz. Les fragments sont ensuite transférés vers la C-Trap en appliquant un gradient de potentiel avant d'être détectés dans l'Orbitrap [figure tirée du manuel de l'utilisateur du Q Exactive, Thermo Scientific].

IV. Le couplage avec la chromatographie

Dans des analyses protéomiques exploratoires, le but est d'identifier un maximum de protéines. Or du fait de la complexité des échantillons biologiques, une analyse en spectrométrie de masse seule n'est pas suffisante pour analyser plusieurs milliers de composés en même temps. Depuis les années 80, le couplage de la chromatographie liquide à la spectrométrie de masse à source d'ionisation à pression atmosphérique est très utilisé pour l'analyse des biomolécules. Aujourd'hui, le challenge est d'accroitre la capacité de pic, c'est-à-dire obtenir le maximum de composés séparés dans une fenêtre de temps donnée [140]. Ce couplage permet l'identification de centaines, voire de milliers de protéines, dans un mélange en routine [59].

IV.1 Généralités

La chromatographie liquide (LC) fut découverte par Tswett en 1903 [141], pour la séparation de pigments. Elle est aujourd'hui utilisée en analyse protéomique comme **technique analytique qui permet de séparer les peptides** (solutés) obtenus après digestion des protéines contenues dans les cellules et tissus issus des biopsies dans notre cas. L'échantillon, solubilisé dans un solvant, est transporté à travers la phase stationnaire (colonne chromatographique) par une phase mobile (éluent). **La séparation se fait suivant un partage d'interactions soluté/phase stationnaire/phase mobile.**

La chromatographie liquide est bien adaptée à la séparation des composés biologiques car ce sont des mélanges complexes et cette technique permet d'analyser des composés indépendamment de leur polarité et de leur masse moléculaire. C'est la chromatographie à polarité de phases inversée (RPLC, *reverse phase liquid chromatography*) qui est souvent employée en analyse protéomique. La phase stationnaire hydrophobe est composée de particules de silice greffées C_{18}. L'élution se fait par une augmentation du pourcentage de solvant organique de la phase mobile facilitant l'ionisation ESI. Les peptides d'intérêt sont séparés en fonction de leur hydrophobicité en RPLC.

À pH acide (pH 2-3), la charge globale des peptides est positive, favorable au couplage avec la spectrométrie de masse. L'ajout d'un modificateur organique va créer une paire d'ions et rendre les solutés plus hydrophobes, facilitant les échanges avec la phase stationnaire hydrophobe. Afin d'être compatible à l'analyse de masse en mode d'ionisation ESI, des tampons volatils sont utilisés. Ce sont l'acide formique (HCOOH) et l'acide acétique (CH_3COOH) qui sont les plus utilisés. Ces modificateurs permettent non seulement

d'accroître la rétention, par la formation de la paire d'ions, mais également d'accroître la résolution de la séparation en RPLC. Un autre modificateur souvent utilisé en chromatographie, l'acide trifluroacétique (TFA) est déconseillé dans le cas de couplage direct avec la masse car cet acide fort, attaque l'acier et provoque également des phénomènes de suppression d'ions [142].

IV.2 La nano-chromatographie (nanoLC)

Depuis sa découverte, la chromatographie n'a cessé d'évoluer. Tout d'abord avec la haute pression (HPLC, *high-performance liquid chromatography*), l'ultra-chromatographie (UHPLC, *ultra-high-performance liquid chromatography*) mais aussi la **nano-chromatographie (nanoLC,** *nano-liquid chromatography***)**, par une miniaturisation des colonnes de chromatographie et de l'instrumentation (cf. § II.1 page 70).

Cette évolution a rendu possible l'analyse de mélanges de plus en plus complexes en faible concentration grâce à son pouvoir résolutif, sa sensibilité accrue par la diminution des diamètres internes des colonnes chromatographiques et sa facilité de couplage avec un mode d'ionisation efficace en nanoESI.

La diminution des diamètres internes des colonnes engendre une diminution de la dilution chromatographique (Equation 6) et donc une augmentation de la concentration locale de l'analyte par rapport à un système classique d'HPLC.

$$D = \frac{C_0}{C_{max}} = \frac{\varepsilon_T \pi r^2 (1+k)\sqrt{2\pi r L H}}{V_{inj}} \qquad \text{(Equation 6)}$$

où C_0 : concentration initiale, Cmax : concentration maximale au sommet du pic, ε : porosité de la colonne, r : rayon de la colonne, k : facteur de rétention, L : longueur de la colonne, H : hauteur équivalente à un plateau et V : volume d'échantillon injecté.

Le gain de la sensibilité peut être déterminé par l'équation (Equation7) suivante:

$$f = \frac{d_1^2}{d_2^2} \text{ avec } d_1 > d_2 \qquad \text{(Equation7)}$$

où d_1 : diamètre colonne 1 et d_2 : diamètre colonne 2.

Le gain de sensibilité progresse de presque de 4000 fois entre une colonne classique de 4.6mm et une nano-colonne de 75µm [143]. Cependant, la diminution du diamètre de colonne a un impact sur tout l'assemblage instrumental, discuté dans les paragraphes de l'Annexe 3.

V. Configurations instrumentales

V.1 LIT-FTICR

V.1.1 Généralités

L'instrument LTQ FT Ultra (Thermo Scientific), illustré à la Figure 31, est équipé d'une source nanoESI pour cette étude. Les ions traversent ensuite un ensemble de guide d'ions (*tube lens* et *skimmer*), d'optiques d'ions de focalisation, avant d'arriver au piège à ions linéaire (LIT, Annexe 4) où ils sont fragmentés par collision avec l'azote (cf. §III.2). Les ions fragments sont ensuite éjectés transversalement et détectés par les deux multiplicateurs d'électrons à faibles résolution et précision. Une mesure précise des fragments générés dans le piège peut également s'effectuer dans la cellule ICR cylindrique (Annexe 5), en accumulant les spectres. Cependant le couplage direct avec la chromatographie limite le temps d'analyse du spectromètre et ne permet pas d'envisager cette configuration sur des mélanges complexes. Lorsque le piège agit comme un guide d'ion, les ions sont transmis sans être fragmentés puis sont guidés par une série de multipôles avant d'atteindre la cellule ICR. Les ions précurseurs sont alors mesurés à HR/AM. Dans cette cellule, les ions peuvent subir une dissociation ECD ou IRMPD (cf. §III.2).

Ce spectromètre de masse est particulièrement adapté pour analyser des échantillons complexes, avec le haut débit de séquencage apporté par le piège à ions linéaire, et la HR/AM, apportées par la cellule ICR. Cependant, le défaut majeur de cette configuration de spectromètre de masse réside dans le *cut-off* des basses masses du piège à ions (1/3 de la masse de l'ion précurseur). En dessous de cette masse, les ions décrivent des trajectoires instables et sont éjectés du piège. Ainsi, il ne permet pas d'observer les ions immoniums qui peuvent contribuer à l'identification des acides aminés. De plus, ce seuil proscrit toute quantification à l'aide de ions rapporteurs de basse masse (cf. §III.1.2).

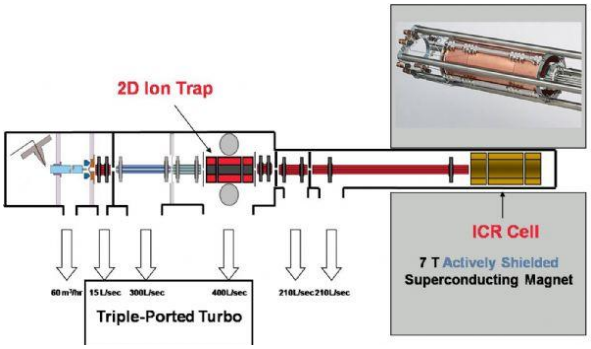


Figure 31 : Spectromètre de masse LTQ FT Ultra de Thermo Scientific.

V.1.2 Data-dependent acquisition (DDA)

En routine, des acquisitions de type data dépendant (DDA, *data-dependent acquisition*) sont réalisées sur ce spectromètre de masse. Les ions précurseurs sont tout d'abord pré-mesurés dans la cellule ICR puis les 7 ions les plus intenses (TOP 7) sont fragmentés dans le piège LIT. Pendant ces fragmentations MS/MS, le rapport m/z des ions précurseurs est mesuré avec une bonne justesse dans la cellule ICR à haute résolution (50000 @ m/z 400, 5ppm). Tout en restant compatible avec la complexité des mélanges biologiques, il faut moins d'une seconde pour réaliser ces acquisitions MS et les 7 MS/MS en parallèle. Même si le séquençage se fait à basse résolution (100 ppm), l'exactitude de mesure apportée par le FT-ICR permet d'identifier les peptides avec robustesse. Afin d'augmenter le nombre de protéines identifiées pour couvrir au mieux le protéome, une liste d'exclusion dynamique des ions précurseurs déjà fragmentés est utilisée pour favoriser la fragmentation des ions moins intenses.

Ce mode d'acquisition DDA est très utilisé en protéomique car il permet de réaliser une analyse globale de l'échantillon pour identifier de façon systématique les protéines dans le mélange biologique.

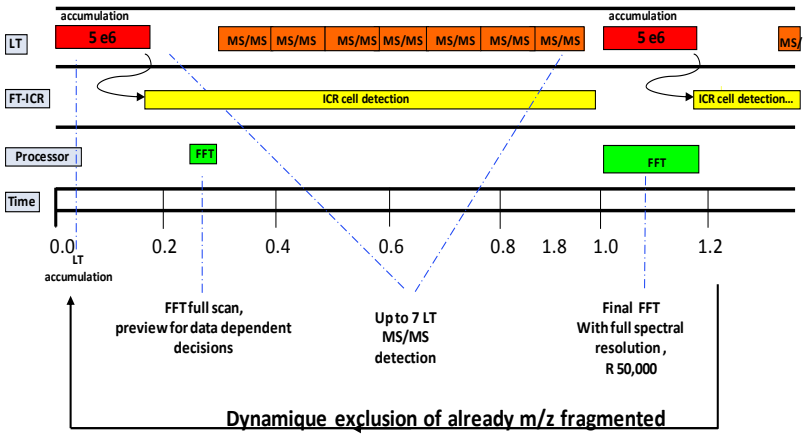


Figure 32 : Schéma temporel de l'acquisition en parallèle dans le LIT/FT-ICR. Les ions précurseurs sont d'abord accumulés dans le piège à ions puis éjectés vers la cellule FT-ICR où les ions sont mesurés à HR/AM (50 000, 5ppm). Pendant l'acquisition du transient d'ion des peptides, une fragmentation des 7 ions les plus intenses se fait à basse précision dans le piège à ions (100ppm) en moins d'une seconde (LTQ FT Ultra, Thermo Scientific).

V.2 Triple-Quadripôle (QqQ)

V.2.1 Généralités

Nous disposons au laboratoire d'un triple quadripôle TSQ Vantage (Thermo Scientific), illustré à la Figure 33, qui est constitué de trois quadripôles (Annexe 4) en série. Après ionisation en mode nanoESI, les ions sont transmis via des lentilles dans une optique appelée S-Lens. Cette dernière est constituée d'une série d'électrodes circulaires plates sur lesquelles une radiofréquence est appliquée. Les ions sont ensuite guidés vers dans le premier quadripôle

(Q1) qui peut agir comme un guide d'ions (ou filtre de masse selon le mode d'acquisition) par une série de lentilles (Q00) et d'un quadripôle (Q0). Ensuite, les ions rencontrent une série de lentilles (L21, L22, L23) avant d'arriver au deuxième quadripôle (Q2), appelé cellule de collision, dans lequel ils peuvent être fragmentés par collision avec un gaz, ici l'argon (cf.III.3). En sortie du deuxième quadripôle, les ions sont refocalisés par autre une série de lentilles (L31, L32, L33). Ils parviennent au troisième quadripôle (Q3) qui agit en tant que guide d'ions ou filtre de masse selon le mode d'acquisition. Ils sont finalement détectés par un multiplicateur d'électron (Annexe 4).

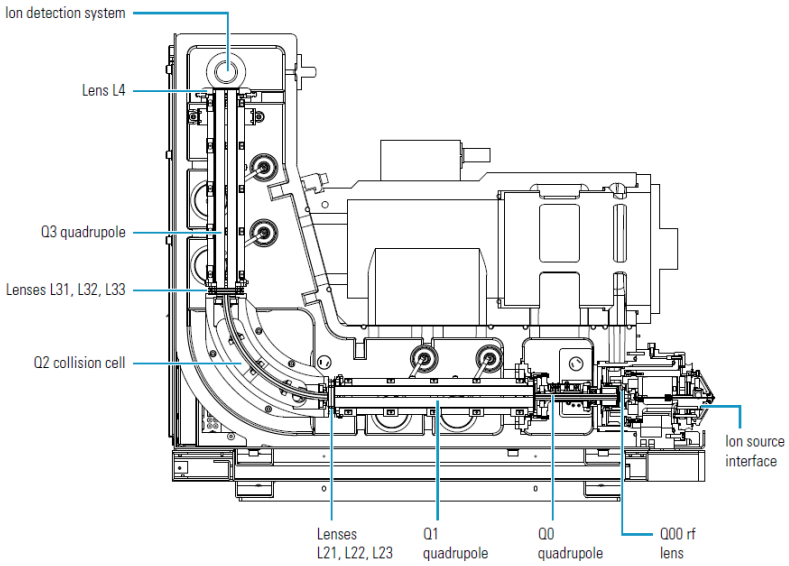


Figure 33 : Schéma détaillé du TSQ Vantage (Thermo Scientific).

V.2.2 *Mode de balayage*

Plusieurs modes de balayage sont possibles sur ce type d'appareil : le balayage normal, le balayage d'ions fils, le balayage de l'ion parent, le balayage de perte de neutre et le balayage d'une transition parent-fils. Seul le dernier mode a été utilisé et sera donc détaillé dans les paragraphes suivants.

V.2.2.1 *Selected reaction monitoring (SRM)*

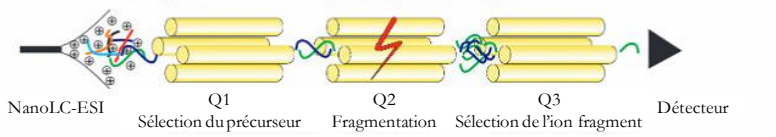


Figure 34 : Balayage des transitions père-fils sur un triple quadripôle (figure adaptée de Lange *et al.* [144]).

Dans ce mode ciblé, appelé SRM (*selected reaction monitoring*), le Q1 et le Q3 agissent tous les deux comme filtres de masse. Le Q1 sélectionne le précurseur à fragmenter dans la cellule de

collision (Q2) et le Q3 sélectionne un ion fragment déterminé. On alors parle de transition père→fils. **Ce mode d'acquisition, qui est le plus sensible, est également très sélectif**, car il n'y a pas de balayage et les électrodes sont à une tension fixe. En outre, il est possible d'alterner la détection des transitions pour augmenter le nombre de peptides détectés, appelé *multiplexe SRM*. Ces divers atouts en font une méthode de choix dans la quantification des protéines. Cette méthode sera évaluée plus loin au cours de projet (chapitre 5).

Cependant afin d'augmenter la sensibilité de la méthode, des optimisations sont nécessaires. Tout d'abord, il faut choisir les peptides et donc les transitions optimales à suivre. En général, ce sont les peptides (entre 2 et 4) et les transitions (entre 2 et 4) les plus intenses, qui sont choisis [144]. Le choix de ces peptides/transitions peut se faire à l'aide des données *shotgun* acquises sur un autre spectromètre de masse en mode DDA (cf. §V.1.2 et V.3.2.1) au sein du laboratoire, ou par d'autres équipes qui ont déposé leurs données dans une plate-forme de partage des données, telle que Peptide Atlas [145] ou PRIDE [146]. Cependant, il faut garder à l'esprit que les conditions d'ionisation peuvent affecter la distribution des états de charge du peptide. De plus, les paramètres des spectromètres peuvent également influencer sur le profil de spectre MS/MS, surtout dans le cas de comparaison entre un piège à ions et un quadripôle [144]. En général, les ions y sont préférés aux ions b, ces derniers étant moins stables dans un triple quadripôle. Concernant les peptides, il est conseillé d'exclure les peptides présentant une coupure de digestion manquée (*missed cleavage*) et des acides aminés susceptibles de subir une modification tels que la méthionine (oxydation) ou encore la cystéine (réaction chimique d'alkylation). Pour que la méthode soit spécifique, il est recommandé de choisir des peptides spécifiques d'une protéine, appelés peptides protéotypiques [144]. De plus pour que la méthode soit la plus robuste et sensible possible, il est préconisé de recourir à des protéines/peptides standards pour les analyser sur le triple quadripôle dédié à la quantification ciblée en mode SRM. Souvent, ce sont des peptides marqués par des isotopes stables, relativement couteux, qui sont choisi pour faire l'optimisation des transitions. Ces peptides marqués ont des propriétés physico-chimiques semblables au peptide endogène, mais des masses différentes.

Pour améliorer la méthode, il faut également optimiser les paramètres du spectromètre pour produire les ions: la tension appliquée, la température du capillaire du transfert, la pression de gaz qui sont les mêmes pour tous les peptides et l'énergie de collision qui est propre à chaque transition. Cette dernière est donc un paramètre important à optimiser pour gagner en sensibilité. Certains logiciels proposent de faire un calcul théorique de l'énergie de collision à partir du rapport m/z du précurseur. L'équation est souvent de la forme :

$$Collision\ Energy = a * m/z_{precurseur} + b \quad \text{(Equation 8)}$$

Cette équation peut donner une idée de l'énergie de collision à appliquer mais reste toutefois une valeur théorique. Une série d'énergies de collision doit être évaluée expérimentalement pour chaque transition. Pour cela, le recours au peptide standard est alors incontournable.

Enfin, une fois les transitions choisies, il faut optimiser le temps accordé à la détection de chaque transition, appelé *dwell time*. Ce paramètre peut être augmenté tout en conservant à l'esprit la compatibilité avec la séparation chromatographique. En effet, pour décrire au mieux un pic chromatographique, nous avons considéré qu'un minimum de 8 à 10 points est nécessaire. En moyenne la largeur à 10% de la hauteur des pics chromatographiques est de 20s. Chaque cycle (cycle time) doit durer 2s, c'est à dire que toutes les transitions doivent être suivies dans un délai maximum de 2s (Equation 9). Ainsi, le temps consacré à suivre une transition doit être au maximum 2s divisé par le nombre total de transitions ciblées. Ce qui donne (Equation 10) :

$$cycle\ time = \frac{largeur\ \text{à}\ 10\%\ de\ la\ hauteur\ du\ pic}{10} \quad \text{(Equation 9)}$$

$$dwell\ time = \frac{cycle\ time}{nombre\ total\ de\ transitions\ \text{à}\ suivre} \quad \text{(Equation 10)}$$

Généralement, pour obtenir une sensibilité correcte, un temps de 10ms est nécessaire par transition (*dwell time*). Ce temps correspond à 100 peptides observables avec 2 transitions. Si chaque protéine doit être quantifiée avec 2 peptides, alors il y aura un maximum de 50 protéines quantifiables avec cette méthode SRM. Si l'on recherche une excellente sensibilité, alors c'est un *dwell time* de 100ms qui sera utilisé. Dans ce cas, seules 5 protéines pourront être quantifiées, avec 2 peptides pour 2 transitions. [144].

V.2.2.2 *Time-scheduled SRM (tSRM)*

Le nombre de transitions peut croitre en faisant du *time-scheduled* SRM (*tSRM*) [147], c'est à dire que les transitions vont être suivies dans un intervalle de temps donné. Le *dwell time* est alors dépendant de la complexité en chromatographie. Si on considère une distribution homogène tout au long d'une analyse chromatographique de 60 min et un intervalle de temps de 2min par transition, plus de 1000 protéines peuvent être peuvent être quantifiées avec un *dwell time* de 15ms [144]. Cette possibilité de suivre plusieurs transitions dans un intervalle de temps plus court et avec une bonne sensibilité (attomole) rendent cette technique très puissante pour la quantification des protéines. Elle a été d'ailleurs nommée méthode de

l'année 2012 par la revue scientifique Nature [148]. L'équipe de Picotti [149] a ainsi pu quantifier de façon absolue en *tSRM* toute une voie signalétique grâce à des peptides marqués par des isotopes stables. (cf. § III.2.1).

V.2.2.3 *Intelligent SRM (iSRM)*

Un autre mode, non évalué dans ce projet, existe : le mode *intelligent selected reaction monitoring* (iSRM), introduit récemment par l'équipe d'Aerbersold et Domon [150]. Il permet d'augmenter le nombre de transitions à suivre sans sacrifier le nombre maximum de peptides cibles possible. Cette méthode intègre le *tSRM* et repose sur une combinaison de deux types d'événements : (1) un événement primaire spécifique où seulement deux transitions (primaires) sont suivies pour la quantification, et (2) un évènement secondaire où six autres transitions (secondaires) sont suivies uniquement lorsque les ions primaires ont été observés (*data-dependent*). Ce dernier événement permet d'accroitre le nombre de transition suivies pour confirmer la présence du peptide, au lieu de faire la validation ultérieure par une acquisition DDA ou de suivre en continu un plus grand nombre de transitions SRM (auquel cas le temps d'acquisition augmenterait au détriment du nombre de peptides cibles pour que le temps de cycle reste inférieur à 2s).

V.3 QqOrbitrap

V.3.1 Généralités

Le QqOrbitrap (Q Exactive, Thermo Scientific), illustré à la Figure 35, est composé d'une source nanoESI, puis d'une succession d'un ensemble de guide d'ions (*S-Lens* et *S-Lens exit lens*) et d'une série d'optiques d'ions (*flatapole*, *bent flatapole* et *lens*) avant d'arriver au quadripôle (Annexe 4). Contrairement au LTQ FT Ultra, il n'y a pas de *cut-off* des basses masses et la gamme de masse du quadripôle est de m/z 50 - 4000. Puis, les ions sont transmis via une optique spécifique (C-Trap et Z-Lens) avant d'être détectés dans l'Orbitrap (Annexe 5) à HR/AM. Les ions peuvent également subir des fragmentations dans la cellule de collision HCD (cf. §III.3) et les ions fragments sont également détectés à HR/AM dans l'Orbitrap.

Bien que la mesure des ions précurseurs et fragments se déroule dans l'Orbitrap, si le réglage de la méthode est bien défini, cet instrument permet de réaliser des événements MS et MS/MS en parallèle. Dans cette configuration, lorsque les premiers ions N sont détectés dans l'Orbitrap, un autre paquet d'ions N+1 peut au même moment être accumulé dans la C-Trap,

rendant le temps d'un cycle d'analyse limité par le temps d'acquisition du transient dans l'Orbitrap (Tableau 4), et/ou par le temps de remplissage de la cellule de collision.

La configuration de ces analyseurs dans l'espace permet en outre de faire du multiplexage, c'est à dire d'accumuler plusieurs ions en même temps avant de mesurer la masse exacte dans l'Orbitrap. Le processus de traitement de signal eFT est également intégré dans cet instrument (Annexe 5).

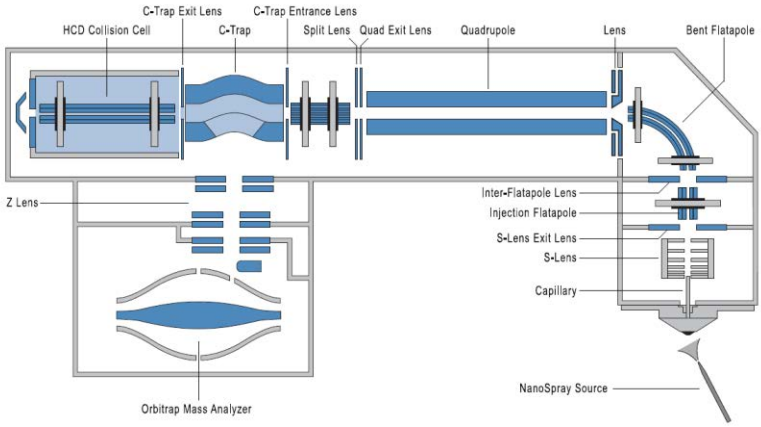


Figure 35 : Schéma détaillé du Q Exactive (Thermo Scientific). Les ions ionisés en nanoESI sont focalisés et transmis vers le quadrupole qui laisse passer soit tous les ions ou un type d'ions spécifiques. Ils traversent ensuit la C-Trap pour être soit éjectés vers l'Orbitrap pour la mesure exacte des ions précurseurs ou fragmentés dans la cellule HCD avant de revenir dans le C-Trap et enfin éjectés vers l'Orbitrap pour une mesure juste des ions fragments.

Tableau 4 : Temps d'acquisition du transient [151].

Résolution (à *m/z* 200)	**Fréquence d'acquisition DDA**	**Temps d'acquisition du transient**
17 500	12 Hz	64 ms
35 000	7 Hz	128 ms
70 000	3 Hz	256 ms
140 000	1.5 Hz	512 ms

*V.3.2 **Modes d'acquisition***

V.3.2.1 ***Data-dependent acquisition (DDA)***

En routine, des acquisitions de type DDA (Figure 36) sont réalisées sur le QqOrbitrap. Les ions précurseurs sont tout d'abord mesurés dans l'Orbitrap à haute résolution (70 000 @ *m/z* 200) et grande exactitude (10 ppm). Pendant l'acquisition du transient, les ions correspondant à l'espèce la plus abondante sont fragmentés dans la cellule de collision HCD puis transférés vers la C-Trap. Ensuite, pendant que les premiers fragments N sont analysés

dans l'Orbitrap à haute résolution, la seconde espèce N+1 est fragmentée et accumulée dans la C-Trap. Le temps d'accumulation dans la C-Trap est proche du temps nécessaire pour générer le transient d'ions des fragments, soit 64 ms à 17 500 de résolution (Tableau 4). Ainsi, un maximum de 10 ions les plus intenses (TOP 10) peuvent être fragmentés et analysés dans l'Orbitrap en 1s à haute résolution. Contrairement au LTQ FT Ultra, le Q Exactive permet la mesure des fragments à haute résolution, ce qui augmente la robustesse de l'analyse.

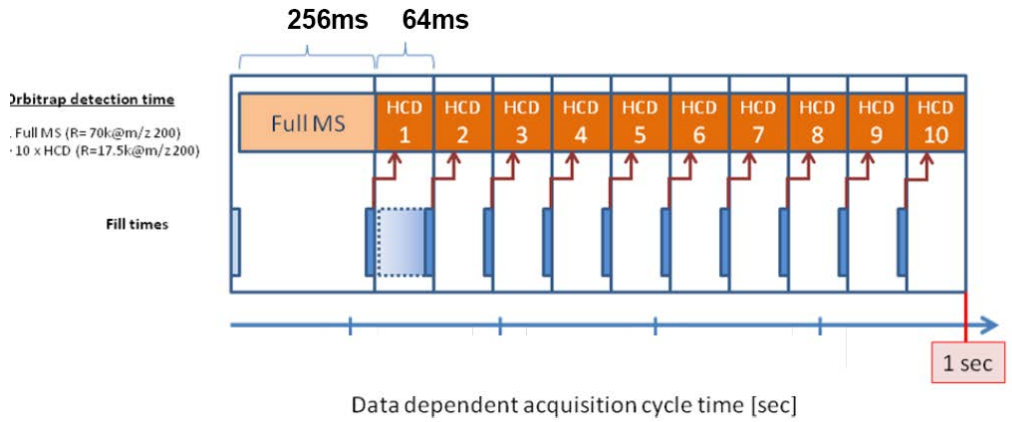


Figure 36 : Acquisition data-dépendant (DDA) sur le QqOrbitrap. Les ions précurseurs sont d'abord accumulés dans la C-Trap puis éjectés vers l'Orbitrap où les ions sont mesurés à HR/AM (70 000, 10ppm). Pendant l'acquisition du transient d'ions, un paquet d'ions correspondant aux ions les plus intenses est sélectionné par le quadripôle pour être fragmenté dans la cellule de collision HCD puis transféré dans la C-Trap. En même temps que la détection de ces ions les plus intenses, un deuxième paquet d'ions est fragmenté et accumulé dans la C-Trap. La tenue de ces événements en parallèle permet de faire une fragmentation des 10 ions les plus intenses à HR/AM (17 500, 20ppm) en 1 seconde (Q Exactive, Thermo Scientific).

V.3.2.2 *Parallel reaction monitoring (PRM)*

Le *parallel reaction monitoring* (PRM) dérive du SRM historiquement. L'ion précurseur est sélectionné dans le quadripôle avec une fenêtre d'isolation de 0,4 à 10 Da. Les ions sont ensuite stockés dans la C-Trap puis transférés dans la cellule de collision HCD pour être fragmentés. Les fragments sont ensuite transférés dans la C-Trap pour être refocalisés avant d'être éjectés et détectés dans l'Orbitrap. Les transitions père→fils sont reconstituées *a posteriori* par un logiciel informatique. Le troisième analyseur est donc ici un Orbitrap. Ce dernier permet d'avoir accès à une haute résolution et grande exactitude de mesure des fragments comme des précurseurs, ce qui minimise les interférents [152].

Un certain nombre d'optimisations est nécessaire pour augmenter la sensibilité de la méthode. En outre des paramètres propres à l'ionisation et la transmission des ions (tension du capillaire, température du tube de transfert et tension de la *S-Lens*), un choix sur les précurseurs les plus intenses est aussi nécessaire. Les données *shotgun* obtenues au sein du laboratoire, ou mises à disposition dans des plateformes de partage peuvent être utilisées pour

guider le choix des précurseurs (m/z et état de charge). Les fragments étant tous détectés par l'Orbitrap, seule une énergie de collision globale pour le précurseur et non plus tous fragments est nécessaire, ce qui diminue considérablement le processus d'optimisation de la méthode.

Quant au nombre de peptides quantifiables, il dépend de la durée d'acquisition du transient et donc de la résolution des ions fragments désirée (Tableau 4). Pour un temps de cycle de 2s (pour un minimum de 8 à 10 points par pic chromatographique) et une résolution de 17 500 soit 64ms, 31 peptides peuvent être fragmentés (Equation 11). Si en plus, des fenêtres de temps de rétention de 2 min sont utilisées (*time-scheduled PRM*), ce nombre augmente et permet de cibler environ 900 peptides par séparation chromatographique de 60min.

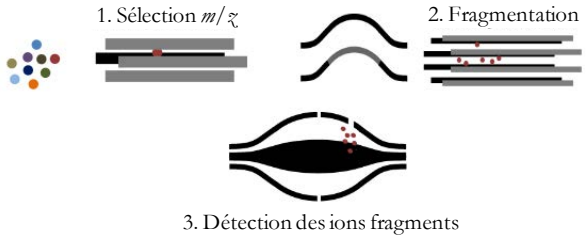


Figure 37 : Mode d'acquisition PRM sur le Q Exactive. L'ion d'intérêt est sélectionné dans le quadripôle puis accumulé jusqu'à un nombre d'ions désiré avant d'être analysé dans l'Orbitrap à HR/AM (figure adaptée de Michalski *et al.*[151]).

V.3.2.3 ***Multiplexage PRM***

Le QqOrbitrap permet de faire du multiplexage en PRM, augmentant la capacité du spectromètre de masse à suivre des peptides et ainsi quantifier plus de protéines. Une fois les ions sélectionnés dans le quadripôle, les ions sont accumulés dans la C-Trap puis transférés vers la cellule de collision HCD où ils sont fragmentés. Ces fragments restent ensuite piégés dans la cellule pendant que d'autres ions subissent une fragmentation. Un maximum de 10 ions de valeurs m/z différents peuvent être fragmentés et accumulés dans la cellule HCD avant d'être renvoyés simultanément vers la C-Trap pour être refocalisés et éjectés vers l'Orbitrap, où ils seront détectés ensemble (Figure 38).

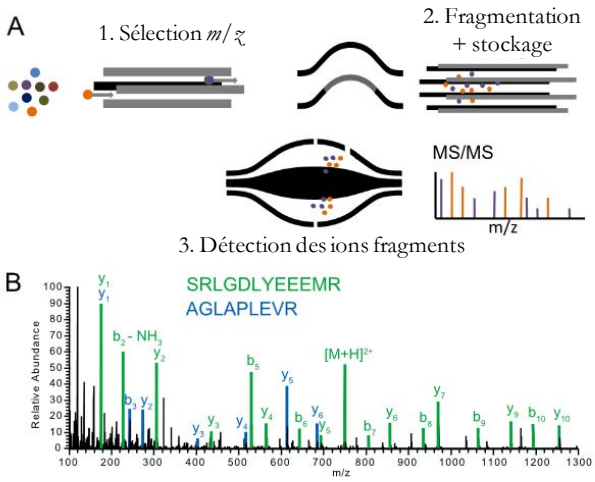


Figure 38 : Mode d'acquisition multiplexe PRM sur le Q Exactive. La tension du quadripôle est fixée pour ne laisser passer qu'un type d'ions de m/z donné. Après accumulation du nombre d'ions désiré dans la C-Trap, les ions sont envoyés vers la cellule de collision HCD pour être fragmentés puis stockés dans cette même cellule. Pendant la fragmentation, le quadripôle change rapidement de tension pour sélectionner une autre masse et accumuler ces ions dans la C-Trap. Ces étapes peuvent être répétées jusqu'à dix fois (multiplexe 10). Tous les ions fragments stockés dans la cellule sont ensuite transférés dans la C-Trap pour être éjectés et détectés en même temps dans l'Orbitrap à HR/AM. Sur cet exemple est présenté le profil de fragmentation de deux ions de rapport m/z différents, détectés simultanément (figure adaptée de Michalski *et al.*[151]).

Le nombre de peptides quantifiables est donc le rapport du temps de cycle (*cycle time*) sur le temps du transient multiplié par le degrés de multiplexage [153].

$$\#peptides = \frac{cycle\ time \times degrès\ multiplexage}{temps\ transient} \quad \text{(Equation 11)}$$

$$Max.\ fill\ time = \frac{temps\ transient}{degrès\ multiplexage} \quad \text{(Equation 12)}$$

Cette méthode PRM, de plus en plus employée dans les laboratoires de recherche en protéomique, sera surement à l'avenir la méthode de choix pour une quantification ciblée puisqu'elle emploie la haute résolution et une grande exactitude de mesure, et nécessite moins d'optimisation que le mode SRM. De plus, grâce à cette nouvelle possibilité de multiplexage pouvant aller jusqu'à un degré de 10, le nombre de peptides quantifiables s'approche ici de celui de la SRM.

V.4 QqOrbitrap HF (Q Exactive HF)

Une nouvelle configuration commerciale modifiée d'analyseur hybride QqOrbitrap a été introduite [154] à la Conférence de l'American Society for Mass Spectrometry en juin 2014 (62nd ASMS Conference on Mass Spectrometry and Allied Topics 2014, Baltimore, MD, USA.), le Q Exactive HF (Thermo Scientific) présenté sur la Figure 39.

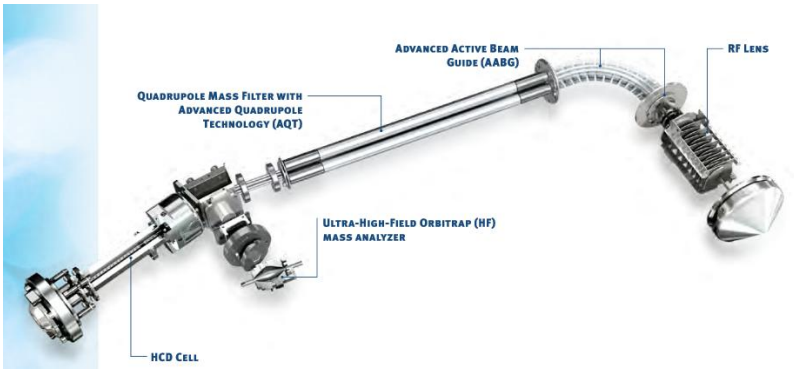


Figure 39 : Schéma du Q Exactive HF (Thermo Scientific). Les ions ionisés en nanoESI sont focalisés et transmis vers le quadrupole segmenté qui laisse passer soit tous les ions ou un type d'ions spécifiques. Ils traversent ensuite la C-Trap pour être soit éjectés vers l'Orbitrap à haut champ (HF) pour la mesure précise des ions précurseurs ou fragmentés dans la cellule HCD avant de revenir dans le C-Trap et enfin éjectés vers l'Orbitrap pour une mesure exacte des ions fragments.

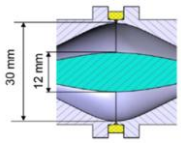
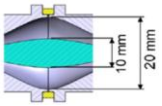
Le fonctionnement reste le même que le Q Exactive que nous possédons au laboratoire avec quelques améliorations technologiques annoncées (récapitulées dans le Tableau 5).

Tout d'abord au niveau de l'Orbitrap dont les dimensions sont réduites comme dans le cas de l'Orbitrap Elite (Annexe 5). En outre, la tension appliquée sur l'électrode centrale est de 5kV au lieu de 3,5kV. La fréquence d'acquisition est augmentée passant à 18 Hz pour une résolution de 15 000 (à m/z 200) contre 12 Hz pour une résolution de 17500 (à m/z 200) sur le Q Exactive. Pour une résolution proche, la vitesse d'acquisition est deux fois plus élevée (le transient passe est de 32 ms au lieu de 64 ms). 20 peptides peuvent être sélectionnés pour une fragmentation MS/MS en moins de 1s (mode d'acquisition DDA). En PRM, nous pourrons suivre deux fois plus de peptides dans un même intervalle de temps. Si nous travaillons au même temps de transient (64 ms par exemple), nous pouvons gagner en résolution (35 000 au lieu de 17 500) et donc en spécificité. Ces deux paramètres seront évalués dans le chapitre 7.

Outre l'Orbitrap, le quadripôle est modifié dans le Q Exactive HF, ce qui était déjà le cas dans la génération intermédiaire, le Q Exactive plus. Le quadripôle est composé de trois parties de quatre barres de section hyperbolique. Les deux parties externe fonctionnent en mode RF only : Une radiofréquence (RF) est appliquée sur les deux parties extrêmes comme sur le piège à ions linaire (Annexe 4) permettant une meilleure focalisation des ions, dans la partie centrale tout d'abord puis dans la C-Trap ensuite, avec un champ plus homogène avec réduction des effets de bord Les profils de transmission de fenêtre d'ions se rapprochent plus du créneau et ce d'autant plus que la fenêtre de sélection est large. La transmission des ions est améliorée et ce phénomène est d'autant plus important que le rapport m/z est élevé. Le gain en termes d'identifications de protéine est annoncé pour être d'autant plus prononcé que les échantillons sont faiblement abondants.

Un troisième changement est fait au niveau des optiques de la source d'ionisation. Pour une utilisation plus robuste sans encrassement de la source, les éléments suivants ont été rajoutés : un filtrage des ions injectés grâce à un guide d'ion permettant de modifier la trajectoire des ions avec un champ axial tout en élimant les contaminants neutres qui ne seront pas déviés, le *bent flatapole,* qui de plus est extractible pour une maintenance de routine.

Tableau 5 : Tableau comparatif entre le Q Exactive et le Q Exactive HF (Thermo Scientific)

Paramètres	Q Exactive	Q Exactive HF
Dimension Orbitrap		
Tension Orbitrap	3.5kV	5kV
Résolution (@ *m/z* 200) / Temps transient / Fréquence d'acquisition	17 500 / 64 ms / 12 Hz 35 000 / 128 ms / 7 Hz 70 000 / 256 ms / 3 Hz 140 000 / 512 ms / 1.5 Hz	15 000 / 32ms / 18 Hz 30 000 / 64 ms / 12 Hz 60 000 / 128ms / 7 Hz 120 000 / 256 ms / 3 Hz 240 000 /512 ms / 1.5 Hz
Quadrupole		
Optique ion	-	- Filtrage des ions injectés - Champ axial dans le *bent flatapole* - Guidage du faisceau d'ions

VI. Méga données (Big Data)

La volonté des protéomistes de vouloir identifier et quantifier toujours plus de protéines pour couvrir tout le protéome d'un organisme donnée, a conduit à des améliorations technologiques des spectromètres de masse. Ainsi, les instruments sont de plus en plus rapides, plus sensibles et plus résolutifs. Ceci se traduit par plus d'événements MS/MS déclenchés et des spectres enregistrés en profil (MS et MS/MS) pour décrire au mieux le pic chromatographique. De plus, pour couvrir au mieux le protéome, des gradients longs sont utilisés sur les systèmes chromatographiques.

Cela engendre alors un volume accru de données (*big data*) de sorte que les outils classiques de travail ne suffisent plus :

- la taille de fichier est de plus en plus grande. Il est donc important aujourd'hui d'avoir de grande capacité de stockage. Pour des analyses en DDA TOP10 sur le Q Exactive, chaque fichier pèse en moyenne 1 Go. À raison d'une analyse en triplicatas, un seul échantillon pèse 3 Go. Ramené à l'ensemble des échantillons analysés pour un projet complet, ceci aboutit très rapidement à plusieurs To de données.
- Le processus d'archivage doit être repensé permettant un accès rapide, tout en conservant une capacité de sauvegarde essentielle pour le suivi des patients. Les données seront certainement amenées à être retraitées avec les mises à jour constantes des banques de données.
- Il faut s'assurer d'une veille technologique régulière au niveau du parc informatique pour le retraitement des données. Les outils d'interrogations dans les banques de recherche ont dû également être adaptés pour traiter ces données volumineuses.
- Les outils statistiques dédiés aux *big data* sont aujourd'hui nécessaires pour valider les protéines identifiées et mettre en évidence les changements d'expression des protéines entre différentes conditions. Par exemple, la méthode de *clustering* hiérarchique (chapitre 5) pour regrouper les groupes d'individus semblables peut également être utilisée ou encore l'analyse en composantes principales (ACP) qui est un outil facile à mettre en œuvre pour distinguer différentes cohortes de patients (chapitre 6). Cette ACP consiste à transformer des variables liées entre elles en de nouvelles variables décorrélées les unes des autres appelées « composantes principales ». Lors du dernier congrès HUPO 2014 à Madrid, toute la communauté scientifique s'est accordée sur l'importante de ces outils statistiques pour légitimer les résultats et comprendre les observations biologiques.

Partie II : Résultats et discussions

Chapitre 4 : *Analyse protéomique globale d'agrégats de protéines insolubles : application aux amyloses*

I. Introduction

Au cours de la dernière décennie, l'analyse protéomique globale, appelée *shotgun proteomics*, de type LC-MS/MS est devenue la méthode de choix pour étudier le protéome d'un système biologique dans un environnement donné. Comme nous l'avons vu au Chapitre 2, les protéines subissent une digestion enzymatique afin de générer des peptides de plus faible masse qui seront analysés par chromatographie liquide couplée à la spectrométrie de masse à HR/AM (*bottom-up proteomics*).

Il s'agit d'identifier de la façon la plus exhaustive possible, les protéines présentes dans un échantillon pour trouver des descripteurs protéiques les plus spécifiques et discriminants possibles comme biomarqueurs potentiels de maladie. En second lieu, une quantification différentielle de ces protéines entre deux populations d'échantillons, le contrôle et l'échantillon d'intérêt, est nécessaire pour valider les biomarqueurs.

Nous avons vu que, dans le cadre de cette étude, la pathologie étudiée est la maladie des amyloses. Cette pathologie, peu connue, est considérée comme une famille de maladies orphelines, hormis le cas de la maladie d'Alzheimer. La caractéristique commune de ces maladies est la formation de dépôts de protéines insolubles de différentes natures biochimiques selon le type d'amylose considéré.

Si l'analyse protéomique de protéines solubles est devenue une analyse de routine dans de nombreux laboratoires, l'analyse d'agrégats de protéines insolubles est plus délicate. En effet les protocoles classiques impliquent une solubilisation et s'avèrent inefficaces. Par ailleurs, il est indispensable de travailler sur de très faibles quantités de matériel brut prélevé et conservé dans des conditions peu contrôlables. En effet, les prélèvements sont réalisés par des praticiens selon des modes opératoires divers, sur des tissus de nature variée. La qualité de la prise en charge du patient étant un aspect fondamental, nous n'avons pas souhaité interférer dans ces étapes cliniques. Notre objectif dans cette première étape indispensable était donc de développer un protocole adapté à la multiplicité des supports obtenus en clinique (biopsies ou prélèvement de tissus divers, congelés ou paraffinés, fixés ou non, stockage de quelques jours

à quelques années, etc.). Il fallait, car il s'agissait de cas réels de patients, obtenir des informations précises et cohérentes avec les données d'immunohistochimie très rapidement. Cette première étape a été réalisée avec le partenariat de l'Association française contre l'Amylose que nous remercions ici chaleureusement.

Techniquement, les objectifs de cette partie sont donc:

- d'évaluer l'extraction et la digestion des protéines insolubles contenues dans des tissus frais ou fixés au formol et inclus en paraffine (FFPE) en utilisant une sonde qui va délivrer des ultrasons intenses et focalisés (HIFU) vers l'échantillon.
- d'évaluer le diagnostic des amyloses par une analyse protéomique *shotgun* à très haute résolution.

II. Passer de l'agrégat insoluble à la solution de peptides protéolytiques

II.1 Les échantillons

Les 12 échantillons pathologiques et 15 échantillons contrôles ont été collectés entre 2009 et 2013 au service d'anatomie pathologique de l'hôpital Tenon à Paris pour un diagnostic des amyloses. Les dépôts amyloïdes ont été caractérisés en immunohistochimie par le service de pathologie. Ils proviennent de divers tissus (glande salivaire, rein, poumon, humeur vitrée, testicule et rate, résumé au Tableau 6) et ont été traités différemment selon des dates et lieux de prélèvement (paraffine, ultracongélation). Certains échantillons très peu enrichis en dépôts amyloïdes ont pu être micro-disséqués par LCM.

Tableau 6 : Liste des échantillons analysés au cours du projet.

N°Echantillon	Origine tissus	Pathologie (IHC)	Fixation	% amylose	Surface mm²	coupes 10 µm soit 0,01 mm	volume mm³
11H00439	Poumon	témoin	oui	-	266.10	0.06	15.966
11H00560	Poumon	témoin	oui	-	316.30	0.06	18.978
11H00556	Poumon	témoin	oui	-	250.80	0.06	15.048
11H00637	Poumon	témoin	oui	-	371.10	0.06	22.266
11E00051	Poumon	2 possibles	oui	-	10.80	0.38	4.104
08E00235	Poumon	blind test	oui	-	9.50	0.69	6.555
10H012063	Bronches	inconnu	oui			minuscule fragment entier	
10H08306	Rein	témoin	oui	-	239.70	0.06	14.382
10H08895	Rein	témoin	oui	-	386.60	0.06	23.196
10H09354	Rein	témoin	oui	-	313.70	0.06	18.822
08E00253	Rein	témoin (LGM)	oui	-	fragment		
10H09471	Rein	témoin	oui	-	266.60	0.06	15.996
10H09813	Rein	témoin	oui	-	296.60	0.06	17.796
10E00578 LCM	Rein (LCM)	AL	oui	>90%	1.82	0.12	0.2184
10E00578 raw	Rein (Raw)	AL	oui	50%	Fragment 3mm		
07E00532	Rein	AA	oui	50%	1/2 fragment biopsie 1.5mm		
10E00579	Rein	AL	oui	<10%	fragment 2mm		
10H09569	BGSA	témoin	oui	-	7.50	0.38	2.85
11h00799	BGSA	témoin	oui	-	17.30	0.7	12.11
11H00935	BGSA	témoin	oui	-	17.40	0.32	5.568
11H00890	BGSA	témoin	oui	-	20.30	0.32	6.496
07H02968	BGSA	TTR	oui		7.10	0.23	1.633
10E00443	BGSA	AA	oui	20-30%	Fragment 5mm		
08E00980	Testicule	ApoA1	oui		94.20	0.05	4.71
10H11423	Humeur vitrée	inconnu	non		non mesurable	gel œil en entier	non dilué
11H06267	Rate (raw)	AA	non	>90%	fragment 3x4x1.5mm		18
RT1.01.1	Rate (raw)	témoin	non	-	fragment 1x1x1mm		

II.2 Traitement des échantillons

La première étape de notre étude s'est centrée sur l'extraction efficace des constituants protéiques caractéristiques, afin de déterminer la nature des protéines amyloïdogènes dans les tissus affectés. La nature insoluble des protéines d'intérêt nous a rapidement orientés vers une autre approche. Nous savons en effet que :

- Il est très difficile d'extraire efficacement les protéines intactes des agrégats,
- La plupart de ces protéines existent sous forme soluble dans l'organisme,
- Nous cherchons à générer les peptides protéolytiques issus de ces protéines.

Puisque les protéines sont extraites entières avec un faible rendement et que nous souhaitons analyser ces peptides, pourquoi ne pas extraire directement les peptides en protéolysant les agrégats de façon plus globale ? Il faut alors travailler à augmenter le rendement de protéolyse et de la solubilisation des peptides.

Nous avons vu qu'une technique efficace pour optimiser les rendements mais aussi pour solubiliser les échantillons faisait appel à l'utilisation d'onde ultrasonores (cf. chap.1 § III.2.1 page 44). C'est à partir de ces considérations que nous avons entrepris cette étape. Afin d'augmenter l'accessibilité des protéines pour l'endoprotéolyse, la microtomie, déjà mise en œuvre pour l'histologie, est intéressante. Nous avons donc tout d'abord conservé ce traitement, naturellement disponible pour les échantillons cliniques, auquel nous avons adjoint la protéolyse assistée par ultrasons. Lorsqu'elle n'était pas réalisée, les tissus ont été finement découpés.

Pour chaque échantillon, des coupes de 10 µm d'épaisseur sur les tissus FFPE ou 1mm^3 sur tissus frais ont été réalisées en pathologie et conditionnés dans un tube contenant un minimum de tampon de 50mM d'ammonium bicarbonate pH 7,5 (entre 100 et 500µl selon le volume tissulaire, ce volume étant généralement de 200µl). Pour améliorer l'accès aux protéines insolubles contenues dans les tissus, les échantillons sont d'abord broyés sous ultrasons (MicroSon XL) pendant 1 à 2min avec une puissance de 10 à 15kW selon les échantillons. Les protéines subissent ensuite une réduction des ponts disulfure au tris-(2-carboxylethyl)phosphine (TCEP, concentration finale 2.5mM, 30min à 50°C). Puis, les thiols libérés par ces réductions, sont carbamidométhylés avec de l'iodoacétamide (IAM, concentration finale 10mM, 1h à température ambiante) pour empêcher les ponts disulfure de se reformer spontanément. Une fois les protéines dénaturées, elles subissent une digestion tryptique sous ultrasons pendant un minimum de 90sec. Cette fois encore la puissance

appliquée oscille entre 10 à 15kW, selon la densité de la solution à digérer. L'ajout d'une solution d'acide formique (AF) aqueux concentré à 50% (v/v), permet d'arrêter la digestion, tout en limitant le volume apporté.

Avec ce traitement particulier, nous obtenons une solution de peptides protéolytiques proche des solutions obtenues classiquement en protéomique *bottom-up*. Il serait possible d'augmenter les rendements en ajoutant des tensioactifs aidant à la solubilisation comme l'octyl-beta-glucopyranoside. Il nous faut aussi optimiser l'étape de réduction/alkylation sous ultrasons. Il reste bien sûr de nombreuses particules insolubles, qui doivent être éliminées. Pour cela, une filtration sur microfiltres fonctionalisés en phase octadécyle C_{18} est systématiquement réalisée (Proxeon C_{18} Stage-tips™, Thermo Scientific). Les peptides sont ensuite élués avec une solution à 30% en acétonitrile (ACN). Enfin, la solution est séchée sous vide puis reconstituée dans une solution à 0.1% d'acide trifluoroacétique (TFA) pour être analysée en nanoLC-MS/MS (LTQ/FT-ICR).

En outre, afin de permettre une comparaison ultérieure entre chaque échantillon, le volume à injecter de digestat purifié avant injection, est normalisé en fonction du volume initial de tissus traité.

III. Du mélange protéolytique à la séquence des peptides

Tous les échantillons sont analysés en triplicatas sur un système automatique de nano-chromatographie liquide couplé à un spectromètre de masse à HR/AM (Ultimate 3000, LTQFT Ultra, Thermo Scientific).

III.1 Optimisations chromatographiques

Les échantillons peptidiques issus de la digestion enzymatique des protéines contenues dans les biopsies sont extrêmement complexes. Ils comportent un grand nombre de protéines digérées qui de plus, présentent une gamme de concentration étendue. Avant toute analyse, il était nécessaire d'optimiser les paramètres chromatographiques. Le but étant d'accroître la capacité de pic, c'est à dire obtenir le maximum de composés séparés dans une fenêtre de temps donné, pour couvrir au mieux le protéome étudié. Les optimisations ont été réalisées sur un même mélange peptidique provenant d'une biopsie de rein témoin. Toutes les analyses ont été effectuées en triplicatas et seules les protéines identifiées de façon reproductible sont retenues dans les résultats: au minimum 2 peptides par protéine avec un FDR inférieur à 1%, les protéines devant être détectées dans les trois réplicas.

III.1.1 ***Optimisation du gradient d'élution***

Un moyen simple d'accroître la capacité de pics sans toucher au système chromatographique, est d'augmenter la durée du gradient d'élution. Le but étant de donner plus de temps aux peptides pour se séparer sur la colonne chromatographique et plus de temps au spectromètre de masse tandem de faire les événements MS pour la quantification et MS/MS pour l'identification.

Ainsi, nous avons voulu tester plusieurs durées de gradients d'élution chromatographique, sur une colonne de 50 cm pour commencer. Une même quantité d'échantillon témoin a été injectée pour chaque condition. Le gradient d'élution adopté est de 2% à 40% de tampon B (eau/acétonitrile, 90/10 (v/v) + 0.1% AF) sur **1h, 3h, 5h, 10h ou 15h**. Les profils chromatographiques sont donnés à la Figure 40 et varient de façon homothétique.

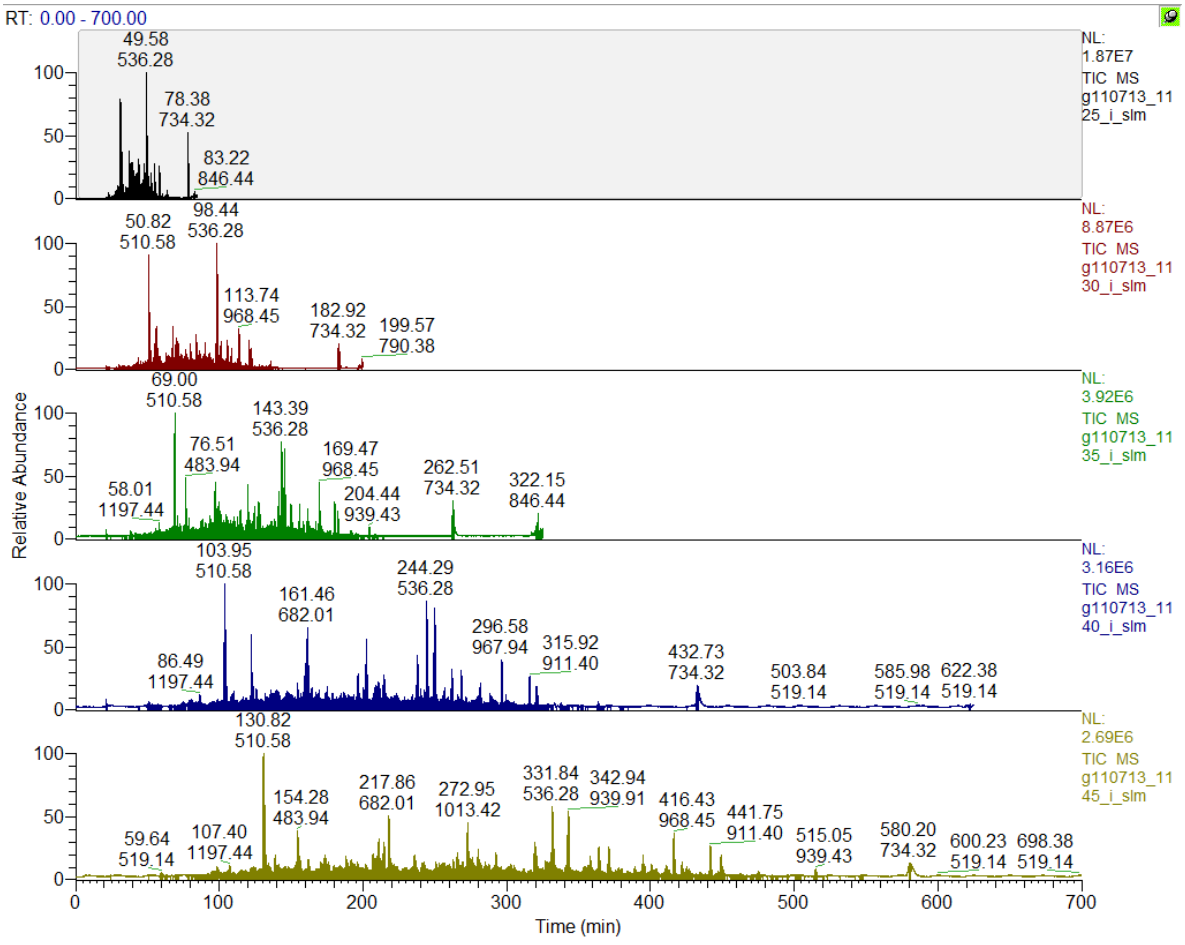


Figure 40 : Superposition des chromatogrammes de courant d'ions total obtenus avec différents gradients d'élution de 2 à 40% de tampon B. De haut en bas : 1h, 3h, 5h, 10h et 15h.

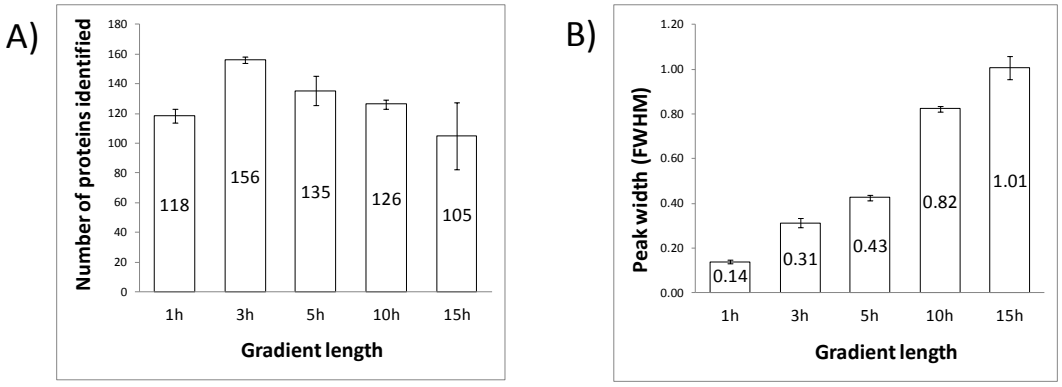


Figure 41 : Comparaison de la largeur de pic à mi-hauteur (FWHM) et le nombre de protéines identifiées avec 2 peptides pour différent temps d'élution chromatographique.

Comme le montre la Figure 41 A (ci-dessus) une meilleure couverture de l'échantillon est obtenue pour un gradient de 3 heures, avec 156 protéines identifiées de façon reproductible et avec une faible déviation standard (CV<2%). En augmentant le temps d'élution, le temps donné aux peptides pour se partager entre la phase stationnaire et la phase mobile augmente. Toutefois la largeur des pics augmente simultanément (Figure 41 B). Pour une même quantité d'échantillon, l'intensité des pics diminue jusqu'à descendre en dessous du seuil de détection efficace pour un séquençage informatif. Il s'agit ici de trouver un compromis entre le rapport signal sur bruit (S/B) et la durée de ce signal. Un S/B élevé donne des spectres plus informatifs, mais il faut avoir le temps de les générer. Pour cette application biologique, un gradient de 3h est le plus adapté.

III.1.2 ***Longueur de la colonne***

Augmenter la longueur des colonnes permet d'accroître le pouvoir résolutif de la séparation. En effet, en rajoutant de la phase stationnaire, les échanges entre le soluté et la phase stationnaire/phase mobile augmentent également. Cela se traduit par une augmentation du nombre de plateaux théoriques N (Equation 13). Plus ce nombre est grand, plus la séparation des composés est efficace.

$$N \propto \frac{L}{d_p} \qquad \text{(Equation 13)}$$

où N : nombre de plateaux théoriques, L : longueur de la colonne et d_p : diamètre des particules.

Cependant, une colonne longue de 50 cm engendre un surplus au niveau du coût de sa mise en œuvre, notamment de son conditionnement de 30 min au lieu de 15 min pour une colonne classique. De plus, une augmentation de la longueur entraine une perte de charge plus élevée (Equation 14) qui peut ne pas être supportée par un système classique de

chromatographie à haute pression. Nous avons par la suite, évalué l'utilisation d'un gradient de 3h sur une colonne classique de 15 cm et l'avons comparé aux résultats précédents.

$$\text{loi de Darcy : } \Delta P = \phi \frac{\eta L u}{d_p^2} \qquad \text{(Equation 14)}$$

où ⊿P : perte de charge dans la colonne, Φ: facteur de remplissage, η: viscosité dynamique, L : longueur de la colonne, u : vitesse linéaire de la phase mobile et d_p : diamètre des particules.

Pour vérifier cette théorie, nous avons également testé le gradient de 3h sur une colonne de 15 cm présentant la même nature de phase stationnaire (C_{18}), diamètre des particules (3µm) et diamètre interne de colonne (75µm). Dans chaque condition, la même quantité de mélange peptidique a été injectée dans le système chromatographique.

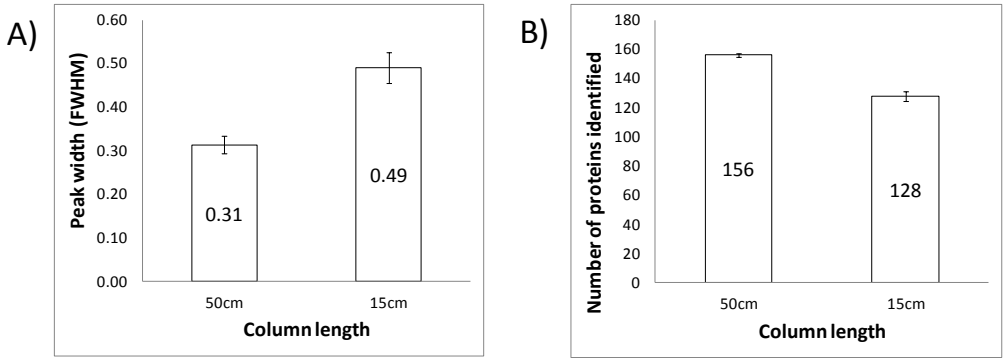


Figure 42 : Comparaison de A) la largeur de pic à mi-hauteur (FWHM) et B) le nombre de protéines identifiées avec 2 peptides entre une colonne de 15cm et 50 cm, avec deux peptides / protéine et un FDR<1%.

À quantité et gradient égaux, la séparation des peptides est meilleure sur une colonne de 50 cm que sur une colonne de 15cm, ce qui est en adéquation avec la définition du nombre de plateaux théoriques (Equation 13). Selon l'équation de Purnell (Equation 15), l'augmentation du nombre de plateaux théoriques engendre une meilleure résolution des pics. Les pics sont alors plus fins (Figure 42 A). Etant donné que les peptides sont mieux séparés, pour une même quantité injectée, les pics sont plus intenses et le nombre de déclenchement d'événements MS/MS est plus élevé. Le nombre d'identification de protéines est donc plus important avec une colonne de 50 cm avec 156 protéines identifiées contre 128 protéines pour la colonne de 15cm (Figure 42B). Il reste toutefois du même ordre de grandeur et est supérieur à ce qui était obtenu avec la colonne de 50 cm pour des gradients de 1h et 10h. Nous tolérons donc une perte de 18% du nombre de protéines identifiées avec 2 peptides distincts, ces protéines comptent parmi les espèces mineures, ce qui ne doit pas être rédhibitoire pour le diagnostic et la classification des amyloses reposant sur l'identification des composants majeurs des tissus.

$$R = \frac{\sqrt{N_B}}{4} \times \frac{\alpha - 1}{\alpha} \times \frac{k_B + 1}{k_B} \quad \text{(Equation 15)}$$

où R : résolution, N_B : nombre de plateaux théoriques, a : sélectivité et k : facteur de rétention.

III.1.3 ***Conclusion***

L'augmentation de la longueur de colonne et du gradient d'élution chromatographique sans toutefois diluer l'échantillon et perdre l'information de séquençage, permet d'accroître le nombre d'identifications de peptides et de protéines associées et d'augmenter la couverture en protéome des espèces étudiées.

Malgré une diminution du nombre de protéines identifiées pour une colonne de 15 cm, il reste largement supérieur au nombre maximum de protéines spécifiques à rechercher pour caractériser le type de pathologie.

Dans ce qui suit, nous avons finalement choisi de travailler avec une colonne de 15cm et un gradient de 3h pour conserver des conditions d'analyse robustes et rapides, transposables aux études cliniques.

III.2 Conditions nanoLC-MS/MS finale

Brièvement, les peptides sont tout d'abord chargés sur une pré-colonne C_{18} (Acclaim PepMap 100, C_{18}, 5µm, 300µm d.i. x 5mm L., Dionex) à un débit de 15µl/min avec un tampon A' (eau/ACN : 98/2 (v/v) + 0.05% (v/v) TFA). Après 6 min de dessalage, les peptides sont transférés et séparés sur une colonne nano à polarité de phases inversée C_{18} (Acclaim PepMap 100, C_{18}, 3µm, 75µm d.i. x 15cm L., Dionex) par un gradient d'élution chromatographique de 2% à 40% de tampon B en 170 min puis de 40% à 50%B en 10min à un débit constant de 220nl/min vers le spectromètre de masse (tampon B: eau/ACN : 90/10 (v/v) + 0.1% FA ; tampon A: eau/ACN : 98/2 (v/v)+ 0.1% FA). Les échantillons séparés par chromatographie liquide sont ensuite analysés sur un spectromètre de masse à piège ionique linéaire couplée à une cellule FT-ICR équipé d'une puce multi-émetteurs nanoESI de 5µm de diamètre interne (TriversaNanomate, Advion).

La méthode employée pour les analyses MS est le DDA Top 7 : une acquisition en mode FT-ICR MS suivie de 7 acquisitions en mode LTQ MS/MS sur les 7 précurseurs les plus intenses, avec une exclusion dynamique de 60s. Les ions précurseurs (peptides) de m/z 400 à 2000 ont été mesurés avec une résolution de 50 000 sur le FT-ICR.

IV. Traitement des données

IV.1 Recherche dans les banques de données

IV.1.1 Critères de recherche

L'identification et la quantification des protéines des biopsies, sans marquage des données MS et MS/MS obtenus sur le LTQ-FT Ultra (Thermo Scientific), ont ensuite été confrontées aux banques de données via un logiciel constructeur (Proteome Discoverer v1.4). Ce logiciel offre la possibilité de faire des recherches d'identification en utilisant parallèlement deux algorithmes de recherche indépendants : Mascot (Matrix Science) [78] et Sequest (Thermo Scientific) [79], et de comparer les échantillons entre eux sous forme de tableau.

Nous avons utilisé des critères de recherche classiques qui découlent directement des spécifications de nos analyses, afin de générer les listes de candidats potentiels. Les spectres expérimentaux MS/MS ont été comparés aux spectres de fragmentation théoriques prédits pour chaque peptide issu de la digestion des protéines contenues dans des banques de données. Pour la recherche Mascot, la banque SwissProt filtrée sur la taxonomie humaine a été utilisée. Dans le cas de la recherche Sequest, pour des raisons de configuration logiciel, il faut directement télécharger la banque sur le site d'Uniprot. Or pour la taxonomie humaine, seul le fichier regroupant les banques SwissProt et TrEMBL est disponible. Au final, nous ne retenons que les protéines de SwissProt et seuls les peptides tryptiques avec au plus 2 sites de coupures manqués sont pris en compte. L'oxydation des méthionines et la carbamidométhylation des cystéines ont été renseignées comme modifications dynamiques. La masse du précurseur étant mesurée par l'analyseur à HR/AM de type FT-ICR, sa tolérance en masse est de 5 ppm. Pour les fragments MS/MS obtenus dans le piège à ions, la tolérance en masse est de 800mDa.

IV.1.2 Critères de validation

Pour un diagnostic des amyloses par analyse protéomique globale sans *a priori*, les critères de validation des peptides et protéines appliqués sur les données sont assez stricts. Seuls les peptides ayant **un taux de faux positifs (FDR) inférieur à 1% et de rang d'identification de 1** sont considérés. De plus, seules les protéines identifiées **dans les deux moteurs de recherche** et **avec au moins 3 peptides sont conservés**. Les résultats finalement sont exportés sous forme de tableau pour un traitement ultérieur.

IV.2 Quantification des protéines

L'abondance des peptides a été évaluée en appliquant la méthode sans marquage *Top 3 Protein Quantification*, détaillée dans la partie I de ce manuscrit (cf. chap.2 § III.3.2 page 64). La quantification se fait par extraction du chromatogramme d'ions total (XIC) avec une tolérance d'erreur de justesse de 2ppm puis une intégration de l'aire sous la courbe (AUC). Cette méthode permet de corréler la moyenne des AUCs des trois peptides les plus intenses à la quantité de protéine dans le mélange biologique. Pour le diagnostic, les profils protéiques des échantillons contrôles sont ensuite comparés à ceux des patients. **Seuls les échantillons présentant les deux biomarqueurs de la maladie des amyloses, le composant P amyloïde (SAMP) et l'apolipoprotéine E (ApoE), sont considérés comme candidats potentiels présentant une amylose**. Enfin, **l'amylose est classifiée selon l'aire relative des protéines amyloïdogènes** listées dans le Tableau 7 ci-dessous.

Lorsqu'une protéine amyloïdogène n'est pas quantifiée dans un échantillon donné, la moyenne des aires minimales quantifiées dans chacun des trois réplicas techniques lui est attribuée. La protéine amyloïdogène majoritaire, à part les protéines SAMP et ApoE, est par défaut considérée comme responsable des dépôts amyloïdes.

Tableau 7 : Listes des protéines amyloïdogènes pris en compte dans cette étude.

Accession number	Amyloidoses	Protein description
P02743	non specif	SAMP_HUMAN Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 SV=2
P02649	non specif	APOE_HUMAN Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1
P02766	ATTR	TTHY_HUMAN Transthyretin OS=Homo sapiens GN=TTR PE=1 SV=1
P04279	ASemI	SEMG1_HUMAN Semenogelin-1 OS=Homo sapiens GN=SEMG1 PE=1 SV=2
Q9UKY0	AprPsc	PRND_HUMAN Prion-like protein doppel OS=Homo sapiens GN=PRND PE=1 SV=2
Q86SH4	AprPsc	PRNT_HUMAN Putative testis-specific prion protein OS=Homo sapiens GN=PRNT PE=5 SV=1
P04156	AprPsc	PRIO_HUMAN Major prion protein OS=Homo sapiens GN=PRNP PE=1 SV=1
P01236	Apro	PRL_HUMAN Prolactin OS=Homo sapiens GN=PRL PE=1 SV=1
Q08431	Amed	MFGM_HUMAN Lactadherin OS=Homo sapiens GN=MFGE8 PE=1 SV=2
P61626	Alys	LYSC_HUMAN Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1
P0CG06	All	LAC3_HUMAN Ig lambda-3 chain C regions OS=Homo sapiens GN=IGLC3 PE=1 SV=1
P0CG05	All	LAC2_HUMAN Ig lambda-2 chain C regions OS=Homo sapiens GN=IGLC2 PE=1 SV=1
P0CG04	All	LAC1_HUMAN Ig lambda-1 chain C regions OS=Homo sapiens GN=IGLC1 PE=1 SV=1
P0CF74	All	LAC6_HUMAN Ig lambda-6 chain C region OS=Homo sapiens GN=IGLC6 PE=4 SV=1
A0M8Q6	All	LAC7_HUMAN Ig lambda-7 chain C region OS=Homo sapiens GN=IGLC7 PE=1 SV=2
P01834	Alk	IGKC_HUMAN Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1
O14960	Alect2	LECT2_HUMAN Leukocyte cell-derived chemotaxin-2 OS=Homo sapiens GN=LECT2 PE=1 SV=2

P02788	Alac	TRFL_HUMAN Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6
Q15582	Aker	BGH3_HUMAN Transforming growth factor-beta-induced protein ig-h3 OS=Homo sapiens
P01308	Ains	INS_HUMAN Insulin OS=Homo sapiens GN=INS PE=1 SV=1
P10997	AIAPP	IAPP_HUMAN Islet amyloid polypeptide OS=Homo sapiens GN=IAPP PE=1 SV=1
P01871	Ahm	IGHM_HUMAN Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3
P01857	AHg	IGHG1_HUMAN Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1
P06396	Agel	GELS_HUMAN Isoform 2 of Gelsolin OS=Homo sapiens GN=GSN
P06396	Agel	GELS_HUMAN Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1
P02671	Afib	FIBA_HUMAN Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2
P01034	Acys	CYTC_HUMAN Cystatin-C OS=Homo sapiens GN=CST3 PE=1 SV=1
P01258	Acal	CALC_HUMAN Calcitonin OS=Homo sapiens GN=CALCA PE=1 SV=2
P05067	Abeta	A4_HUMAN Amyloid beta A4 protein OS=Homo sapiens GN=APP PE=1 SV=3
P61769	Ab2M	B2MG_HUMAN Beta-2-microglobulin OS=Homo sapiens GN=B2M PE=1 SV=1
P06727	AApoAIV	APOA4_HUMAN Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3
P02652	AApoAII	APOA2_HUMAN Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1
Q8NCW5	AApoAI	AIBP_HUMAN Isoform 2 of Apolipoprotein A-I-binding protein OS=Homo sapiens GN=APOA1BP
P02647	AApoAI	APOA1_HUMAN Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1
P01160	AANF	ANF_HUMAN Natriuretic peptides A OS=Homo sapiens GN=NPPA PE=1 SV=1
P02735	AA	SAA_HUMAN Serum amyloid A protein OS=Homo sapiens GN=SAA1 PE=1 SV=2
P0DOJI8	AA	SAA1_HUMAN Serum amyloid A-1 protein OS=Homo sapiens GN=SAA1 PE=1 SV=2
P0DJI9	AA	SAA2_HUMAN Serum amyloid A-2 protein OS=Homo sapiens GN=SAA2 PE=1 SV=1
P35542	AA	SAA4_HUMAN Serum amyloid A-4 protein OS=Homo sapiens GN=SAA4 PE=1 SV=2

V. ***Résultats***

L'analyse protéomique *bottom-up,* par digestion tryptique sous ultrasons, permet de faire un diagnostic rapide des amyloses à partir de tissus frais ou fixés. Les résultats sont cohérents avec ceux issus des analyses immunohistochimiques obtenus indépendamment à l'aide d'une série d'anticorps dirigés contre les protéines amyloïdogènes les plus fréquentes. Les exemples de biopsies de rein, de rate et de poumons témoins ainsi que celles des patients illustrent les résultats.

V.1 Cas des reins

V.1.1 Témoins

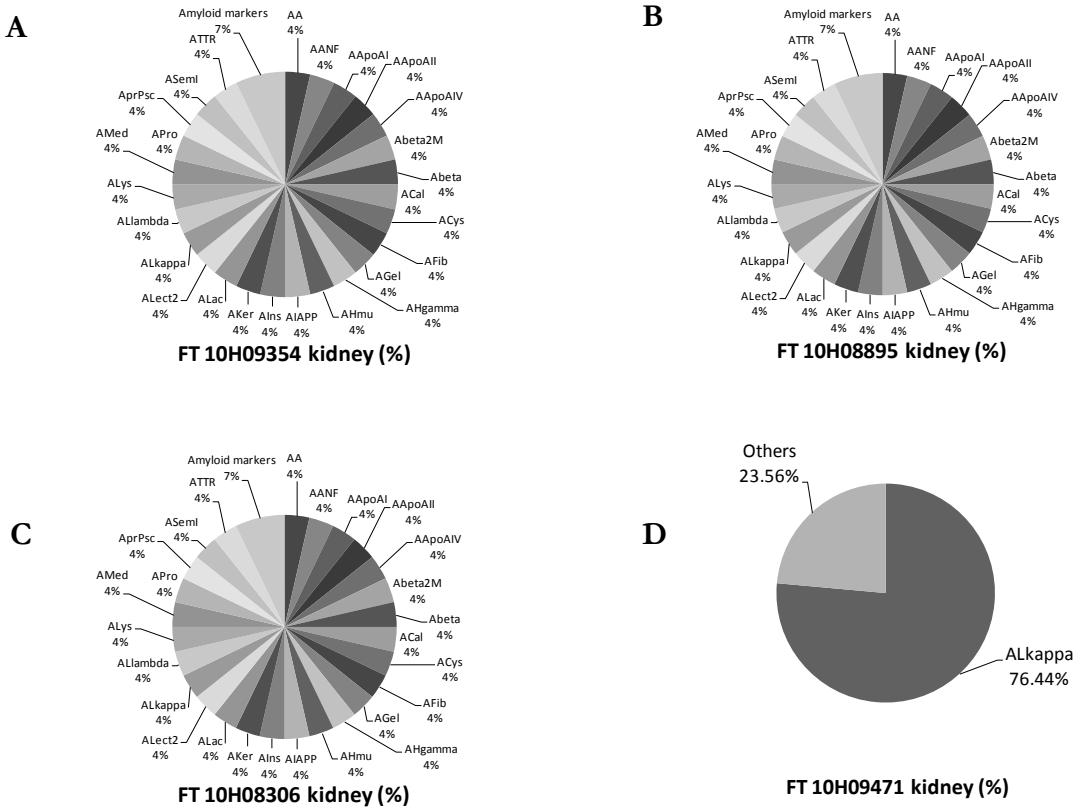


Figure 43 : Biopsies de rein de contrôles négatifs. (A, B, C) aucune protéine amyloïdogène n'a été quantifiée. (D) Ce témoin présente une proportion non négligeable de protéine Ig kappa mais ne présente par les protéines SAMP et ApoE, marqueurs des amyloses. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

Les témoins analysés sont appelés témoins négatifs, dans la mesure où les patients sur lesquels ont été réalisées les biopsies ne sont pas atteints de la maladie des amyloses mais d'une autre pathologie. En effet un témoin déclaré comme sain est difficile à obtenir puisqu'une extraction d'un morceau d'organe est nécessaire. Ce type d'échantillon est donc extrêmement peu probable, puisque la biopsie est pratiquée par définition chez les patients. Le contrôle idéal, d'un point de vue opératoire, serait de pratiquer une biopsie sur le même patient dans le même tissu mais dans une zone ne contenant aucun dépôt amyloïde. Cependant, cette pratique est peu souhaitable pour le patient puisqu'il s'agit d'un second prélèvement tout aussi invasif. Pour l'intérêt scientifique, il serait possible de produire ce type de contrôle lors de la microdissection LCM en prélevant spécifiquement les zones exemptes de dépôt.

Il se peut donc que des protéines d'intérêt soient même retrouvées chez les témoins, comme dans l'échantillon 10H09471, Figure 43 D. Cependant, dans ce cas, les deux biomarqueurs des fibrilles amyloïdes (SAMP et ApoE) ne sont pas détectés de façon significative. Ce profil est par ailleurs compatible avec une autre pathologie de type glomérulaire qui se caractérise par un enrichissement en chaines légères d'immunoglobuline kappa (Ig kappa) [155]. Les autres témoins (Figure 43 A, B, C) ne présentent pas de façon détectable de protéine amyloïdogène accumulée dans les reins.

V.1.2 ***Patients amyloses***

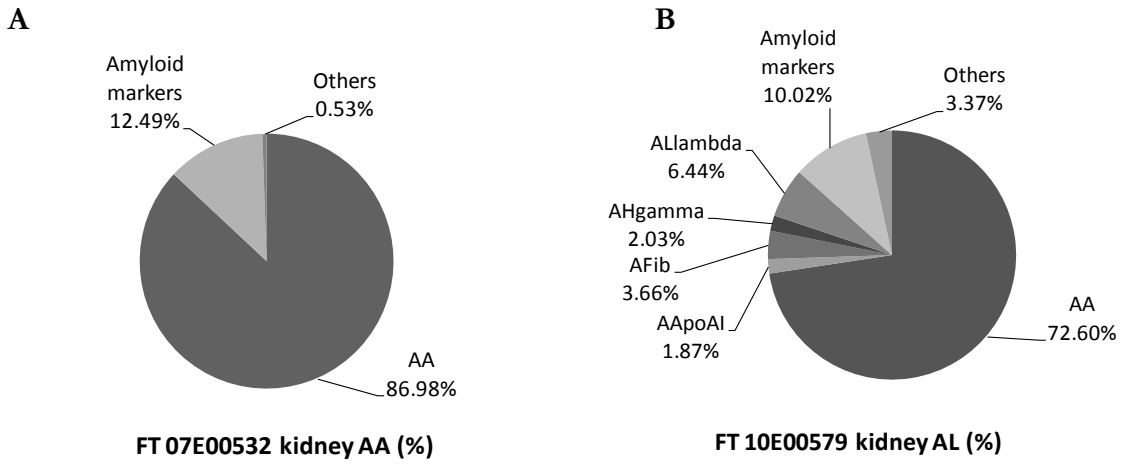


Figure 44 : Biopsies de rein de patients atteints d'une amylose. (A) Amylose de type AA avec 80% de dépôt fibrillaire (B) Amylose de type AL avec 10% de dépôt fibrillaire. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

L'échantillon dont les résultats sont exposés en Figure 44A, ne présente aucune ambigüité de diagnostic. Il montre clairement les protéines SAMP et ApoE à hauteur de 12 ,5% (notées *amyloid markers* sur la figure) ainsi qu'une forte teneur en protéine Sérum amyloïde A (SAA), dont le pourcentage calculé à partir des protéines amyloïdogènes répertoriées dans le Tableau 7, atteint 86,98%. Ce patient présente une amylose de type AA, diagnostic confirmé par le pathologiste.

Les résultats de l'échantillon exposés en Figure 44B, révèlent également une amylose puisque 10.02% des protéines amyloïdogènes correspondent aux protéines SAMP et ApoE. Ce patient présente une forte accumulation des protéines SAA (AA, 72.60%) en plus des chaines légères d'Ig Lambda (AL lambda, 6.44%), Ig Gamma 1 (AH gamma, 2.03%), fibrinogène alpha (AFib, 3.66%) et d'apolipoprotéine A I (AApoA1, 1.87%). L'analyse en immunohistologie a de son coté, révélé une amylose de type AL (Ig Lambda) et non une

amylose de type AA. Il faut noter que dans ce cas, le test n'a pu être effectué qu'avec un nombre restreint d'anticorps et non l'ensemble des anticorps disponibles pour les caractérisations d'amyloses. Grâce à l'analyse sans *a priori* en spectrométrie de masse, nous montrons que ce patient présente vraisemblablement une amylose mixte de type AA/AL.

V.1.3 ***Intérêt de la microdissection***

Pour ce dernier cas, une micro-dissection par capture laser (LCM) est nécessaire pour lever l'ambigüité. Cette technique est utilisée en routine dans certains départements cliniques américains pour collecter uniquement la partie amyloïdogène sans contamination des tissus alentours [49,51–54,14]. Cependant, ce procédé nécessite un équipement spécial absent à l'hôpital Tenon, et un savoir-faire particulier. De plus, cette étape peut être chronophage et certains tissus qui présentent des micro-hétérogénéités ne peuvent pas être disséqués par LCM. Nous avons alors évalué la possibilité de nous passer de ce mode de préparation en comparant un échantillon de rein avec et sans traitement au LCM.

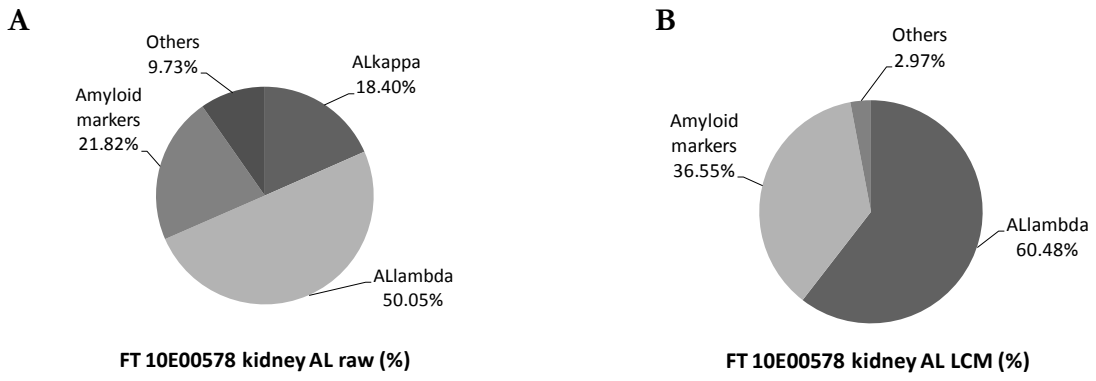


Figure 45 : Biopsies de rein de patients atteints d'une amylose. (A) Tissu non fixé (frais) avec 50% de dépôt fibrillaire, présentant une amylose de type AL (B) Tissu fixé découpé par laser (LCM) avec 50% de dépôt fibrillaire présentant une amylose de type AL. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

Dans le cas de la Figure 45 A, les tissus fixés ne sont pas coupés par microtomie et le dépôt amyloïde ne représente que 50% du tissu. Cet échantillon est positif pour une amylose (21.82% de protéines SAMP et ApoE). Cependant, le typage est ambigu. En effet, même si un dépôt de chaines légères d'Ig Lambda (50.05%) est le plus abondant, un dépôt de chaines légères Ig kappa (18.40%) est détectable. On peut rappeler ici que dans le cas du témoin rein numéro 10H09471 (Figure 43 D), un dépôt d'Ig kappa a pu être également détecté. Le diagnostic penche ici en faveur d'une amylose de type AL, confirmé par les analyses histologiques.

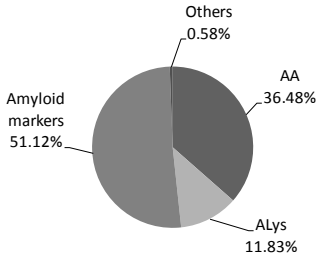
Pour lever cette ambigüité de typage, une dissection par LCM a été réalisée sur ce même échantillon présenté à la Figure 45 B. Les tissus collectés après dissection contiennent désormais plus de 90% de dépôts amyloïdes. Dans ce cas, le diagnostic est clair. Le patient présente une amylose de type AL : 36.55% de protéines SAMP et ApoE, et seules les chaines légères d'Ig lambda (60,48%) sont détectées.

V.2 Cas de la rate

Le traitement sous ultrasons permet également de faire un typage des amyloses sur différents organes tel que la rate. Les résultats d'analyse des échantillons de rate malade et témoin ni fixés et ni microtomés sont présentés sur la Figure 46A et B.

Pour ce tissu le diagnostic est sans ambiguïté : les protéines SAMP et ApoE sont détectées à un pourcentage supérieur à 51%. La protéine AA majoritaire est estimée à plus de 36% et le lysozyme C à plus de 11% (Figure 46A). Or, le lysozyme C qui est un antimicrobien, est constitutif du tissu. En effet, il est détecté en majorité dans l'échantillon contrôle présenté à la Figure 46B. Ce patient présente alors une amylose de type AA confirmée par immunohistologie.

A



B

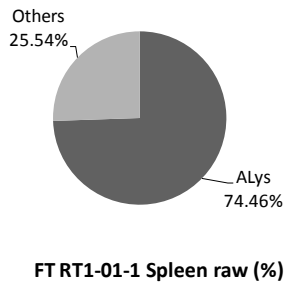


Figure 46 : Biopsie de rate. (A) Échantillon présentant une amylose de type AA avec 90% de dépôt fibrillaire. (B) Témoin rate. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

V.3 Cas des bronches/poumons

V.3.1 Témoins

Tous les témoins du tissu poumon présentent le même profil. Aucune protéine amyloïdogène de la liste n'a été détectée de façon significative. Deux exemples de biopsies négatives aux amyloses sont montrés à la Figure 47 A et B.

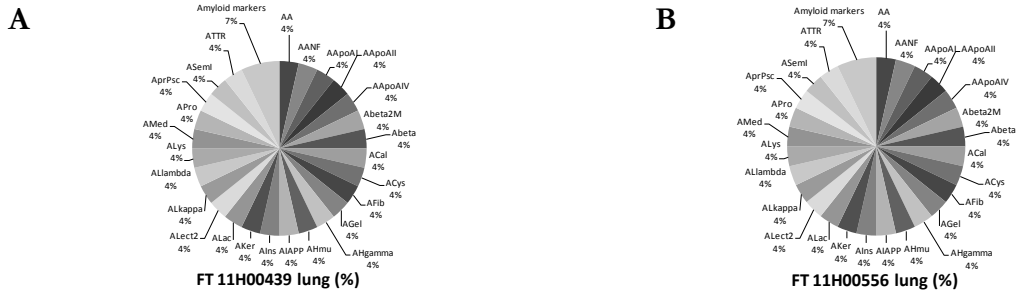


Figure 47 : Biopsies contrôles de poumon. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses.

V.3.2 ***Patient amylose***

Seule une biopsie de bronches d'un patient susceptible de présenter une amylose nous a été confiée pour une analyse en spectrométrie de masse.

Ce patient (Figure 48) présente bien une amylose (SAMP et ApoE 49.53%). Les protéines lysozyme C (26.04%) et Ig gamma 1 (20.16%) sont détectées. Or le lysozyme est une défense naturelle contre les infections dans les voies respiratoires [156,157]. Le diagnostic par spectrométrie de masse s'oriente alors vers une amylose de type AHg (Ig gamma 1). Malheureusement ce tissu est trop diffus pour être diagnostiqué en immunohistologie. Cet exemple montre bien les limitations de cette dernière technique et la nécessité d'une technique complémentaire telle que celle que nous proposons.

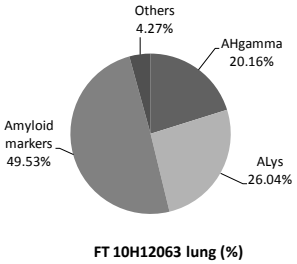


Figure 48 : Biopsie de bronches présentant une amylose de type AH. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

V.3.3 ***Patients non amyloses***

Des échantillons de poumons présentant une pathologie proche des amyloses ont été également analysés au laboratoire avec notre traitement aux ultrasons.

Ces deux patients présentent un dépôt non amyloïde, puisqu'aucun marqueur des fibrilles amyloïdes n'a été détecté de façon significative. En revanche, ces dépôts sont très enrichis en Ig kappa, 90.29% pour l'échantillon présenté à la Figure 49 A, et 85.85% pour Figure 49 B.

Ces deux patients ont développé une maladie de dépôt de chaines légères (LCDD, *light-chain deposition disease*) décrit récemment dans une analyse par spectrométrie de masse par nos collaborateurs [42,43].

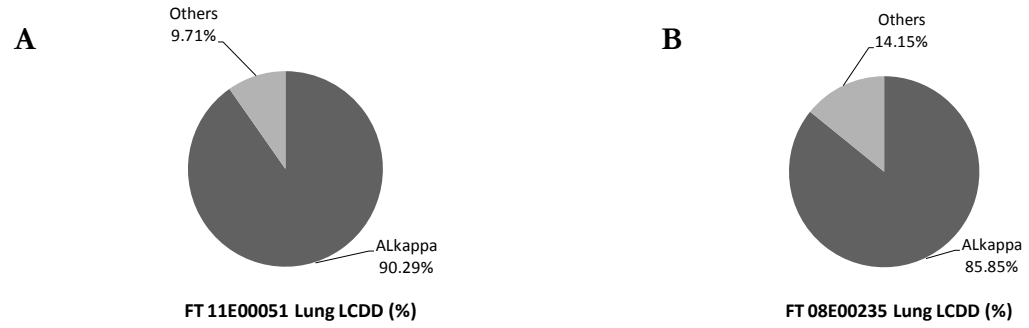


Figure 49 : Biopsie de poumon malade présentant des maladies de dépôt de chaines légères Ig kappa non amyloïdes (LCDD). La catégorie « Others » regroupe toutes les autres protéines amyloïdogènes non détectés.

V.4 Conclusion

Le traitement des biopsies par les ultrasons a permis d'extraire les protéines amyloïdogènes insolubles des tissus. La classification des amyloses peut se faire par quantification relative sans marquage en spectrométrie de masse tandem à HR/AM.

Les deux exemples (échantillon numéro 10H00579 à la Figure 44B et 10H00578 à la Figure 45A) de dépôts amyloïdes représentant moins de 50% des tissus montre la limitation de notre approche qui nécessite des échantillons concentrés en dépôts amyloïdes pour un diagnostic sans ambigüité. Lorsque ce cas se présente, une micro-dissection LCM s'avérera utile pour confirmer le diagnostic lorsque cela est possible.

Lorsque ce pourcentage s'élève à plus de 50% le diagnostic obtenu par notre approche est robuste et fiable. De plus il peut être employé sur différents tissus (rein, rate, poumon ou encore glandes salivaires accessoires, testicules, humeur vitrée non montrés dans cette partie).

Notre approche permet également de distinguer les dépôts amyloïdes des dépôts non amyloïdes comme la maladie des chaines légères (LCDD) qui ne présentent pas les biomarqueurs des fibrilles amyloïdes, les protéines SAMP et ApoE.

VI. Data Mining: visualisation graphique globale

Afin de proposer une visualisation globale de tous les résultats, un tableau général en code couleur appelé cartographie *heatmap*, rassemble l'intégralité des données de quantification

relative des protéines issues du traitement enzymatique sous ultrasons des biopsies de différents organes (glandes salivaires accessoires, reins, poumons, rate, testicules et humeur vitrée).

VI.1 Cartographie heatmap sous un tableur

VI.1.1 Traitement des données

La cartographie *heatmap* est générée après avoir collecté toutes les abondances des peptides quantifiés dans les analyses de protéomique *shotgun* par le logiciel Proteome Discoverer. Seuls les peptides appartenant aux protéines amyloïdogènes du Tableau 7 identifiés avec un FDR de moins de 1% et un classement de rang 1 dans l'identification sont considérés.

Les peptides de même séquence mais différemment modifiés sont regroupés sous une même valeur somme. Tous les AUCs des peptides sont ensuite normalisées par l'AUC du chromatogramme du courant d'ions total (TIC) de l'expérience correspondante. L'aire reportée dans la cartographie heatmap correspond à la moyenne des AUCs normalisées des trois réplicas, à condition que le peptide soit vu dans les trois expériences. Lorsque le peptide amyloïdogène n'est pas détecté de façon significative dans un échantillon donné, la plus faible valeur lui est alors attribuée. Elle correspond au rapport de la plus petite aire intégrée sur l'aire maximale des TICs de toutes les expériences confondues.

VI.1.2 Résultats

La cartographie *heatmap*, montrée à la Figure 50, est obtenue à l'aide d'un programme sous Visual basic appelé *sparklines* compatible avec les tableurs [158].

Chaque colonne correspond aux différents échantillons analysés, classés par nature des tissus et des pathologies. Chaque ligne correspond à l'intensité normalisée des peptides protéotypiques spécifiques à une protéine amyloïdogène. Ces peptides sont regroupés par protéine d'intérêt pour une visualisation plus claire. Plus la cellule est verte claire, plus le peptide associé est abondant. Or, les peptides sont les éléments descripteurs des protéines. Donc plus cette cellule tend vers le vert clair, plus la protéine correspondante est abondante dans l'échantillon.

Globalement, nous pouvons constater que les mêmes peptides protéotypiques sont détectés de façon reproductible dans la plupart des expériences, indépendamment de la nature

des tissus, de la fixation et du traitement par LCM. Le traitement par ultrasons est donc assez reproductible d'un échantillon à un autre. Ces peptides sont alors de bons candidats pour développer une méthode ciblée pour un typage des amyloses avec un appareil plus robuste et adapté aux départements cliniques. Ce travail qui fait suite à ces analyses protéomiques *shotgun*, sera détaillé dans le chapitre suivant.

Les deux premières protéines répertoriées sur la cartographie *heatmap* sont les protéines marqueurs des amyloses, SAMP et ApoE. Moins de 3 peptides par échantillon appartenant à ces deux protéines marqueurs sont détectés dans les échantillons non amyloses (colonnes contrôles de 13 à 27) alors qu'un signal massif est enregistré pour les dépôts amyloïdes (colonnes patients de 1 à 10).

Cette cartographie *heatmap* permet donc de classifier facilement les échantillons amyloses des échantillons non amyloses. Le typage est alors donné par la protéine de la liste en dessous des deux biomarqueurs possédant plusieurs peptides les plus abondants. Globalement presque tous les échantillons contiennent du lysozyme C en quantité plus ou moins abondante dans l'échantillon. Pour ce type d'amylose (ALys), il faudra faire de la micro-dissection LCM pour être sûr que la présence de la protéine identifiée soit responsable de la pathologie des amyloses et non d'une surproduction due à l'infection.

Le typage pour chaque échantillon analysé en triplicata par spectrométrie de masse *shotgun* à HR/AM peut être décrit de la façon suivante :

- Les **échantillons 1, 3, et 8** sont des **amyloses de type AA**. Plusieurs peptides abondants appartenant aux protéines SAA1 et SAA2 codant pour cet amylose, y sont ainsi détectés.
- L'é**chantillon 2** montre plusieurs peptides de la transthyrétine (TTHY) et seulement deux peptides de chaines légères d'Ig lambda (LACs) dont un très intense. Ce patient a **une amylose de type ATTR**.
- L'**échantillon 4** présente plusieurs peptides appartenant aux protéines SAA1/SAA2, Ig kappa (IGKC), Ig lambda (LACs), Ig gamma 1 (IGHG1), transthyrétine (TTHY), apolipoprotéine A1 (APOA1) et fibrinogène alpha (FIBA). Ce cas, ambigüe, a été présenté à la Figure 44B. Le pourcentage de dépôt amyloïde est inférieur à 10%. Cependant, les peptides de la SAA1/SAA2 sont les plus intenses. Le diagnostic penche plutôt pour **une amylose de type AA** avec des dépôts secondaires d'autres protéines citées précédemment.

- Les **échantillons 5 et 6** sont des **amyloses de type AL** même si l'échantillon 6 contient des dépôts secondaires de chaines légères d'Ig kappa (IGKC) et chaines lourdes d'Ig gamma 1 (IGHG1). Ce résultat est cohérent puisque l'échantillon 5 correspond à une dissection par LCM de l'échantillon 6 montré aux Figure 45 A et B.
- L'**échantillon 7** contient des peptides des protéines lysozyme C (LYSC) et Ig gamma 1 (IGHG1). Or, le lysozyme est un antimicrobien sécrété dans le cas d'infection pulmonaire [156,157]. Ce patient a vraisemblablement **une amylose de type AHg** présenté à la Figure 48.
- L'**échantillon 9** présente des peptides appartenant majoritairement aux protéines apolipoprotéine A1 (APOA1) et lactotransferrine (TRFL). Ce dernier est une glycoprotéine transporteur du fer mais est également un antimicrobien impliqué dans la protection du liquide séminal [159–161], ce qui semble en corrélation avec la nature du tissus. Ce patient présente alors **une amylose de type ApoA1**.
- L'**échantillon 10** présente plusieurs peptides intenses de la transthyrétine (TTHY) et cystatine C (CYTC). Ce dernier a été décrit comme protéine constitutive de l'humeur vitrée [162]. Ce patient est atteint d'**une amylose de type ATTR**.
- Enfin, deux échantillons patients se distinguent du lot : les **échantillons 11 et 12**. Ils ne contiennent ni de SAMP ni d'ApoE mais des dépôts majoritaires de chaines légères d'Ig kappa (IGKC). Ces deux patients présentent une maladie de **dépôt de chaines légères non amyloïdes** (LCDD) [42,43].

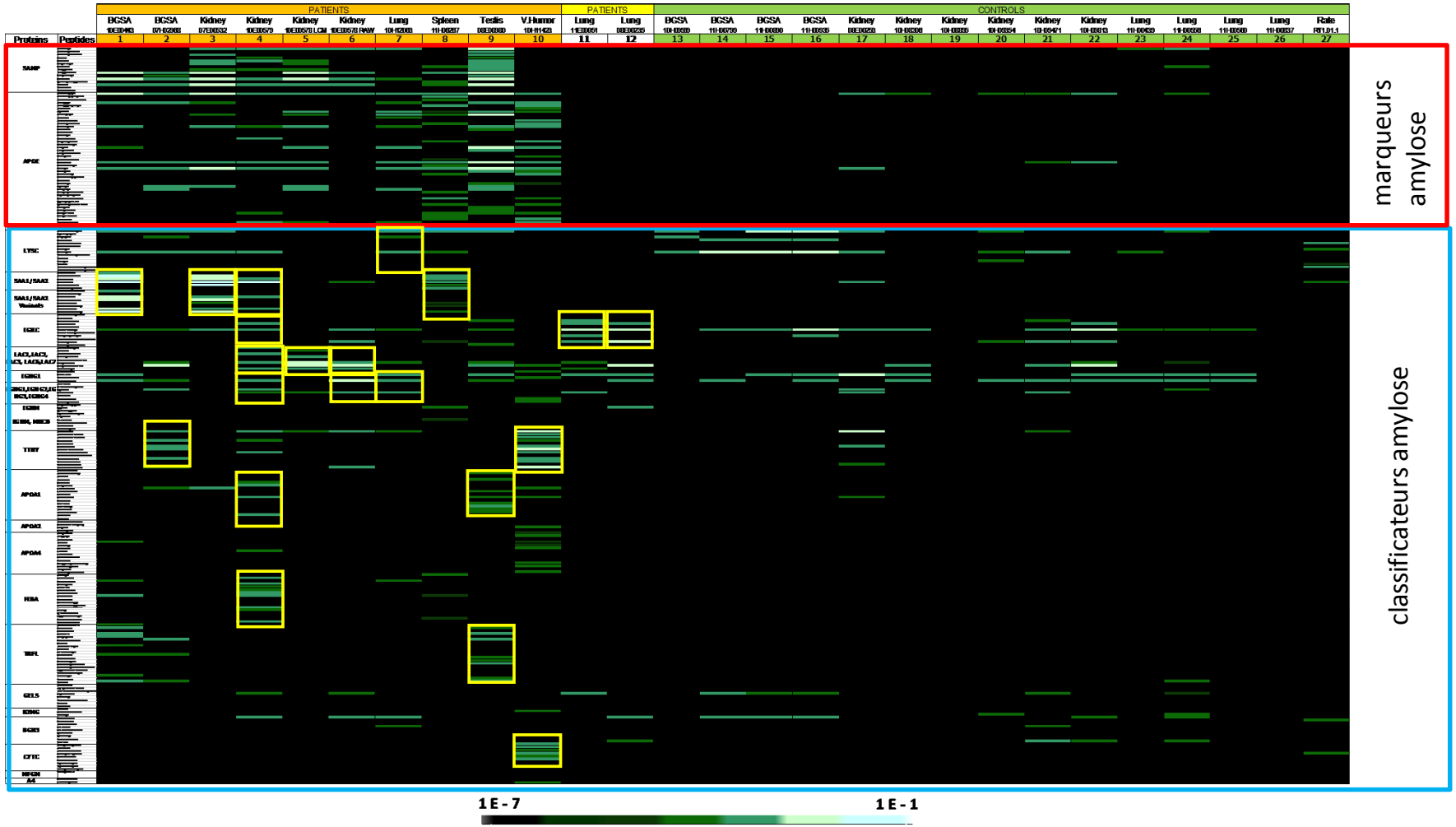


Figure 50 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans les expériences protéomiques *shotgun* de différents tissus avec Proteome Discoverer. La couleur des cellules reflète l'intensité des signaux.

VI.2 Visualisation plus automatisée avec MaxQuant et Perseus

Le regroupement hiérarchique (*clustering*) précédent a été réalisé manuellement. Afin d'évaluer la faisabilité de notre méthode à discriminer les échantillons présentant une amylose des autres pathologies, un *clustering* automatisé de nos données par une méthode à l'aveugle est nécessaire. Cela permettra de proposer une méthode de retraitement des données plus simple pour le praticien hospitalier. Plusieurs logiciels sont disponibles sur le marché. Nous nous sommes intéressés au logiciel MaxQuant qui est un logiciel libre de droit, capable de faire de l'identification et de la quantification sans marquage à partir d'extraction du chromatogramme d'ions (XIC) des données MS et M/MS. Contrairement au logiciel Proteome Discoverer, lorsqu'un peptide a déjà été identifié en MS/MS, le peptide peut être identifié et quantifié dans un autre réplica ou échantillon par son rapport m/z et son temps d'élution chromatographique. Le résultat est ensuite retraité par sa suite logiciel, Perseus, pour générer une cartographie *heatmap*[163,164].

VI.2.1 Recherche dans les banques de données

VI.2.1.1 *Paramètres expérimentaux*

Les données MS et MS/MS sont soumises à l'algorithme de recherche Andromeda [80], seul algorithme compatible et inclus dans le logiciel MaxQuant (v. 1-3-0-5). Afin de proposer une visualisation globale exploitable, il est préférable de restreinte la recherche à une banque ne contenant que les protéines d'intérêt. Une banque de données interne (157 séquences), regroupant les différentes protéines amyloïdogènes répertoriées dans les publications, est alors utilisée pour faire les recherches.

Les paramètres de recherche et de validation sont résumés dans le Tableau 8 et comparés aux paramètres utilisés dans la partie précédente sur Proteome Discoverer.

Tableau 8 : Paramètres de recherche pour les analyses shotgun avec les Proteome Discoverer et MaxQuant. La dernière colonne indique si les conditions sont strictement identiques.

	Proteome Discoverer (v1.4)	MaxQuant (v. 1-3-0-5)	Identique?
Algorithme de calcul	Mascot et Sequest	Andromeda	Non
Banque de données	SwissProt	Amyloses	Non
Enzyme	Trypsine	Trypsine	Oui
Nombre de coupures manqués	2	2	Oui
Etat de charge maximum	-	5	Non
Modifications dynamique	Carbamydométylation (Cystéine) Oxydation (Méthionine)	Carbamydométylation (Cystéine) Oxydation (Méthionine)	Oui
Tolérance MS	5 ppm	6 ppm *	Non
Tolérance MS/MS	0.8 Da	0.5 Da *	Non
Critères de validation	FDR <1% Rang 1 Peptides vus 3/3 replicas	FDR <1% Peptides vus 3/3 replicas Erreur alignement 3 min	Non

*** paramètres recommandés pour l'utilisation de MaxQuant**

VI.2.2 ***Résultats de la classification par regroupement hiérarchique***

L'abondance d'un peptide est estimée par la valeur d'AUC du massif isotopique complet obtenu à partir du XIC. Toutes ces valeurs issues de la recherche précédente avec MaxQuant sont ensuite importées sous Perseus (v1.4.1.3). Pour la suite seuls les peptides identifiés dans les trois réplicas techniques sont considérés. Les valeurs manquantes sont affectées de la valeur minimale associée à une identification toutes expériences confondues. Après passage en échelle logarithmique (log2), un regroupement hiérarchique en utilisant la distance euclidienne (Equation 16) comme mesure de dissimilarité des 27 échantillons selon les colonnes et les lignes est réalisé pour visualiser les données sous forme de cartographie *heatmap* avec des dendrogrammes.

$$||a - b||_2 = \sqrt{\sum_i (a_i - b_i)^2} \qquad \text{(Equation 16)}$$

Ce dendrogramme permet de visualiser l'arrangement des groupes d'échantillons générés par le regroupement hiérarchique. Les échantillons sont ainsi classés indépendamment de la nature des tissus par Perseus. Une représentation en heatmap est proposée de manière analogue (Figure 51). Les colonnes représentent les différents échantillons et les lignes, les peptides quantifiés. Plus la case est rouge, plus le peptide est abondant dans l'échantillon donné.

Bien que le dendrogramme ne permette pas de distinguer deux groupes principaux amyloïde/non amyloïde, les échantillons sont bien triés de telles sorte qu'on retrouve les deux groupes : à gauche les patients amyloses (PATH) et de l'autre les échantillons contrôles (CTRL) ne présentant point de pathologie de type amylose. Seuls deux échantillons de type

PATH se retrouvent dans les CTRL : ces deux échantillons particuliers proviennent de patients ne souffrant pas d'amylose qui présentent toutefois des dépôts de protéines non amyloïdes (échantillons PATH_Lung_11E00051 et 08E00235). Il faut noter que le typage à partir de cette analyse automatisée n'est pas aisé puisque les peptides ont été indépendamment clustérisés sans être regroupés par protéine d'intérêt. Il est à noter cependant que le classement obtenu reste cohérent avec la cartographie *heatmap* manuelle générée précédemment (Figure 50).

Les colonnes de gauche à droite correspondent aux colonnes de la Figure 50 : échantillon 10 (ATTR/ALac), échantillon 9 (AApoA1/ACys), échantillons 8-1-3 (AA), échantillon 4 (AA/AL), échantillon 2 (ATTR), échantillons 6 et 5 (AL) et échantillon 7 (AHg/Alys).

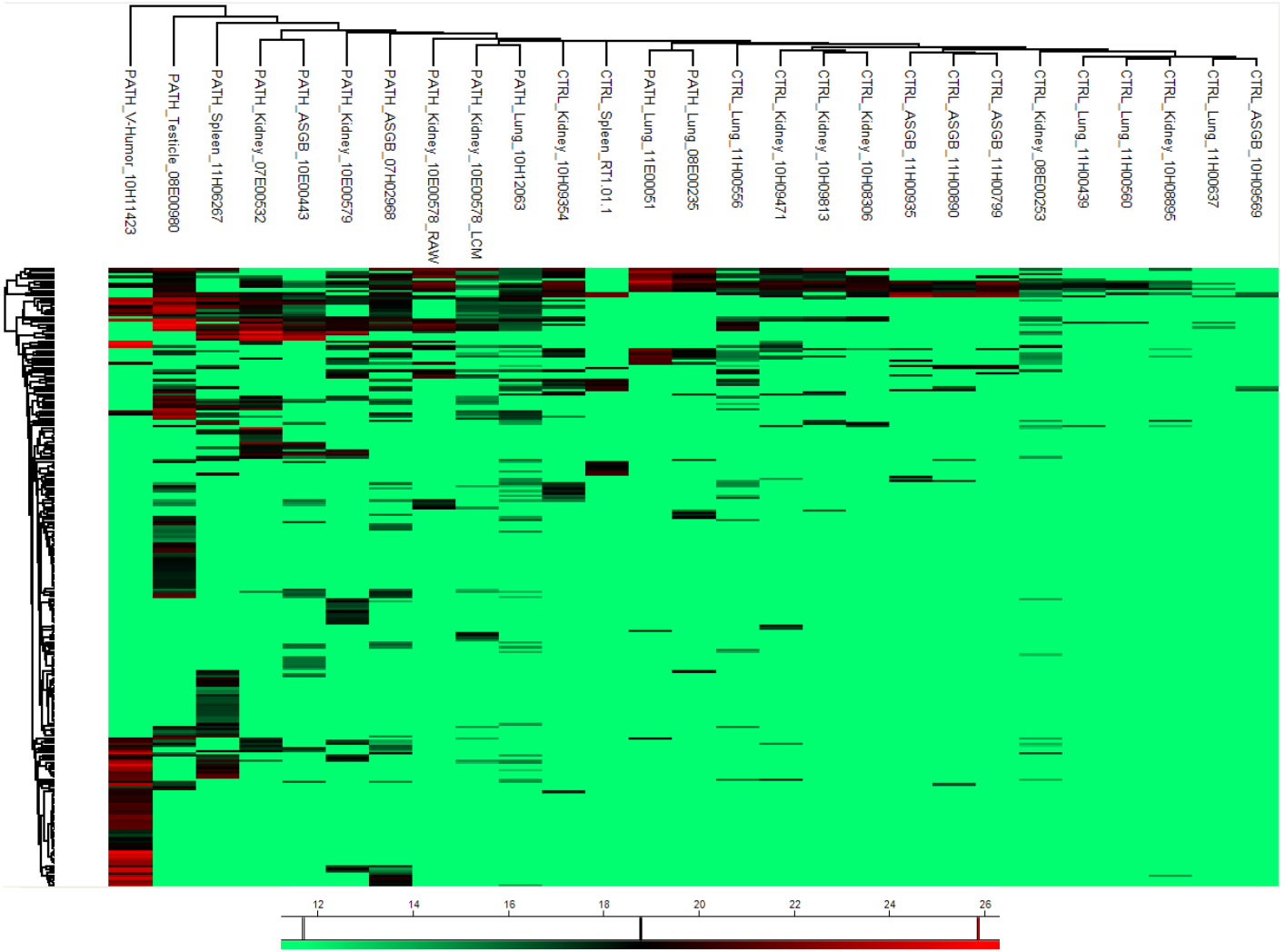


Figure 51 : Visualisation sous forme de cartographie *heatmap* des moyennes des intensités des peptides quantifiés dans les expériences protéomiques *shotgun* de différents tissus avec MaxQuant et Perseus. **La couleur des cellules reflète l'intensité des signaux selon l'échelle indiquée**.

*VII. **Conclusion***

L'analyse globale sans *a priori* par spectrométrie de masse *shotgun* de biopsies est possible. Après traitement enzymatique sous ultrasons, elle permet de typer différents dépôts amyloïdes présents dans différents organes. Les résultats obtenus sont en accord avec ceux obtenus indépendamment en immunohistologie. Contrairement à cette dernière technique utilisée en routine pour le typage des amyloses dans les départements cliniques français, aucun anticorps n'est nécessaire. Toutes les protéines sont visibles à conditions qu'elles soient assez abondantes et que les conditions d'ionisation ESI soient favorables à leur ionisation. Notre approche développée sans *a priori* met alors en évidence des dépôts amyloïdes non typés par immunohistochimie et des dépôts secondaires. De plus, les dépôts amyloïdes sont également distinguables des dépôts non amyloïdes. Grâce à l'usage des ultrasons pour la digestion enzymatique et l'analyse en gradient long sur un spectromètre à haute résolution de type FT-ICR, la préparation et l'analyse de l'échantillon nécessitent une journée et le retraitement informatique deux jours.

Il s'agit désormais de développer une approche ciblée pour la détection des protéines amyloïdogènes. La méthode sera plus spécifique et la détection sera meilleure en ciblant l'analyte. En revanche, cette méthodologie ferme la possibilité de découvrir de nouvelles protéines amyloïdogènes car seuls les candidats connus sont alors ciblés. Cette approche pourra être facilement transposée sur des appareils de spectrométrie de masse, déjà disponible dans certains laboratoires cliniques, qui sont plus robustes et moins onéreux.

Chapitre 5 : *Analyse protéomique ciblée en mode SRM : application aux amyloses*

I. *L'intérêt des approches ciblées*

Comme présenté au Chapitre 4, l'analyse protéomique globale est la méthode de choix pour étudier en première approche le protéome d'un système biologique dans un environnement donné. Les protéines sont identifiées grâce à des appareils à HR/AM. Cependant, elle n'est pas adaptée à la quantification de ces biomarqueurs du fait de sa faible sélectivité et de sa gamme dynamique limitée (Figure 52A). Son manque de reproductibilité, qui résulte du caractère stochastique de la sélection des précurseurs devant être fragmentés en mode DDA, fait que seules les protéines les plus abondantes sont bien souvent quantifiables de façon robuste. Pour éliminer ce caractère aléatoire, des approches alternatives ciblées ont émergé ces dernières années. Ainsi en ciblant un sous-ensemble d'analytes d'intérêt, l'analyse n'introduit pas de biais de quantification favorisant les protéines les plus abondantes, et elle permet une quantification précise des analytes au sein même d'une matrice biologique complexe.

L'analyse ciblée repose sur une étude préalable de l'échantillon, dans lequel un certain nombre de composés ont été sélectionnés comme descripteurs d'intérêt. Seuls ces composés seront recherchés au cours de l'analyse. Dans le cas de la protéomique, il s'agit d'isoler les composés de même rapport m/z que les peptides protéotypiques représentant les protéines cibles puis à les fragmenter afin d'en obtenir une signature spécifique. Seul un petit nombre de fragments est ensuite suivi pour la détection et quantification. Ce mode particulier permet d'augmenter de façon significative la spécificité, la sensibilité, la reproductibilité, la limite de détection et la gamme dynamique de l'analyse (Figure 52B). Originairement, cette approche appelée SRM (*selected reaction monitoring*) a été développée sur un spectromètre de masse de type triple-quadripôle à basse résolution. Aujourd'hui, une approche similaire, la PRM (*parallel reaction monitoring*), est possible sur un spectromètre de masse à haute résolution de type quadripôle-Orbitrap. Ces deux types d'approche ciblée, détaillées dans la partie I du manuscrit, sont utilisées dans ce projet.

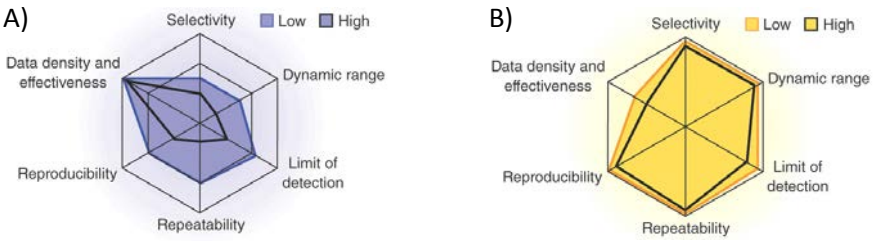


Figure 52 : Comparaison entre les performances des analyses shogun (A) et les analyses ciblées (B). Les analyses *shotgun* sont utilisées en première approche pour identifier le plus de protéines possible avec une limite de détection et une gamme dynamique moindres. Les analyses ciblées, plus sensibles, plus robustes et plus sélectives sont utilisées pour la quantification des protéines potentiellement marqueurs de maladies.[165].

II. *Analyse ciblée en mode Single Reaction Monitoring (SRM)*

Aujourd'hui, les laboratoires des départements cliniques des hôpitaux français sont, pour la plupart, équipés de spectromètres de masse de type triple-quadripôle. Ces instruments sont robustes, moins couteux et nécessitent peu d'entretien. Cependant, leur mode de fonctionnement à balayage en fait des configurations en général moins sensibles que les appareils à piège à ions, temps de vol ou à analyseur FTMS. Il y a cependant certains modes d'analyse ciblée ne nécessitant pas de balayage systématique, et dans ce type d'analyse les configurations à triple quadripôle sont extrêmement performantes. C'est le cas du mode SRM. Sensible et spécifique, ce mode consiste à exécuter une sélection en deux étapes de rapports m/z d'intérêt grâce aux différents quadripôles. Pour chaque peptide d'intérêt, le rapport m/z du peptide intact est sélectionné, puis cette espèce est fragmentée et un sous-ensemble de m/z, correspondant aux fragments principaux spécifiques attendus pour cette séquence, est recherché dans un second analyseur. C'est le mode multiplexe SRM (aussi parfois appelé MRM pour *Multiple Reaction Monitoring*). Si le temps de rétention de ce peptide dans les conditions de séparation de l'analyse est connu, il est possible de limiter dans le temps cette recherche sur une fenêtre centrée autour de ce temps de rétention. C'est le mode *time-scheduled* SRM (*tSRM*), ou encore SRM programmé en temps. L'objectif de cette partie est d'évaluer les possibilités de l'analyse en mode *t*SRM sur les prélèvements de tissus, avec comme but un transfert de technologie rapide, par la suite, dans les départements cliniques.

Les échantillons cliniques étant précieux et disponibles en très petites quantités, nous ne disposions que d'un seul échantillon (PATH_Spleen_11H06267) en quantité suffisante pour reproduire plusieurs expériences avec les différentes approches que nous souhaitions tester (en dehors des triplicatas opératoires). Seul cet échantillon a donc été analysé en *tSRM* sur le triple-quadripôle dont nous disposons au laboratoire, le TSQ Vantage (Thermo Scientific).

II.1 Stratégie expérimentale pour l'analyse tSRM

II.1.1 Conditions nanoLC-MS/MS

Dans le cadre d'une application clinique, les analyses en gradient de 3 heures sont trop longues. Notre première hypothèse de travail est que le gain en sensibilité et en spécificité doit nous permettre de conserver la même qualité de couverture de l'échantillon en un temps moindre. Afin d'établir une quantification ciblée pour un diagnostic des amyloses en clinique à plus haut débit, la durée du gradient d'élution de 2 à 40% de tampon B (eau/ACN 90/10 (v/v) + 0,1% (v/v) FA) a été réduite de 3h à 1h sur le même système chromatographique que les analyses *shotgun* (Ultimate 3000, Dionex), bien que nous avons vu dans le chapitre 4 qu'un gradient de 3h permettait d'identifier plus de protéines qu'un gradient de 1h par une augmentation de la capacité de pic du système.

Concernant la longueur de la colonne, nous avons conservé le même matériel de 15 cm utilisé pour les expériences *shotgun* (cf. chap.4 § III.2 page 107).

II.1.2 Choix des peptides cibles

II.1.2.1 Acquisition de nouvelles données

Désormais, les conditions chromatographiques sont différentes des analyses *shotgun*. Pour la quantification en mode *tSRM*, une série d'échantillons patients est tout d'abord analysée en DDA pour constituer une librairie de spectres MS/MS des peptides amyloïdogènes avec les nouveaux temps de rétention associés aux nouvelles conditions chromatographiques. Dans ce but, les échantillons séparés par nanoLC sont alors analysés sur un spectromètre de masse constitué d'un quadripôle couplé à un Orbitrap (Q Exactive, Thermo Scientific) équipé d'une source nanoESI classique avec des aiguilles en silice recouvert d'un matériau conducteur (360 µm d.e., 20 µm d.i., 10 µm d.i au bout de l'aiguille, New Objective).

Dans cette configuration, les ions peptidiques de 400 à 2000 Da sont mesurés dans des conditions d'analyse classique, avec une résolution de 70 000 (à m/z 200), un AGC de 1e6 et un temps d'injection maximum de 250 ms suivi d'une fragmentation des 10 ions les plus intenses (*TOP 10*) avec une exclusion dynamique de 30s à une résolution de 17500 (à m/z 200), un AGC de 5e4 et un temps d'injection maximum de 120 ms.

Ces échantillons analysés en triplicatas sont ensuite retraités via le logiciel Proteome Discoverer pour une recherche avec les algorithmes Mascot et Sequest dans la banque de

données SwissProt. Les paramètres de recherche renseignés sont les mêmes que pour les expériences *shotgun* sur le LTQ FT Ultra à l'exception des tolérances en masse (résumés en Tableau 9). Les peptides sont mesurés sur le LTQ FT Ultra avec meilleure justesse que sur le Q Exactive. À contrario, les fragments détectés dans l'Orbitrap ont une meilleure justesse de masse en mode MS/MS.

Tableau 9 : Paramètres de recherche des expériences *shotgun* sur les instruments LTQ FT Ultra et Q Exactive. La dernière colonne indique si les conditions sont strictement identiques.

	LTQ FT Ultra	QExactive	Identique?
Algorithme de calcul	Mascot et Sequest (Proteome Discoverer)	Mascot et Sequest (Proteome Discoverer)	Oui
Banque de données	SwissProt (Humaine)	SwissProt (Humaine)	Oui
Enzyme	Trypsine	Trypsine	Oui
Nombre de coupures manqués	2	2	Oui
Etat de charge maximum	-	-	Oui
Modifications dynamique	Cam (C) Ox(M)	Cam (C) Ox(M)	Oui
Tolérance MS	5 ppm	10 ppm	Non
Tolérance MS/MS	0.8 Da	20 mmu	Non

Cam (C) : carbamidométhylation des cystéines; Ox(M) : oxydation des méthionines

II.1.2.2 *Critères de sélection des peptides*

Une fois établies les nouvelles listes d'identification avec les nouveaux temps de rétention, les résultats en mode DDA (format msf) sont importés dans le logiciel Pinpoint (v1.4, Thermo Scientific) pour constituer la librairie de spectres et permettre une comparaison des peptides identifiés. Tous les peptides identifiés dans les expériences protéomiques *shotgun* sur le LTQ FT Ultra sont également considérés comme peptides cibles. Les temps de rétention associés seront calculés dans la suite grâce à leur facteur d'hydrophobicité.

Les peptides les plus intenses identifiés sur le Q Exactive et le LTQ FT Ultra, correspondant aux 28 protéines amyloïdogènes répertoriées dans le Tableau 10, sont choisis en priorité. Un maximum de 10 peptides par protéine est sélectionné.

Dans le Tableau 11, une comparaison est réalisée entre les peptides appartenant à une des deux protéines marqueurs des amyloses identifiées avec nos deux spectromètres de masse, la protéine SAMP. Ces peptides (identifiés avec FDR < 1%) sont classés par ordre d'intensité décroissante, c'est-à-dire par ordre décroissant des moyennes d'AUCs des trois réplicas techniques. Les peptides de cette protéine répertoriés dans la banque de données Peptide Atlas, qui regroupe différentes données mises à disposition par différents chercheurs de la

communauté scientifique [145,166], classés par score décroissant, sont également renseignés dans ce tableau. Globalement, les mêmes peptides sont détectés par le LTQ FT Ultra avec un gradient de 3h et le Q Exactive avec un gradient de 1h. La réduction du temps d'élution chromatographique ne semble donc pas être une limitation pour la quantification ciblée en *tSRM* sur ces espèces majeures. De plus, ces peptides identifiés expérimentalement concordent avec ceux renseignés dans la banque de données MS de Peptide Atlas. Notre traitement par ultrasons n'induit donc pas de biais au niveau de la digestion. Nous pourrons donc nous appuyer sur ces banques de données pour cibler des peptides qui n'auraient pas été détectés expérimentalement sur nos deux spectromètres de masse.

Concernant le choix des peptides, plusieurs conditions peuvent conduire à un clivage non effectif couramment accepté [4–6] : deux résidus basiques successifs (RR, KK, RK, KR) ainsi qu'un environnement acide par la présence d'acide aspartique (D) ou acide glutamique (E). D'autre part, un résidu lysine ou arginine suivi d'une proline (P) crée un encombrement stérique qui rend la zone peu accessible au site actif de la trypsine. Lorsque le cas se présente, les peptides vus avec et sans site de coupure manqué sont conservés pour la PRM s'ils sont intenses. Si l'on se penche sur le Tableau 11, nous constatons que le premier peptide le plus abondant détecté par nos deux spectromètres de masse, présente un site de coupure non clivé (IVLGQEQDSYGG**K**FDR). De plus, il est positionné en 7ème position dans la banque de données Peptide Atlas avec 32071 observations. Ce peptide ainsi que son homologue, dont l'intensité est non négligeable (IVLGQEQDSYGGK), seront gardés dans la liste finale (extrait donné en Tableau 12).

Tableau 10 : Liste des protéines amyloïdogènes sélectionnées pour la quantification en mode time-scheduled PRM.

Numéro accession	Nom accession	Nom protéine
P02647	APOA1_HUMAN	Apolipoprotéine A-I
P02652	APOA2_HUMAN	Apolipoprotéine A-II
P06727	APOA4_HUMAN	Apolipoprotéine A-IV
P02649	APOE_HUMAN	Apolipoprotéine E
P01034	CYTC_HUMAN	Cystatine C
P02671	FIBA_HUMAN	Fibrinogène chaine alpha
P06396	GELS_HUMAN	Gelsoline
P01857	IGHG1_HUMAN	Chaine constante Ig gamma-1
P01859	IGHG2_HUMAN	Chaine constante Ig gamma-2
P01860	IGHG3_HUMAN	Chaine constante Ig gamma-3
P01861	IGHG4_HUMAN	Chaine constante Ig gamma-4
P01834	IGKC_HUMAN	Chaine constante légère Ig kappa
P0CG04	LAC1_HUMAN	Chaine constante Ig lambda-1
POCG05	LAC2_HUMAN	Chaine constante Ig lambda-2
P0CG06	LAC3_HUMAN	Chaine constante Ig lambda-3
P0CF74	LAC6_HUMAN	Chaine constante Ig lambda-6
A0M8Q6	LAC7_HUMAN	Chaine constante Ig lambda-7
P01871	IGHM_HUMAN	Chaine constante lourde Ig mu
P04220	MUCB_HUMAN	Protéine chaine lourde Ig mu pathologique
Q08431	MFGM_HUMAN	Lactadhérine
P02788	TRFL_HUMAN	Lactotransferrine
P61626	LYSC_HUMAN	Lysozyme C
P0DJI8	SAA1_HUMAN	Sérum amyloïde A-1
P0DJI9	SAA2_HUMAN	Sérum amyloïde A-2
P35542	SAA4_HUMAN	Sérum amyloïde A-4
P02743	SAMP_HUMAN	Composant P du sérum amyloïde
Q15582	BGH3_HUMAN	Protéine facteur de croissance de transformation, beta induit ig-h3
P02766	TTHY_HUMAN	Transthyrétine

Tableau 11 : Comparaison des peptides de la protéine SAMP obtenus sur les spectromètres de masse LTQ FT Ultra et Q Exactive et Peptide Atlas.

LIT/FT-ICR (LTQ FT Ultra)			QqOrbitrap (Q Exactive)			Peptide Atlas		
Sequence	Score Mascot	Xcorr Sequest	Sequence	Score Mascot	Xcorr Sequest	Sequence	Score	# observation
IVLGQEQDSYGGKFDR	97	4.50	IVLGQEQDSYGGKFDR	91	4.11	VGEYSLYIGR	0.74	30468
VGEYSLYIGR	76	3.95	VGEYSLYIGR	50	2.92	AYSLFSYNTQGR	0.70	26459
QGYFVEAQPK	55	2.44	QGYFVEAQPK	40	2.72	QGYFVEAQPK	0.69	26290
AYSLFSYNTQGRDNELLVYKER	69	5.06	AYSLFSYNTQGRDNELLVYKER	56	4.79	GYVIIKPLVWV	0.65	21922
GYVIIKPLVWV	39	2.86	AYSLFSYNTQGR	45	2.70	DNELLVYK	0.57	14016
AYSLFSYNTQGR	85	3.93	IVLGQEQDSYGGK	61	3.07	IVLGQEQDSYGGK	0.55	12037
DNELLVYKER	51	3.42	DNELLVYKER	49	3.13	IVLGQEQDSYGGKFDR	0.51	32071
IVLGQEQDSYGGK	93	3.61	KGLRQGYFVEAQPK	38	2.95	AYSDLSR	0.44	1795
AYSLFSYNTQGRDNELLVYK	43	3.77	QGYFVEAQPKIVLGQEQDSYGGKFDR	60	4.84	AYSLFSYNTQGRDNELLVYK	0.40	17617
KGLRQGYFVEAQPK	-	3.21	AYSLFSYNTQGRDNELLVYK	37	4.59	ESVTDHVNLITPLEKPLQNFTLCFR	0.39	535
AYSDLSR	31	-	GLRQGYFVEAQPK	47	3.12	FPAPVHICVSWESSSGIAEFWINGTPLVK	0.36	239
GLRQGYFVEAQPK	39	3.23	ERVGEYSLYIGR	36	3.29	SQSFVGEIGDLYMWDSVLPPENILSAYQGTPLPANILDWQALNYEIR	0.34	38
ERVGEYSLYIGR	55	3.33	VIEKFPAPVHICVSWESSSGIAEFWINGTPLVKK	51	4.23	DNELLVYKER	0.31	4486
VIEKFPAPVHICVSWESSSGIAEFWINGTPLVKK	78	5.73	VGEYSLYIGRHK	31	-	AYSLFSYNTQGRDNELLVYKER	0.30	3660
VIEKFPAPVHICVSWESSSGIAEFWINGTPLVK	93	3.89				HTDLSGKVFVFPR	0.29	993
						ERVGEYSLYIGR	0.29	969
						GLRQGYFVEAQPK	0.27	650
						QGYFVEAQPKIVLGQEQDSYGGKFDR	0.26	1102
						VIEKFPAPVHICVSWESSSGIAEFWINGTPLVK	0.25	539
						KGLRQGYFVEAQPK	0.25	192
						QGYFVEAQPKIVLGQEQDSYGGK	0.24	190
						DNELLVYKERVGEYSLYIGRHKVTSK	0.24	48
						VGEYSLYIGRHKVTSK	0.24	23
						AYSDLSRAYSLFSYNTQGR	0.24	186
						KGLRQGYFVEAQPKIVLGQEQDSYGGKFDR	0.23	37
						GLRQGYFVEAQPKIVLGQEQDSYGGK	0.23	34
						GLRQGYFVEAQPKIVLGQEQDSYGGKFDR	0.23	27
						ERVGEYSLYIGRHK	0.23	20
						VIEKFPAPVHICVSWESSSGIAEFWINGTPLVKK	0.23	15
						KGLRQGYFVEAQPKIVLGQEQDSYGGK	0.23	13
						DNELLVYKERVGEYSLYIGR	0.23	8
						AYSDLSRAYSLFSYNTQGRDNELLVYK	0.22	2
						VGEYSLYIGRHK	0.22	2

Les peptides verts sont vus par les trois méthodes. Les peptides rouges ont été vus par le LTQ FT Ultra et Peptide Atlas seulement. Les peptides bleus ont été vus avec le Q Exactive et Peptide Atlas seulement. Dans l'ensemble, les peptides observés sur nos instruments sont identiques à ceux identifiés par les autres laboratoires de la communauté scientifique.

Les peptides modifiés sont également pris en considération pour établir la méthode. Lorsqu'un peptide portant une méthionine (oxydée ou non), est régulièrement détecté avec une forte intensité, le même peptide avec et sans oxydation est alors sélectionné. Les peptides contenant une ou plusieurs cystéines sont exclusivement conservés sous leur forme carbamidométhylés. Seul le peptide modifié sera préféré. Un exemple de protéine présentant ces deux cas est donné dans le Tableau 12.

Enfin, le choix des peptides en fonction de leur état de charge donne priorité au plus abondant. Quand deux états de charges sont d'intensité équivalente pour un même peptide, les deux états charges sont conservés. Dans ce cas, ils comptent pour deux peptides.

Tableau 12 : Exemple de sélection de peptides (protéines SAMP et ApoE).

Numéro accession	Nom accession	Nom protéine	Peptides
P02743	SAMP_HUMAN	Composant P sérum amyloïde	AYSLFSYNTQGR
			AYSLFSYNTQG**R**DNELLVY**K**ER
			DNELLVYK
			DNELLVY**K**ER
			GLRQGYFVEAQPK
			GYVIIKPLVWV
			IVLGQEQDSYGGK
			IVLGQEQDSYGG**K**FDR
			QGYFVEAQPK
			VGEYSLYIGR
P02649	APOE_HUMAN	Apolipoprotéine E	AATVGSLAGQPLQER
			A**K**LEEQAQQIR
			AYKSELEEQLTPVAEETR
			DRLDEV**K**EQVAEVR
			LAVYQAGAR
			LGAD**M**[Ox]EDV**C**[Cam]GR
			LGAD**M**EDV**C**[Cam]GR
			LGPLVEQGR
			LQAEAFQAR
			LSKELQAAQAR

II.1.2.3 ***Validation des peptides expérimentaux***

Les rapports m/z des trois premiers isotopes des peptides (précurseurs) sélectionnés sont générés par le logiciel Pinpoint. Les données brutes (fichiers raw) correspondant aux identifications (fichiers msf) utilisées pour créer la méthode sont importées dans le logiciel pour une quantification par la méthode des AUCs puis validation des peptides choisis. Les temps de rétention des peptides intégrés peuvent ensuite être reportés (ordonnées) en

fonction des temps expérimentaux de la librairie (abscisses) sur un graphique (Figure 53). Tout point en dehors de la droite est ainsi facilement visualisable pour être corrigé. Les AUCs sont réintégrées si besoin. Les taux de variations des AUCs entre les réplicas doivent être inférieurs à 15%.

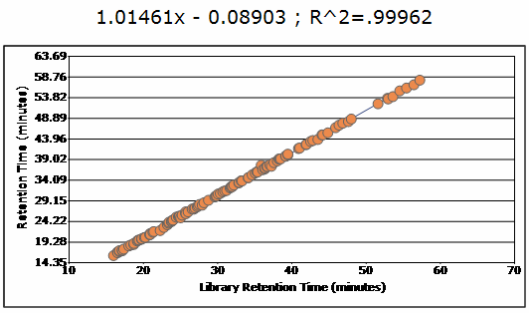


Figure 53 : Temps de rétention en fonction du temps de rétention de la librairie.

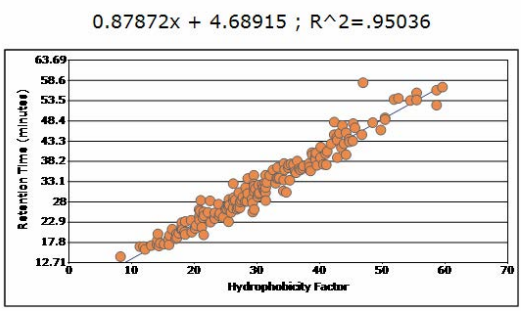


Figure 54 : Temps de rétention en fonction du facteur d'hydrophobicité prédit par l'algorithme SSRCalc.

Les temps de rétention intégrés peuvent également être reportés en fonction du facteur d'hydrophobicité calculé selon l'algorithme SSRCalc (*Sequence Specific Retention Calculator*) développé par Krokhin et Spicer [169,170]. En principe, le coefficient de corrélation (R^2) de la fonction générée doit être proche de 1 (bonne corrélation entre l'hydrophobicité et le temps de rétention). Ce qui est à l'exception de quelques points, notre cas sur la Figure 54. Cette fonction (y = 0.87872 x + 4.68915) est sauvegardée pour prédire les temps de rétention des peptides qui seront rajoutés *a posteriori* et qui n'ont jamais été détectés dans nos analyses *shotgun* avec un d'intervalle de temps d'élution chromatographique de +/- 1,5 min. Cela permettra de compléter notre liste de peptides des protéines déjà identifiées mais également celles des protéines amyloïdogènes ne faisant pas parties de la liste des protéines contenues dans les échantillons testés au laboratoire. Ainsi, dans le cas d'un transfert technologique de notre méthode en clinique, les peptides de ces protéines pourrait être facilement intégré pour un ciblage par spectrométrie de masse.

Avec cet algorithme, le temps de rétention théorique pour différentes conditions opératoires de chromatographie les plus usuelles, peut être également calculé via le lien internet suivant : http://hs2.proteome.ca/SSRCalc/SSRCalcX.html.

II.1.2.4 ***Ajout des peptides des banques de données***

Nous avons choisi de considérer 10 peptides par protéine cible. Quand moins de 10 peptides par protéine ont été identifiés dans les expériences *shotgun*, la liste est complétée par des peptides vus expérimentalement dans d'autres laboratoires de la communauté protéomique, disponibles sur des plates-formes informatiques. Lorsque les protéines présentent des isoformes, un nombre plus élevé de peptides est alors autorisé. Dans ce projet, seul Peptide Atlas [145,166] est utilisé pour rajouter des peptides à la liste d'inclusion des précurseurs pour une méthode ciblée en SRM (disponible sous http://www.peptideatlas.org/). Mais d'autres plates-formes existent comme PRIDE [146,171] ou GPM Proteomics [172].

Ces peptides sélectionnés selon leur score et le nombre d'observations sont inclus manuellement dans Pinpoint. Grâce à la fonction calculée précédemment, une prédiction du temps de rétention est réalisée pour chaque peptide selon leur facteur d'hydrophobicité calculé par l'algorithme SSRCalc.

II.1.3 ***Choix des transitions***

Un nombre limité de fragments des peptides peut être choisis pour la quantification en *tSRM*. **Généralement, ce sont les fragments de type y, ions les plus stables dans un triple-quadripôle, qui sont pré-sélectionnés. Les ions y les plus intenses, d'une gamme de masse de 150 à 1500 *m/z*, obtenus dans les expériences *shotgun* sont choisis pour la quantification**. Pour que la quantification des protéines se déroule avec des transitions spécifiques, un minimum de deux transitions par peptide est conseillé [15,144]. Nous avons choisi de suivre trois transitions par peptide, pour obtenir une quantification plus robuste.

Lorsque les peptides n'ont pas été identifiés dans la librairie de spectres des expériences *shotgun*, des listes de transitons sont mises à dispositions dans les plates-formes informatiques tel que SRM Atlas [173] (disponible sur http://www.srmatlas.org/). Dans le cas ultime où les transitions ne sont pas disponibles dans ces plates-formes partagées, ce sont les transitions les plus intenses calculées dans Pinpoint qui sont choisies. La liste des transitions est fournie dans le tableau en Annexe 6.

Une fois toutes les transitions sélectionnées, il faut définir la durée de suivi de chaque transition (*dwell time*). Il doit être suffisamment long pour être sensible sans toutefois être une limitation pour le nombre de transitions à suivre (cf. chap.3 §V.2.2.1 page 86). Arbitrairement,

le temps donné pour suivre chaque transition sélectionnée est de 30 ms. Ce temps est un bon compromis entre la sensibilité et la rapidité de balayage des différentes transitions pour un même temps de rétention.

II.1.4 ***Vérification du nombre de transitions sélectionnées***

Pour décrire au mieux le pic chromatographique et ainsi avoir une quantification représentative, il faut au minimum 8 à 10 points par pic chromatographique d'une largeur d'environ 20s à 10% de la vallée. Pour que le cycle soit inférieur à 2s, seules 66 transitions pourront être suivies avec un *dwell time* de 30ms dans une même fenêtre de temps d'élution chromatographique (Equation 17).

$$\#peptides = \frac{cycle\ time}{dwell\ time} \qquad \text{(Equation 17)}$$

Sur Pinpoint, un module permet de voir la répartition des transitions en fonction du temps d'élution chromatographique, montrée sur la Figure 55. La plupart des peptides sélectionnés pour la *tSRM* sont élués entre 5 et 60 min avec une partie dense entre 15 et 40 min. Un agrandissement sur la partie la plus dense permet de voir qu'au maximum 81 transitions sont suivies simultanément. Ce nombre est alors supérieur aux conditions définies au paragraphe précédent. Il faudrait dans ce cas diminuer le nombre de transitions à suivre à 2 par peptide, ce qui ferait au maximum 54 transitions qui sont éluées simultanément. Nous avons préféré dépasser ce temps de cycle et conserver un nombre de transitions suffisant pour valider la quantification. Le temps de cycle est alors de 2,43s au maximum, ce qui reste raisonnable en considérant les largeurs de pic chromatographique expérimentales obtenues.

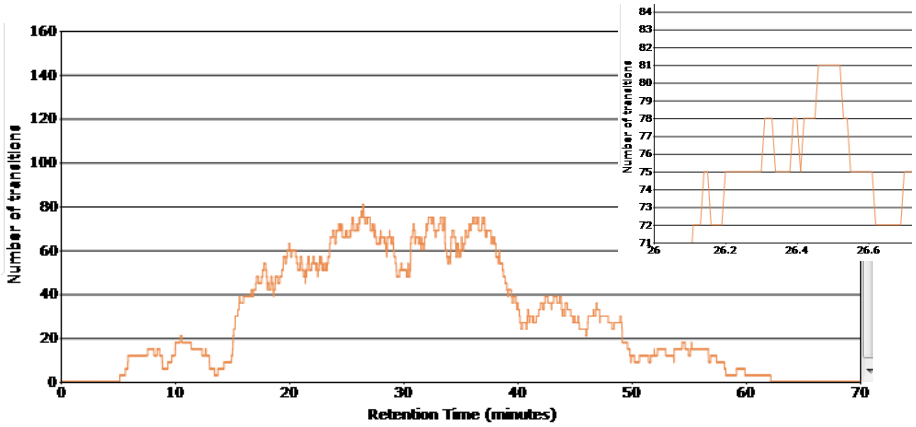


Figure 55 : Nombre de transitions en fonction du temps d'élution chromatographique. Un élargissement sur la partie dense entre 26.4 et 26.6 min est montré dans l'encart en haut à droite.

II.1.5 Choix des énergies de collision

En SRM, une optimisation de l'énergie de collision pour chaque transition est nécessaire pour obtenir une bonne sensibilité de détection. Toutefois, cette étape nécessite du temps et le recours à des peptides standards, souvent onéreux, est presque incontournable.

Le but de cette partie étant d'évaluer la possibilité de faire du diagnostic ciblé, nous n'avons pas optimisé toutes les énergies de collision. Les valeurs d'énergies de collision appliquées pour fragmenter les peptides sont calculées théoriquement par Pinpoint selon l'équation (Equation 18) ci-dessous, avec une pression d'argon de 1mTorr.

$$Collision\ Energy = 0{,}034 * m/z_{precurseur} + 3{,}314 \quad \text{(Equation 18)}$$

II.2 Traitements des données

II.2.1 Quantification des peptides

L'étape de quantification consiste à corréler l'abondance relative d'un peptide avec l'AUC du pic associé. Les chromatogrammes d'ions (XIC) sont extraits par Pinpoint. **Pour évaluer la méthode ciblée en mode *tSRM*, les AUCs des pics correspondant aux transitions listées en** Annexe 6 **sont intégrées automatiquement par Pinpoint sans correction manuelle**. Une largeur de pic de 0,25 min est renseignée. Une erreur d'alignement des analyses de 3 min est tolérée. Seuls les pics présentant un signal dont l'intensité est supérieure au bruit de fond dans les expériences *tSRM*, c'est-à-dire supérieure à 5e2, sont pris en considération. La quantification des peptides se fait alors par addition de toutes les AUCs des transitions correspondantes. Les résultats pour chaque triplicata technique sont exportés forme de tableau.

II.2.2 Quantification des protéines

La quantification des protéines est menée des deux manières suivantes :

- soit en prenant la moyenne de la somme des AUCs de tous les peptides détectés des trois réplicas (Figure 56 A),
- soit en prenant en compte la moyenne de la somme des AUCs des trois peptides les plus intenses des trois réplicas, similaire à la méthode T3PQ utilisée dans les expériences *shotgun* (Figure 56B).

Seules les protéines vues dans les trois réplicas sont prises en compte. Lorsque la protéine n'a pas été détectée de façon significative, la plus petite valeur de quantification des peptides lui est attribuée.

II.3 Interprétation des données

Comme pour les expériences *shotgun, l*e résultat de la quantification relative de toutes les protéines amyloïdogènes est représenté sous forme de camembert, sur la Figure 56. Nous pouvons observer que la variation de la quantification est faible malgré le changement du mode de calcul, utilisant trois peptides au lieu de tous les peptides. La méthode T3PQ rend donc bien compte de la quantité de protéine dans l'échantillon, ce qui conforte *a posteriori* nos choix initiaux en mode DDA (Chapitre 4). Par la suite en revanche, pour simplifier les retraitements des données, la quantification se fera en prenant en compte la somme de tous les peptides, car elle est directement donnée par le logiciel Pinpoint.

A) B)

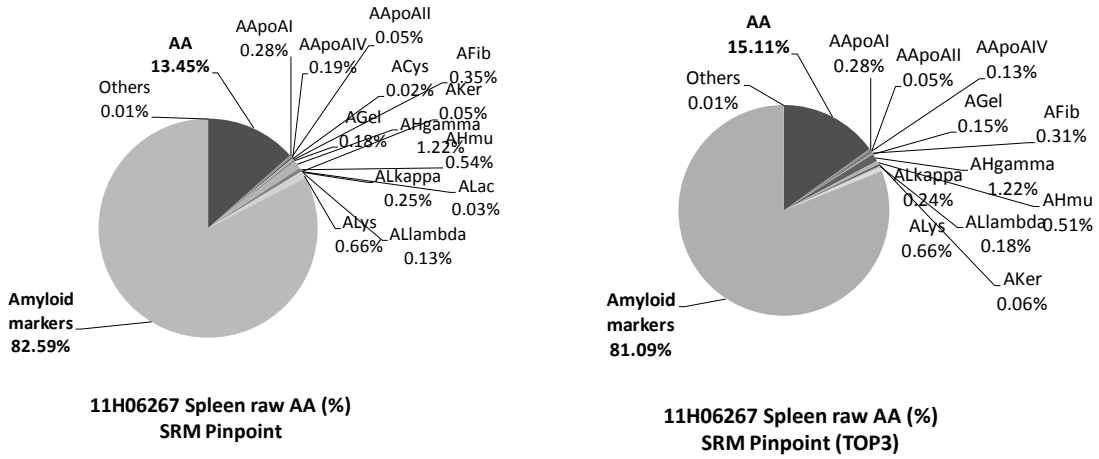


Figure 56 : Biopsie de rate atteint d'une amylose de type AA analysée en *time-scheduled* SRM. A) Quantification avec tous les peptides. B) Quantification avec les trois peptides les plus intenses. La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

Globalement, la méthode *tSRM* est plus sensible puisque nous détectons plus de protéines qu'en protéomique *shotgun* où seules les protéines SAMP, ApoE, SAA et LysC ont été détectées de manière significative (cf. chap.1 figure 7A p. 15).

Pour le diagnostic en mode *tSRM*, quelque soit la méthode de traitement de données utilisée, nous quantifions bien les deux marqueurs fibrillaires, SAMP et ApoE (sous la désignation « *Amyloid markers* »). Le patient présente donc manifestement une amylose. Pour le

typage, la protéine majoritaire qui apparait ensuite dans la liste est le sérum amyloïde A. **L'échantillon de rate 11H06267 présente une amylose de type AA**.

III. Analyses ciblées en mode tSRM et tPRM, analyse shotgun

Le spectromètre de masse hybride quadripôle-Orbitrap (Q Exactive) à haute résolution, disponible au laboratoire, est capable d'effectuer des analyses ciblées similaires au SRM : c'est le mode PRM (*parallel reaction monitoring*). Les résultats encourageants obtenus en SRM nous ont poussés à reproduire la même expérience sur le Q Exactive.

III.1 Méthode tPRM

La quantification en mode PRM, repose sur la mesure des AUCs de certains fragments spécifiques. Ils sont issus de la fragmentation des peptides sélectionnés dans le quadripôle, et sont mesurés à HR/AM en mode MS/MS grâce au troisième analyseur, l'Orbitrap. Ces caractéristiques de ce troisième analyseur permettent alors de minimiser les erreurs d'intégration dues aux autres ions peptidiques co-isolés quasi-isobares et augmente ainsi la spécificité de la méthode pour une quantification plus précise [174]. Comme en SRM, pour augmenter le nombre de peptides cibles, la dimension temporelle est rajoutée aux paramètres donnant lieu à une expérience appelée *time-scheduled PRM* (*tPRM*).

III.1.1 Vérification du nombre de peptides sélectionnés

Comme pour la méthode *tSRM*, en *tPRM*, il faut s'assurer que le temps nécessaire pour effectuer la détection des transitions pour les peptides d'intérêt est compatible avec la largeur d'un pic chromatographique pour que la quantification soit représentative. Pour qu'un maximum de peptides puisse être suivi, les événements (acquisition du transient et accumulation des ions dans la C-Trap) sont réalisés en parallèle. Toutes les transitions sont enregistrées simultanément pour un peptide donné. C'est pourquoi, il faut que le temps d'injection dans la C-Trap soit inférieur au temps d'enregistrement du transient d'ions. Le temps nécessaire pour enregistrer le transient d'ions fragments est fixé à 64 ms pour une résolution de 17500 (à m/z 200). Pour un temps de cycle inférieur à 2s, sans faire de multiplexage, on en déduit que seuls 31 peptides pourront être suivis dans une même fenêtre de temps d'élution chromatographique (Equation19).

$$\#peptides = \frac{cycle\ time \times degrès\ multiplexage}{temps\ transient} \qquad \text{(Equation19)}$$

Les peptides déjà retenus pour le mode tSRM, donnés en Annexe 6, sont sélectionnés pour la quantification en *tPRM*. La répartition des peptides est illustrée sur la Figure 55 . Il faut noter que la valeur reportée sur le graphe concerne ici le nombre de transitions et non le nombre de peptides. Le zoom montre qu'un maximum de 81 transitions est suivi simultanément. Or, en *tPRM*, tous les fragments sont détectés simultanément. Il n'y a donc pas de limitation au niveau du nombre de transitions. La seule limitation concerne le nombre de peptides sélectionné. En suivant un raisonnement inverse, le nombre maximum de peptides suivis pour un temps de rétention donné est donc de 27 peptides (=81/3 transitions par peptide), ce qui respecte les conditions citées ci-dessus.

III.1.2 ***Paramètres d'acquisition***

L'échantillon rate (11H06267) est à nouveau analysé en triplicata sur le Q Exactive, selon une méthode constituée de deux événements:

- l'acquisition Full Scan MS à une résolution de 35 000 (à m/z 200), un AGC de 3e6 et un temps maximum d'injection de 120ms.
- l'acquisition *tPRM* qui cible 220 peptides dans une fenêtre d'élution de +/- 1,5min. Les peptides sont d'abord sélectionnés par le quadripôle avec une fenêtre de masse de +/- 2 m/z puis accumulés dans la C-Trap (AGC de 1e5 et un temps maximum d'injection de 60 ms) sans multiplexage. Ils subissent ensuite une fragmentation dans la cellule de collision HCD avec une énergie normalisée de 30. Enfin, les fragments sont analysés dans l'Orbitrap à une résolution de 17 500 (à m/z 200).

III.1.3 ***Traitement des données***

Les XICs sont extraits avec les paramètres rassemblés au Tableau 13. Ici, les ions sont mesurés à HR/AM. La tolérance en masse est donc de 10 ppm pour les précurseurs (MS) et 20 ppm pour les fragments (MS/MS). De même, la quantification des peptides se fait par addition de toutes les AUCs des transitions correspondantes. La quantification des protéines est ensuite réalisée en prenant la moyenne de la somme des AUCs de tous les peptides détectés des trois réplicas. Seules les protéines vues dans les trois réplicas sont prises en compte. Lorsque la protéine n'est pas détectée de façon significative, la plus petite valeur calculée à partir des minima de quantification des peptides lui est attribuée.

Tableau 13 : Paramètres d'extractions des données pour la méthode tSRM et tSRM par Pinpoint. La dernière colonne indique si les conditions sont strictement identiques.

	tSRM	tPRM	Identique?
Logiciel	Pinpoint (v1.4)	Pinpoint (v1.4)	Oui
Fenêtre d'élution	+/- 1.5 min	+/- 1.5 min	Oui
Largeur des pics à la base	0.25 min	0.25 min	Oui
Erreur alignement	3 min	3 min	Oui
Signal seuil	5.10^{2}	10^{4}	Non
Tolérance MS	*	10 ppm	Non
Tolérance MS/MS	*	20 mmu	Non
Fenêtre d'isolation	*	+/- 2 *m/z*	Non

* : extrait automatiquement sans paramètrage à mettre

III.1.4 ***Résultats***

Globalement, pour cet échantillon de biopsie de rate, les mêmes protéines sont détectées avec des pourcentages proches en *tPRM* et *tSRM*. Les protéines SAMP et ApoE représentent plus de 73% des protéines amyloïdogènes. Ceci est cohérent avec le fait que le patient exprime une amylose. D'autre part, en dehors de ces protéines, la protéine majoritaire est le sérum amyloïde A avec plus de 22%. La quantification en *tPRM* permet donc bien de poser le même diagnostic pour l'échantillon de **rate 11H06267 : le patient présente une amylose de type AA**.

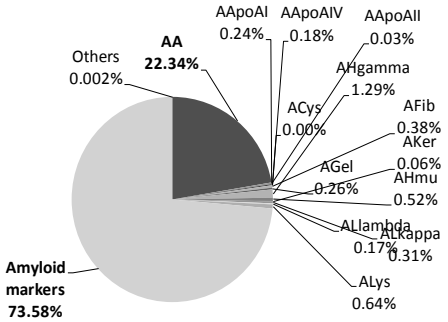


Figure 57 : Biopsie de rate atteint d'une amylose de type AA analysée en *time-scheduled* PRM. La quantification est réalisée avec tous les peptides sélectionnés pour la protéine. . La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

III.2 Comparaison sous forme de cartographie heatmap

III.2.1 ***Traitement des données***

III.2.1.1 ***Quantification des peptides***

Afin de proposer une visualisation de la comparaison globale des résultats de la quantification ciblée en mode *tSRM* et *tPRM,* ainsi que la quantification globale (*shotgun*), l'intégralité des abondances des peptides, listés en Annexe 6, est rassemblée sous forme d'une cartographie *heatmap* présentée en Figure 58.

Le traitement est identique à celui appliqué précédemment pour les expériences *tSRM* et *tPRM*.

Pour les expériences *shotgun*, deux traitements de données différents sont réalisés :

- les chromatogrammes d'ions (XIC) ont été extraits par Pinpoint sans correction manuelle sur une fenêtre de temps d'élution de +/- 1,5 min avec une tolérance en masse de 10 ppm pour les précurseurs (MS) et une fenêtre d'isolation de +/- 2 m/z. Les réglages incluent une largeur de pic de 0,25 min ainsi qu'une erreur d'alignement des analyses de 3 min afin de se rapprocher au cas du *tPRM*. Seuls les pics présentant un signal dont l'intensité est supérieure au bruit de fond dans les expériences *shotgun*, c'est à dire supérieur à 1e4, sont pris en considération.
- l'abondance des peptides a également été quantifiée par le logiciel Proteome Discoverer avec une tolérance d'erreur de 2 ppm. Seuls les peptides appartenant aux protéines amyloïdogènes du Tableau 10, identifiées avec un FDR < 1% et un classement de rang 1 dans l'identification, sont considérés.

III.2.1.2 ***Quantification des protéines***

Dans toutes les expériences, la quantification des peptides se fait par addition de toutes les AUCs des transitions correspondantes. Les résultats pour chaque triplicata technique sont exportés forme de tableau.

Les AUCs des peptides présentant un ou plusieurs acide(s) aminé(s) modifié(s) sont regroupées avec leur homologue non modifié. Tous les AUCs des peptides sont ensuite normalisées par le total de l'AUCs correspondant des peptides de l'expérience. L'aire reportée dans la cartographie *heatmap* correspond à la moyenne des AUCs normalisées des trois réplicas, si le peptide a été vu dans ces trois expériences. Lorsque le peptide amyloïdogène

n'est pas détecté de façon significative dans un échantillon donné, la plus faible valeur lui est alors attribuée. Elle correspond au rapport entre la plus petite aire intégrée et l'aire maximale du TIC des peptides de toutes les expériences confondues.

III.2.2 ***Résultats***

Les résultats sont compilés dans la cartographie *heatmap* générée par le programme *sparklines* [158] et présentée en Figure 58. Les échantillons sont disposés en colonne et les peptides spécifiques, regroupés par protéine amyloïdogène d'intérêt, en ligne. Plus une cellule est verte claire, plus la protéine correspondante est abondante dans l'échantillon.

Dans le cadre des expériences *shotgun*, l'extraction des données montre un plus grand nombre de peptides par le traitement sous Pinpoint (colonne 3 « Area TOP10 PP ») que par Proteome Discoverer (colonne 4 « Area TOP10 PD »). En effet pour ce dernier, la quantification des peptides est possible si et seulement si une identification en MS/MS a eu lieu. Dans le cas de Pinpoint, les données sont extraites à partir d'une masse et de temps de rétention. Un peptide peut alors être quantifié même si aucun spectre MS/MS n'a été effectué ou même lorsque la qualité de fragmentation n'est pas optimale. Si le spectromètre mesure des masses avec une bonne justesse, la masse du peptide et le temps de rétention peuvent alors être suffisants pour le valider. Cet exemple montre la limitation des expériences *shotgun* DDA où l'identification d'un peptide peut être manquée. En effet, en acquisition DDA, la sélection des précurseurs s'établit en fonction de l'intensité, c'est pourquoi certains ions pourtant présents ne seront pas systématiquement fragmentés, selon l'abondance relative des espèces co-éluées. Cette approche stochastique implique qu'il est fréquent que les peptides sélectionnés pour la fragmentation soient différents d'un réplica à un autre.

Dans les expériences ciblées (colonnes 1 et 2 nommées « Area SRM » et « Area PRM »), le nombre de peptides identifiés est inférieur à celui obtenu au cours des expériences *shotgun* (colonne 3), dont les données sont toutes extraites à partir des masses calculées par Pinpoint. La différence s'explique dans la spécificité de la méthode ciblée où l'extraction des données s'opère par rapport aux précurseurs et aux fragments. Or plus la sélection est s'effectuée à des niveaux multiples, plus la spécificité de l'analyte est élevée et moins il n y'a de chance de contamination par un autre peptide ou une autre molécule capable d'être retenue sur une colonne C_{18}. Certains peptides quantifiés dans la colonne 3 peuvent donc être des interférents. Par exemple, dans le cas des peptides attribués à la protéine transthyrétine (TTHY) (en bas à gauche sur Figure 58) où seul ce traitement souligne sa présence alors qu'ils ne sont pas détectés lors des analyses ciblées, plus spécifiques. La HR/AM en *tPRM* apporte également

un autre niveau de sélection. Ici, un nombre de protéines quasi équivalent est quantifié en *tSRM* et *tPRM*. Pour savoir si les peptides vus en *tSRM* et non en *tPRM* sont des interférents ou estimer si la méthode *tSRM* sur un appareil de type triple-quadripôle est plus sensible, une quantité connue de peptides standards marqués pourrait par exemple être incorporée à l'échantillon. En effet ces standards marqués par des isotopes stables homologues aux peptides étudiés permettent d'établir des gammes de dilution. D'après l'équipe de Domon [174], la sensibilité des méthodes *tPRM* et *tSRM* sont proches, avec dans la plupart du temps une minimisation des interférences dans le cas de la *tPRM*. Nous n'avons pas pu tester ce type d'approche dans le cadre de cette étude.

Dans tous les cas, ces quatre expériences permettent de poser un diagnostic. Malgré la quantité importante de peptides détectés, trois protéines se distinguent de l'ensemble des candidats : les deux protéines marqueurs des amyloses, SAMP et ApoE et la protéine sérum amyloïde A (isoformes SAA1 et SAA2). Le diagnostic est donc fort heureusement ici le même dans tous les modes d'analyse : **le patient a une amylose de type AA.**

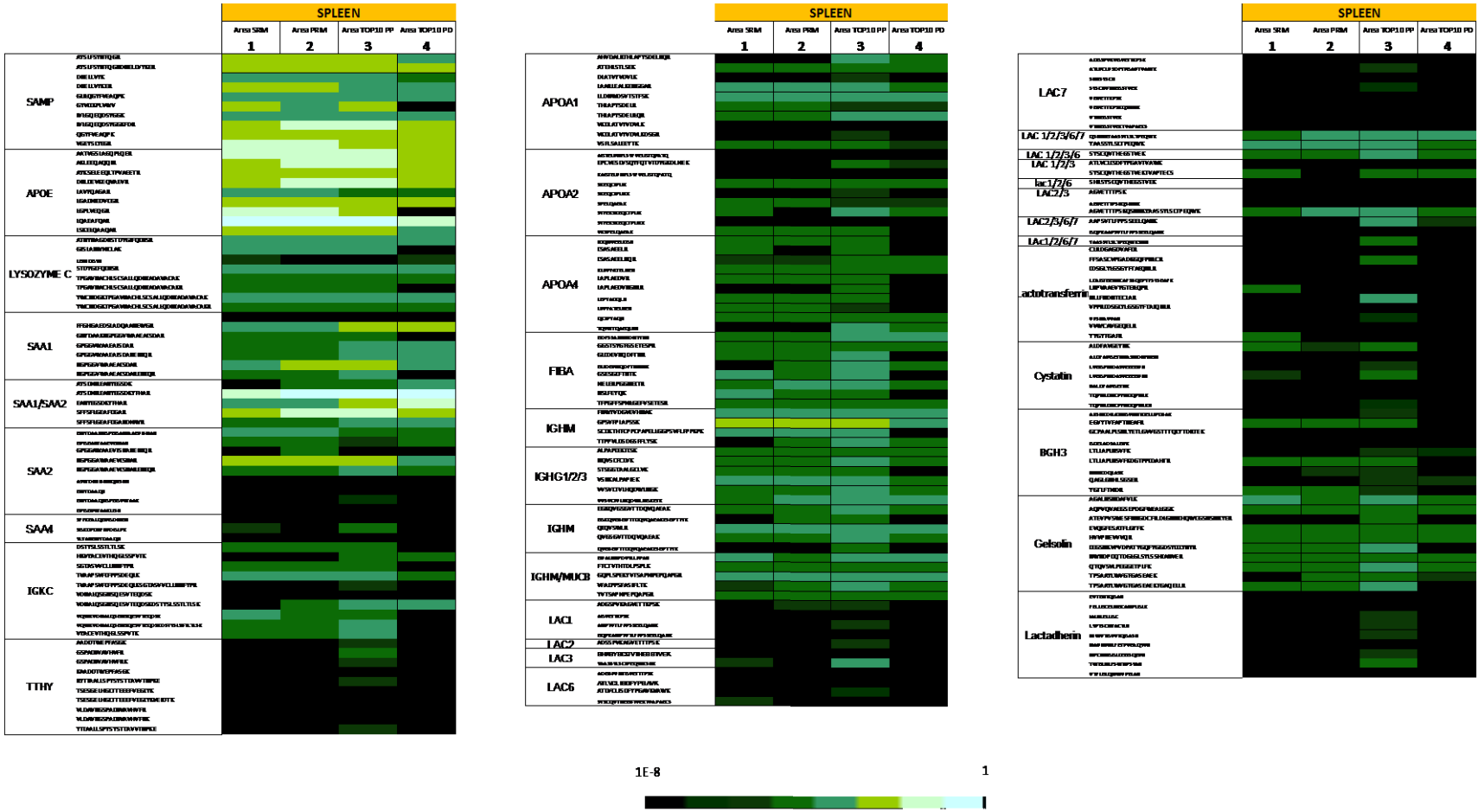


Figure 58 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans un échantillon de biopsie de rate. Plusieurs expériences sont comparées : les expériences protéomiques ciblées en tSRM (colonne 1) et *tPRM* (colonne 2), tous les deux traitées avec Pinpoint.; et les expériences shotgun traité avec Pinpoint (colonne 3) et Proteome Discoverer (colonne 4). La couleur des cellules reflète l'intensité des signaux. La cartographie a été séparée en trois par souci de clarté, pour une meilleure visualisation.

IV. Optimisation de la méthode tSRM

Jusqu'à présent la méthode *tSRM* utilisée ne présente aucune optimisation au niveau du choix des transitions et de l'énergie de collision, à l'exception du calcul théorique proposé par Pinpoint. Il s'agit dans cette partie de montrer l'importance de ces optimisations pour établir une méthode fiable, sensible et robuste pour une implémentation d'un diagnostic des amyloses par analyse *tSRM* dans les départements cliniques français.

N'ayant pas de protéine/peptide standard à notre disposition, nous avons choisi, pour optimiser la méthode, de travailler sur une protéine abondante, le composant P (SAMP), et sur l'échantillon de rate numéro 11H06267, dont nous disposions en plus large quantité.

IV.1 Optimisation des transitions

Pour des raisons techniques, les transitions précédentes ont été choisies arbitrairement parmi les données obtenues sur le Q Exactive, et non sur le TSQ Vantage, sur lequel elles sont transférées par la suite. En effet la sensibilité de l'analyse en mode sans *a priori* est infiniment supérieure. Ceci implique qu'il y a un risque que les transitions obtenues ne soient pas les plus adaptées pour l'appareil triple-quadripolaire. Une optimisation du choix des transitions est donc effectuée puis comparée aux résultats précédents.

IV.1.1 Méthode

Pour chaque peptide de la protéine SAMP listé dans le Tableau 12, tous les ions fragments y de y_3 à $y_{m/z\ <1500}$ sont sélectionnés. L'énergie de collision appliquée est calculée théoriquement par Pinpoint selon l'équation précédemment établie au paragraphe II.1.5 (Equation 18). Chaque transition est suivie pendant 30 ms (*dwell time*). Avec un maximum de 78 transitions atteint vers 27 min (Figure 59), le temps de cycle maximum reste convenable : 2,34s.

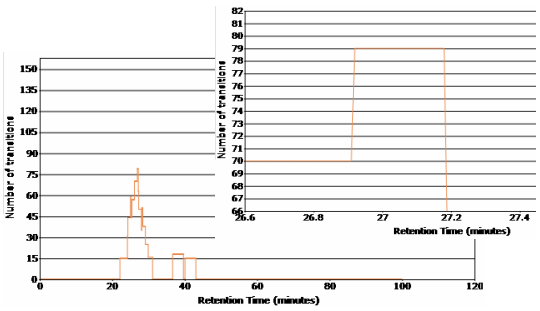


Figure 59 : Nombre de transitions en fonction du temps d'élution chromatographique. Un zoom sur la partie dense vers 27 min est montré dans l'encart en haut à droite.

IV.1.2 ***Résultats***

Les AUCs des transitions sont intégrées automatiquement par Pinpoint. Les données mesurées en triplicatas sont reproductibles (CV<10%). Un exemple est donné à la Figure 60.

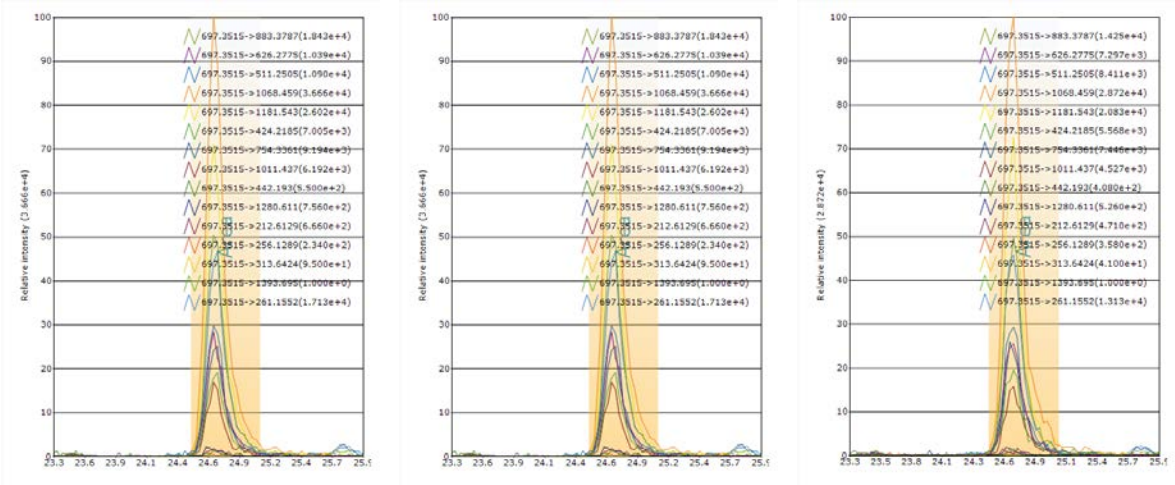


Figure 60 : Optimisation des transitions. Exemple d'un peptide : IVLGQEQDSYGGK analysé en triplicata.

La moyenne de la somme des AUCs des trois transitions les plus intenses des réplicas sont reportés sur la Figure 61. Globalement, les transitions les plus intenses sont les mêmes sur le QqOrbitrap et le TSQ Vantage (histogrammes bleu et rouge respectivement, Figure 61). **Cependant, un gain non négligeable peut avoir lieu en optimisant le choix de certaines transitions**. Par exemple, dans le cas du peptide IVLGQEQDSYGGK, l'évaluation de l'aire passe de 4,09e5 à 8,8e5, soit une augmentation de 50% (Tableau 14 et Tableau 15).

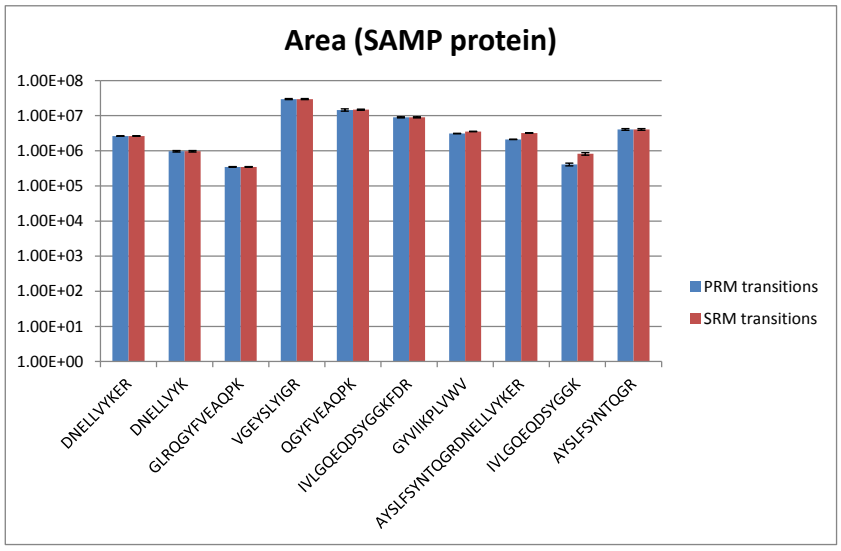


Figure 61 : Comparaison des aires de pics pour les 10 peptides de la protéine SAMP sélectionnés pour une expérience tSRM. Les histogrammes bleus (PRM Transitions) correspondent à la moyenne de la somme des trois transitions choisies avec les expériences *shotgun* sur le Q Exactive. Les histogrammes rouges (SRM transitions) correspondent à la moyenne de la sommes des trois transitions les plus intenses sur le TSQ Vantage.

Tableau 14 : Listes des transitions avant optimisation sur le triple-quadripôle. Les transitions sont issues des données *shotgun* obtenues sur le Q Exactive.

PRM transitions						
Peptide	Average area	Precursor ion (m/z)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)
AYSLFSYNTQGR	4.06E+06	703.8	y6,y7,y8	738.4	825.4	972.5
AYSLFSYNTQGRDNELLVYKER	**2.11E+06**	**667.3**	**y5,y6,y7**	**694.4**	**807.5**	**920.6**
DNELLVYK	9.70E+05	497.3	y3,y4,y5	409.2	522.3	635.4
DNELLVYKER	**2.62E+06**	**426.9**	**y3,y4,y5**	**432.3**	**595.3**	**694.4**
GLRQGYFVEAQPK	3.48E+05	498.3	y4,y5,y6	443.3	572.3	671.4
GYVIIKPLVWV	**3.09E+06**	**643.9**	**y5,y6,y7**	**613.4**	**741.5**	**854.5**
IVLGQEQDSYGGK	**4.09E+05**	**697.4**	**y5,y6,y8**	**511.3**	**626.3**	**883.4**
IVLGQEQDSYGGKFDR	8.97E+06	604.6	y6,y8,y9	679.4	929.4	1044.5
QGYFVEAQPK	**1.45E+07**	**583.8**	**y4,y5,y6**	**443.3**	**572.3**	**671.4**
VGEYSLYIGR	2.97E+07	578.8	y4,y6,y7	508.3	708.4	871.5

Tableau 15 : Listes des transitions les plus intenses après optimisation sur le triple-quadripôle. Les transitions qui diffèrent de celles établies dans le tableau précédent sont surlignées en gras.

SRM transitions						
Peptide	Average area	Precursor ion (m/z)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)
AYSLFSYNTQGR	4.06E+06	703.8	y6,y7,y8	738.4	825.4	972.5
AYSLFSYNTQGRDNELLVYKER	**3.20E+06**	**667.3**	**y4,y5,y17**	**595.3**	**694.4**	**1043.0**
DNELLVYK	9.70E+05	497.3	y3,y4,y5	409.2	522.3	635.4
DNELLVYKER	**2.64E+06**	**426.9**	**y4,y5,y6**	**595.3**	**694.4**	**404.2**
GLRQGYFVEAQPK	3.48E+05	498.3	y4,y5,y6	443.3	572.3	671.4
GYVIIKPLVWV	**3.55E+06**	**643.9**	**y6,y7,y8**	**741.5**	**854.5**	**967.6**
IVLGQEQDSYGGK	**8.18E+05**	**697.4**	**y8,y10,y11**	**883.4**	**1068.5**	**1181.5**
IVLGQEQDSYGGKFDR	8.97E+06	604.6	y6,y8,y9	679.4	929.4	1044.5
QGYFVEAQPK	**1.48E+07**	**583.8**	**y5,y6,7**	**671.4**	**572.3**	**818.4**
VGEYSLYIGR	2.97E+07	578.8	y4,y6,y7	508.3	708.4	871.5

IV.2 Optimisation des énergies de collision

Une fois les transitons les plus intenses sur le triple-quadripôle choisies, une optimisation de l'énergie de collision est nécessaire pour encore augmenter la sensibilité de la méthode *tSRM*.

IV.2.1 Méthode

Pour chaque transition listée dans le Tableau 15, sept énergies de collision par palier de 2 eV autour de la valeur calculée théoriquement par Pinpoint sont testées pour déterminer l'énergie optimale de chaque transition. Un *dwell time* de 30ms est employé pour chaque transition, soit un cycle maximum de 3,15s (105 transition visibles sur la Figure 62).

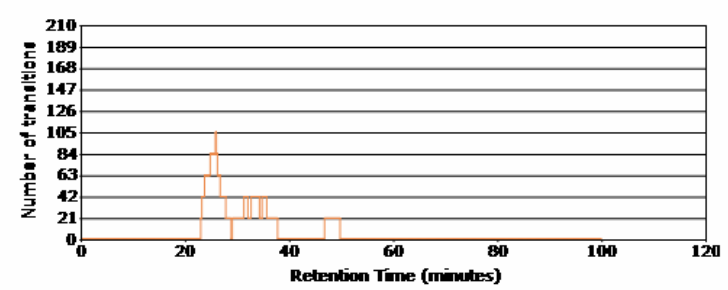


Figure 62 : Nombre de transitions en fonction du temps d'élution chromatographique.

IV.2.2 ***Résultats***

Une courbe d'efficacité de fragmentation en fonction de l'énergie de collision peut être représentée pour visualiser les énergies de collision optimales à appliquer pour suivre les transitions comme dans l'exemple du peptide IVLGQEQDSYGGK sur la Figure 63.

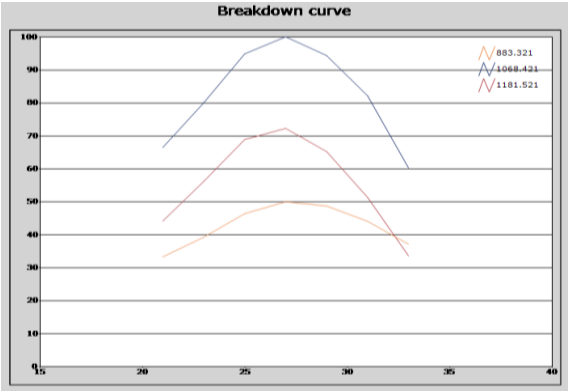


Figure 63 : Courbe d'efficacité de fragmentation des transitions sélectionnées pour le peptide IVLGQEQDSYGGK. Le sommet des courbes représente l'énergie de collision optimale pour produire les ions sélectionnés pour le suivi en tSRM.

La moyenne des AUCs des transitions mesurées dans les trois réplicas est reportée sur la Figure 64, pour chaque transition sélectionnée pour la protéine SAMP. **L'optimisation de l'énergie de collision permet d'augmenter l'aire des transitions de façon considérable**. Par exemple, l'aire de la transition 426.9 → 404.2 du peptide DNELLVYKER (premier peptide de la Figure 64) augmente de plus de 70% grâce à l'optimisation de son énergie de collision (attention l'ordonnée est affichée en log10).

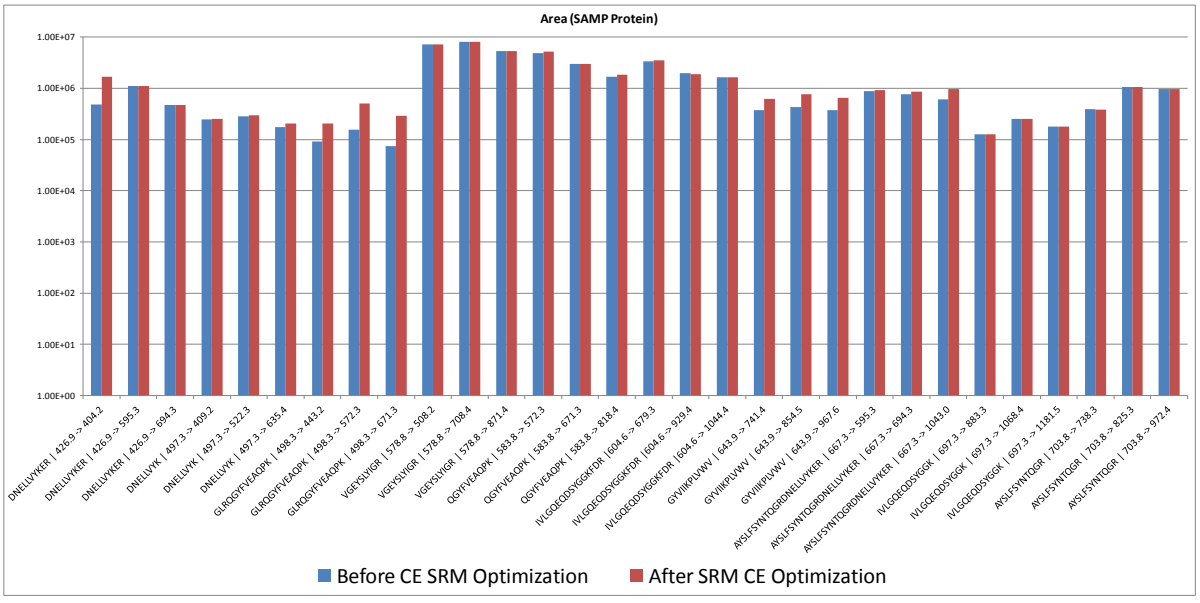


Figure 64 : Comparaison des aires de pics (en log10) avant et après optimisation des énergies de collisions pour les 3 transitions sélectionnées par peptide de la protéine SAMP. Les histogrammes bleus (Before CE SRM Optimization) correspondent à la moyenne de l'aire de la transition obtenue avec l'énergie de collision théorique calculée par Pinpoint. Les histogrammes rouges (Before CE SRM Optimization) correspondent à la moyenne de l'aire de la transition obtenue avec l'énergie de collision optimale.

Ces optimisations permettent également d'affiner la fonction permettant de calculer les énergies de collision théoriques pour les ions bichargés (Equation 20) et trichargés (Equation 21).

Ions bichargés : $Collision\ Energy = 0.04005 * m/z_{precurseur} - 2{,}40023$ (Equation 20)

Ions trichargés : $Collision\ Energy = 0{,}06438 * m/z_{precurseur} - 10{,}82893$ (Equation 21)

IV.3 Comparaison tSRM /tPRM à leur optimum

IV.3.1 Méthode tSRM

L'échantillon rate 11H06267 est à nouveau analysé en *tSRM* avec une liste de transitions comprenant les 10 peptides de la protéine SAMP, sélectionnés sur une fenêtre de +/- 1.5min, et leurs énergies de collision optimisées. Les premier et troisième quadripôles sélectionnent les masses avec une fenêtre d'isolation de 0,7 Da. **Chaque transition est suivie pendant 200 ms au lieu de 30 ms pour gagner en sensibilité**. Le temps de cycle est porté à 3s avec une valeur maximale de transitions au même moment de 15 (Figure 65). Nous avons postulé que ce temps même supérieur à 2s, ne devrait pas en théorie altérer la quantification.

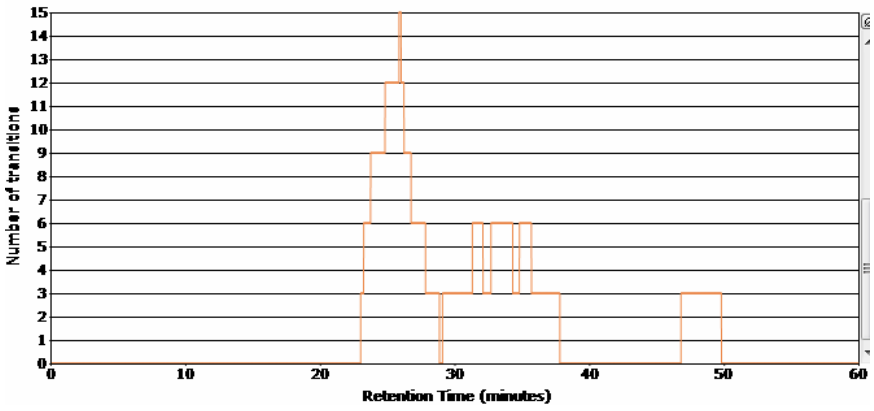


Figure 65 : Nombre de transitions en fonction du temps d'élution chromatographique.

IV.3.2 Méthode PRM

La même liste de peptides est sélectionnée pour une quantification en *tPRM* de l'échantillon rate 11H06267, avec une fenêtre de +/- 1.5 min. Les précurseurs sont isolés avec une fenêtre de 4 m/z et mesurés à 35 000 (à m/z 200) de résolution, un AGC de 3e6 et un temps d'injection de 120ms (acquisition Full MS). Les précurseurs sont sélectionnés par le quadripôle avec une fenêtre d'isolation de 4 m/z puis fragmentés avec une énergie de collision normalisée de 30. La mesure des fragments est réalisée à une résolution de 17500, un AGC de

1e6 et un temps d'injection maximum de 500ms (acquisition PRM). Le temps de cycle en *tPRM* est alors de 2,5 s au maximum pour 5 précurseurs élués simultanément.

Les résultats des acquisitions en mode tSRM et *tPRM* à leur optimum sont extraits automatiquement par Pinpoint avec les paramètres décrits au Tableau 16 ci-dessous.

Tableau 16 : Paramètres d'extractions des données pour la méthode tSRM et tSRM à leur optimum par Pinpoint. La dernière colonne indique si les conditions sont strictement identiques.

	tSRM	tPRM	Identique?
Logiciel	Pinpoint (v1.4)	Pinpoint (v1.4)	Oui
Fenêtre d'élution	+/- 1.5 min	+/- 1.5 min	Oui
Largeur des pics à la base	0.25 min	0.25 min	Oui
Erreur alignement	1 min	1 min	Oui
Signal seuil	5.10^{2}	10^{4}	Non
Tolérance MS	*	10 ppm	Non
Tolérance MS/MS	*	20 mmu	Non
Fenêtre d'isolation	*	+/- 2 *m/z*	Non

* : extrait automatiquement sans paramètrage à mettre

IV.3.3 Comparaison des résultats

Les valeurs d'AUCs obtenues, sont nettement plus élevées par la méthode *tPRM* que par la méthode *tSRM* (Figure 66). Cependant, il faut tenir compte du fait que l'enregistrement des signaux est différent entre les deux spectromètres de masse. Par exemple, l´intensité des signaux correspondant aux polycyclosiloxanes présents dans l'environnement atmosphérique et couramment utilisés comme étalons internes [131] est relevé à 1e6 sur le Q Exactive alors qu'elle est de 1e3 sur le TSQ Vantage. Pour chaque peptide, ce ne sont donc pas les aires qui doivent être comparées mais le signal sur bruit (S/B) (Figure 67). Une estimation de la valeur du S/B peut être calculée par Pinpoint. Elle correspond au carré du ratio du signal à l'apex du pic sur le bruit qui est le médian de l'intensité des pics à un intervalle de ±2 x largeurs de pic à la base autour du pic d'intérêt.

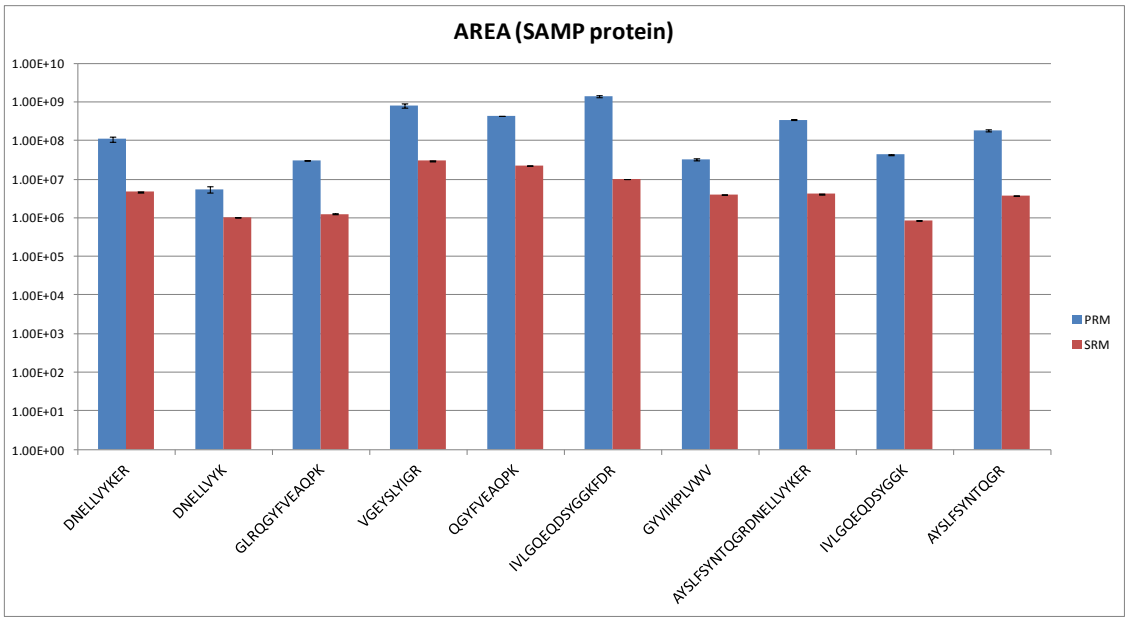


Figure 66 : Comparaison des aires de pics (en log10) des peptides sélectionnées pour la protéine SAMP obtenues sur le Q Exactive et le TSQ Vantage à leur optimum. Les histogrammes bleus correspondent à la moyenne des aires des peptides obtenus avec un temps d'accumulation dans la C-Trap de 500 ms maximum sur le Q Exactive en mode *tPRM*. Les histogrammes rouges correspondent à la moyenne des aires des peptides obtenus avec un *dwell time* de 200ms maximum sur le TSQ Vantage en mode tSRM.

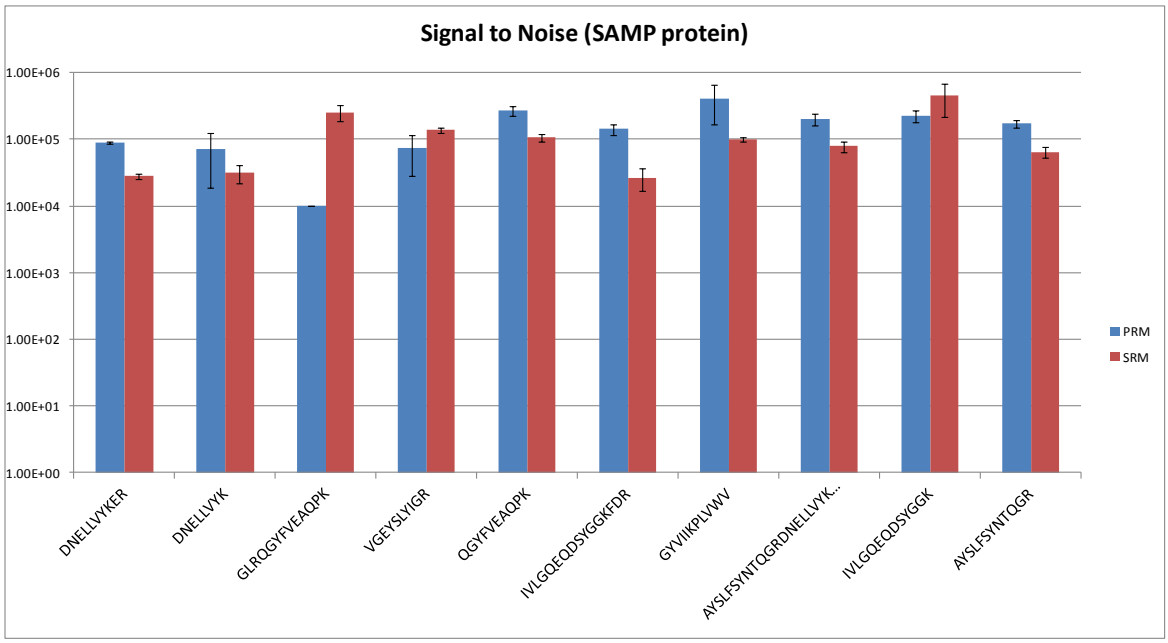


Figure 67 : Comparaison du signal sur bruit (S/B) de pics des peptides sélectionnées pour la protéine SAMP obtenues sur le Q Exactive et le TSQ Vantage à leur optimum. Les histogrammes bleus correspondent à la moyenne des S/B des peptides obtenue avec un temps d'accumulation dans la C-Trap de 500 ms maximum sur le Q Exactive en mode *tPRM*. Les histogrammes rouges correspondent à la moyenne des S/B des peptides obtenue avec un

dwell time de 200ms maximum sur le TSQ Vantage en mode tSRM.Globalement le S/B est plus élevé dans les expériences *tPRM* que *tSRM*. À l'inverse, pour certains peptides (tels que GLRQGYFVEAQP**K**, VGEYSLYIG**R** et IVLGQEQDSYGG**K**), le S/B apparait nettement plus élevé avec le triple-quadripôle. Concernant la reproductibilité des analyses, l'analyse du signal sur bruit montre que la méthode *tSRM* est plus reproductible avec un coefficient de variation d'en moyenne de 1% entre les réplicas contre 7% pour les expériences PRM.

Pour travailler sur la limite de quantification entre ces deux méthodes, il aurait été intéressant de travailler avec des peptides standards marqués par des isotopes stables, en réalisant une gamme de dilution, ce dont nous ne disposions pas pour cette étude.

Les deux méthodes tSRM et *tPRM* sont donc équivalentes dans l'ensemble. Pour certains peptides, le signal est meilleur en *tSRM* mais pour d'autres il l'est en *tPRM*. La méthode *tSRM* est bien établie dans les départements cliniques. Les résultats montrent que dorénavant la *tPRM* peut être également utilisée pour la quantification des protéines à HR/AM.

V. Conclusion

La méthode *time-scheduled SRM* est considérée comme la méthode de choix pour une quantification ciblée des biomarqueurs. Elle peut être effectuée avec un spectromètre de masse de type triple-quadripolaire qui est peu couteux, robuste et facile d'entretien. Cependant, la mise en place de la méthode est délicate. De façon conjointe au choix des peptides, il est fondamental d'optimiser leur énergie de collision pour obtenir des ions fragments les plus intenses possibles. Cette pratique pour être menée à bien nécessite l'utilisation de protéines/peptides standards généralement onéreux. Ces optimisations sont essentielles pour établir une méthode sensible. En outre, pour que la méthode soit universelle et transférable d'un laboratoire à un autre, une correction du temps de rétention est nécessaire en *tSRM*. Cette correction est dorénavant implémentée dans les nouvelles versions des logiciels pilotes de l'instrument. Elle repose sur l'ajout de peptides marqués par des isotopes stables dont le temps de rétention couvre toute la gamme d´élution chromatographique. Une correction peut alors être effectuée en direct (*on fly*) et minimiser ainsi les décalages de temps entre différentes analyses [175]. Pour plus de rigueur, il aurait fallu parler de facteur de rétention (k') qui est le vrai descripteur d'un analyte en chromatographie et non du temps de rétention (tr). Cependant, les logiciels de retraitement ne sont pas capables de gérer ce paramètre puisqu'il faut définit le t_0, temps référence d'un composé non retenu dans la colonne.

Bien que la fenêtre de sélection du quadripôle soit assez petite (0,7 da), des interférents de masse proche des analytes peuvent être co-isolés dans les quadripôles et fausser la quantification. C'est pour cela qu'un intérêt pour la quantification ciblée à HR/AM a émergé ces dernières années et s'est concrétisé notamment par l'introduction d'un spectromètre de masse hydride combinant le quadripôle avec son pouvoir filtrant sans *cut-off* de masse et l'Orbitrap pour la mesure des ions précurseurs et fragments à HR/AM (Q Exactive). Les résultats obtenus sur cet appareil, en mode *tPRM*, montrés dans les paragraphes précédents, sont similaires au *tSRM*. Il s'agit dans le chapitre suivant d'évaluer la méthode ciblée, *tPRM,* pour du diagnostic.

Chapitre 6 : *Analyse protéomique ciblée en mode PRM : application aux biopsies amyloïdes*

I. *Analyse ciblée à haute résolution : vers une application clinique ?*

Les spectromètres de masse de type quadripolaire restent aujourd'hui les instruments les plus utilisés dans les départements cliniques français pour le screening à haut débit d'analytes cibles en mode SRM (*Selected Reaction Monitoring*). Cependant, l'introduction récente d'un nouvel appareil de type quadripôle-Orbitrap, dont la prise en main est plus aisée que les appareils à haute résolution pour des applications « recherche », semble peu à peu changer le parc instrumental de ces départements. Ceci permet d'envisager que dans un avenir à plus ou moins court terme, la haute résolution sera alors disponible dans les services cliniques. Comme détaillé dans le Chapitre 5, une approche similaire au SRM, le PRM (*Parallel Reaction Monitoring*) est possible avec ce modèle de spectromètre de masse. Les résultats sont de qualité équivalente et la mise en œuvre est plus simple car elle nécessite moins d'optimisation. Dans ce cas, la mesure des ions précurseurs et fragments par dans un analyseur de type Orbitrap permet d'établir une quantification plus précise à haute résolution. Les ions précurseurs sont sélectionnés avec une erreur de justesse de 10 ppm et les ions fragments de 20 ppm au lieu d'une centaine de ppm pour le triple-quadripôle. Si l'investissement pour un tel équipement reste à l'heure actuelle plus lourd, il offre une réelle économie de fonctionnement puisqu'il est possible de s'affranchir de nombreuses étapes d'optimisation avec des réactifs de synthèse couteux. Ce mode de quantification PRM a été est évalué et comparé au SRM dans le chapitre précédent sur un exemple particulier. Dans la suite de ce chapitre, nous avons généralisé son application sur divers type de prélèvement cliniques, afin de mieux identifier ses avantages et ses limitations.

II. Stratégie analytique pour un diagnostique robuste

II.1 Choix des peptides

II.1.1 Caractéristiques de rétention chromatographiques

Pour contrôler les temps de rétention des analytes et disposer de données fiables, plusieurs mois après les premières injections, l'échantillon de biopsie de rate (numéro 11H06267) est choisi pour une nouvelle série d'analyses. Les mêmes conditions chromatographiques que celles employées pour la campagne de la SRM sont employées (voir Chap. 5, § II.1.1 page 129)

Les résultats présentés à la Figure 68 montrent que le profil chromatographique a changé. Des différences de temps de rétention sont notables entre les pics (chromatogramme noir vs rouge). Après un décalage du déclenchement de l'acquisition MS (chromatogramme vert), le profil est proche de celui d'origine (chromatogramme noir).

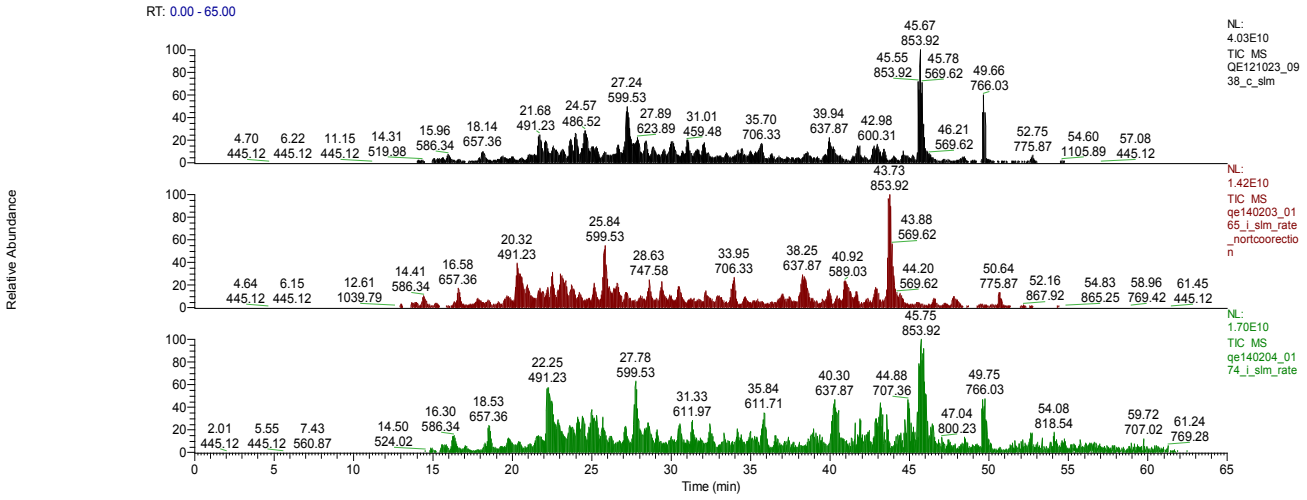


Figure 68 : Profil chromatographique d'un échantillon de biopsie de rate analysé à différents temps. En haut, en noir : Analyse de référence. Au milieu, en rouge : Analyse 15 mois plus tard. En bas, en vert : Nouvelle analyse après correction du temps de départ d'acquisition.

Cette variabilité se vérifie d'autant plus lorsque plusieurs colonnes de même longueur et nature sont comparées. Les temps de rétention peuvent alors être décalés (Figure 69) car chaque colonne possède des propriétés qui lui sont propres.

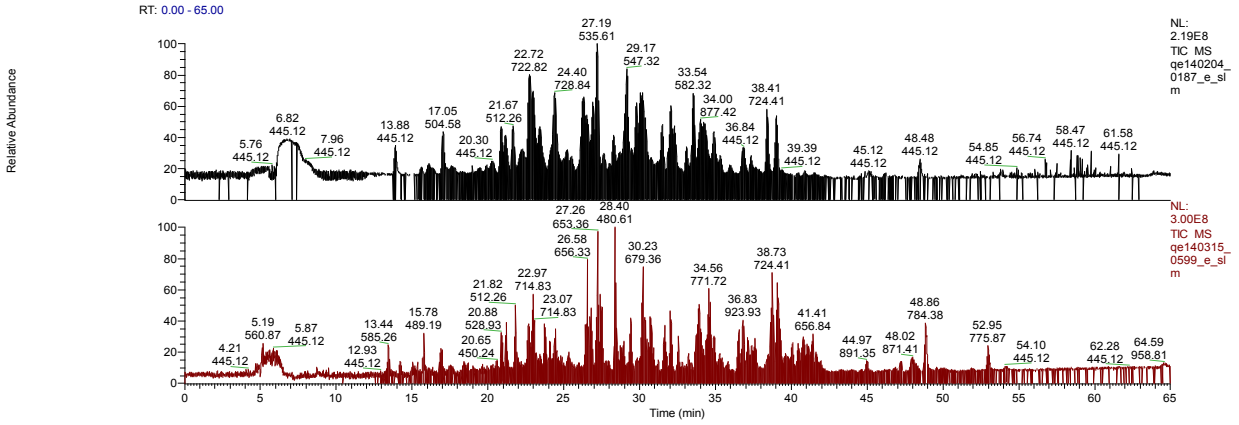


Figure 69 : Profil chromatographique d'un échantillon peptidique de protéines standards sur deux colonnes de 15 cm de même nature. Les temps de rétention varient de façon homothétique. Cette variation devra être prise en compte lors du développement de la méthode pour un criblage à haut débit où les colonnes peuvent être facilement amenées à être changées lorsqu'elle ne présente plus un pouvoir de séparation suffisante.

Il faudrait à l'avenir analyser un mélange standard au début de chaque série d'analyse pour vérifier les temps de rétention comme sur la Figure 69. Cette analyse permettra également de s'assurer du bon fonctionnement du système nanoLC-MS/MS. Nous préconisons d'analyser un standard tous les deux jours au minimum.

À défaut de pouvoir entièrement maitriser ces paramètres, une première solution consiste à ouvrir la fenêtre de temps de sélection des transitions pour ne pas « passer à côté » de certains peptides. Cette ouverture n'est pas sans conséquence puisqu'elle réduit le nombre total de peptides qui peuvent être suivis au cours de l'analyse.

Une sélection d'échantillon a donc été analysée à nouveau en DDA sur le Q Exactive pour créer une nouvelle méthode en *tPRM*. Elle ciblait les 29 protéines listées dans le Tableau 17, avec une fenêtre d'intervalle de temps plus large.

Tableau 17 : Liste des protéines amyloïdogènes sélectionnées pour la quantification en mode *time-scheduled PRM*.

Numéro Accession	Nom accession	Nom protéine
P01034	CYTC_HUMAN	Cystatine C
P01834	IGKC_HUMAN	Chaines légères Ig kappa, région constante
P01857	IGHG1_HUMAN	Ig gamma-1, région constante
P01859	IGHG2_HUMAN	Ig gamma-2, région constante
P01860	IGHG3_HUMAN	Ig gamma-3, région constante
P01861	IGHG4_HUMAN	Ig gamma-4, région constante
P01871	IGHM_HUMAN	Ig mu, région constante
P04220	MUCB_HUMAN	Protéine chaine lourde Ig mu pathologique
P02647	APOA1_HUMAN	Apolipoproteine A-I
P02649	APOE_HUMAN	Apolipoproteine E
P02652	APOA2_HUMAN	Apolipoproteine A-II
P02671	FIBA_HUMAN	Fibrinogene, chaine alpha
P02743	SAMP_HUMAN	Composant P du sérum amyloïde
P02766	TTHY_HUMAN	Transthyretine
P02788	TRFL_HUMAN	Lactotransferrine
P06727	APOA4_HUMAN	Apolipoproteine A-IV
P0CG04	LAC1_HUMAN	Ig lambda-1, région constante n
POCG05	LAC2_HUMAN	Ig lambda-2, région constante
P0CG06	LAC3_HUMAN	Ig lambda-3, région constante
P0CF74	LAC6_HUMAN	Ig lambda-6, région constante
A0M8Q6	LAC7_HUMAN	Ig lambda-7, région constante
P0DJI8	SAA1_HUMAN	Sérum amyloïde A-1
P0DJI9	SAA2_HUMAN	Sérum amyloïde A-2
P35542	SAA4_HUMAN	Sérum amyloïde A-4
P61626	LYSC_HUMAN	Lysozyme C
P61769	B2MG_HUMAN	Bêta-2-microglobuline
Q08431	MFGM_HUMAN	Lactadhérine
Q15582	BGH3_HUMAN	Protéine facteur de croissance de transformation, beta induit ig-h3
P06396	GELS_HUMAN	Gelsoline

II.1.2 ***Critères de choix***

Les fichiers de recherches DDA TOP10 générés au format msf sont importés dans le logiciel Pinpoint pour constituer la librairie de spectres et pour une comparaison des peptides identifiés. Tous les peptides identifiés dans les expériences précédentes sont aussi considérés comme peptides cibles. Les peptides les plus intenses identifiés sur le Q Exactive et le LTQ FT Ultra, correspondant aux 29 protéines répertoriées dans le Tableau 17, sont choisis en priorité. Le nombre maximum de peptides par protéines est réduit à 8 pour pouvoir ajouter une protéine amyloïdogène (bêta-2-microglobuline) et augmenter la fenêtre d'intervalle de temps de suivi des transitions. Lorsque les protéines présentent des isoformes, un nombre plus élevé de peptides est alors autorisé (jusqu'à 24 peptides).

II.1.3 Validation des peptides expérimentaux

Comme pour la SRM, une vérification de la cohérence entre l'hydrophobicité des peptides choisis et le temps de rétention dans le système chromatographique est réalisée pour valider les peptides sélectionnés.

Les rapports m/z des trois premiers isotopes des peptides sont générés par le logiciel Pinpoint. Une quantification par la méthode des AUCs est effectuée sur les données utilisées pour créer la méthode. Les AUCs sont réintégrées manuellement si besoin. Les taux de variations des AUCs entre les réplicas doivent être inférieurs à 15%. Les temps de rétention théoriques sont reportés en fonction du facteur d'hydrophobicité, calculé selon l'algorithme SSRCalc [169,170,176]. Nous obtenons un facteur de corrélation (R^2) de 0,95207, soit proche de 1 (Figure 70).

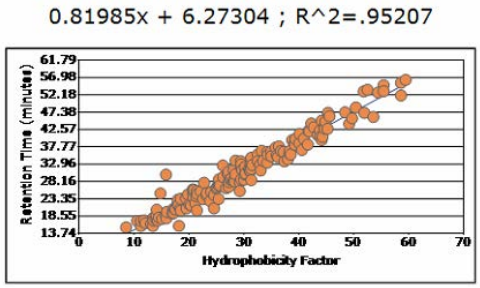


Figure 70 : Temps de rétention prédit en fonction du facteur d'hydrophobicité prédit par l'algorithme SSRCal pour les peptides choisis pour le *time-scheduled PRM*.

Cette fonction (y = 0.81985x + 6.27304) est sauvegardée pour prédire les temps de rétention des peptides qui seront rajoutés par la suite, avec un intervalle de temps d'élution de +/- 2 min.

II.1.4 Ajout des peptides des banques de données

Nous appliquons la même démarche que présentée au Chapitre 5. Quand moins de 8 peptides par protéines sont identifiés dans les expériences *shotgun*, la liste est complétée par des peptides vus expérimentalement par d'autres laboratoires, disponibles à partir de la plate-forme informatique Peptide Atlas [145,166]. Ces peptides, sélectionnés selon leur score et le nombre d'observations, sont inclus manuellement dans Pinpoint. Grâce à la fonction calculée précédemment, une prédiction du temps de rétention est déterminée pour chaque peptide selon leur facteur d'hydrophobicité calculé par l'algorithme SSRCalc.

II.2 Optimisation des paramètres d'acquisitions

II.2.1 Vérification du nombre de peptides sélectionnés

Une fois le processus de sélection des peptides achevé, il faut régler la méthode afin que la détection en Orbitrap et l'accumulation des prochains peptides puissent se dérouler en parallèle l'un de l'autre, pour suivre un maximum de peptides. Nous avons vu dans le chapitre précédent qu'une acquisition à une résolution de 17500 (@ m/z 200) dure 64 ms, c'est-à-dire que pour respecter les 2 s de temps de cycle, seuls 31 peptides peuvent être suivis simultanément sans faire de multiplexage.

Ainsi, la répartition en fonction du temps d'élution chromatographique, des peptides validés pour une quantification en *tPRM* est montrée à la Figure 71. La plupart des peptides choisis sont élués entre 15 et 50 min avec un pic vers 32 min. Un agrandissement sur ce sommet permet de voir que 93 transitions sont suivies simultanément à ce temps d'élution. Il faut cependant noter ici que les transitions affichées correspondent en réalité aux 3 premiers isotopes de chaque précurseur. Le nombre maximum de peptides suivis au même temps de rétention est donc de 31 peptides (=93/3), ce qui respecte les conditions citées au-dessus.

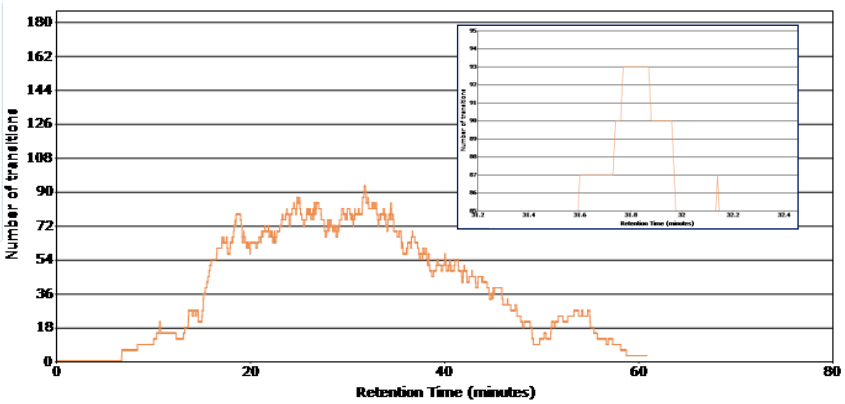


Figure 71 : Nombre de transitions en fonction du temps d'élution chromatographique. Un élargissement sur la partie dense entre 31.2 et 32.4 min est montré.

L'ensemble des échantillons est à nouveau analysé en triplicata sur le Q Exactive. Les peptides amyloïdogènes sont désignés pour le réglage des analyses en mode *tPRM*, cependant à ce stade, aucune transition parent→fils n'est nécessaire.

La méthode est constituée de deux événements:

- l'acquisition Full MS à une résolution de 35 000 (à m/z 200), un AGC de 3e6 et un temps maximum d'injection de 120ms.

- l'acquisition *tPRM* qui cible 193 peptides dans une fenêtre d'élution de +/- 2min (cf. tableau en Annexe 7). Les peptides (précurseurs) sont d'abord sélectionnés par le quadripôle avec une fenêtre de masse de +/- 2 *m/z* puis accumulés dans la C-Trap (AGC de 5e5 avec un temps maximum d'injection de 60 ms) sans multiplexage. Ils subissent ensuite une fragmentation dans la cellule de collision HCD avec une énergie normalisée de 30. Enfin, les fragments sont analysés dans l'Orbitrap à une résolution de 17 500 (à *m/z* 200).

II.2.2 ***Choix des transitions***

Tous les fragments des peptides peuvent être choisis pour la quantification en *tPRM*. Mais généralement, ce sont **les fragments les plus intenses de type y (mono, bi, tri et quadrichargés) d'une gamme de masse allant de 150 à 1500 *m/z*,** obtenus dans les expériences *shotgun*, qui sont retenus pour la quantification. Arbitrairement 5 ions les plus intenses seront suivis pour chaque peptide.

Pour compléter la librairie de spectres, les données issues des expériences PRM peuvent également être incluses. En effet, contrairement au SRM qui ne cible qu'un nombre limité de transitions, les données PRM contiennent aussi les informations nécessaires à l'identification: la masse des précurseurs et le spectre entier de fragmentation, avec une haute résolution et une grande exactitude de masse (HR/AM).

De façon analogue au Chapitre 5, lorsque les peptides ne sont pas identifiés dans la librairie de spectres issus des expériences *shotgun* et PRM, la liste des transitions est complétée par les données mises à dispositions dans les plates-formes informatiques telles que SRM Atlas [173] (disponible sous http://www.srmatlas.org/). La liste des transitions est fournie dans le tableau en Annexe 7.

III. Résultats

Deux logiciels ont été utilisés et comparés au cours de ce projet : le logiciel constructeur Pinpoint avec lequel la méthode *tPRM* a été créée et un logiciel open-source Skyline (v.2.5.0.6079, MacCoss Lab) utilisable sous licence GNU (*General Public License*) [177].

III.1 Quantification des transitions par Pinpoint

III.1.1 Traitement des données

Les XICs sont extraits avec les paramètres utilisés au Chapitre 5 (cf. § IV.3.2 page 151) avec une fenêtre d'élution de +/- 2 min. La quantification des peptides se fait ici encore par addition de toutes les AUCs des transitions correspondantes. Les résultats pour l'ensemble des triplicatas techniques sont exportés forme de tableau.

III.1.2 Data mining

III.1.2.1 Etude qualitative et cartographie heatmap

Comme pour les expériences *shotgun*, une visualisation globale regroupe l'ensemble des données de quantification relative des protéines amyloïdogènes, sous forme d'une cartographie *heatmap* pour une classification des amyloses. Cette cartographie, présentée en Figure 73, est générée par le programme *sparklines* [158]. Les échantillons sont disposés en colonne et les peptides spécifiques regroupés par protéine amyloïdogène d'intérêt, en ligne.

Les AUCs des peptides présentant un ou plusieurs acide(s) aminé(s) modifié(s) sont regroupées avec leur homologue non modifié. Tous les AUCs des peptides sont ensuite normalisées par le total de l'AUC du TIC de l'expérience correspondant. Ce traitement est rendu possible puisqu'en PRM le premier événement enregistré est une mesure des masses des précurseurs. L'aire reportée dans la cartographie *heatmap* correspond à la moyenne des AUCs normalisées des trois réplicas si le peptide est vu dans ces trois expériences. Lorsque le peptide amyloïdogène n'est pas détecté de façon significative dans un échantillon donné, la plus faible valeur normalisée, toutes expériences confondues, lui est attribuée.

Globalement, la sensibilité apparait meilleure en *tPRM* qu'en *shotgun* puisque plus de peptides spécifiques sont quantifiés. Les peptides des deux protéines marqueurs des amyloses, SAMP et ApoE, sont identifiés en abondance chez les patients présentant la pathologie des amyloses (colonnes 1 à 10). Ces deux protéines sont également quantifiées chez les contrôles (colonnes 13

à 27) mais la quantité et l'intensité avec lesquelles ils sont caractérisés sont très en deçà de celle des échantillons pathologiques. Cela est également le cas pour les deux patients LCDD, colonnes 11 et 12 où des peptides ont été identifiés en mode ciblé et non dans les analyses *shotgun*.

Le typage pour chaque échantillon analysé en triplicata par spectrométrie de masse ciblée à HR/AM peut être décrit de la façon suivante:

- Plusieurs peptides spécifiques des protéines SAA1 et SSA2 sont identifiés dans **les échantillons 1, 3, et 8**. Ces patients présentent une pathologie d'**amylose de type AA**
- Dans les **échantillons 2 et 10,** tous les peptides cibles de la transthyrétine (TTHY) sont détectés avec une forte intensité. Ces patients ont **une amylose de type ATTR**.
- L'**échantillon 4** exprime plusieurs peptides appartenant aux protéines apolipoprotéine A4 (APOA4) et fibrinogène alpha (FIBA), et dans une moindre mesure SAA1/SAA2, Ig kappa (IGKC) et Ig lambda (LACs). **Le diagnostic pour ce patient est ambigu.**
- Très peu de peptides amyloïdogènes sont identifiés dans l'**échantillon 5**. **Aucune pathologie ne peut être diagnostiquée ici.**
- **L'échantillon 6** présente plusieurs peptides appartenant à différentes protéines: chaines légères d'Ig kappa (IGKC), transthyrétine (TTHY), apolipoprotéine A4 (APOA4), fibrinogène alpha (FIBA) et lactotransferrine (TRFL). **Le diagnostic pour ce patient est ambigu.**
- L'**échantillon 7** contient très peu de protéines amyloïdogènes. **Aucune pathologie ne peut être diagnostiquée ici.** Le prélèvement dont nous disposions pour ce patient était en quantité limitée. De plus sur les spectres de masse obtenus en mode DDA TOP 7 sur le LTQ FT Ultra, nous pouvons noter une forte proportion de polymères (Figure 72) préjudiciables à l'analyse.

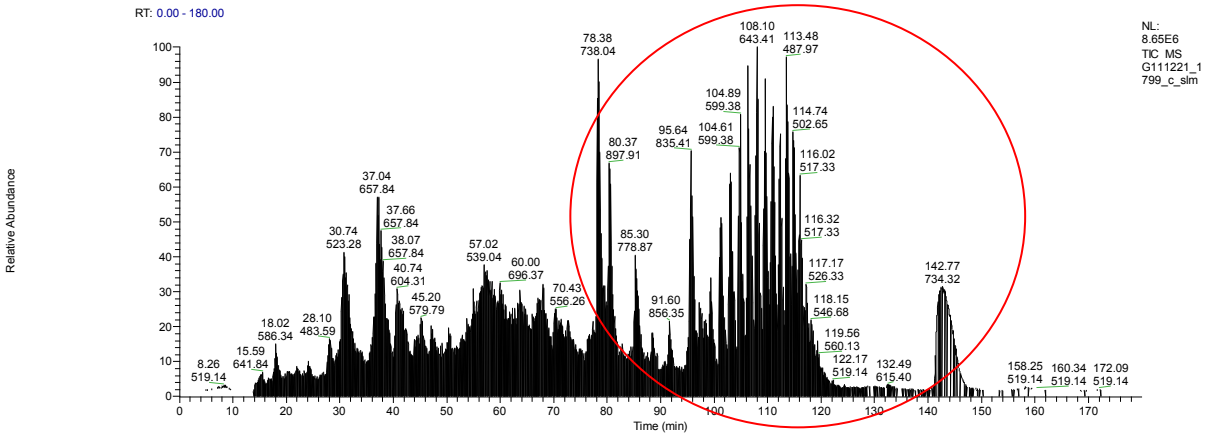


Figure 72 : Spectre en acquisition DDA (TOP7) sur le LTQ FT Ultra de l'échantillon bronche 10H12063. L'échantillon, peu intense, est par ailleurs pollué par des polymères.

- Dans l'**échantillon 9,** les peptides les plus intenses détectés appartiennent à la lactotransferrine (TRFL), constitutive du tissus [159–161]. Les peptides dont l'intensité quoique légèrement inférieure, mais tout aussi abondants, sont attribués à l'apolipoprotéine A1 (APOA1). Ce patient présente alors **une amylose de type ApoA1**.
- Enfin, deux échantillons patients se distinguent dans cette cartographie: les **échantillons 11 et 12**, se situent à l'interface entre les **contrôles** et les échantillons patients. En effet, dans ces tissus, les protéines SAMP et d'ApoE ont une abondance proche des échantillons contrôles. Néanmoins les peptides de l'Ig kappa (IGKC) sont majoritaires. Ces deux patients présentent une maladie de **dépôt de chaines légères non amyloïdes** (LCDD) [42,43].

Les échantillons présentant une pathologie d'amylose de type AL (échantillons 4, 5 et 6) n'ont pas pu être diagnostiqués à l'issue de la *tPRM* par cette visualisation des données, traitées automatiquement par Pinpoint. Une observation plus approfondie des résultats permet de distinguer la présence dans ces pistes de quelques peptides supplémentaires appartenant aux chaines légères d'Ig Lambda. Pour nous permettre d'émettre une conclusion par comparaison, il aurait été souhaitable de pouvoir analyser d'autres échantillons présentant ce type d'amylose avec un fort pourcentage de dépôt amyloïde.

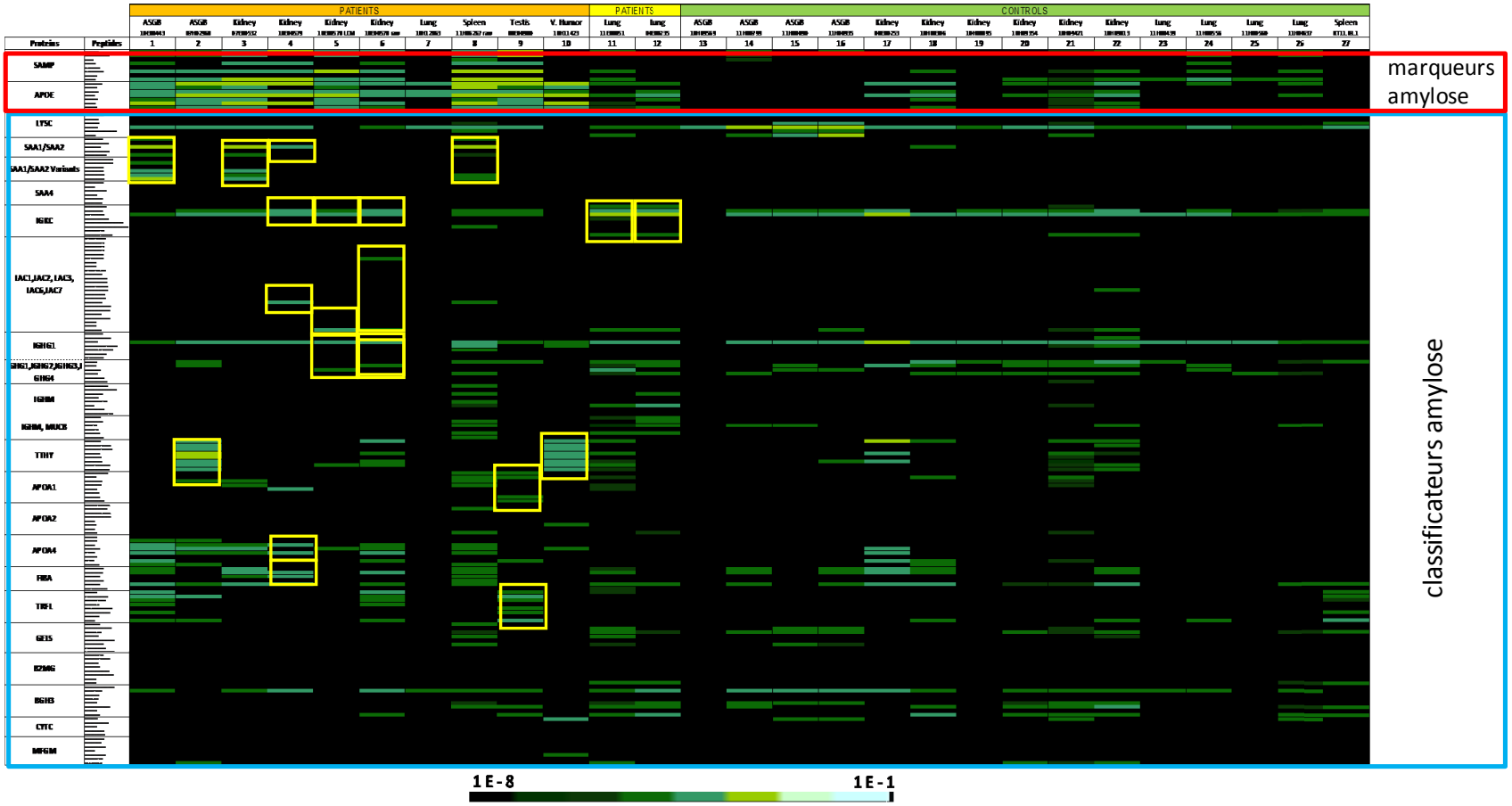


Figure 73 : Visualisation sous forme de cartographie *heatmap* des moyennes d'aires normalisées des peptides quantifiés dans les expériences protéomiques ciblées *time-scheduled PRM* de différents tissus avec Pinpoint. La couleur des cellules reflète l'intensité des signaux. Plus le peptide est vert clair, plus il est abondant.

III.1.2.2 ***Analyse en composantes principales***

Les données précédentes sont exploitées dans le logiciel open-source Tanagra, développé par Rakotomalala à Lyon [178], pour une analyse en composantes principales (ACP). Logiciel disponible via le lien internet : http://eric.univ-lyon2.fr/~ricco/tanagra/fr/tanagra.html.

Les meilleures composantes (PCA 1 et PCA 2) sont reportés sous forme de nuage de points de corrélation afin de distinguer les profils d'expression des peptides entre les patients amyloses et les contrôles non amyloses.

Contrairement aux données *shotgun* (Figure 74), dans le cas de l'analyse ciblée, une clustérisation des échantillons en deux groupes a pu se faire avec les données *tPRM* (Figure 75) : d'une part les patients amyloses et de l'autre les témoins. De plus, dans le groupe des témoins se trouvent les deux patients particuliers présentant des dépôts de chaines légères non amyloïdes (LCDD) correspondant aux colonnes 11 et 12 de la Figure 73.

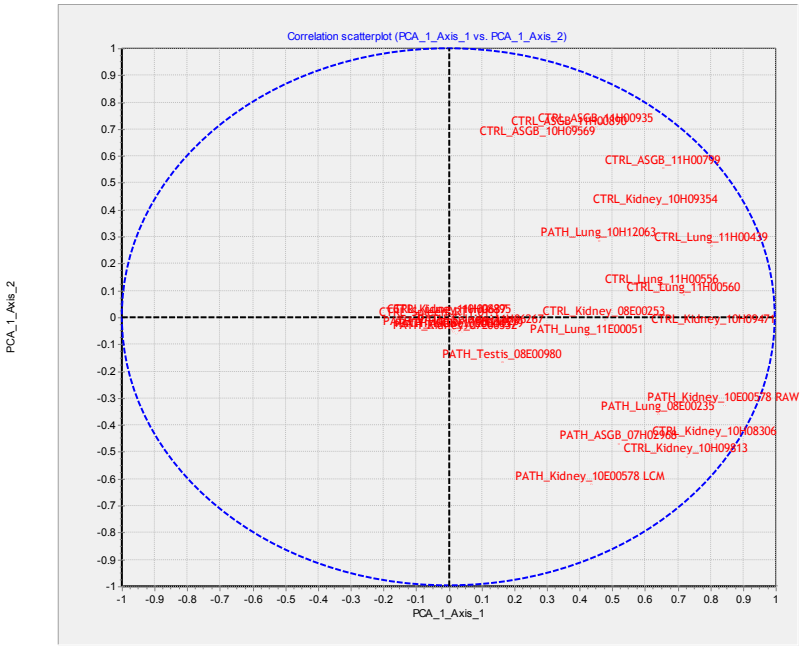


Figure 74 : Analyse en composantes principales des données *shotgun* après traitement via Proteome Discoverer. **Malheureusement, aucun groupe ne peut être distingué par l'ACP avec les données collectées provenant de l'analyse globale.**

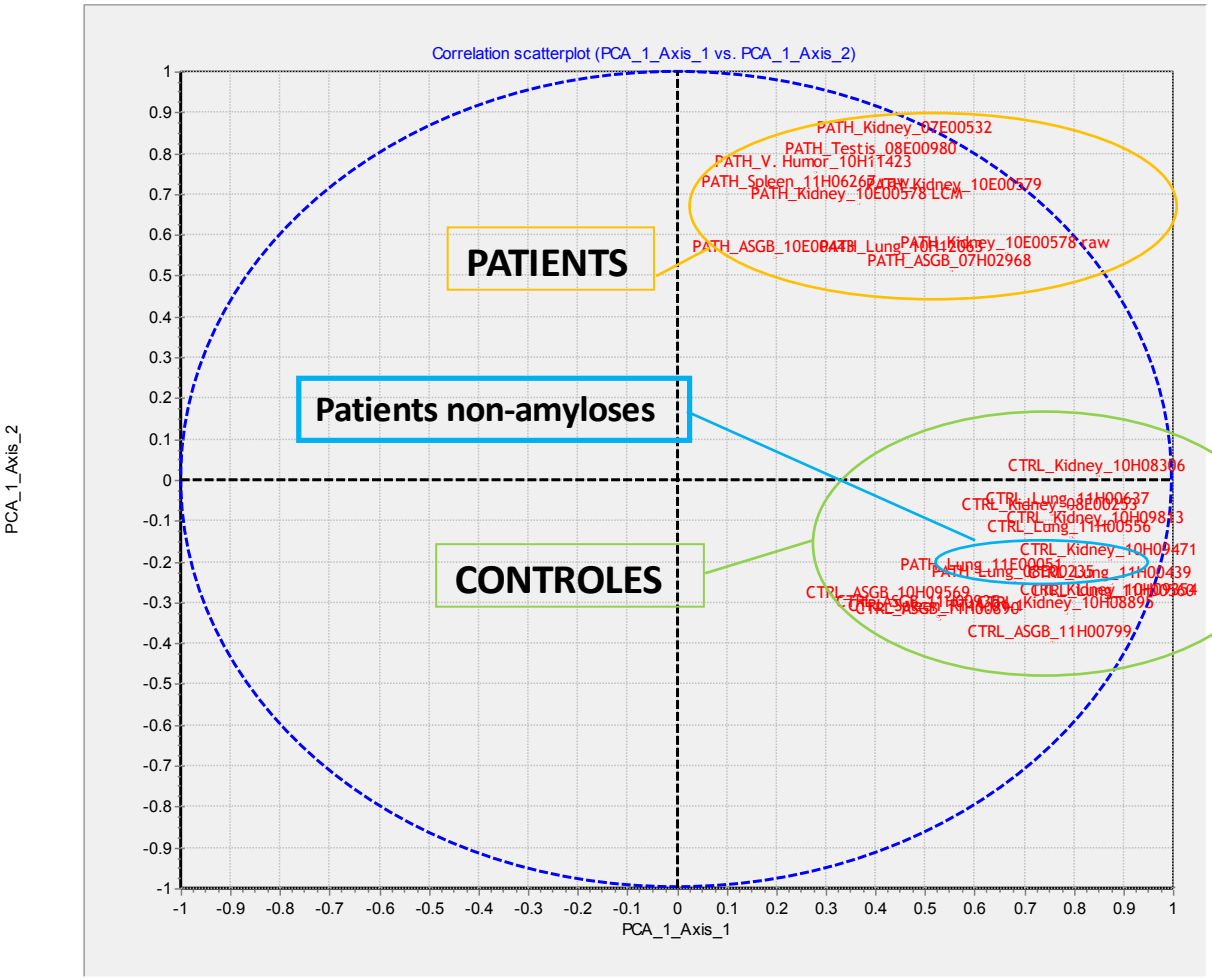


Figure 75 : Analyse en composantes principales des données *time-scheduled* PRM après traitement via Pinpoint. **Deux groupes se distinguent. D'un coté les patients amyloses et de l'autre les témoins ne présentant pas de pathologie amylose. Dans le groupe des témoins se trouvent les deux patients LCDD qui présentent une maladie de dépôt de protéines non amyloïde.**

Cette analyse en composantes principales permet de distinguer les patients atteints d'une pathologie amylose d'une pathologie non amylose. Afin de classifier le type d'amylose, seules les données des patients amyloses (groupe orange) sans les peptides appartenant aux deux marqueurs fibrillaires sont conservées pour la suite. Une deuxième ACP, montrée à la Figure 76 ci-dessous, est effectuée avec ces dernières données.

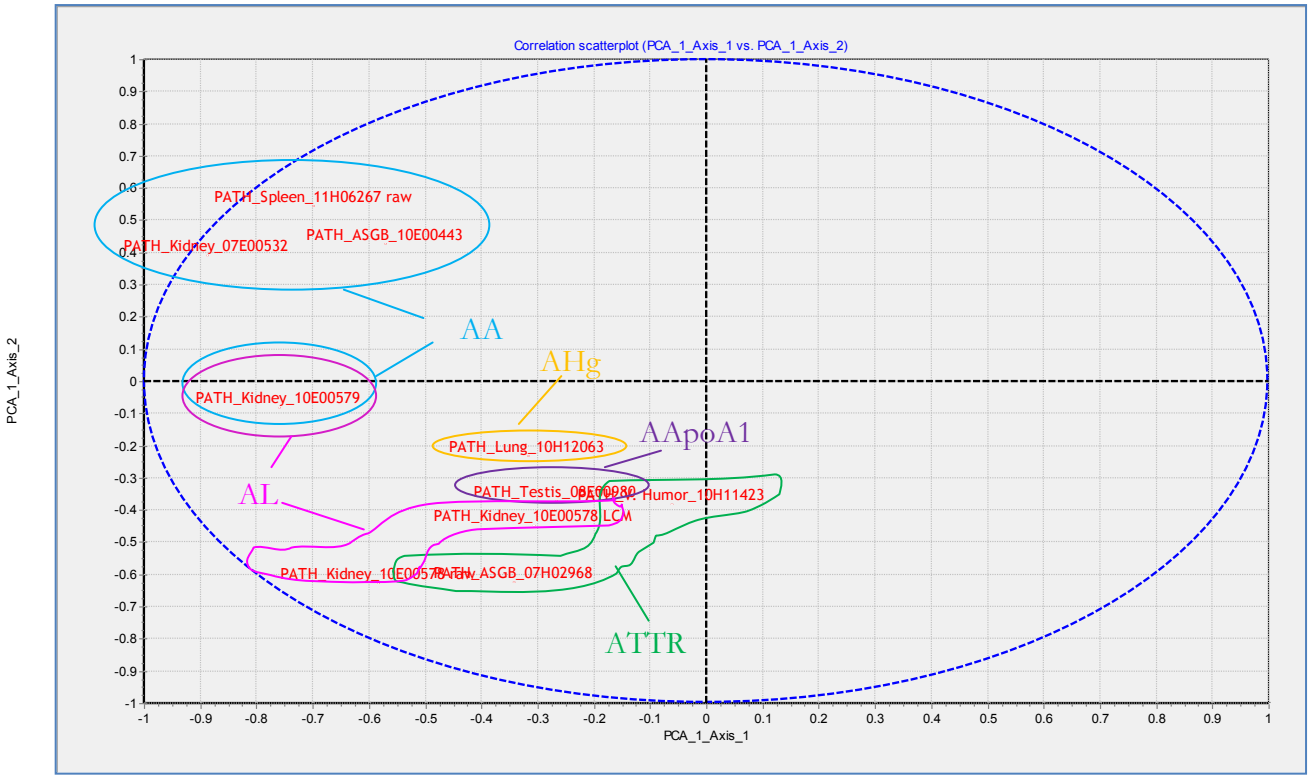


Figure 76 : Analyse en composantes principales des données *time-scheduled* PRM des patients présentant une amylose sans les peptides appartenant aux deux marqueurs fibrillaire (SAMP et ApoE) après traitement via Pinpoint.

Il est clair que nous manquons ici d'individus de chaque type pour pouvoir réellement effectuer un traitement statistique. Plusieurs groupent se distinguent dans cette ACP:

- Les patients présentant une amylose de type AA sont regroupés ensembles (en bleu, PATH_Spleen_11H06267_raw, PATH_ASGB_10E00443 et PATH_Kidney_07E00532).
- Les deux patients amyloses de type AL sont proches (en rose, PATH_Kidney_10E00578_RAW et PATH_Kidney_10E00478_LCM).
- L'échantillon PATH_Kidney_10E00579 qui présente une amylose de type AA et AL se distingue entre les deux groupes.
- Les deux échantillons de patients atteints d'une amylose de type ATTR sont proches (en vert, PATH_V.HUMOR_10H11423 et PATH_ASGB_07H02968).
- Concernant les échantillons présentant une amylose de type AApoA1 (PATH_Testis_08E00980, en violet) et AHg (PATH-lung_10H12063, en orange), seul un représentant de chaque type nous a été confié. Les deux types sont donc isolés.

Une classification des amyloses serait probablement réalisable avec cette méthode d'analyse en composantes principales. Le groupe des amyloses de type AA se détachent clairement des

autres groupes. Pour une clustérisation automatique des échantillons plus significative, il serait cependant nécessaire d'analyser plus d'échantillons de patients amyloses et non amyloses.

III.1.2.3 ***Analyse statistique et cartographie heatmap sous R***

Une clustérisation hiérarchique sous le logiciel statistique open-source R est également évaluée pour classer les amyloses. Les données utilisées dans le paragraphe précédent (présentées en Figure 76), sont transformées en logarithme (base 2) avant traitement avec ce logiciel.

Une normalisation en colonne pour distribuer les données sur une échelle de -1 à 1 est d'abord effectuée (Equation 22) avant de faire une clustérisation en colonne sous R avec le package *gplots* représenté à la Figure 77.

$$nombre\ normalisé = 1 - 2\left(\frac{nombre - valeur\ minimum\ colone}{valeur\ maximum\ colonne - valeur\ minmum\ colonne}\right) \quad \text{(Equation 22)}$$

Les échantillons sont disposés en colonne et les peptides en ligne. Ces derniers sont regroupés en fonction de la protéine dont ils sont issus pour une meilleure visibilité. Plus la cellule du peptide est rouge, plus son abondance dans l'échantillon est élevée. La lecture s'effectuant des colonnes de gauche à droite, le typage se décline de la manière suivante:

- Les patients **PATH_ASGB_07H02968** et **PATH_V.humor_10H11423** sont clustérisés ensemble et présentent une **amylose de type ATTR** (cas numéros 2 et 10 sur le *heatmap* à la Figure 73).
- L'échantillon provenant du patient **PATH_Testis_08E00980,** montre des peptides intenses de la protéine lactotransferrine (TRFL) constitutive du tissus [159–161]. Les peptides les plus abondants repérables ensuite, proviennent de l'apolipoprotéine A1 (APOA1). Ce patient présente donc **une amylose de type ApoA1** (cas numéros 9 sur le *heatmap* à la Figure 73).
- L'échantillon PATH_Kidney_10E00578_Raw est constitué de plusieurs protéines d'intensité égales. Le diagnostic est ambigu pour ce patient. (cas numéros 6 sur les *heatmap* à la Figure 73).
- Les patients **PATH_ASGB_10E00443**, **PATH_Spleen_11H06267_raw** et **PATH_Kidney_07E00532** ont une pathologie **amylose de type AA** et sont regroupés ensemble (cas numéros 1,3 et 8 sur le *heatmap* à la Figure 73).

- L'échantillon **PATH_Kidney_10E00579** comporte plusieurs peptides intenses de différentes protéines. Le diagnostic n'est pas distinct dans ce cas, car l'échantillon de tissus fournit ne présentait que 10% de dépôt amyloïde (cas numéro 4 sur le *heatmap* à la Figure 73).
- Le patient **PATH_Kidney_10E00578_LCM** manifeste une amylose de type AL mais probablement aussi AHg et AK. Cette expérience PRM reste cependant peu convaincante (cas numéros 5 sur le *heatmap* à la Figure 73).
- Les résultats pour le dernier patient **PATH_Lung_10H12063 ne permettent pas d'aboutir à une validation** car trop peu de peptides sont détectés dans ce cas (cas numéro 7 sur le *heatmap* à la Figure 73).

Comme montré précédemment dans la section III.1.2.1 (page 164), le diagnostic des échantillons porteurs d'amyloses de type AL reste peu concluant d'une méthode à l'autre. Cependant, les résultats de l'analyse en composantes principales sont très satisfaisants car ils permettent largement de discriminer les patients atteints de la maladie des amyloses d'une autre pathologie. Cette perspective de visualisation *heatmap* ouvre les portes d'un diagnostic plus direct qui se limiterait aux seules données des patients sans inclure les peptides appartenant aux deux biomarqueurs fibrillaires.

*III.1.3 **Interprétation générale des données***

L'abondance des peptides amyloïdogènes présent dans les différents tissus, doivent être en premier lieu extrait par le logiciel Pinpoint. Les AUCs sont ensuite normalisés par le TIC total de l'expérience correspondant et moyennées sur les trois réplicas techniques. L'ensemble des données sont ensuite soumis au logiciel Tanagra pour séparer la population en deux groupes distinct : patients amyloses des patients/témoins non amyloses. Les patients (sans les peptides des biomarqueurs) doivent ensuite être soumis au logiciel statistique R pour générer un regroupement hiérarchique (*heatmap*). La protéine amyloïdogène dont les peptides sont détectés en abondance (en rouge) est alors considérée comme responsable du dépôt amyloïde.

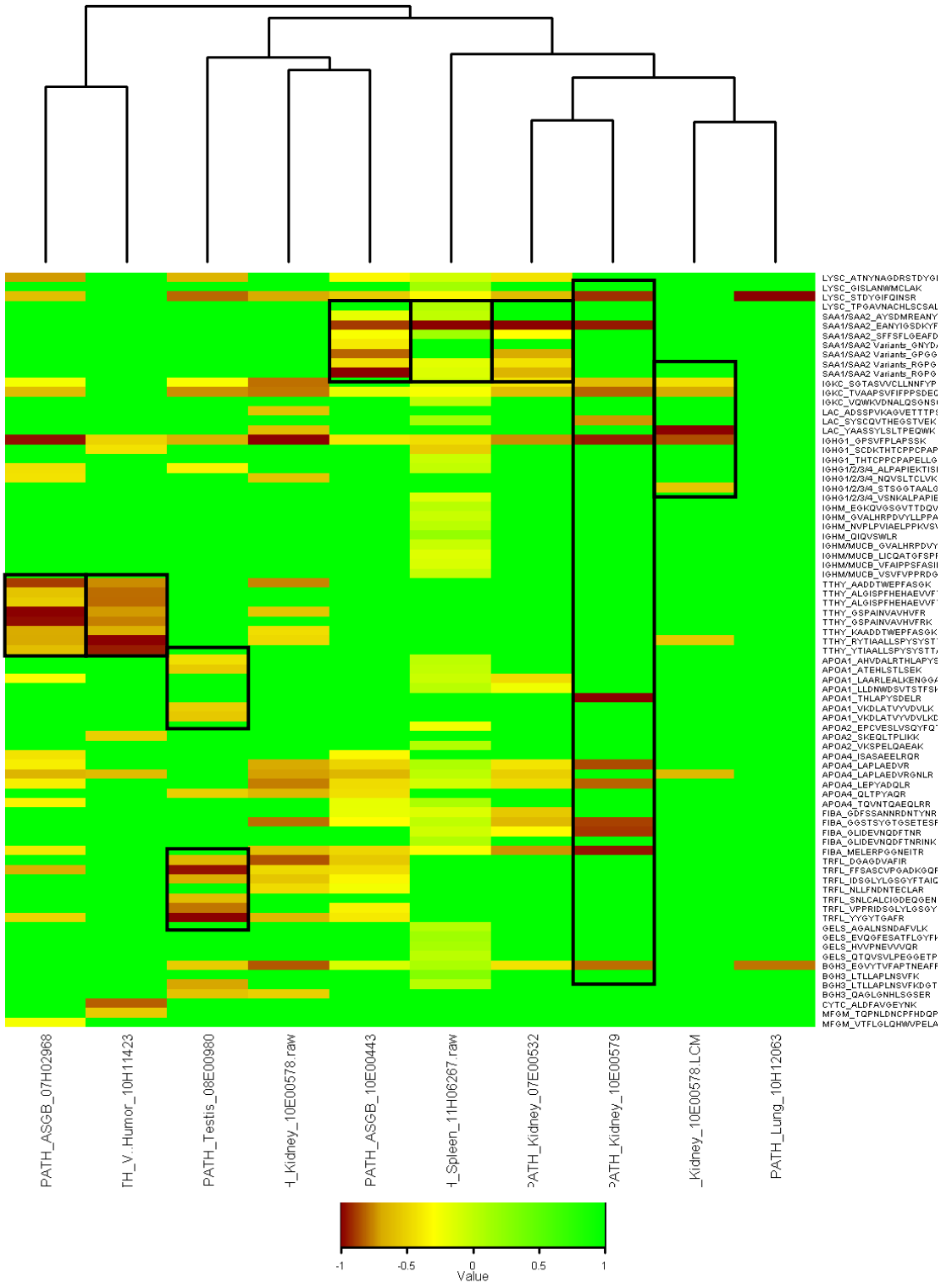


Figure 77 : Visualisation sous forme de cartographie *heatmap* des données *time-scheduled* PRM normalisées verticalement, des patients présentant une amylose sans les peptides appartenant aux deux marqueurs fibrillaires (SAMP et ApoE) après traitement via Pinpoint. **La couleur des cellules reflète l'intensité des signaux. Plus le peptide est rouge, plus il est abondant.**

III.2 Quantification des transitions par Skyline

Les données issues des analyses en mode *tPRM* sur le Q Exactive sont également traitées avec le logiciel open-source Skyline, développé par l'équipe de MacCoss aux États-Unis [177]. Ce logiciel est disponible via le lien suivant : http://proteome.gs.washington.edu/software/skyline

III.2.1 Traitement des données

Les peptides choisis sur Pinpoint sont importés sous Skyline. Les 5 ions les plus intenses de type y, générés par le logiciel Pinpoint, listés en Annexe 7, sont repris. Pour évaluer la méthode ciblée en mode *tPRM* avec Skyline, **les AUCs des pics correspondant aux transitions sont intégrées automatiquement sans correction manuelle**. Ce logiciel n'impose pas de minimum de signal, ni de largeur de pic comme paramètres d'intégration. **Pour valider les transitions, le score *dot product* (dotp) est utilisé. Il rend compte de la similitude entre l'aire des ions fragments mesurée et l'intensité de ces ions fragments dans les spectres de la libraire**. Pour constituer une libraire de spectres MS/MS la plus vaste possible, plusieurs librairies ont été créées sous Skyline. La première regroupe les spectres issus des expériences *tPRM*. La deuxième rassemble les spectres des expériences *shogun*. Et la troisième est constituée d'une banque de spectres publiques de plasma humain [179], disponible via le lien internet : http://www.peptideatlas.org/speclib/. Seuls les peptides ayant un score *dotp* supérieur ou égal à 0.7 sont pris en considération.

La quantification des peptides se fait par addition de toutes les AUCs des transitions correspondantes. Les résultats pour chaque réplica technique sont exportés sous forme de tableau comme pratiqué pour les données sous Pinpoint.

III.2.2 Data mining

III.2.2.1 *Cartographie heatmap sous tableur*

Une cartographie *heatmap* des données Skyline est également générée par le programme *sparklines* [158] présentée à la Figure 78. Les mêmes traitements de données Pinpoint détaillés au paragraphe III.1.2.1 à la page 164 du manuscrit, sont réalisés sur ces données Skyline.

Globalement, le nombre de peptides quantifiés de façon automatique est supérieur avec Skyline comparé à Pinpoint, car aucun seuil de détection n'est renseigné sur Skyline. Les peptides des deux protéines marqueurs des amyloses, SAMP et ApoE, sont très abondants chez les patients atteints de la maladie des amyloses (colonnes 1 à 10). Ces deux protéines sont également

détectées et quantifiées chez les contrôles (colonnes 13 à 27) ainsi que chez les deux patients LCDD (colonnes 11 et 12) mais en quantité et intensité moindre.

Les données Skyline pour chaque échantillon analysé en triplicata sont néanmoins plus compliquées à interpréter. Cependant le type de pathologie pour certains échantillons peut être clairement identifié.

- Plusieurs peptides des protéines SAA1 et SSA2 sont identifiés dans les **échantillons 1, 3, et 8**. Ils présentent une pathologie d'**amylose de type AA.**
- Plusieurs peptides de différentes protéines amyloïdogènes sont détectés dans l'**échantillon 2.** Cependant, ce sont les peptides de la transthyrétine (TTHY) qui sont les plus fortement quantifiés. **Le patient présente donc une amylose de type ATTR.**
- De même, plusieurs peptides des protéines amyloïdogènes sont identifiés dans l'**échantillon 4**. **Le diagnostic pour ce patient est ambigu.** On note cependant une forte concentration en apolipoprotéine A4 (APOA4), fibrinogène alpha (FIBA), sérum amyloïde (SAA1 et SAA2) et transthyrétine (TTHY).
- L'**échantillon 5** présente également plusieurs peptides de diverses protéines amyloïdogènes mais les peptides les plus intenses détectés et quantifiés sont issus des chaines légères d'Ig Lambda (LACs). **Au vu de ces résultats le diagnostic pour ce patient est une amylose de type AL.** Ce diagnostic est moins distinct pour l'**échantillon 6** qui provient du même prélèvement que l'échantillon numéro 5, préparé sans dissection LCM. Cet échantillon 5 ne présentait que 50% de tissu amyloïde. Ces deux exemples montrent bien l'apport de la dissection LCM dans le cas d'échantillons présentant peu de dépôts amyloïdes.
- Le **diagnostic ne peut être posé pour l'échantillon 7**, car il présente peu de peptides amyloïdogènes.
- Un certain nombre de peptides de protéines différentes est détecté dans **l'échantillon 9.** Toutefois les peptides appartenant aux protéines de lactotransferrine (TRFL) et apolipoprotéine A1 (ApoA1) sont les plus intenses. En première approche le diagnostic serait attribué à une co-amylose de type ALAC/AApoA1. Or la lactotransferrine étant constitutive du tissu, ce patient présente alors une **amylose de type AApoA1**.
- L'**échantillon 10** contient plusieurs de peptides de la transthyrétine (TTHY) et de l'apolipoprotéine A4 (ApoA4). En absence de témoin, il est difficile de conclure sur les effets de l'ApoA4. Cependant les peptides de la protéine TTHY étant les plus intenses, ce patient est potentiellement porteur d'une **amylose de type ATTR**.

- Enfin, le cas des **échantillons 11 et 12**. Les protéines SAMP et d'ApoE sont bien détectées dans ces biopsies mais en quantité proche des échantillons contrôles. Par contre les peptides provenant des chaines légères d'Ig kappa (IGKC) ainsi que les chaines lourdes d'Ig gamma (IGHG1) sont clairement visibles. Ces dernières sont aussi très présentes dans les tissus « témoins poumons », et peuvent donc être écartés. Ces deux patients présentent une maladie de **dépôt de chaines légères non amyloïdes** (LCDD) [42,43].

Avec une détection des peptides de façon plus sensible en PRM qu'en *shotgun*, il est possible classifier l'amylose du patient à l'exception du cas du poumon (échantillon 9) où très peu d'échantillon nous ont été donné et a été contaminé par les polymères (Figure 9).

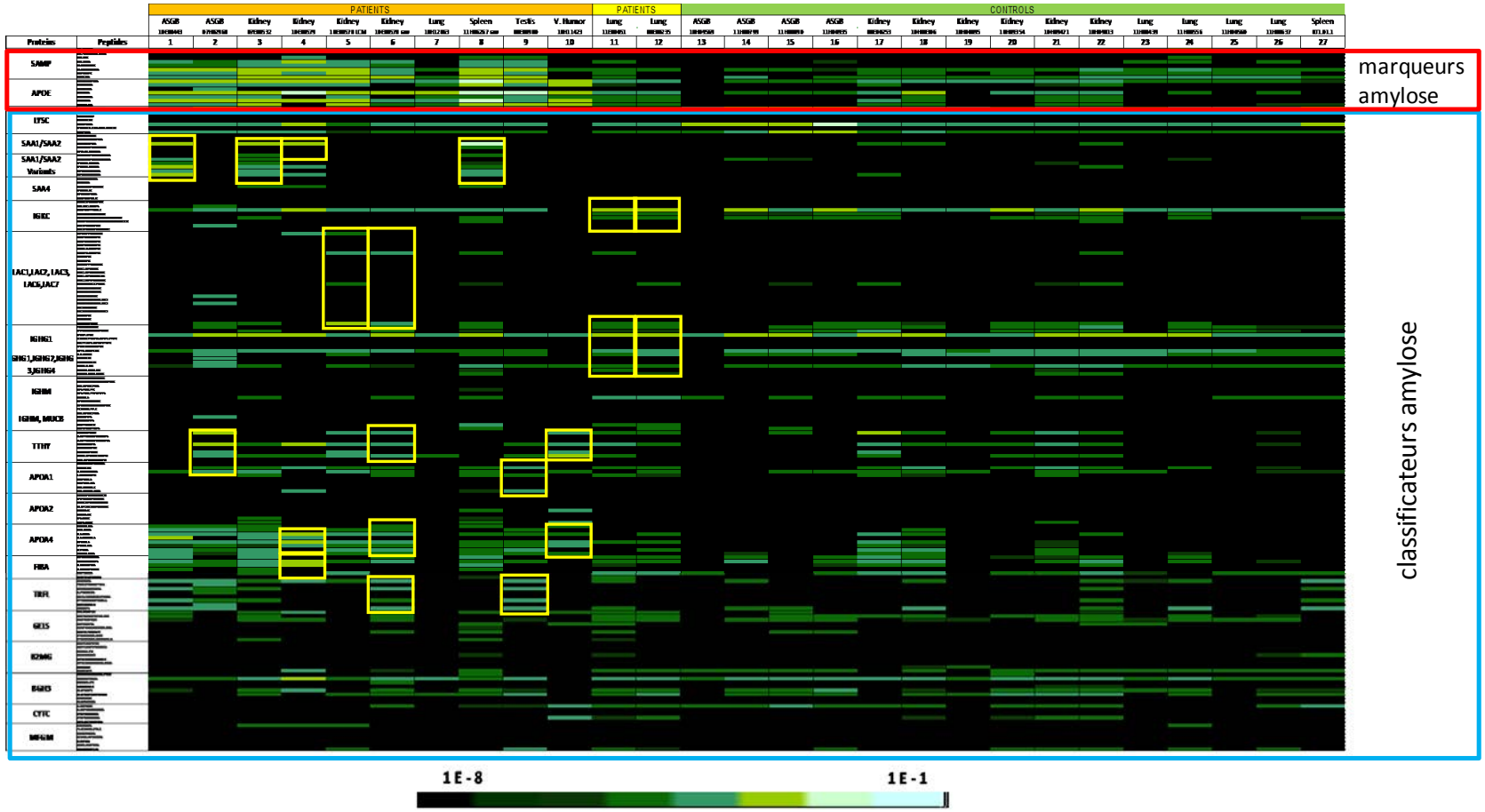


Figure 78 : Visualisation sous forme de cartographie *heatmap* des moyennes d'aires normalisées des peptides quantifiés dans les expériences protéomiques ciblées de différents tissus avec Skyline. La couleur des cellules reflète l'intensité des signaux. Plus la cellule est verte claire, plus le peptide est abondant.

III.2.2.2 ***Analyse en composantes principales***

Les données précédentes sont transférées dans le logiciel Tanagra [178] pour une analyse en composantes principales. Les meilleures composantes (PCA 1 et PCA 2) sont reportées sur la Figure 79.

À l'image des données *tPRM* traitées avec Pinpoint, une clustérisation automatique des échantillons est mise en avant après traitement des données par Skyline, regroupant les patients d'une part et les témoins d'autre part. Les deux patients particuliers (dépôts de chaines légères non amyloïdes, LCDD) sont également associés au groupe des témoins (colonnes 11 et 12 de la Figure 79). Nous distinguons clairement une séparation patients amyloses vs. patients non amyloses.

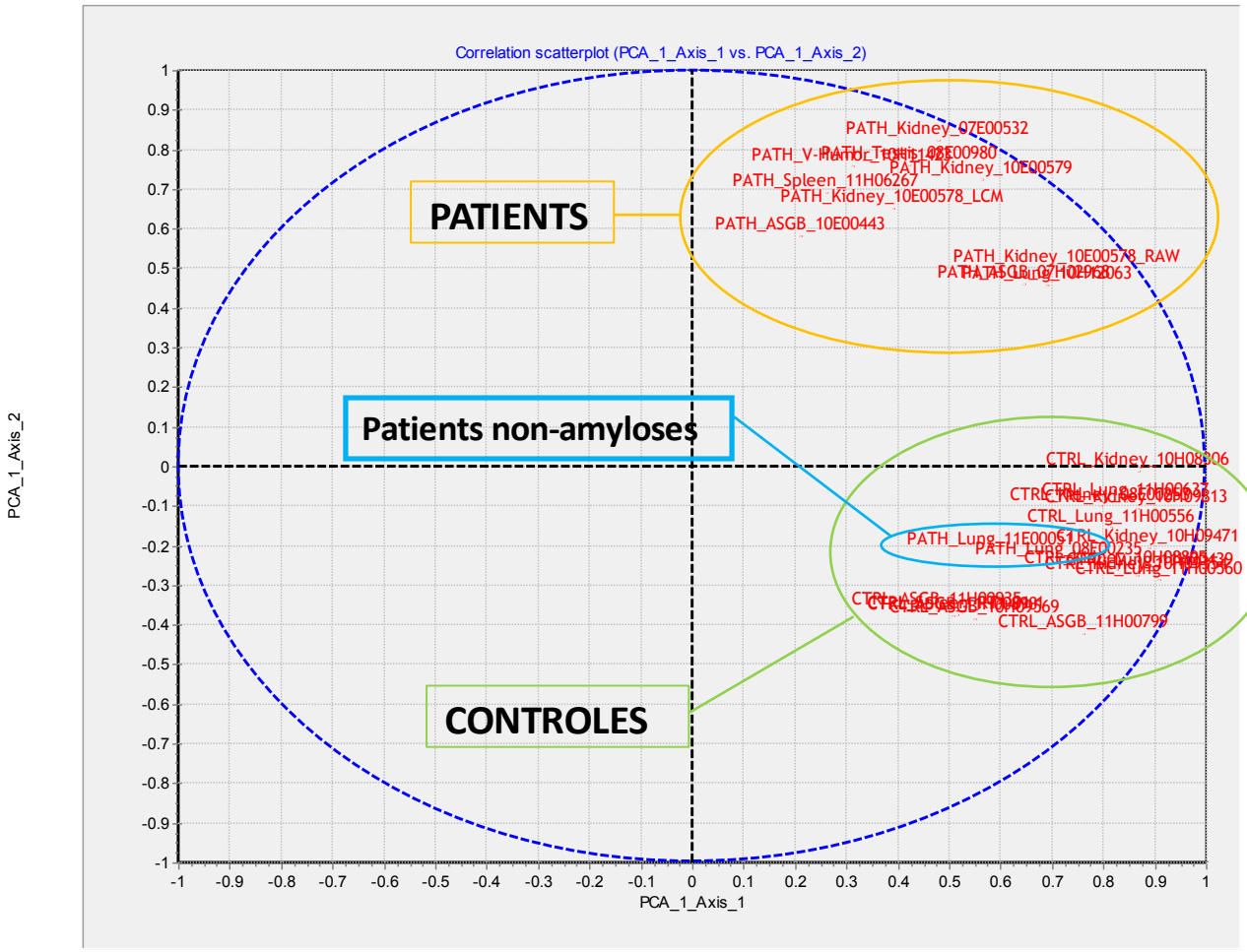


Figure 79 : Analyse en composantes principales des données *time-scheduled* PRM après traitement via Skyline. **Deux groupes se distinguent. D'un coté les patients amyloses et de l'autre les témoins ne présentant pas de pathologie amylose. Dans le groupe des témoins se trouvent les deux patients LCDD qui présentent une maladie de dépôt de protéines non amyloïde.**

Afin de classifier le type d'amylose, seules les données des patients amyloses (groupe orange) sans les peptides appartenant aux deux marqueurs fibrillaires sont conservées. Le résultat de la deuxième analyse en composantes principales est montré Figure 80 ci-dessous.

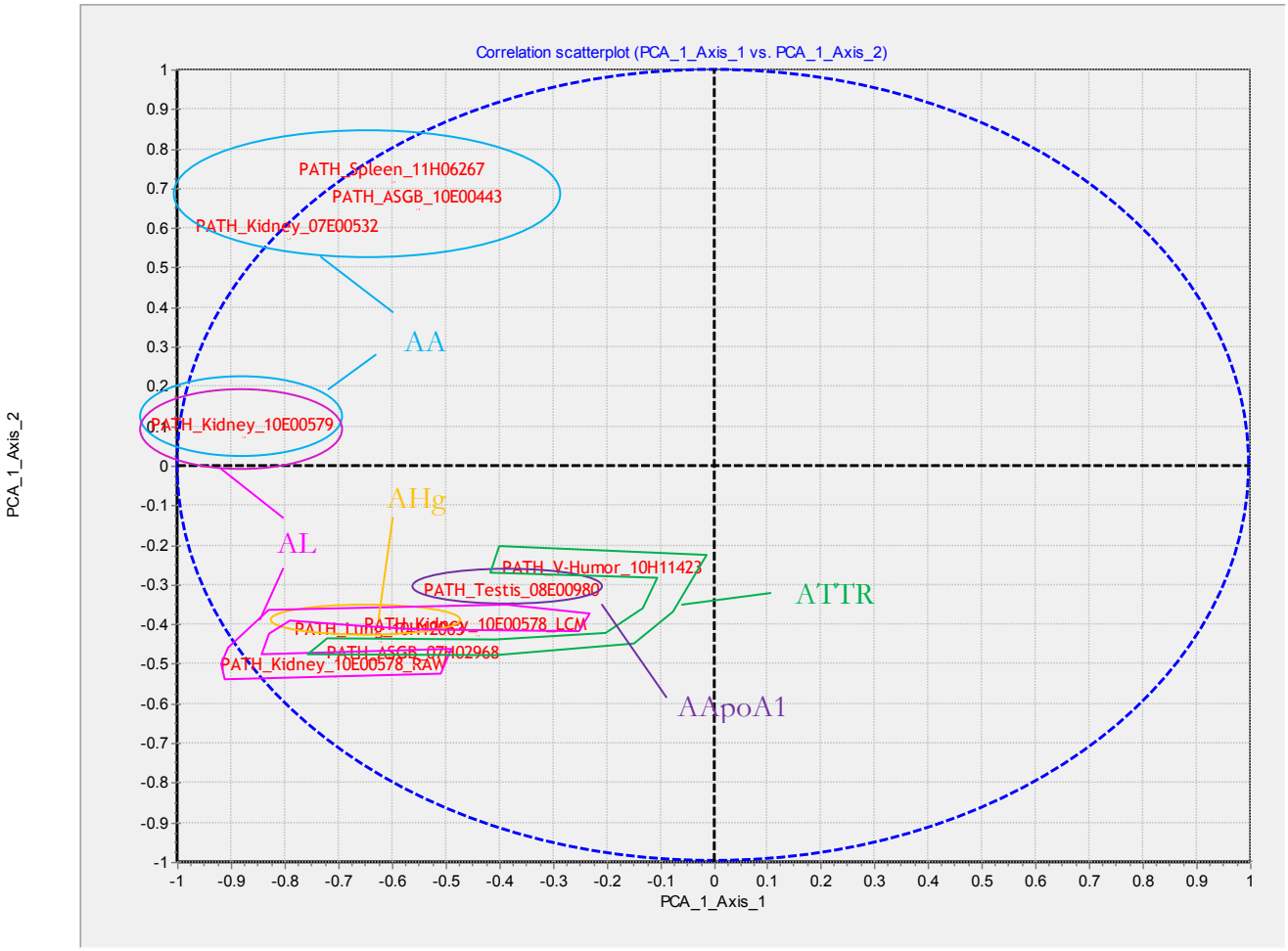


Figure 80 : Analyse en composantes principales des données *time-scheduled* PRM des patients présentant une amylose, sans les peptides appartenant aux deux marqueurs fibrillaires (SAMP et ApoE) après traitement via Skyline.

De même, un nombre trop faible d'échantillons ne permet pas d'effectuer un traitement statistique correct. Plusieurs groupent se distinguent dans cette ACP : Les patients présentant une amylose de type AA sont regroupés dans la même zone (en bleu, PATH_Spleen_11H06267_raw, PATH_ASGB_10E00443 et PATH_Kidney_07E00532) et constituent le premier grand groupe. Le deuxième ensemble réunit des patients pourtant porteurs d'amyloses de types différents, ce qui rend la classification assez délicate. Deux zones se distinguent toutefois au sein de ce deuxième espace. Les patients présentant une amylose de type AApoA1 (en violet, PATH_Testis_08E00980) et AHg (en orange, PATH-lung_10H12063) se répartissent entre les deux groupes d'amylose de type AL (PATH_Kidney_10E00578_RAW et PATH_Kidney_10E00478_LCM, en rose) et de type

ATTR (PATH_V.HUMOR_10H11423 et PATH_ASGB_07H02968 en vert). Finalement, l'échantillon PATH_Kidney_10E00579 qui montre une amylose de type AA et AL se positionne entre les deux grands groupes principaux.

L'amylose de type AA peut clairement être diagnostiquée avec ces méthodes, ce qui ne semble pas être le cas des autres types d'amylose. Une clustérisation plus significative serait peut être possible si plus d'échantillons sont intégrés à l'étude.

III.2.2.3 ***Cartographie heatmap sous R***

Une clustérisation sous R avec le package *gplots* est également évaluée pour classer les amyloses. Les données utilisées sont les mêmes que celles de la deuxième analyse en composantes principales présentée Figure 80 mais transformées au préalable par le logarithme de base 2 puis normalisées de la même façon que ceux du paragraphe III.1.2.3 (p.171). Le résultat est présenté en Figure 81.

Les échantillons sont disposés en colonne et les peptides en ligne (le nom de la protéine associée au peptide est inclus à la séquence de ce dernier pour faciliter le traitement). Plus la cellule du peptide est rouge, plus son abondance dans l'échantillon est élevée. Le typage peut alors se faire de la manière suivante, des colonnes de gauche à droite:

- Les patients **PATH_Spleen_11H06267_raw, PATH_Kidney_10E00579**, **PATH_Kidney_07E00532** et **PATH_ASGB_10E00443** ont une pathologie **amylose de type AA** et sont regroupés ensemble (cas numéros 8, 4, 3 et 1 sur le *heatmap* à la Figure 78).

  Pour le patient PATH_Kidney_10E00579 qui présente plusieurs types d'amylose, cette clustérisation penche pour une amylose de type AA. Ce qui est en concordance avec le résultat obtenu dans les analyses *shotgun* où le sérum amyloïde A est majoritaire.
- Le résultat de l'échantillon **PATH_Kidney_10E00578_Raw** est constitué de plusieurs protéines d'intensité égales. Le diagnostic est ambigu pour ce patient. (cas numéros 6 sur le *heatmap* à la Figure 78).
- L'échantillon du patient **PATH_Testis_08E00980** présente plusieurs peptides intenses des protéines lactotransferrine (TRFL) constitutive du tissu, apolipoprotéine A1 (APOA1), apolipoprotéine A4 (APOA4) et chaines lourdes d'Ig gamma (IGHG1). Ce patient présente des amyloses potentielles de type **ApoA1 ou et ApoA4 ou AHg**. (cas numéros 9 sur le *heatmap* à la Figure 78).

- Le patient **PATH_ASGB_07H02968** présente une **amylose de type ATTR** (cas numéro 2 sur le *heatmap* à la Figure 78).
- Le patient **PATH_Kidney_10E00578_LCM** est porteur d'une amylose de type AL mais aussi AHg (cas numéros 5 sur le *heatmap* à la Figure 78).
- Le patient **PATH_V.humor_10H11423** est atteint d'une **amylose de type ATTR** (cas numéro 10 sur le *heatmap* à la Figure 78).
- L'échantillon du dernier patient **PATH_Lung_10H12063** montre la présence du lysozyme (LYSC) et des chaines légères Ig gamma (IGHG1). Le lysozyme étant un facteur antimicrobien, le diagnostic ici penche plutôt pour **une amylose de type AHg**, ce qui corrobore les résultats obtenus lors des expériences *shotgun* (cas numéro 7 sur le *heatmap* à la Figure 78).

Avec un nombre plus élevé de peptides identifiés de façon significative avec le logiciel Skyline, un typage des amyloses est toutefois possible avec cette visualisation *heatmap* créée sous R. Les patients atteint d'une amylose de type AA sont tous clustérisés ensembles. Pour les autres pathologies amyloses, la clustérisation n'est pas clairement explicable. Il faudrait probablement une cohorte de patients plus importante pour parvenir à une meilleure clustérisation.

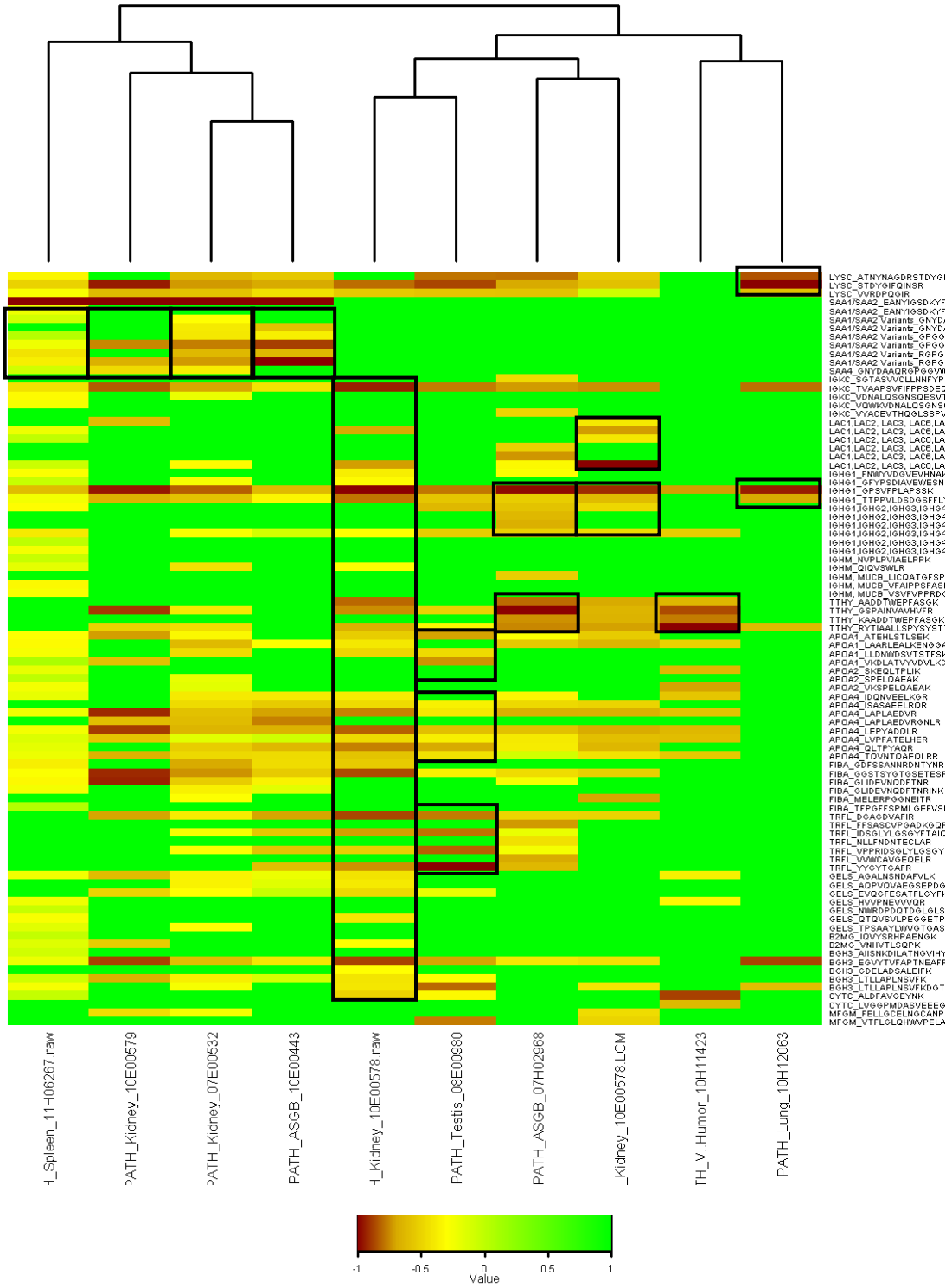


Figure 81 : Visualisation sous forme de cartographie *heatmap* des données *time-scheduled* PRM normalisées verticalement des patients présentant une amylose sans les peptides appartenant aux deux marqueurs fibrillaires (SAMP et ApoE) après traitement via Skyline. La couleur des cellules reflète l'intensité des signaux. Plus le peptide est rouge, plus il est abondant.

IV. Conclusion : proposition d'un protocole

La quantification des protéines amyloïdogènes peut se faire avec le mode *tPRM* sur un instrument hybride à haute résolution de type quadripôle-Orbitrap. Deux solutions logicielles sont disponibles : le logiciel constructeur, Pinpoint, ou un logiciel open-source, Skyline. Ils présentent des résultats assez comparables avec une légère complémentarité.

En ciblant les analytes, la spécificité de détection est améliorée. Ce qui se traduit par un nombre de peptides d'intérêt détectés plus élevé. Les peptides des deux marqueurs fibrillaires sont dorénavant détectés de façon significative chez les témoins et patients non amyloïdes. Cependant, l'abondance relative mesurée est plus faible dans les contrôles « négatifs » que chez les patients atteints d'une amylose, ce qui concorde bien avec les observations du pathologiste. De plus, à l'inverse des approches DDA globales (*shotgun*), l'analyse en composantes principales de ces données permet de bien distinguer des échantillons analysés selon deux groupes distincts : les échantillons présentant un dépôt amyloïde et les échantillons négatifs à la pathologie des amyloses.

Une limitation pour l'analyse automatique concerne la stabilité des conditions chromatographiques. Puisque l'un des paramètres majeurs de la méthode *tPRM* est le temps de rétention, il est important de contrôler la robustesse de ce paramètre avant de faire les analyses. Plutôt que de corriger les temps avant chaque série d'analyse, comme pratiqué dans notre cas, une correction en direct au cours de l'analyse peut être théoriquement réalisée comme en SRM. Cependant, cette option n'est pour l'instant disponible que dans un seul laboratoire situé au Luxembourg [153], mais nul doute que ce paramétrage sera dans l'avenir intégré dans l'ensemble des logiciels constructeurs.

Afin de cribler plus de peptides, un multiplexage peut être réalisé avec le spectromètre de masse Q Exactive. Plusieurs peptides co-élués peuvent alors être sélectionnés et fragmentés individuellement puis stockés dans la cellule de collision HCD pour ensuite être analysés en même temps dans l'Orbitrap. Le spectre obtenu est donc la combinaison de différents fragments de plusieurs peptides. Le nombre de peptides peut croître d'un facteur maximum de 10 avec un déca-multiplexage. Cependant, il faut rester très prudent quant au choix des transitions. En effet des ions fragments quasi-isobares de peptides éluant à des temps de rétention très proches peuvent être co-élués et se retrouver à tord attribués à un précurseur donné. La quantification pourrait en être biaisée.

Chapitre 7 : *Comparaison des analyses sur deux appareils de différente génération*

I. Introduction

Les solutions technologiques en spectrométrie de masse ne cessent d'évoluer. Cette évolution extrêmement rapide résulte d'une part de la pression scientifique toujours plus exigeante, mais aussi des progrès technologiques constants dans les domaines de l'électronique et de traitement de données. Dans le domaine commercial cette évolution continue est aussi la conséquence d'une politique de marketing agressif. Chaque nouveauté doit donc être évaluée avec circonspection et il n'est pas rare que les différences de performance ne soient réellement significatives que sur plusieurs générations d'instrument, selon le type d'application envisagé. Lors du dernier congrès de l'ASMS (*American society of mass spectrometry*) en juin dernier, une nouvelle génération de Q Exactive a été présentée à la communauté scientifique, le Q Exactive HF [154]. En comparaison avec l'instrument disponible au laboratoire il s'agit de la génération N+2 de notre appareil. Pour mettre à jour l'évaluation de notre méthode nous avons donc décidé de reprendre notre stratégie sur cette nouvelle configuration.

Ce dernier modèle possède une optique et un quadripôle améliorés permettant d'augmenter le pourcentage de transmission des ions et un champ électrostatique plus intense (dit HF*, high filed*) dans un Orbitrap plus petit pour augmenter la fréquence d'analyse et la résolution. Par rapport à l'instrument sur lequel les analyses au cours des chapitres précédent ont été réalisées, cet instrument est alors capable de faire des analyses à haut débit pour une même résolution ou des analyses à plus haute résolution pour une même fréquence (cf. chap.3 §V.4 page 93).

L'objectif de cette partie est de chercher à mettre en évidence les gains apportés par cette nouvelle génération d'instrument en mode d'acquisition *shotgun* DDA et en mode ciblée PRM.

II. Analyse shotgun

II.1 Méthodologie

Afin de comparer les performances des deux spectromètres de masse, les échantillons de rate contrôle (numéro RT1.01.1) et de patient (numéro 11H06267) qui ont été notre modèle d'étude dans le Chapitre 5, sont de nouveaux repris pour être analysés en parallèle en *shotgun* sur le Q Exactive au sein de notre laboratoire et sur le Q Exactive HF, mis à disposition dans un laboratoire d'application (Hemel, GB). Les instruments sont tous deux couplés à une nanoLC (Ultimate 3000 RSLC, Thermo Scientific). Nous avons ici quelque peu changé notre protocole d'analyse pour nous approcher au mieux des conditions d'étude préconisée par le laboratoire d'application du constructeur. D'une façon générale, le principe de séparation en LC et d'acquisition MS change peu. Les paramètres et les tampons A et B (Tableau 18) sont légèrement différents de nos conditions analytiques. Notre montage chromatographique et notre spectromètre de masse étant différents du laboratoire de test, nous avons adapté notre méthode pour nous rapprocher au maximum des conditions analytiques de ce laboratoire. La méthode choisie pour les analyses MS est le DDA Top 20 (1 MS, suivie de 20 MS/MS sur les ions les plus intenses non exclus). Reprenons ci-dessous ces paramètres.

Tableau 18 : Comparaison des conditions chromatographiques entre les deux laboratoires, la dernière colonne indique si les conditions sont strictement identiques

	Q Exactive HF (Hemel, GF)	QExactive (Paris, France)	Identique?
Précolonne	Acclaim PepMap 100, C_{18}, 5µm, 100µm d.i. x 2cm L., Dionex	Acclaim PepMap 100, C_{18}, 5µm, 300µm d.i. x 5mm L., Dionex	Non
Dessalage	6µl/min ; 5min	15µl/min, 5 min	Non
Tampon A'	eau/ACN 98 :2 (v/v) ; 0.05% TFA	eau/ACN 98 :2 (v/v) ; 0.05% TFA	Oui
Colonne	Acclaim PepMap 100, C_{18}, 3µm, 75µm d.i. x 15cm L., Dionex	Acclaim PepMap 100, C_{18}, 3µm, 75µm d.i. x 15cm L., Dionex	Oui
Gradient	300nl/min, 3% à 35% tampon B en 60 min	300nl/min, 2.7% à 31.1% de tampon B en 60 min	Oui
Tampon A	eau + 0.1% FA	eau/ACN 98 :2 (v/v) + 0.1% FA	Non
Tampon B	eau/ACN: 80/20 + 0.1% FA	eau/ACN 90 :10 + 0.1% FA	Non
Température	40°C	40°C	Oui

TFA acide trifluoroacétique; FA : acide formique ; ACN : acétonitrile.

Il faut donc noter (voir Tableau 18) que la colonne de piégeage utilisée est de plus grande capacité, ce qui nous amène à travailler avec un volume équivalent colonne plus important

(cependant le rapport des volumes de colonne est 9, le rapport des volumes de solvant est 2,5). Les gradients sont programmés différemment pour compenser les différences de composition de tampon ce qui correspond finalement à des séparations analytiques identiques.

En ce qui concerne les conditions d'analyse en spectrométrie de masse (voir Tableau 19) nous avons travaillé à une résolution de 35000 au lieu de 60000 en mode MS car le Q Exactive de première génération est équipé d'un détecteur plus lent à résolution égale, et nous avons toléré des temps de remplissage légèrement supérieurs car la l'efficacité transmission est légèrement inférieure. En revanche nous ne sommes pas limités pour le mode MS/MS dont la résolution est même légèrement supérieure dans le cas de l'appareil de l'ESPCI.

Tableau 19 : Comparaison des conditions de détection entre les deux laboratoires, la dernière colonne indique si les conditions sont strictement identiques

	Q Exactive HF (Hemel, GF)	**QExactive (Paris, France)**	**Identique?**
méthode	DDA TOP 20	DDA TOP 20	Oui
Gamme de masse	400 à 2000 Da	400 à 2000 Da	Oui
Résolution MS	60 000 (à m/z 200)	35 000 (à m/z 200)	Non
AGC MS	1e6	1e6	Oui
temps d'injection max MS	100 ms	120 ms	Non
Fenêtre de sélection	1,4 m/z,	1,4 m/z,	Oui
Résolution MS/MS	15000 (à m/z 200)	17500 (à m/z 200)	Non
AGC MS/MS	5e4	5e4	Oui
temps d'injection max MS/MS	100 ms	100 ms	Oui
Energie collision normalisée	28	28	Oui
Exclusion dynamique	30s	30s	Oui

II.2 Résultats

Ces échantillons analysés en triplicatas sont ensuite soumis au logiciel Proteome Discoverer pour une recherche dans la banque de données SwissProt. Les paramètres de recherche renseignés sont les mêmes que pour les expériences *shotgun* sur le Q Exactive définis au cours du chapitre 5. Seuls les peptides identifiés avec un FDR inférieur à 1% et de rang 1 sont considérés.

Les chromatogrammes d'ions obtenus dans chaque laboratoire sont montrés à la Figure 82 pour le témoin (numéro RT1.01.1) et Figure 83 pour le Patient (numéro 11H06267). Globalement, le profil chromatographique est similaire avec un écart d'environ 10 min entre

les deux systèmes LC, à l'exception d'un large massif qui apparait entre 75 et 80 min sur les profils chromatographiques obtenus avec le Q Exactive HF. Ce décalage de temps de rétention est dû à la différence de volume mort entre les deux couplages et non une différence de temps de départ puisque l'acquisition est déclenchée au bout de 5 min dans les deux laboratoires. Nous avons aussi imaginé que le massif supplémentaire observé à Hemel et qui n'a jamais été observé à l'ESPCI provient d'une contamination du système chromatographique non dédié. Seule la partie informative est alors conservée pour les recherches dans les banques de données: soit de 6 à 71 minutes pour les analyses sur le Q Exactive et de 12 à 77 minutes pour les analyses sur le Q Exactive HF.

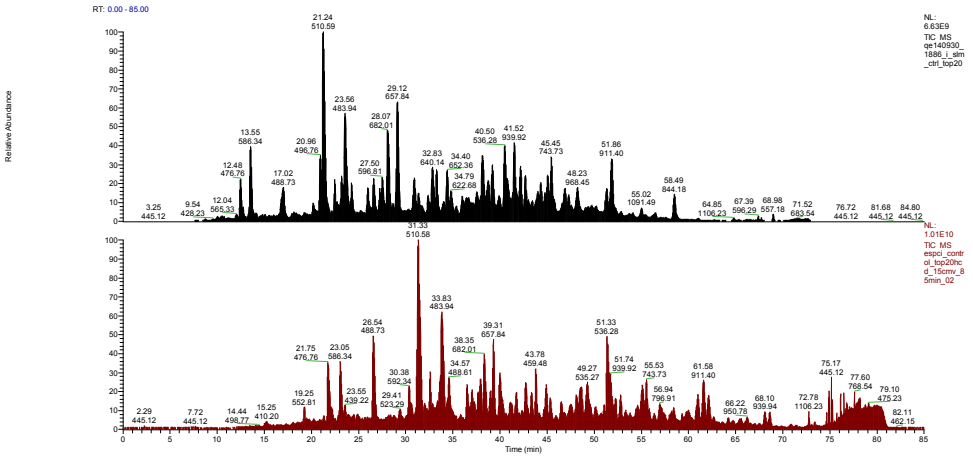


Figure 82 : Profil chromatographique des analyses *shotgun* DDA TOP 20 de l'échantillon contrôle (numéro RT1.01.1) sur le Q Exactive (en haut en noir) et sur le Q Exactive HF (en bas en rouge).

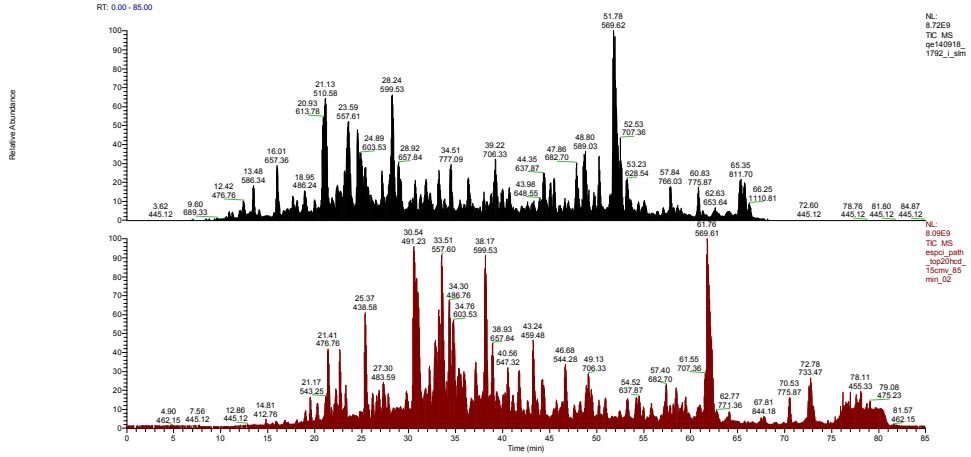


Figure 83 : Profil chromatographique des analyses *shotgun* DDA TOP 20 de l'échantillon patient (numéro 11H06267) sur le Q Exactive (en haut en noir) et sur le Q Exactive HF (en bas en rouge).

Le nombre de peptides uniques (identifiés par Mascot et Sequest) et de groupes de protéines identifiés (avec au minimum 3 peptides uniques identifiés par Mascot et Sequest) sont résumés pour chaque instrument sous forme d'histogramme sur la Figure 84. Les tendances ne sont pas exactement les mêmes pour les deux échantillons.

Pour l'échantillon contrôle (CTRL de la Figure 84), le Q Exactive HF apporte près de 40% d'identification de peptides uniques en plus pour une même durée d'élution chromatographique (60 min). Cet écart entre les deux systèmes se réduit à 30% au niveau des groupes de protéines identifiés avec toutefois un nombre plus élevé avec le Q Exactive HF.

En revanche, concernant l'échantillon pathologique (PATH de la Figure 84), le nombre de peptides uniques et groupes de protéines sont sensiblement identiques avec les deux configurations. Cette différence peut s'expliquer en observant l'intensité relative (NL, *normalized level*) des profils chromatographiques. En effet, l'intensité normalisée globale de l'échantillon témoin sur le Q Exactive (NL:6e9) est presque deux fois moins élevée que sur le Q Exactive HF (NL:1e10) alors que pour le patient il y a peu de différence (NL:8.7e9 pour le Q Exactive et NL:8.1e9 pour le Q Exactive HF). Il est possible qu'il ait eu une différence de quantité injectée pour cet échantillon. Ceci nous permet de montrer l'influence de la quantité d'échantillon sur la robustesse de l'analyse et ce indépendamment de l'appareil utilisé.

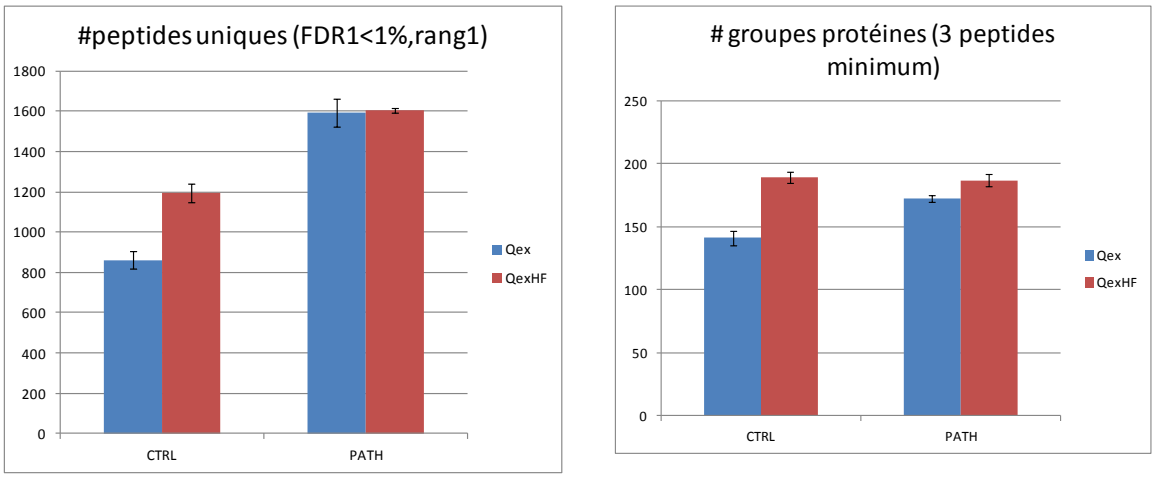


Figure 84 : Comparaison du nombre de peptide uniques identifiés et le nombre de groupes de protéines avec au minimum 2 peptides uniques entre le Q Exactive (en bleu) et le Q Exactive HF (en rouge). **Pour l'échantillon témoin, cette nouvelle génération d'instrument apporte un nombre significatif d'identification en plus. Cependant pour l'échantillon patient, l'apport est réduit.**

Concernant le diagnostic des amyloses, il peut être posé sans difficulté bien sûr pour les deux instruments. L'échantillon témoin rate RT1.01.1 exprime en majorité du lysozyme C, protéine antimicrobienne naturellement présent dans ce type de tissus (Figure 85 A et C) et aucun marqueur fibrillaire n'est détecté de façon significative. L'échantillon patient 11H06267 (Figure 85 B et D) présente bien une amylose (protéines SAMP et ApoE sous la dénomination *Amyloid markers*) de type AA (protéine AA).

Il faut cependant noter une différence en termes d'abondance relative des différentes protéines d'intérêt entre les deux spectromètres pour le diagnostic des amyloses en mode DDA. Le but étant de mettre en place une méthode ciblée pour un criblage à haut débit dans les départements cliniques français, nous nous sommes donc tournés vers la quantification des protéines amyloïdogènes en protéomique ciblée à HR/AM.

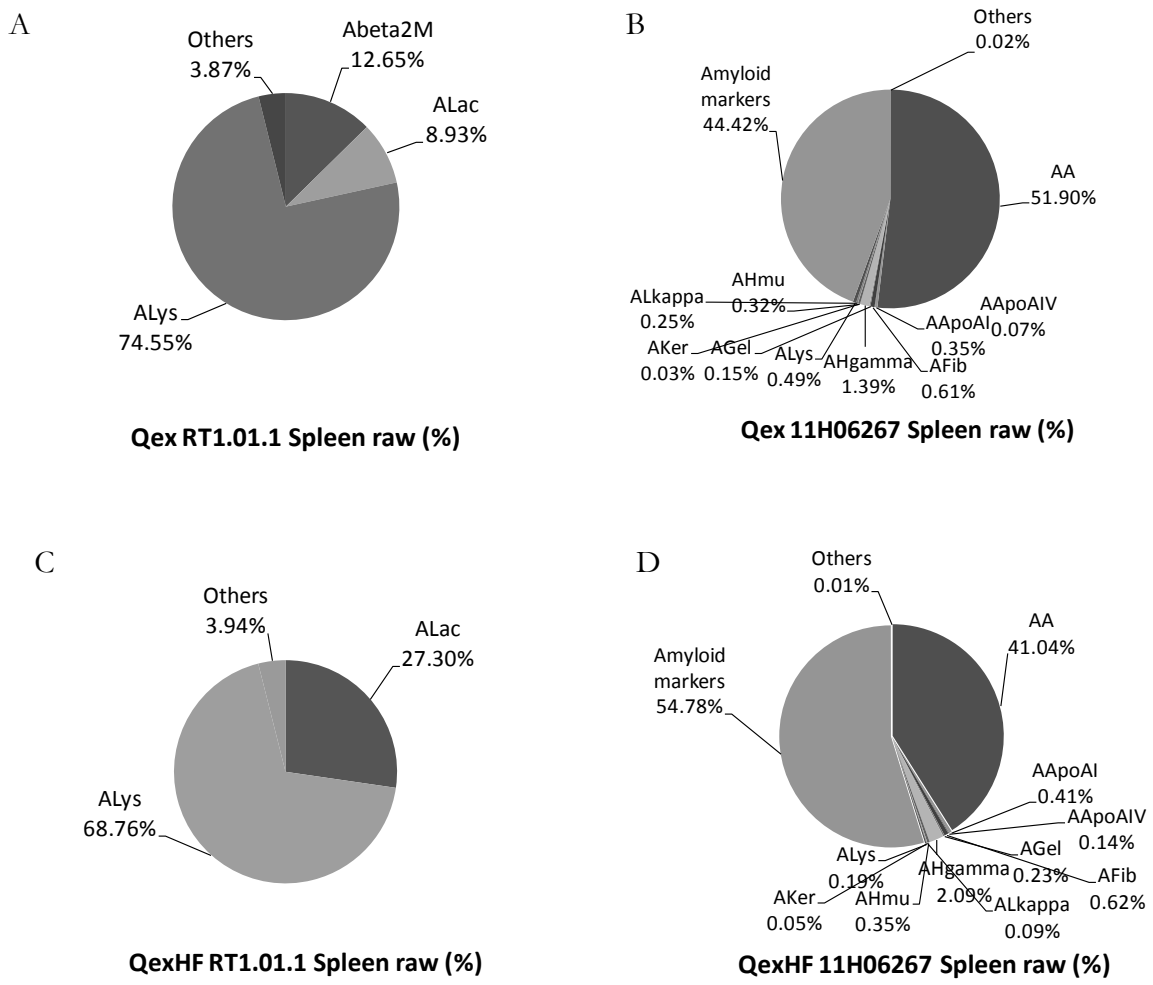


Figure 85 : Biopsie de rate. Témoin rate analysé sur le Q Exactive (A) et le Q Exactive HF (C). Échantillon présentant une amylose de type AA avec 90% de dépôt fibrillaire analysé sur le Q Exactive (B) et le Q Exactive HF (D). La catégorie « *Amyloid markers* » cumule en fait l'apport des deux protéines communes caractéristiques des amyloses. La catégorie « *Others* » regroupe toutes les autres protéines amyloïdogènes non détectées.

III. Analyse ciblée en mode PRM

A l'instar du Q Exactive, une analyse ciblée à haute résolution en mode PRM est possible sur le Q Exactive HF.

III.1 Méthodologie

III.1.1 Choix des peptides

La liste des peptides identifiés dans les échantillons témoin et patient en mode DDA TOP20 sur le Q Exactive HF est confrontée à la liste du fichier Pinpoint créé dans le chapitre 6 précédent. Les peptides de cette liste non détectés sur le Q Exactive HF sont exclus pour la PRM (environ 30%). Lorsque le nombre de peptides est insuffisant, les peptides identifiés présentant un site de coupure manqué sont inclus. La liste finale des peptides suivis est listée en Annexe 8. Pour chaque peptide, 6 ions de type y et b, parmi les plus intenses, sont sélectionnés par Pinpoint pour la quantification en *tPRM*. Cette méthode préconisée par le laboratoire de tests, nous donne un maximum de 78 transitions au lieu de 93 dans notre méthode originelle suivies en même temps (Figure 86) soit 13 peptides au lieu de 31 dans notre méthode originelle. Ceci va permettre d'obtenir plus de points descriptifs le long du pic pour assurer une quantification plus robuste.

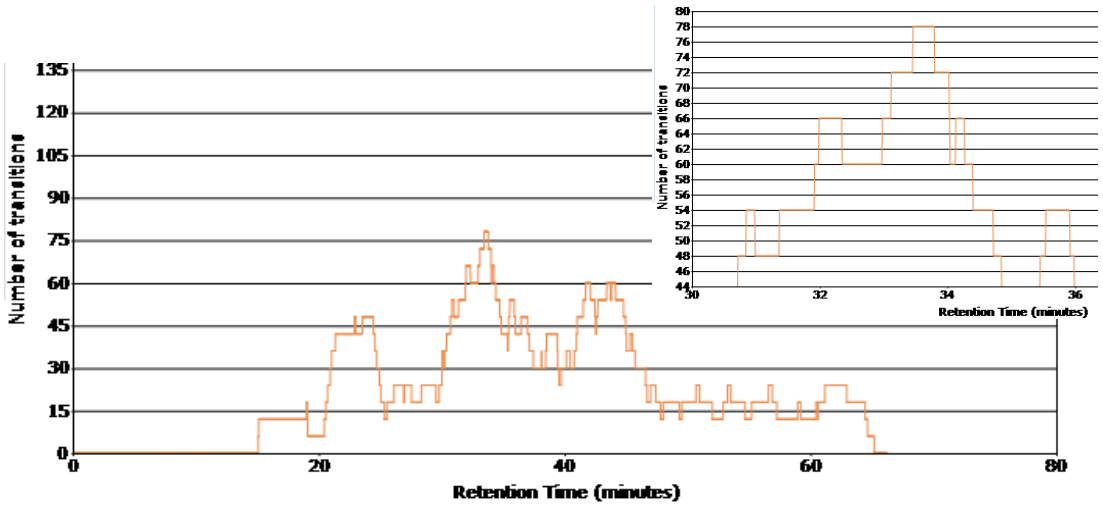


Figure 86 : Nombre de transitions en fonction du temps d'élution chromatographique. Un grossissement de la partie dense entre 30 et 36 min est montré dans l'encart en haut à droite.

III.1.2 Méthodes tPRM

Des triplicatas sont à nouveau obtenus avec les deux échantillons précédents sur le Q Exactive et le Q Exactive HF en mode *tPRM*. La méthode est constituée de 13 événements au total (1MS et 12 *tPRM*, détaillée en Tableau 20).

Le Q Exactive HF est capable d'enregistrer les spectres à plus haute fréquence que le Q Exactive avec une résolution équivalente. D'autre part, il est possible de travailler avec une résolution élevée à une même fréquence. Afin d'évaluer les capacités du nouvel instrument ces deux méthodes sont donc évaluées.

Tableau 20 : Comparaison des conditions de détection entre les deux laboratoires pour l'analyse en mode *tPRM*. La dernière colonne indique si les conditions sont strictement identiques.

	Q Exactive HF (Hemel, GF)	**Q Exactive HF (Hemel, GF)**	**QExactive (Paris, France)**	**Identique?**
méthode	*tPRM*	*tPRM*	*tPRM*	Oui
Gamme de masse MS	400 à 1500 Da	400 à 1500 Da	400 à 1500 Da	Oui
Résolution MS (à *m/z* 200)	30 000	30 000	35 000	Non
AGC MS	3e6	3e6	3e6	Oui
temps d'injection max MS	200ms	200 ms	120 ms	Non
Fenêtre d'élution	4 min	4 min	4 min	Oui
Fenêtre de sélection	2 *m/z*	2 *m/z*	2 *m/z*	Oui
Début de gamme de masse MS/MS	120 *m/z*	120 *m/z*	120 *m/z*	Oui
Résolution MS/MS (à *m/z* 200)	30 000	15 000	17 500	Non
AGC MS/MS	2e5	2e5	2e5	Oui
temps d'injection max MS/MS	60 ms	30 ms	60 ms	Oui
Energie collision normalisée	28	28	28	Oui
Exclusion dynamique	30s	30s	30s	Oui

III.2 Traitement des données

Les XICs sont extraits de la même manière que dans le chapitre 6.

À l'exception de l'expérience à 30 ms de temps d'injection sur le Q Exactive HF où une seule analyse a pu être réalisée. Les AUCs des peptides sont regroupés comme précédemment et la moyenne est reportée sous cartographie *heatmap* présentée à la Figure 87.

III.3 Résultats

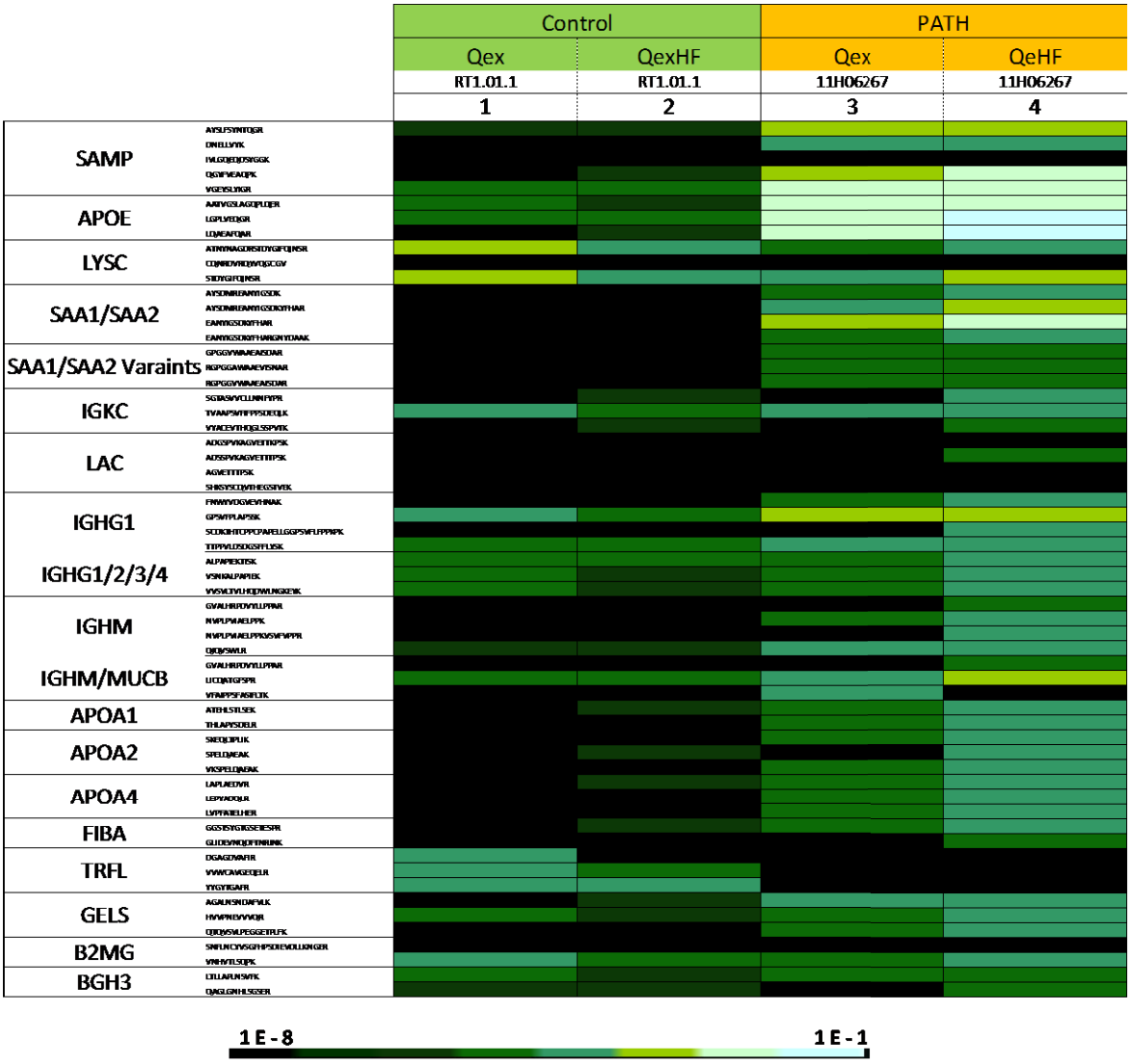


Figure 87 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans un échantillon de biopsie de rate témoin et patient. Les expériences protéomiques ciblées en *tPRM* sont comparées sur le Q Exactive (colonne 1 et colonne 3) et sur le Q Exactive HF (colonne 2 et 4). La couleur des cellules reflète l'intensité des signaux.

Ce *heatmap* confirme la tendance soulevée dans les paragraphes précédents. Quelques peptides supplémentaires sont détectés avec le Q Exactive HF (colonnes 2 et 4), ce qui est plus particulièrement visible dans le cas du témoin (colonne 2) par rapport au Q Exactive. Les expériences *tPRM*, plus sensibles que les expériences *shotgun*, mettent également en évidence chez le témoin (colonnes 1 et 2), de nouveaux les peptides issus des marqueurs fibrillaires (SAMP et APoE). Cependant, ces peptides sont détectés avec une abondance relative bien plus faible que chez le patient (colonnes 3 et 4), très en deçà des seuils observés chez les patients (en moyenne 10^{-6} chez les témoins contre 10^{-3} chez les patients). Le diagnostic est encore une fois sans équivoque, bien que d'autres protéines soient également présentes de

façon assez abondante : la protéine SAA est la protéine majoritaire détectée sur les deux instruments (après les deux marqueurs). Le patient présente bien une amylose de type AA.

Au niveau instrumentation, l'extraction du XIC de certains peptides dont un exemple est visible Figure 7 (peptide LQAEAFQAR appartenant à la protéine ApoE) montre qu'un nombre plus important de MS sont réalisées sur le Q Exactive (**Erreur ! Source du renvoi introuvable.** B) par rapport au Q Exactive HF (Figure 88 A). Cette particularité est favorable si la quantification est réalisée au niveau MS. Cependant pour plus de spécificité, la quantification est préférable au niveau MS/MS. La différence observée ne provient donc pas directement des spectromètres. La méthode employée est la même pour les deux instruments, c'est à dire un full MS couplé à 12 *tPRM* pour un temps d'acquisition du transient d'ions équivalent (64 ms). La distinction s'opère au niveau du logiciel installé sur le Q Exactive HF, dont la version gère différemment les événements. En effet, dans les deux cas, les 12 *tPRM* sont bien exécutés à la suite, mais lorsque moins de 12 ions sont élués dans la même fenêtre de temps, le Q Exactive HF va continuer les événements PRM en re-ciblant les ions déjà fragmentés. A l'inverse, dans le cas du Q Exactive, l'instrument repasse par un événement full MS. Ce réglage permet d'avoir une meilleure estimation des AGC mais est cependant au détriment du nombre de points des transitions. Cette particularité est visible sur les extractions des transitions du peptide montrées sur la Figure 88. Les spectres restent toutefois semblables et de bonne qualité. La nouvelle version sera implémentée au laboratoire prochainement.

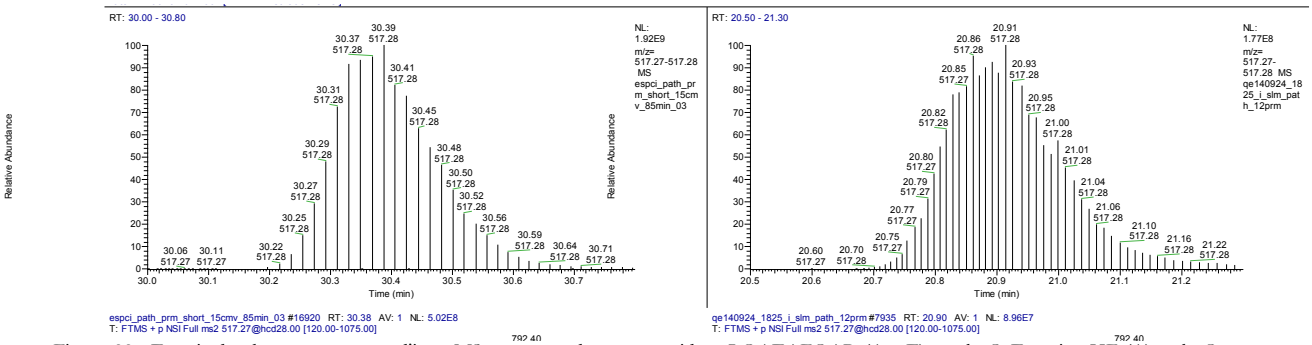


Figure 88 : Extrait du chromatogramme d'ions MS correspondant au peptide LQAEAFQAR (ApoE) sur le Q Exactive HF (A) et le Q Exactive (B). **Moins d'événements full MS sont réalisés sur le Q Exactive HF.**

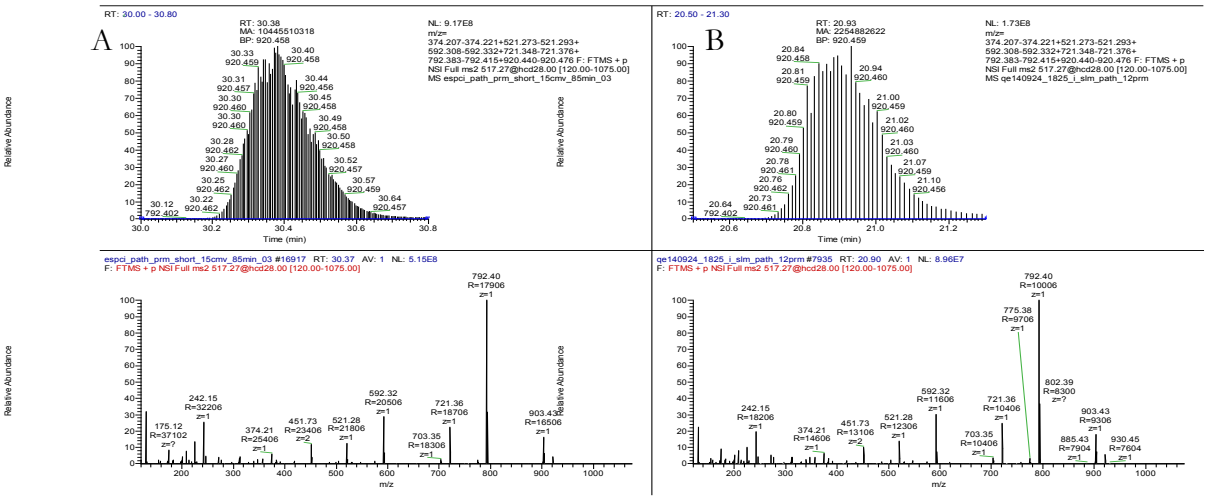


Figure 89 : Extrait du chromatogramme d'ions MS/MS correspondant au peptide LQAEAFQAR (ApoE) sur le Q Exactive HF (A) et le Q Exactive (B). **Moins de points sont mesurés pour les transitions sur le Q Exactive.**

Nous avons également évalué l'impact de la diminution de la résolution sur le Q Exactive HF. Pour que les événements d'accumulation des d'ions N+1 dans la C-Trap et détection des fragments des peptides N dans l'Orbitrap se passent en parallèle, le temps d'injection des ions ne doit donc pas dépasser 30 ms pour une résolution de 15000 (à m/z 200). Cela signifie que le temps donné aux ions pour remplir l'AGC est plus faible. Or, pour l'expérience précédente, lorsqu'on suit la distribution du temps d'injection des d'ions au cours de l'élution chromatographique (Figure 90), nous pouvons remarquer que le temps maximum autorisé, ici 60 ms, est souvent atteint. La C-Trap est donc peu souvent remplie avec le nombre de charges maximales défini par l'AGC dans le Q Exactive HF. S'il on tient compte uniquement de cette particularité, il faut vérifier que la méthode à 30 ms permet d'avoir un remplissage optimal de la C-Trap pour être aussi sensible que lors des analyses à 60ms sur le Q Exactive.

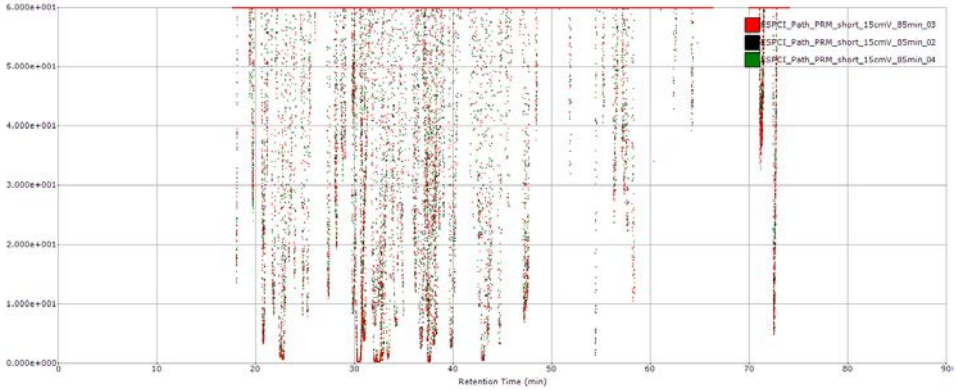


Figure 90 : Distribution du temps d'injection des ions dans le Q Exactive HF au cours de l'élution chromatographique. **En général, le seuil d'AGC n'est pas atteint. Le temps d'injection est souvent toujours au maximum autorisé.**

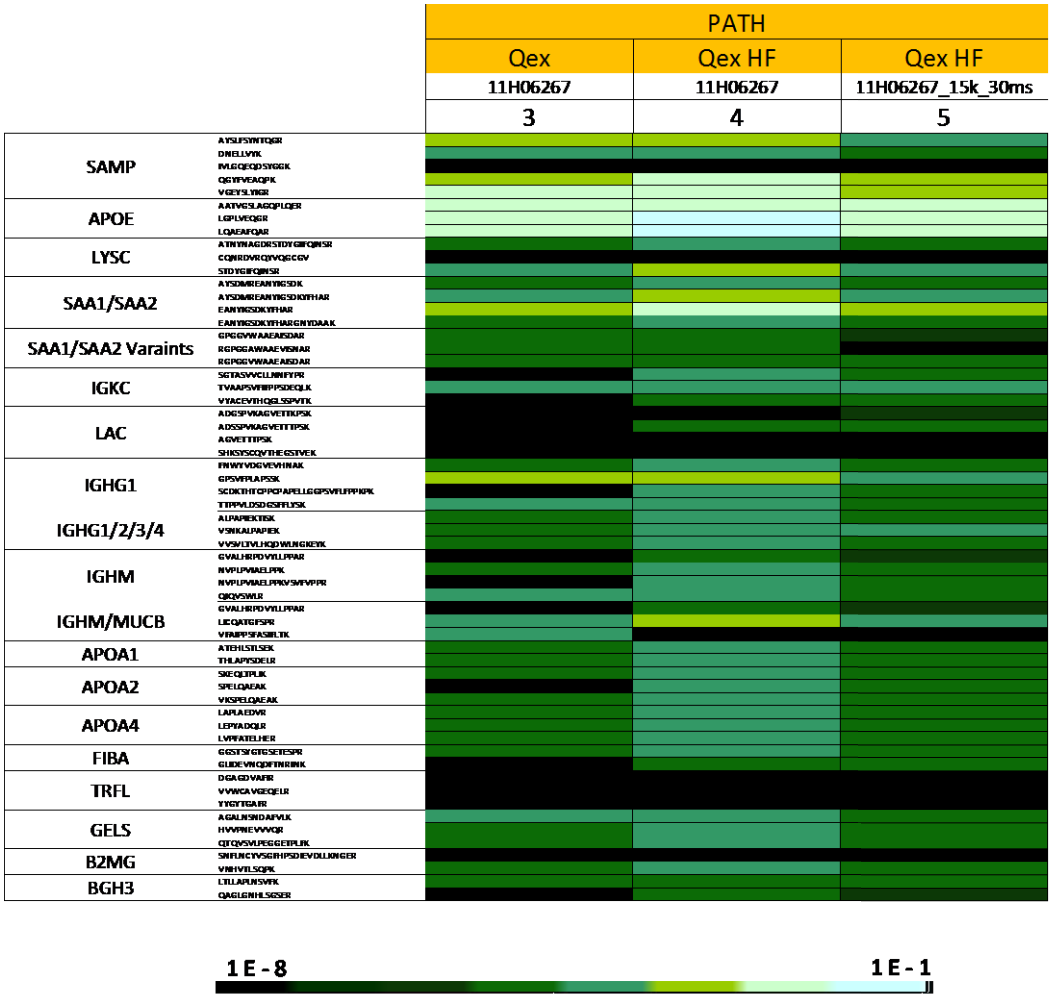


Figure 91 : Visualisation sous forme de cartographie heatmap des moyennes d'aires normalisées des peptides quantifiés dans un échantillon de biopsie de rate patient. Plusieurs expériences sont comparées : les expériences protéomiques ciblées en *tPRM* sur le Q Exactive (colonne 3) et sur le Q Exactive HF (colonne 4 et 5). La couleur des cellules reflète l'intensité des signaux.

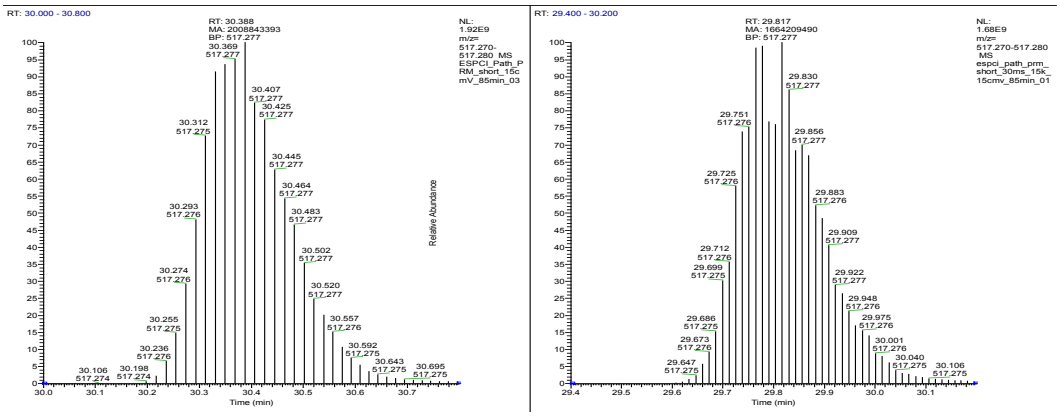


Figure 92 : Extrait du chromatogramme d'ions MS correspondant au peptide LQAEAFQAR (ApoE) acquis sur le Q Exactive HF à 30 000 de résolution en MS/MS (à gauche) et à 15 000 (à droite). Diminuer la résolution permet d'obtenir plus de point pour un même nombre de peptides cibles. L'aire du pic est faiblement impactée. L'enveloppe du pic est définie de manière moins robuste à 15000 vraisemblablement parce que le taux de remplissage n'est pas correct.

Globalement, les peptides détectés avec un temps d'injection maximale de 30 ms sont légèrement moins intenses que ceux détectés avec un temps de 60 ms sur le Q Exactive HF (montrés sur le heatmap à la Figure 91 et les XIC sur la Figure 92). Cependant, plus de peptides sont détectés de façon significative sur le Q Exactive HF (colonne 5) que sur le Q Exactive (colonne 3). Ceci peut s'expliquer par l'amélioration de la transmission des ions sur le Q Exactive HF. Le temps mis pour remplir la C-Trap d'un même nombre de charge sera alors plus faible que sur le Q Exactive. La diminution du temps d'accumulation des ions (30 ms) sur le Q Exactive HF n'est donc pas préjudiciable à la qualité des acquisitions par rapport au Q Exactive de première génération avec 60 ms d'accumulation.

Plus de peptides peuvent être ainsi ciblés sur le Q Exactive HF sans sacrifier la sensibilité par rapport au Q Exactive en maintenant une résolution proche. D'autre part, tout en gardant une bonne sensibilité (60 ms d'accumulation), il est possible d'augmenter la résolution des fragments MS/MS et ainsi augmenter la spécificité de la méthode sur le Q Exactive HF.

*IV. **Conclusion***

Les analyses *shotgun* réalisées sur le Q Exactive et le Q Exactive HF montrent peu de différences.

En analyse ciblée en mode PRM, les deux instruments fournissent des résultats assez similaires avec des intensités plus élevées pour le Q Exactive HF pour un même temps d'injection des ions. Cela va permettre alors de travailler à des résolutions supérieures et une spécificité accrue. Ce nouvel instrument permet également d'accroitre le nombre de peptides cibles à sensibilité et résolution égale à l'ancienne génération.

Pour confirmer ces résultats, une dilution des échantillons dans une matrice est nécessaire pour estimer la limite de détection et quantification des deux spectromètres de masse.

Conclusion générale

Au cours de ce projet nous avons pu mettre en œuvre une méthodologie complète pour la caractérisation de tissus amyloïdes. Les échantillons de biopsies ont été traités avec une méthode originale employant une sonde à ultrasons mise directement en contact avec l'échantillon. Elle délivre des jets intenses de liquide qui permettent de casser les tissus, qu'ils soient fixés au formol et inclus en paraffine ou frais et conservés à -80°C. Ce qui permet de récupérer les protéines en quelques secondes, voire quelques minutes, sans avoir recours au chauffage des échantillons à haute température pendant une durée non négligeable (souvent 90 min) pour le déparaffinage par exemple. Les ultrasons permettent également d'accélérer les réactions telles que la digestion enzymatique nécessaire pour générer des peptides. La protéolyse enzymatique se fait en quelques minutes à l'aide des ultrasons au lieu de 5 heures conventionnellement. Les peptides sont générés de façon reproductibles et en accord avec les données disponibles sur Peptide Atlas.

Les peptides sont ensuite séparés par nano-chromatographie en phase liquide à polarité de phases inversée. Une optimisation a été réalisée en amont des analyses pour augmenter le pourcentage de couverture du protéome. Ainsi, pour cette application biologique, un gradient d'élution chromatographique de 3h est le plus adapté. Bien que la séparation chromatographique soit plus résolutive sur une colonne de 50 cm, nous avons opté pour une colonne classique de 15 cm, pour transposer aisément la méthode en clinique. En effet, un nombre conséquent d'échantillons est amené à être analysé menant à des changements de colonne réguliers. Sa mise en place et son conditionnement doivent être rapides et simples. Les peptides sont ensuite analysés en spectrométrie de masse selon deux méthodes différentes : l'analyse globale appelée couramment *shotgun proteomics* et l'analyse ciblée appelée *targeted proteomics*.

La première méthode protéomique consiste à caractériser de façon la plus exhaustive possible le protéome d'un organisme donné pour trouver des biomarqueurs de la maladie des amyloses. Pour cela, un spectromètre de masse à haute résolution est indispensable. Aujourd'hui, deux types d'analyseurs sont disponibles : la cellule à résonance cyclotronique d'ions (ICR) et le piège orbitalaire électrostatique (Orbitrap). Ces deux instruments utilisent la transformée de Fourier pour convertir le signal de l'espace temporel vers l'espace des fréquences, pour en déduire un spectre de masse en rapport m/z.

La quantification relative des protéines a été réalisée sans *a priori* par la méthode T3PQ (*top 3 protein quantification*) en prenant en compte la moyenne des aires sous la courbe des trois peptides les plus intenses. Les résultats ont été montrés sous deux visualisations différentes: une visualisation échantillon par échantillon en forme de camembert montrant une quantification relative des protéines amyloïdogènes; et une visualisation globale sous forme de cartographie *heatmap* des peptides spécifiques aux protéines d'intérêt. Différents traitements de données ont été évalués. Dans tous les cas, les analyses *shotgun* des échantillons de biopsies permettent d'obtenir un diagnostic de typage de la maladie des amyloses sur différents organes (rein, rate, glandes salivaires accessoires, poumon, testicule, humeur vitrée). Les résultats obtenus sont conformes aux données immunohistologiques quand elles sont disponibles. L'analyse *shotgun* permet en outre, de faire une analyse des protéines sans *a priori* pour ainsi mettre en évidence des amyloses peu fréquentes et des dépôts secondaires, pour lesquels le pathologiste ne dispose bien souvent pas des anticorps. Cependant, lorsque le pourcentage de tissus amyloïdes est inférieur à 50% dans un prélèvement, une ambigüité de diagnostic peut avoir lieu. Pour confirmer le diagnostic, une microdissection par capture laser (LCM) est souhaitable. Bien que ces analyses *shotgun* nécessitent des investissements plus lourds, plutôt adaptés à la Recherche qu'aux applications en Clinique, elles sont indispensables pour caractériser les isoformes des protéines, pour découvrir de nouvelles protéines amyloïdogènes ou isoformes de protéines amyloïdogènes déjà connues.

Cependant, les analyses *shotgun* ne permettent pas toujours de faire une quantification spécifique des protéines dans les matrices biologiques. C'est pour cela qu'une approche complémentaire, ciblée, plus sensible et spécifique, a été développée au cours de mes travaux de thèse. Deux technologies sont actuellement disponibles: le mode SRM (*selected reaction monitoring*) utilisé couramment dans les départements cliniques, et le mode à haute résolution PRM (*parallel reaction monitoring*) proposé aujourd'hui sur les spectromètres de masse de nouvelle génération. Pour augmenter le nombre d'espèces ciblées, les temps de rétention chromatographiques ont été pris en compte (*time-scheduled SRM* ou *time-scheduled PRM*).

La première approche (SRM) est développée sur des instruments de type triple-quadripôles, très robustes et moins onéreux. Pour que la méthode soit la plus sensible possible, des optimisations sont réalisées une fois pour toutes les analyses : au niveau instrumental (tension appliquée, température du capillaire); mais également au niveau des peptides mêmes (choix des peptides, choix des transitions, énergies de collision pour chaque transition). La quantification des protéines a été effectuée comme pour les données *shotgun*, en prenant en compte soit uniquement les trois peptides les plus intenses, soit tous les peptides

identifiés. Les deux méthodes montrent des résultats similaires avec un diagnostic sans ambigüité du type d'amylose. A l'origine, le traitement des données en sélectionnant les trois peptides les plus intenses était manuel. Dans un souci de transfert d'une méthode simple en clinique avec le minimum de retraitement de données, nous avons privilégié la prise en compte automatique de la somme de toutes les aires des peptides, qui est directement fournie par traitement informatique (Pinpoint ou Skyline), dans la suite des analyses. Les résultats obtenus en SRM ont été comparés aux résultats *shotgun* après traitement sur deux logiciels : Proteome Discoverer, où l'identification MS/MS est nécessaire et Pinpoint, qui se base uniquement sur la masse des précurseurs en mode MS. Cette comparaison montre l'apport de la SRM en termes de spécificité de la méthode. Elle montre également le défaut de Proteome Discoverer qui quantifie peu de peptides. Si l'extraction des chromatogrammes d'ions (XIC) par Pinpoint permet d'augmenter le nombre de peptides quantifiés, il est cependant nécessaire de disposer d'un spectromètre à haute résolution et bonne exactitude de masse pour augmenter la robustesse de la quantification. Les mêmes peptides ont également été suivis en *time-scheduled PRM*, sur l'appareil de type quadripôle-Orbitrap. Le résultat est très similaire au cas du SRM. Etant donné que la mise en place des expériences PRM demande moins d'optimisation, tous les échantillons à disposition ont été analysés avec ce nouveau mode d'acquisition.

Les résultats montrent que les expériences PRM sont plus sensibles que les expériences *shotgun*. En effet, plus de peptides ont été détectés de manière significative chez les patients et les témoins. Les deux protéines biomarqueurs des amyloses, les protéines SAMP et ApoE, sont détectées chez les patients mais également chez les témoins, bien que présentes en quantité moins abondante. Une analyse en composantes principales (ACP) des données PRM permet de séparer la population en deux groupes bien distincts : les patients souffrant d'une amylose et les patients et/ou témoins non atteints par cette pathologie. Cette séparation n'a pas pu être réalisée à partir des données *shotgun*. L'ACP ne permet pas en revanche de différencier les classes d'amyloses. Une clustérisation hiérarchique a alors été évaluée mais nous avons abouti au même résultat. En effet ici, nous manquons de point de mesure pour effectuer une réelle analyse statistique. Il faudrait analyser plus d'échantillon de patients atteints de différents types d'amyloses pour avoir une population statistiquement représentative de chaque type amyloïde.

Pour suivre le plus de peptides possible, nous avons établi une méthode qui prend en compte les temps de rétention, *tPRM*. Or, nous avons vu que ces temps peuvent changer d'une analyse à l'autre. Pour que la méthode soit robuste pour une application clinique, il faut

aligner ces temps de rétention entre les expériences *tPRM*, qui peuvent être espacées dans le temps sur des mois. Pour cela, une solution simple est d'analyser un mélange peptidique avant toute expérience ciblée et au milieu de série. Ce standard permet également de contrôler le système analytique (contrôle qualité primordial en clinique). Le déclenchement de l'acquisition MS pourra alors être modulé en fonction des variations du temps de rétention de ces peptides standards. Une autre méthode, plus onéreuse, est l'ajout de peptides marqués par des isotopes stables qui sont élués tout au long du gradient chromatographique et permettent de corriger les temps de rétention en temps réel. Les peptides se trouvant avant le premier peptide marqueur auront un intervalle de temps de suivi plus large que les suivants. Nous pourrions aussi faire du multiplexage des échantillons en choisissant des transitions spécifiques pour chaque protéine d'intérêt. Cependant, le multiplexage se fera au détriment de la sensibilité. En effet, la durée d'accumulation des ions multiplexés doit être inférieure au temps d'acquisition du transient d'ions. Plus le nombre de multiplexage est grand, plus le nombre de peptides suivis par cycle sera grand mais la sensibilité de la méthode sera moindre. L'injection des ions dans la C-Trap se faisant séquentiellement, le temps propre à chaque type d'ion peptidique sera donc divisé par le nombre total de peptides multiplexés. Par contre, pour un même nombre de peptides qu'en simplex, le multiplexage permet d'avoir une plus haute résolution.

Nous avons enfin comparé nos résultats obtenus sur le Q Exactive à son homologue de nouvelle génération, le Q Exactive HF. Globalement les résultats sont presque similaires avec une sensibilité légèrement supérieure avec le nouvel instrument. Il permet d'augmenter la spécificité de la méthode grâce à une résolution accrue par un Orbitrap plus petit et un champ électrostatique plus intense. Du fait de l'amélioration des optiques et du quadripôle, la transmission des ions est améliorée et permet de diminuer le temps d'accumulation des ions sans perte de sensibilité par rapport à la première génération. Un nombre plus élevé de peptides peut alors être suivi, mais ceci a un coût.

Aujourd'hui émerge une alternative au PRM, le DIA (*data independant analysis*) ou SWATH™ (ABSciex). Dans ce cas, les peptides sont isolés et fragmentés par paquet de petites fenêtres de masse (en général 10 à 25 Da) et tous les fragments sont analysés dans l'Orbitrap. Cette expérience n'a cependant pas pu être mise en place au laboratoire, compte tenu du temps imparti. La méthode sera moins sensible puisqu'il n'y a pas de filtrage de précurseurs, mais elle permet d'augmenter le nombre de peptides cibles et de quantifier *a posteriori* une nouvelle protéine qui ne serait pas dans la liste des protéines conventionnelles.

Pour cibler spécifiquement les isoformes des protéines amyloïdogènes, une autre banque ciblée peut être réalisée. Pour cela, l'emploi des peptides standards, tels que les peptides AQUA, seront indispensables pour confirmer les isoformes détectés.

À l'avenir, il est possible d'envisager que les hôpitaux seront équipés de spectromètres de masse de type quadripôle-Orbitrap, dits de "paillasse", et auront la possibilité d'alterner facilement entre analyses *shotgun*, DIA et PRM sans changer d'instrument. Les résultats obtenus sont encourageants et les méthodes mises en place au laboratoire sont facilement transférables dans un laboratoire clinique. De plus, la méthode ciblée en mode PRM, sensible et spécifique, nécessite peu d'optimisation par rapport à la méthode classique en mode SRM. Ce mode d'acquisition sera de plus en plus utilisé tant par les chercheurs que par les praticiens hospitaliers. Les logiciels évalués pour le retraitement des données au cours de cette thèse sont pour la plupart libres de droits et peuvent facilement être utilisés sur n'importe quel système informatique.

Le protocole mis en place au cours de ma thèse est dorénavant intégré dans le processus du diagnostic clinique des amyloses. Nous sommes convaincus que la spectrométrie de masse peut être un outil diagnostique clinique de référence pour d'autres maladies.

Annexes

Annexe 1: Principe de l'électrospray

Contrairement au MALDI, en ionisation électrospray (ESI), l'échantillon sous forme liquide est introduit à **pression atmosphérique** dans un capillaire métallique avec un orifice de l'ordre de 100µm. Entre ce capillaire placé entre 1 à 3 cm de distance d'une contre-électrode, est appliquée une différence de potentiel induisant un **fort champ électrique,** typiquement entre +/- 3 et +/- 5 kV selon la polarité. Le processus d'ESI se définit en trois étapes: (1) formation des gouttelettes chargées, (2) évaporation/fissions des gouttelettes chargées en plus petites gouttelettes puis (3) formation de l'ion en phase gazeuse. Les ions formés sous pression atmosphérique sont alors focalisés puis accélérés à travers une série de lentilles et multipôles vers l'analyseur où règne un vide poussé (Figure 93).

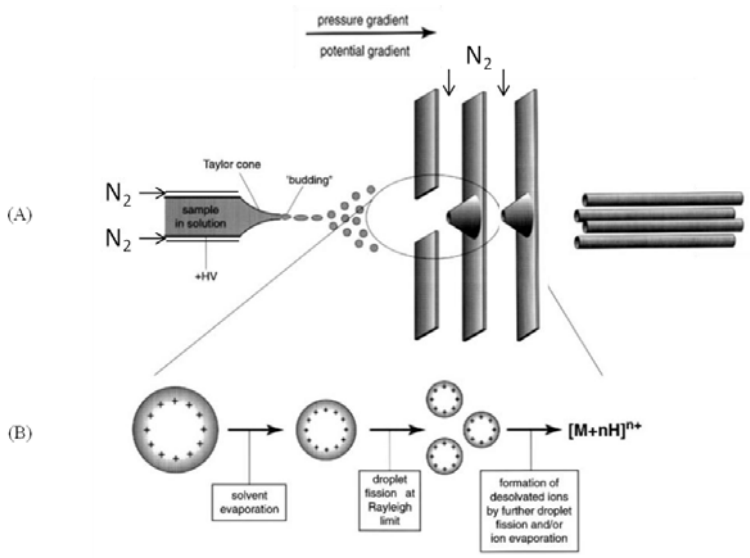


Figure 93: Processus d'ionisation en électrospray**. Schéma adapté de Gaskell [16]**

I. Formation des gouttelettes chargées

En 1991, le groupe de Kebarle [180], démontre qu'en plus des charges apportées dans la solution d'analytes, généralement acide, un apport non négligeable de protons a lieu à l'interface du métal et de la solution dans le capillaire par oxydation du solvant comme l'eau (Equation 23), le méthanol (Equation 24) ou l'acétonitrile (Equation 25) très utilisés en ESI. L'excès de protons généré va alors se fixer sur les sites protonables des analytes [180]. Une réaction de réduction se produit alors au niveau de la contre électrode pour contre balancer l'excès de charges afin d'assurer une production continue de gouttelettes chargées [181]. Cette propriété de cellule électrochimique de l'ESI peut être exploitée pour augmenter le signal de

l'analyte protoné en reliant également l'aiguille en métal de la seringue à la terre [182]

Oxydation de l'eau : $\mathbf{2\ H_2O \rightarrow 4\ H^+ + 4\ e^- + O_2}$ (Equation 23)

Oxydation du méthanol : $\mathbf{CH_3OH \rightarrow HCHO + 2H^+ + 2\ e}$ (Equation 24)

Oxydation de l'acétonitrile : $\mathbf{2\ CH_3CN \rightarrow NCCH_2CH_2CN + 2\ H+ + 2\ e^-}$ (Equation 25)

La séparation des ions dans ce capillaire d'ESI est essentiellement due à un mouvement électrophorétique [183], c'est à dire que les ions positifs et négatifs vont migrer dans des directions opposées. Les ions de même polarité que le champ électrique migrent vers la contre-électrode, soit vers l'extrémité du capillaire. L'accumulation de toutes ces charges à l'extrémité déstabilise la surface du liquide qui va alors former un cône dynamique à la sortie du capillaire, appelé cône de Taylor. Lorsque la force électrostatique est suffisamment élevée pour dépasser la tension de surface, le cône s'allonge formant un filament liquide (jet) qui se rompt sous forme de gouttelettes (nébulisât ou « spray ») très chargées [180,184,185]. L'introduction d'un gaz neutre, généralement de l'azote, coaxialement au capillaire (« *sheath gas* ») peut grandement faciliter la nébulisation. Cette aide est dorénavant incorporée dans tous les appareils commerciaux. Son intérêt varie selon le débit de la solution, composition du solvant et la polarité du champ électrique appliqué. À ce stade, le diamètre des gouttelettes est de l'ordre du μm [184] et varie en fonction de nombreux paramètres tels que le champ électrique ou la propriété du solvant [186]. Ces gouttelettes vont ensuite traverser un gradient de champ électrique et un gradient de pression dans la direction de l'analyseur du spectromètre de masse (Figure 93A).

I.1 Évaporation/fission des gouttelettes chargées

Deux forces dans ces gouttelettes s'y opposent: **la tension de surface** qui permet de retenir ces charges sous forme de gouttelettes rondes et **la force de répulsion électrostatique de Coulomb** de ces charges à la surface qui essaye de déstabiliser la forme sphérique [184]. Lorsque ces deux forces se compensent, la limite de stabilité de Rayleigh (Equation 26) est atteinte.

Équation de Rayleigh : $$q_R = 8\pi\ \sqrt{\varepsilon_0 \gamma R^3}$$ (Equation 26)

où ε_0 est la perméabilité, γ est la tension de surface et R *le rayon de l'analyte.*

Pendant le transport vers l'analyseur, la taille des ces gouttelettes diminue progressivement par évaporation du solvant tout en gardant une charge constante. L'évaporation peut être

facilitée par un autre flux gazeux (« *curtain gas* ») ou en chauffant l'interface [187]. En diminuant leur taille jusqu'à un rayon légèrement supérieur à la limite de Rayleigh, la densité de charge augmente jusqu'à ce que les forces de répulsions électrostatiques coulombiennes excèdent la tension superficielle et provoquent une explosion coulombienne formant des gouttelettes secondaires plus petites. Ce phénomène est communément appelé fission coulombienne à la limite de Rayleigh.

Les travaux de Rayleigh en 1882 [188] puis Gomez et Tang en 1914 [189] montrent que les gouttelettes primaires ne se séparent pas en deux gouttelettes secondaires plus petites de charges et de tailles quasi égales mais qu'elles vibrent alternativement de la forme aplatie (oblate) à allongée (prolate) créant l'émission d'une vingtaine de gouttelettes secondaires beaucoup plus petites et monodisperses ($1/10^{ème}$ du rayon de la gouttelette mère) similaire au cône de Taylor, illustrée sur des photographies flashs en Figure 94. La masse de la totalité des ces gouttelettes n'excède pas 2% de la gouttelette mère. En outre elles ne possèdent que 15% de sa charge. Les nouvelles gouttelettes formées ont alors un ratio masse sur charge élevé où la répulsion coulombienne est plus petite, la surface étant plus grande [184]. Le phénomène d'évaporation se reproduit jusqu'à que ces gouttelettes secondaires atteignent la limite de Rayleigh et subissent à leur tour une fission. Ces étapes évaporation/fission se succèdent pour créer des gouttelettes de plus en plus petites jusqu'à atteindre une taille de l'ordre du nm (Figure 93B).

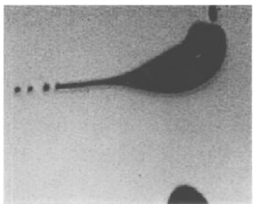


Figure 94: Shadowgraphes flash illustrant la fission coulombienne des gouttelettes d'heptane. La direction du champ électrique est du haut vers le bas. Illustration de photographie flash proposée par l'équipe de Gomez [189].

I.2 Formation d'ions en phase gazeuse

Après ces cycles successifs d'évaporation et fission des gouttelettes, les ions en phase gazeuse sont générés et guidés vers la contre électrode. Malgré un large intérêt de la communauté scientifique pour l'ESI et le nombre de publications croissant sur le sujet, le processus d'émission de ces ions en phase gazeuse reste encore très discuté. Différents mécanismes sont admis pour les protéines sous leur forme native (résidu chargé [190]), les

protéines dépliées (éjection [191]) et petites molécules (évaporation d'ion [192]) et sont abordés dans les paragraphes suivants.

I.3 Modèle du résidu chargé

Le groupe de Dole [190] propose un modèle dans lequel la succession d'évaporations/fissions mène à la dernière gouttelette très chargée contenant plusieurs charges mais une seule molécule d'analyte. Les charges sont alors retenues à la dernière étape de l'évaporation du solvant résiduel par la molécule analyte, ce qui permet l'obtention d'un ion multichargé en phase gazeuse (Figure 95). Ce mécanisme est appelé le modèle du résidu chargé (CRM *charged residue model*).

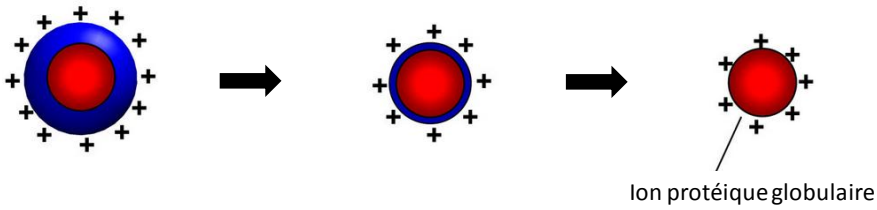


Figure 95 : Modèle d'évaporation de l'ion proposé par le groupe de Dole pour les macromolécules globulaires **(figure adapté de Konermann et al.** [193]**)**

Le groupe de Mora [194] a démontré en 2000 que pour des protéines globulaires natives suivant ce modèle du résidu chargé, le nombre de charges (ou état de charge) observé en ESI est alors très proche de celui déterminé par la limite de Rayleigh (Equation 26) pour une gouttelette d'eau de même dimension (la valeur de masse volumique de la protéine est ici considérée comme proche de celle de l'eau (ϱ=1g.cm^{-3})). Ce modèle permet alors de corréler la masse de l'analyte et la moyenne des états de charges observés (z_{av}) dans les spectres de masse en ESI de l'analyte (Figure 96).

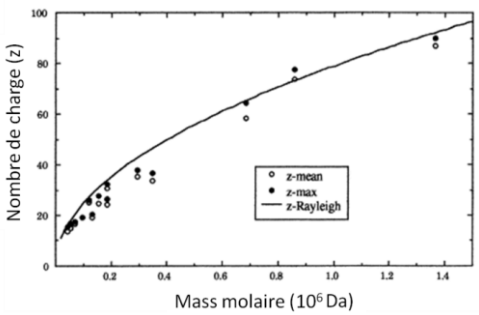


Figure 96 : Etat de charge maximum et moyen en fonction de la masse de différentes protéines qui suivent le modèle du résidu chargé comparé avec la limite de Rayleigh avec ϱ=1g.cm^{-3} et γ=0.072 N.m^{-1} (figure adapté de **Fernandez de la Mora** [194]**).**

I.4 Modèle d'évaporation de l'ion

Quant au groupe de Iribane et Thomson [191], il propose un second modèle appelé l'évaporation d'ion (IEM *ion evaporation model*). Ce modèle stipule que le champ électrostatique émanant de la surface des petites gouttelettes très chargées, issues des explosions coulombiennes successives, est suffisamment élevé pour que des ions solvatés soient éjectés en phase gazeuse. Ce phénomène est dû à la répulsion entre l'ion chargé et les autres charges de la surface (Figure 97). Ce mécanisme se produit quand le rayon des gouttelettes est inférieur à 10 nm où l'émission d'ion domine la fission colombienne pour de petites molécules organiques et inorganiques. Contrairement au modèle du résidu chargé (CRM), ce modèle ne requiert pas des gouttelettes très petites (R≈1nm) où il n'y a qu'une seule molécule d'analyte. La constante de temps d'émission d'un ion peut alors être donnée par l'équation suivante :

Équation d'Iribane et Thomson : $$k_1 = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^0}{RT}\right)$$ (Equation 27)

où k_B est la constante de Boltzmann, T est la température, h est la constante de Planck et ΔG^0 est l'énergie d'activation libre.

Or l'énergie d'activation libre est dépendante de plusieurs paramètres tels que l'attraction entre l'ion éjecté et le solvant, la répulsion coulombienne entre l'ion éjecté et les charges résiduelles sur la surface ou encore l'énergie de désolvatation de l'ion. Ce modèle ne permet pas de prédire l'intensité des ions sur un spectre ESI puisqu'il dépend des paramètres physicochimiques des analytes et des solvants utilisés.

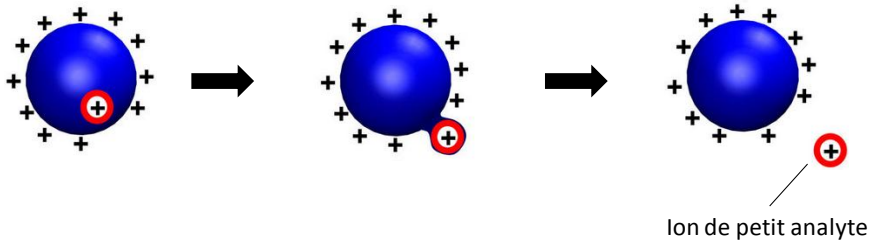


Figure 97 : Modèle d'évaporation de l'ion proposé par le groupe de Iribane et Thomson pour les petites molécules **(figure adapté de Konermann *et al.*[193]).**

I.5 Modèle d'éjection de la chaine hydrophobe

Récemment un dernier mécanisme semblable au modèle d'évaporation d'ion (IEM) pour les protéines dépliées a été proposé par le groupe de Konermann [192], le model d'éjection de la chaine hydrophobe (CEM *chain ejection model).* Alors que dans un solvant aqueux neutre la protéine adopte une forme globulaire compacte (forme native) où la majorité des charges et résidus polaires sont à l'extérieur pour maximiser les interactions avec l'eau, dans un

environnement acide tel que les solvants utilisés en chromatographie, les résidus hydrophobes séquestrés à l'intérieur du noyau deviennent dorénavant accessible au solvant, rendant défavorable la présence de la protéine dans la gouttelette. Lors de la formation des nano-gouttelettes, la chaine dépliée se met immédiatement à la surface de la gouttelette entrainant l'éjection de la chaine hydrophobe. Séquentiellement, la protéine entière est entrainée dans la phase gazeuse (Figure 98). La protonation élevée des ions est due à l'étape d'équilibration des charges montrée sur l'image centrale en Figure 98. La distribution des charges d'une protéine dépliée est alors plus grande que sous sa forme native qui suit le modèle du résidu chargé.

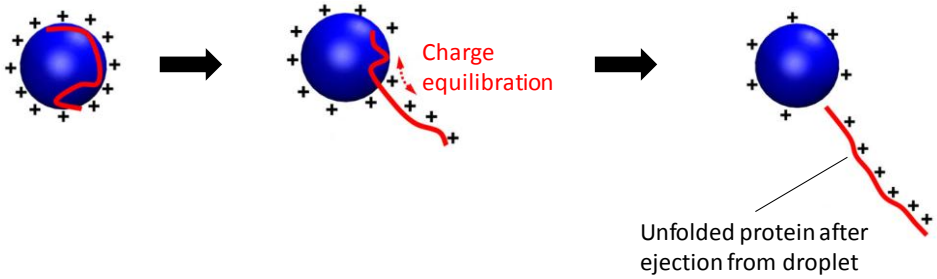


Figure 98: Modèle de l'éjection de chaine proposé par le groupe de Konermann pour les protéines dépliées **(figure adapté de Konermann *et al.*[193]).**

II. *Spectres de masse de protéine en électrospray*

À l'issu du processus, des ions mono et multichargés (dits multiprotonés) de type $[M+nH]^{n+}$ sont obtenus en ESI ; M correspond à la masse moléculaire de la molécule analysée, et n au nombre de charges portées par cette molécule ionisée. L'obtention des ions multichargés permet alors d'analyser des molécules de haut poids moléculaire telles que les protéines sans disposer d'analyseur à gamme de balayage élevée. L'utilisation de différents outils informatiques, permet ensuite de déterminer très simplement, la masse M, par déconvolution du spectre en un spectre intensité/masse de la protéine.

Une déconvolution mathématique peut également se faire à partir du spectre. En effet sur le spectre de masse, les ions multichargés se traduisent par une enveloppe comportant plusieurs pics isotopiques (Figure 100). Ces pics correspondent à une distribution statistique des charges qui dépend de la stabilité des ions multiprotonés en phase gazeuse de l'espèce moléculaire intacte. Cette distribution varie selon plusieurs paramètres tels que la conformation de la protéine, le pH de la solution, la composition du solvant ou encore les paramètres du spectromètre [192,193,195,196].

La masse moléculaire peut être calculée en prenant en compte deux pics consécutifs de rapports m/z mesurés sur le spectre m_1 et m_2, avec $m_2 > m_1$ (Figure 99). En résolvant le

système d'équations à 2 inconnues M et z (Equation 28), sachant que la différence d'état de charge entre les pics est d'environ 1 Da (Equation 29), la masse moléculaire de la protéine est calculée à partir de chaque pic (Equation 30) et (Equation 31). M doit ensuite être calculée pour chaque pic de la série. Le résultat sera la moyenne de toutes les M calculées avec une précision correspondant à la déviation standard. Plus il y a d'ions multichargés et plus l'appareil à disposition est précis, plus la masse pourra être mesurée avec précision.

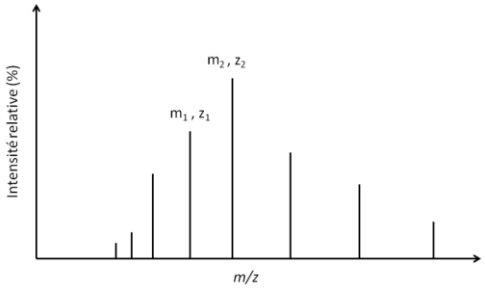


Figure 99: Simulation d'un spectre de masse d'une protéine multichargée en ESI.

$$m_1 = \frac{M + z_1 m_H}{z_1} \qquad m_2 = \frac{M + z_2 m_H}{z_2} \qquad \text{(Equation 28)}$$

avec $m_H \approx 1$, masse monoisotopique de l'hydrogène

$$z_2 = z_1 - 1 \qquad \text{(Equation 29)}$$

$$z_1 = \frac{m_2 - 1}{m_2 - m_1} \qquad \text{(Equation 30)}$$

$$M = (m_1 - 1) \times \frac{(m_2 - 1)}{(m_2 - m_1)} \qquad \text{(Equation 31)}$$

Avec un analyseur à haute résolution, le calcul de la masse moyenne de protéines peut être plus rapide. En effet, la distribution isotopique de chaque pic multichargé sera accessible. Sachant que le défaut de masse apporté par la présence d'un isotope est d'environ 1 Da, la différence de masse est égale à $\Delta m/z$ soit $1/z$ entre deux pics d'un même profil isotopique (zoom sur la Figure 100). Les états de charges de chaque profil isotopiques en découlent alors, ainsi que la masse moyenne de la protéine.

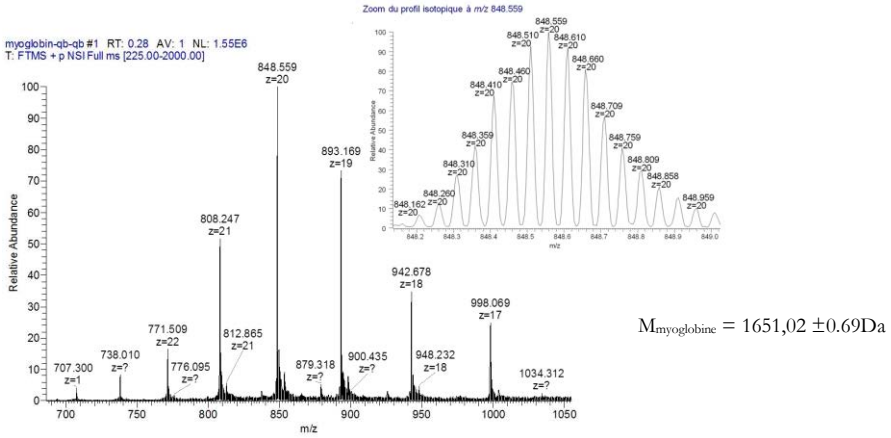


Figure 100: Exemple d'un spectre de masse d'une protéine multichargée (myoglobine à 10pmol/µl dans 50/50 H20/ACN+0.1% acide formique, masse théorique = 1651,499Da) obtenue sur un spectromètre de type FTICR.

Annexe 2 : Fragmentations des peptides

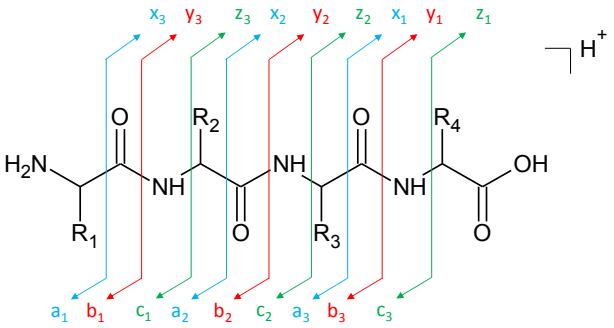


Figure 101: Nomenclature de la fragmentation de la chaine peptidique pour un tétrapeptide.

La rupture d'une liaison dans la chaine peptidique peut avoir lieu à trois endroits différents décrits à la Figure 27. La numérotation se fait quant à elle dans l'ordre croissant à partir du coté terminal.

- ✓ en **C_α-CO**, donnant les ions **a_n** si la charge est retenue du coté N-terminal et **x_n** si la charge est retenue du coté C-terminal, (en bleu sur la Figure 27)

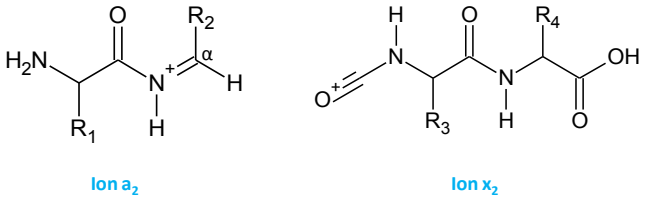


Figure 102: Ions fragments a_2 et x_2 issus de la rupture C-CO.

- ✓ en **CO-NH**, liaison amide peptidique, donnant les ions **b_n** si la charge est retenue du coté N-terminal et **y_n** si la charge est retenue du coté C-terminal, (en rouge sur la Figure 27)

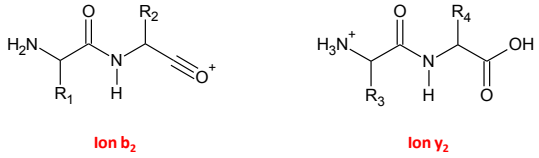


Figure 103: Ions fragments b_2 et y_2 issus de la rupture de la liaison peptidique CO-NH.

- ✓ en **NH-C$_\alpha$H**, donnant les ions **c$_n$** si la charge est retenue du coté N-terminal et **z$_n$** si la charge est retenue du coté C-terminal. (en vert sur la Figure 27)

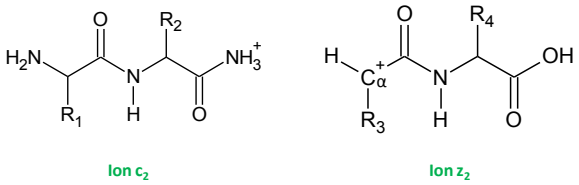


Figure 104: Ions fragments c_2 et z_2 issus de la rupture NH-CH.

Lorsque l'énergie de fragmentation fournie est élevée, **une rupture de la chaine latérale** peut avoir lieu, donnant plus d'information sur la structure des acides aminés mis en jeu [137,138].

- ✓ en **C$_\beta$ - C$_\gamma$**, donnant les ions **d$_n$** si la charge est retenue du coté N-terminal et **w$_n$** si la charge est retenue du coté C-terminal.

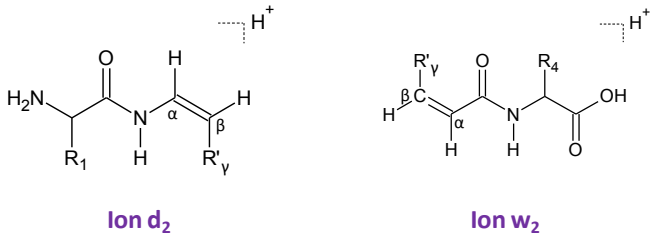


Figure 105: Ions fragments d_2 et w_2 issus de la rupture des chaines latérales en C_β - C_γ.

- ✓ en **Cα - C$_\beta$**, donnant les ions **v$_n$** avec la charge retenue du coté C-terminal.

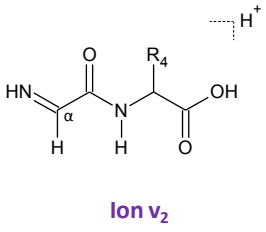


Figure 106: **Ion fragment v_2 issus de la rupture des chaines latérales** en C_α - C_β.

D'autres fragments, appelés fragments secondaires, peuvent également se produire par perte d'eau, d'ammoniaque ou du groupement carbonyle CO.

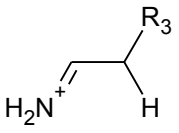


Figure 107: Fragments internes

Annexe 3 : Montage du système nano-chromatographique

Afin de profiter des avantages de la nano-chromatographie, il est important d'optimiser tout le système chromatographique c'est-à-dire, la géométrie et nature des colonnes et capillaires de connexions et le gradient d'élution de phase mobile appliqué.

I. Système de pré-concentration

En nano-chromatographie, l'injection est réalisée en utilisant d'une pré-concentration en amont de la séparation analytique: soit à l'aide une micro-cartouche de pré-concentration contenue dans des cônes de pipettes (*off-line*) soit directement intégrée au système chromatographique (*on-line*). Le principe de montage du système nano-chromatographique utilisé dans le cadre de cette étude est schématisé dans la Figure 108 ci-dessous.

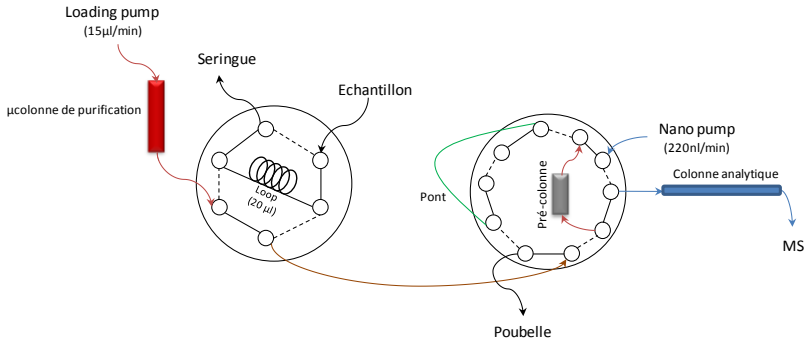


Figure 108 : Schéma du montage nanoLC

Un volume d'injection d'échantillon, en général 2µl, est chargé dans la boucle d'injection de 20µl (*loop*) par aspiration d'une seringue. Pendant cette étape la vanne d'injection de gauche est en position dite de chargement communément appelée *load* (trait plein). Elle bascule ensuite en position d'injection ou *inject* (trait en pointillé).L'échantillon est pré-concentré sur une pré-colonne C_{18} (de capacité de charge plus élevée que la colonne analytique), sur la vanne de droite à l'aide d'une pompe de chargement (*loading pump*) à un débit relativement élevé.

Pendant ce chargement dans la pré-colonne, le flux de tampon aqueux est dirigé vers la poubelle pour dessaler l'échantillon. En parallèle, la colonne analytique est équilibrée en tampon aqueux par la pompe nano-débit. Une fois l'échantillon chargé dans la cartouche de pré-concentration, la vanne de droite bascule (trait plein) pour former un circuit pré- colonne et colonne *on-line*. Les analytes sont alors transférés de la pré-colonne de piégeage vers la

colonne analytique en *back flush* pour être séparés puis détectés en spectrométrie de masse. Cette configuration nécessitant deux pompes, permet d'avoir un débit constant entre la nano-colonne et le spectromètre de masse, augmentant la stabilité du spray ainsi que le temps de vie de l'émetteur ESI.

Afin d'éliminer les contaminants de polymères (notamment les dérivés du polyéthylène glycol, PEG) contenus dans les matériaux du chemin fluidique, une cartouche de purification C_{18} est également ajoutée et mise en ligne avec la pompe de chargement, qui délivre le débit le plus important de 15µl/min [197].

II. Colonnes nanoLC

Les colonnes sont généralement remplies de particules de silice dans des capillaires de silice comportant des frittés à l'intérieur pour minimiser les volumes morts. Les premières colonnes ont été réalisées avec de petits diamètres de 20 à 70 µm pour un débit de 50 à 200nl/min [198,199] mais les colonnes de diamètre interne de 75 µm sont les plus couramment utilisées. Ce diamètre est un compromis entre la sensibilité, la capacité de chargement et la robustesse.

Les diamètres des particules sont similaires à ceux utilisés en UHPLC, à savoir des particules de 5µm, 3 µm et sub-2µm. En diminuant le diamètre des particules, l'efficacité de séparation mesurée par le nombre de plateau théorique N (Equation 32) et donc la résolution R (Equation 33) augmentent mais la pression croit également (Equation 34). Ainsi les particules de 3µm sont les plus couramment utilisées. Des alternatives à ce type de colonnes, telles que les colonnes monolithiques, sont également disponibles sur le marché.

$$N \propto \frac{L}{d_p} \qquad \text{(Equation 32)}$$

avec N : nombre de plateaux théoriques, L : longueur de la colonne et d_p : diamètre des particules.

$$R = \frac{\sqrt{N_B}}{4} \times \frac{\alpha - 1}{\alpha} \times \frac{k_B + 1}{k_B} \qquad \text{(Equation 33)}$$

avec R : résolution, N_B : nombre de plateaux théoriques, a: sélectivité et k : facteur de rétention.

$$\text{Loi de Darcy } \Delta P = \phi \frac{\eta L u}{d_p^2} \qquad \text{(Equation 34)}$$

avec ⊿P : perte de charge dans la colonne, Φ: facteur de remplissage, η: viscosité dynamique, L : longueur de la colonne, u : vitesse linéaire de la phase mobile et d_p : diamètre des particules.

III. Volumes morts et connexions

Miniaturiser le système chromatographique signifie également optimiser le système extra-colonne pour réduire un maximum les élargissements de pics qui réduisent l'efficacité de séparation et donc la sensibilité. Idéalement, les élargissements extra-colonne doivent être inférieurs à l'élargissement dû à la séparation chromatographique. Ces élargissements peuvent avoir deux origines: (1) avant la colonne causant un retard sur la gradient et (2) après la colonne entrainant un retard de gradient et donc un élargissement des pics.

Ainsi pour éviter la dispersion des solutés, décrit par l'équation de Aris-Taylor (Equation 35), la dimension des capillaires de connexion doit être réduite au maximum. Ils doivent être le plus court possible et leur diamètre interne doit être adapté à la nano-chromatographie. En général pour minimiser les élargissements de pics tout en gardant une perte de charge convenable, ce sont des diamètres de capillaires de 20 à 50µm qui sont utilisés.

$$\sigma^2 = \frac{\pi \ d_{cap}^4 \ L_{cap} \ u}{96 \ D_m} \qquad \text{(Equation 35)}$$

avec σ^2 : dispersion du soluté, d et L : diamètre et longueur du capillaire, u : vitesse linéaire de la phase mobile et Dm: coefficient de diffusion molaire.

En outre, la nature des capillaires est très importante car elle peut constituer elle-même une phase stationnaire et interagir avec les composés de même nature. Nous avons privilégié des tubes en silice ou en PEEKSil™. Ce dernier est composé de silice recouverte de polyamide PEEK (*PolyEtherEtherKetone*) qui résiste mieux à de haute pression que le PEEK simple qui lui, est hydrophobe.

Les volumes morts sont aussi à l'origine des élargissements des pics en nano-chromatographie. En général, les capillaires sont connectés aux unions et vannes via un manchon en PEEK et une vis et férule en inox. Si le capillaire n'est pas dans l'alignement direct de la vanne, un volume mort peut être créé et engendrer une dispersion extra-colonne. Récemment de nouvelles connexions, de type Viper™ (Dionex), supportant la haute pression et aisément ajustables à la main, ont été introduites et contribuent ainsi à minimiser les effets de volume mort. Il faut également ajouter que les volumes morts post-colonne dans des unions en PTFE sont souvent à l'origine des élargissements des pics.

Annexe 4 : Principe des analyseurs quadripolaires (Equations de Mathieu)

I. Quadripôle

L'analyseur quadripolaire, introduit par Paul en 1955 [200],est constitué de quatre électrodes métalliques parallèles disposées symétriquement autour d'un axe z ayant idéalement une section hyperbolique mais qui sont le plus souvent circulaire (Figure 109).

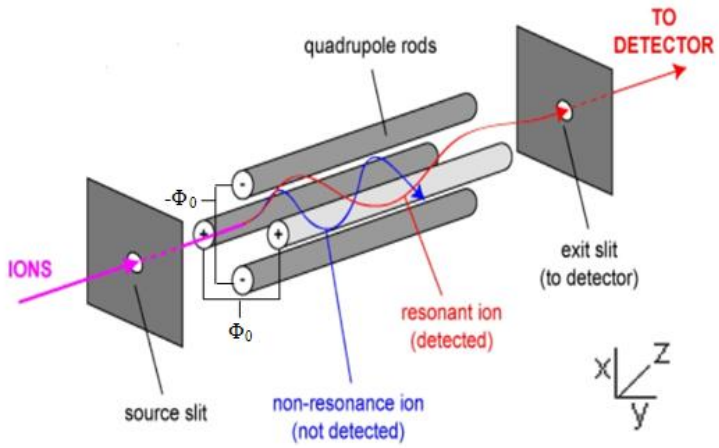


Figure 109: Schéma d'un quadripôle. Les ions rouges passent au travers du quadripôle tandis que les ions bleus sont instables et éliminés du système.

Ces électrodes sont associées électriquement deux à deux. Lorsqu'une paire d'électrodes est soumise à un potentiel (Φ_0), l'autre paire est soumise à un potentiel de même valeur mais de signe opposé ($-\Phi_0$). Ce potentiel Φ_0 est la résultante d'une tension continue U et d'une tension alternative sinusoïdale de radio fréquence $V\cos\omega t$ tel que:

$$\Phi_0 = U - V\cos\omega t \qquad \text{(Equation 36)}$$

avec ω, la pulsation fixée par le constructeur

L'application de cette différence de potentiel crée un champ électrique dans lequel les ions suivent des trajectoires oscillantes obéissants aux équations de Mathieu [201,202] et sont transmis à travers le quadripôle jusqu'au détecteur en fonction de leur stabilité par inversement successif des polarités des électrodes. La résolution des équations de Mathieu donne le mouvement d'un ion de masse m et de charge ze dans le champ quadripolaire sur l'axe x et y peut être décrit par:

$$\frac{d^2x}{d\zeta^2} + x(a_x - 2q_x \cos 2\zeta) = 0 \qquad \text{(Equation 37)}$$

$$\frac{d^2y}{d\zeta^2} - y(a_y - 2q_y \cos 2\zeta) = 0 \qquad \text{(Equation 38)}$$

Des zones de stabilité peuvent être décrites en fonction des valeurs U et V telles que les coordonnées x et y de l'ion restent strictement inférieures à r_0. Les solutions des équations de Mathieu peuvent être représentées sous forme de diagramme de stabilité de a en fonction de *q* (Figure 110).

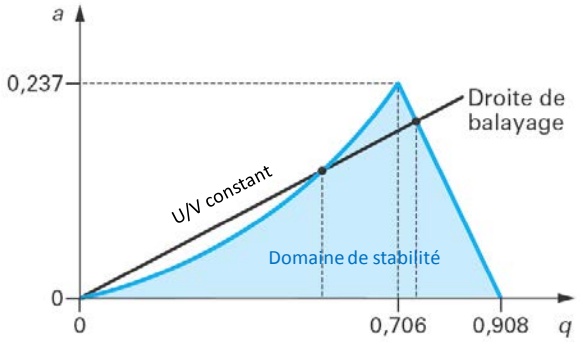


Figure 110: Diagramme de stabilité. Le domaine de stabilité, donné par les solutions des équations de Mathieu, est représenté par la zone en bleu. L'intersection de la droite de balayage et de la zone de stabilité délimite la largeur des pics (figure adaptée de Bouchoux [202]).

En gardant constant le rapport U/V, on obtient une droite de fonctionnement de l'analyseur (en trait plein noir sur la Figure 110). Un balayage de U et V, avec un rapport U/V constant et situé à l'intérieur du domaine de stabilité, permet l'observation successive d'ions. La largeur des pics observés, est proportionnelle à la longueur du segment de la droite de fonctionnement qui coupe le domaine de stabilité. La résolution entre ces ions est donc d'autant plus grande que la pente de la droite est élevée.

Pour séparer les ions en fonction de leur rapport m/z, une valeur particulière de U et V est appliquée, ne laissant passer que certaines masses (ions rouges sur la Figure 109) tandis que les autres ont une trajectoire instable qui ne leur permet pas d'atteindre le détecteur. Ces derniers vont se décharger en heurtant les électrodes ou les parois internes du spectromètre de masse et seront éliminés du système par pompage (ions bleus sur Figure 109). **Le quadripôle agit ainsi comme un filtre de masse**.

En l'absence de tension (U=0), le quadripôle fonctionne en radio fréquence (RF), c'est à dire qu'il laisse passer tous les ions dont le rapport m/z est supérieur à la valeur fixée par V. **Le quadripôle sert alors à focaliser et à transmettre les ions. Il agit dans cette**

configuration en tant que guide d'ions. C'est le cas des filtres quadripolaires positionnés à la sortie d'une source ESI ou à l'entrée d'un autre analyseur.

Le signal des ions à la sortie du quadripôle est ensuite amplifié par un photomultiplicateur porté à un potentiel de l'ordre du kV (Figure 111). Ces ions énergétiques vont se heurter sur une plaque en métal, appelée dynode de conversion, qui va émettre des électrons secondaires. Ces derniers sont ensuite accélérés et focalisés sur une série de dynodes discrètes ou dynode continue. Le nombre d'électrons émis croit à chaque dynode et peut atteindre un facteur de multiplicité d'une valeur d'environ 10^6. Le courant induit est transformé en signal électrique qui sera traduit en signal d'intensité par une analogie digitale.

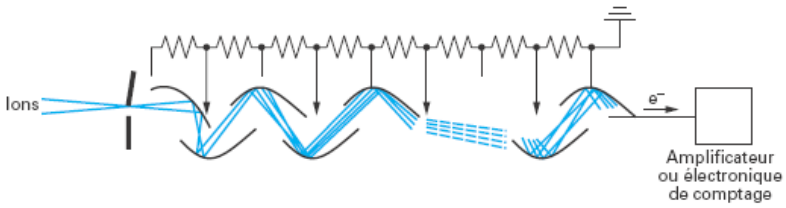


Figure 111: photomultiplicateur à dynodes discrètes. Les ions vont se heurter à une série de dynodes qui va émettre des électrons pour amplifier le signal [202].

Malgré sa résolution unitaire, le filtre de masse quadripolaire est surement l'analyseur le plus utilisé dans le monde du fait de sa facilité de mise en œuvre, sa robustesse, sa vitesse de balayage satisfaisante et sa compatibilité avec la chromatographie gazeuse ou liquide. Il est aujourd'hui très utilisé comme composant d'un spectromètre de masse tandem.

II. Les pièges à ions quadripolaires

II.1 Le piège à ions quadripolaire tridimensionnel (3D)

Le piège à ions est un développement du quadripôle, introduit par Paul et Steinwedel en 1960 [203], qui a valu le prix Nobel de physique en 1989 à Wolfgang Paul [201]. Il est constitué d'un ensemble de 3 électrodes : une électrode annulaire centrale (*ring electrode*) et deux électrodes "chapeaux" d'entrée et de sortie, percées au centre permettant la circulation des ions (Figure 112). Le fonctionnement du piège à ions est également basé sur l'action d'un champ électrique de radiofréquence sur les ions issus de la source d'ionisation. Cependant, dans le cas du piège, les ions vont être confinés dans un espace restreint et non filtrés comme dans un quadripôle.

Comme le quadripôle, un potentiel Φ_0 est appliqué sur l'électrode annulaire et génère un champ électrique quadripolaire qui confère aux ions un mouvement oscillant, dont l'amplitude et la fréquence dépend du rapport m/z et la valeur V appliquée, décrit par les équations de Mathieu [201,202].

Le champ quadripolaire n'étant homogène qu'au centre du piège, les ions ne doivent pas s'approcher des électrodes qui pourraient déstabiliser leur trajectoire et ainsi ne plus être piégés. Une solution pour remédier à ce défaut a été l'introduction en continu d'un gaz tampon dans le piège, l'hélium, pour ralentir les ions par chocs avec ces derniers pour réduire l'amplitude des mouvements des ions et ainsi les regrouper au centre du piège. L'efficacité de piégeage est alors augmentée. Sa capacité à contracter la trajectoire des ions piégés au centre du piège, permet également d'augmenter la résolution.

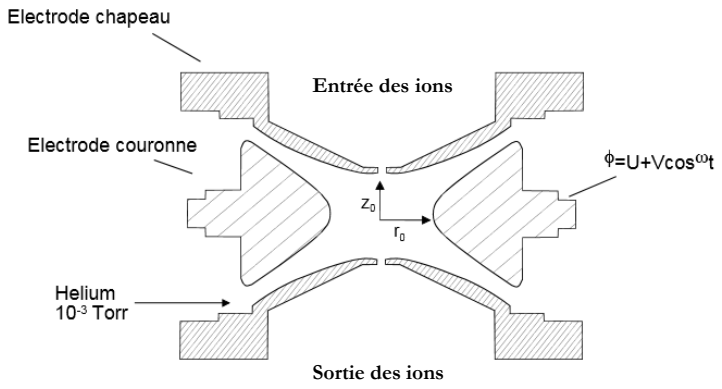


Figure 112 : Piège ionique

Sous l'effet de ce champ, les ions décrivent une trajectoire de la forme d'une courbe complexe appelée courbe de Lissajous (Figure 113). Les solutions des équations de Mathieu (Equation 39) et (Equation 40) peuvent être représentées sous forme de diagramme de stabilité de a en fonction de q (Figure 114).

$$\frac{d^2z}{d\zeta^2} - z(a_z - 2q_z\cos 2\zeta) = 0 \qquad \text{(Equation 39)}$$

$$\frac{d^2r}{d\zeta^2} + r(a_r - 2q_r\cos 2\zeta) = 0 \qquad \text{(Equation 40)}$$



Figure 113: Courbe de Lissajous

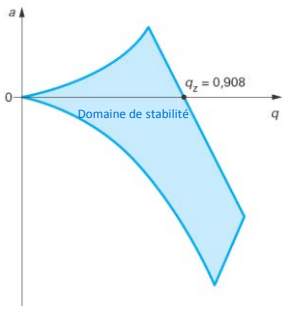


Figure 114: Diagramme de stabilité. Le domaine de stabilité, donnée par les solutions des équations de Mathieu, est représenté par la zone en bleu. (figure réadapté de technique ingénieur).

Bien souvent, les appareils fonctionnent dans un mode plus simple qui consiste à piéger les ions et les éjecter un à un. Dans ces cas-là, la tension continue U étant nulle a=0. Seule la tension alternative $V\cos\omega t$ est alors appliquée. Le diagramme de stabilité est simplifié puisqu'il ne contient que la portion sur l'axe q.

Pour enregistrer un spectre, les ions sont éjectés suivant l'axe z dans l'ordre de leur rapport m/z croissant, en augmentant le potentiel de l'électrode annulaire V de manière à déstabiliser la trajectoire des ions (Figure 112). Les ions éjectés sont ensuite captés par un photomultiplicateur d'électrons comme pour le quadripôle.

II.2 Le piège à ion quadripolaire linéaire (2D)

Le piège à ion linéaire ou 2D est un piège de seconde génération dont la géométrie est semblable au quadripôle. Il est constitué de quatre barreaux de section hyperbolique qui sont chacun divisés en 3 segments (Figure 115). Les ions, introduits de façon axiale selon l'axe z, sont confinés à l'intérieur de l'analyseur en appliquant une tension continue sur des lentilles présentes aux deux extrémités et avec l'utilisation d'un gaz tampon comme pour le piège 3D. L'éjection des ions se fait par augmentation croissante d'une tension de radiofréquence sur tous les barreaux de façon transversale par les deux orifices présents dans deux barreaux sur l'axe x vers deux détecteurs, de manière à accroitre le rapport signal sur bruit (S/B), ou vers un autre analyseur de façon axiale.

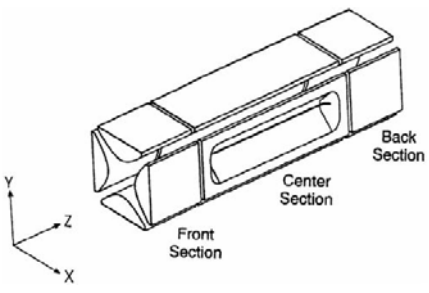


Figure 115: Piège à ions linaire **(LTQ Thermo Scientific)**

Les ions décrivent alors une trajectoire en spirale d'avant en arrière le long de l'axe z, qui peut être décrit par les équations de Mathieu et les solutions de ces équations Les solutions des équations (Equation 41) et (Equation 42) permettent de décrivent la stabilité de ions.

$$\frac{d^2x}{d\zeta^2} + x(a_x - 2q_x \cos 2\zeta) = 0 \qquad \text{(Equation 41)}$$

$$\frac{d^2y}{d\zeta^2} + x(a_y - 2q_y \cos 2\zeta) = 0 \qquad \text{(Equation 42)}$$

L'augmentation du volume de piégeage dans un piège bidimensionnelle par sa géométrie linéaire permet de piéger un nombre d'ions plus important et ainsi accroitre la sensibilité de l'analyseur. Le rendement de piégeage est d'environ 45 à 70% pour un piège 2D alors qu'il n'est que de 5% pour un piège traditionnelle 3D. De plus, les effets de charges sont d'autant plus réduits que le champ à l'intérieur du piège est parfaitement linéaire sans effet de bord (Figure 116). La vitesse de balayage et la résolution sont également accrues avec un piège 2D. Une meilleure efficacité de transmission vers le détecteur est observée.

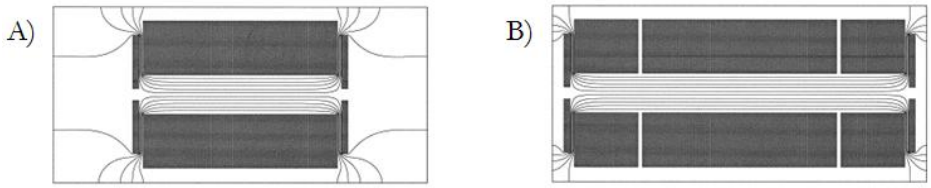


Figure 116: Simulation du champ électrique dans un piège à ions à une section (A) et trois sections (B). **Grace aux sections dans le piège 2D, le champ est plus uniforme sans effet de bord** [204]**.**

II.2.1 ***Le piège à ions quadripolaire linéaire à double pression***

Une nouvelle génération de piège à ions linéaire à double pression a fait son apparition depuis peu [205]. Ce nouveau piège est scindé en deux parties : une partie haute pression qui sert à confiner les ions, isoler et fragmenter les ions précurseurs et une partie basse pression qui sert d'analyseur avant l'envoie des ions vers les deux détecteurs radiaux (Figure 117).

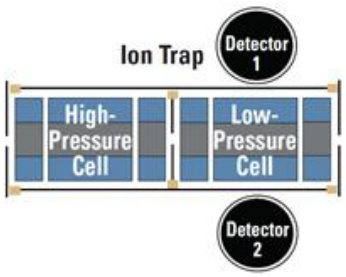


Figure 117 : Piège à ions linaire à double pression **(LTQ Velos™, Thermo Scientific)**

L'augmentation de la pression de gaz, passant de 2,5 à 5mTorr, permet d'accroitre l'efficacité de piégeage et réduire le temps de remplissage, rendant ainsi les cycles d'analyse plus court et augmentant la limite de détection des analytes. L'efficacité de fragmentation est également accrue passant de 69% à plus de 80%, ce qui augmente la sensibilité en analyse MS^n.

L'abaissement de la pression dans la deuxième partie, passant de 2,5 à moins de 1mTorr, permet d'augmenter la résolution. Pour une même vitesse de balayage, la largeur à mi-hauteur (FWHM) passe de 0,7 à 0,45 uma (unité de masse atomique). Ainsi, pour une même résolution, la vitesse de balayage est plus grande d'un facteur 2 à 33 uma/s.

Annexe 5 : Principe des analyseurs FTMS

Pour obtenir une très haute résolution et accroitre l'exactitude de masse, il faut s'orienter vers deux autres types d'analyseurs, de plus en plus utilisés en protéomique: la cellule de résonance cyclotronique à transformée de Fourier et le piège à ions orbitalaire électrostatique. Dans les deux cas, une fréquence d'oscillation des ions de rapport m/z est mesurée pour obtenir une masse précise après transformée de Fourier.

I. La cellule de résonnance cyclotronique d'ions à transformée de Fourier

L'analyseur à résonance cyclotronique d'ions se compose d'une cellule ICR (*Ion Cyclotron Resonance*) qui comporte six électrodes sous tension associées par deux. Chaque paire d'électrode a un rôle bien déterminé : piégeage, excitation et détection (Figure 118).

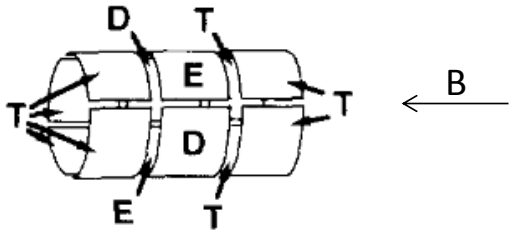


Figure 118 : Cellule ICR. Ions soumis à leur fréquence cyclotronique dans un champ magnétique B entre six électrodes : deux électrodes de piégeage (T) , deux électrodes d'excitation € et deux électrodes de détection (D) (figure d'après [206]).

Le principe décrit par Marshall [207] repose sur la soumission des ions animés d'une vitesse ν, dans un champ magnétique $\vec{B}$ suivant l'axe oz perpendiculaire aux électrodes de piégeage, à la force de Lorentz $\vec{F}$ (Equation 43) :

$$\vec{F} = q\,\vec{v} \wedge \vec{B} \qquad \text{(Equation 43)}$$

avec q :charge, v :vitesse, B : champ magnétique

Les ions sont confinés suivant l'axe z par application d'une faible tension créant un champ électrostatique imposé par les deux électrodes de piégeage dans le champ magnétique $\vec{B}$, appelé piège de Penning. Ce piège a été conçu par F.M. Penning en 1936 [208] et amélioré par H.G. Dehmelt qui a reçu le prix Nobel de physique en 1989 pour ce travail [209].

Une différence de potentiel alternative de fréquence précisément égale à la fréquence cyclotronique des ions considérés, est ensuite appliquée entre les deux électrodes d'excitation

pour mettre tous les ions de même rapport m/z en phase avant leur détection. Ces ions sont alors accélérés et le rayon de leur orbite va augmenter jusqu'à atteindre leur rayon de stabilisation (Equation 44). L'excitation est arrêtée avant que le diamètre des orbitales excède celui de la cellule. Les ions de même m/z sont alors regroupés et sont dit en cohérence.

$$r = (m/z).\frac{v}{eB} \qquad \text{(Equation 44)}$$

avec ze : charge, B : champ magnétique, v : vitesse

Les ions acquièrent alors un mouvement dans le plan (xy) dit cyclotronique, c'est-à-dire un mouvement circulaire uniforme de fréquence f donné par l'équation (Equation 45), indépendamment de leur vitesse initiale.

$$f = \frac{q\ B}{2\pi(m)} = \frac{z\ e\ B}{2\pi(m)} = \frac{1{,}535611.10^{7} B}{(m/z)} \qquad \text{(Equation 45)}$$

avec q : charge élémentaire B : champs magnétique, m/z : rapport masse sur charge

Le courant induit par le mouvement cohérent des ions de même m/z sera mesuré alternativement sur les deux électrodes de détection, générant une sinusoïde amortie de fréquence cyclotronique. **Le signal détecté, appelé transient ou interférogramme (FID *free induction decay*), correspond à la superposition des différents courants induits correspondant aux différents ions de *m/z* différents excités par résonance (signal créneau)**. Du fait de la collision des ions avec le gaz résiduel dans la cellule, les paquets d'ions perdent leur cohérence. Il en résulte une perte d'intensité du signal aboutissant à une sinusoïdale amortie exponentiellement (Figure 119 A). Pour minimiser ces collisions, la cellule doit être maintenue dans un vide très poussé de l'ordre de 10^{-10}mbar. Les différentes fréquences sont ensuite extraites du signal grâce à l'opération mathématique de Fourier, la transformée de Fourier (FT *Fourier Transform*), qui permet de convertir un signal en temps, en un signal en fréquence. La fréquence cyclotron (Equation 45) étant proportionnelle à $1/(m/z)$, l'inverse de la transformée de Fourier du courant induit obtenu, permet ainsi d'aboutir au spectre de masse en m/z (Figure 119 B).

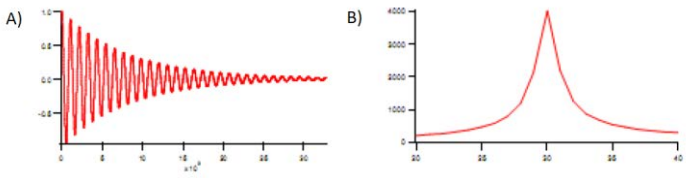


Figure 119: Signaux avant et après transformée de Fourier. A) Transient ou interférogramme correspondant à un signal sinusoïdal amorti exponentiellement (échelle en temps). B) Spectre obtenu après transformation de Fourier (échelle en fréquence) (figure extraite du cours de Guillaume van der Rest, CJSM 2011)

Dans le cas des spectromètres à résonance cyclotronique d'ions, la résolution est dépendante du champ magnétique et du rapport m/z suivant l'équation (Equation 46) :

$$R = \frac{M}{\Delta M} = cte\,\frac{eB}{2\ (m/z)} \qquad \text{(Equation 46)}$$

Où cte est une constante d'amortissement du signal inversement proportionnelle à la pression

Pour atteindre de très haute résolution, plusieurs paramètres peuvent donc être optimisés. Tout d'abord, le champ magnétique $\vec{B}$, créé par un aimant (bobine). Plus ce champ $\vec{B}$ est élevé plus la résolution sera élevée. Actuellement, des aimants de 7 à 12 Tesla sont communément installés dans les laboratoires d'analyse protéomique. Cependant, le champ $\vec{B}$ doit être stabilisé. Il doit donc être maintenu à très basse température grâce à l'utilisation de fluides cryogéniques. Cette très haute résolution permet d'obtenir une bonne exactitude de masse mais cette dernière est également dépendante du nombre d'ions. En effet si la cellule contient trop d'ions, des phénomènes de répulsions électrostatiques statiques apparaissent entre les ions déstabilisant leur trajectoire (effet de charge d'espace). Il en résulte un biais dans la mesure de la fréquence cyclotronique et donc sur la masse. Pour réduire ce biais, la calibration externe doit se faire dans des conditions proches, en termes d'effet d'espace de charge, de l'échantillon. Dans l'instrument que nous avons utilisé (de type LIT/FT-ICR, Thermo Fisher Scientific) cette calibration externe s'effectue avec différents nombre d'ions dans la cellule, ce qui minimise les variations. Une autre solution serait d'utiliser la calibration interne où le calibrant subit le même effet d'espace de charge que l'échantillon à mesurer. Dans ce cas, la mesure de masse devient encore plus précise.

L'analyseur FT-ICR permet en principe d'atteindre une gamme illimitée de masse, à très haute résolution. Mais dans la pratique, cette gamme de masse varie entre 10 à 4000 m/z. De plus, d'après l'équation (Equation 46), le pouvoir de résolution diminue avec l'augmentation de la masse.

Pour résumer, les appareils FTICR utilisés en routine, possèdent une gamme dynamique de l'ordre de 4 ordres de grandeur, et permettent d'atteindre une sensibilité de l'ordre de l'attomole. Leur pouvoir résolutif peut atteindre 1 000 000 à m/z 1000, induisant une erreur de justesse l'ordre de la ppm, voire moins suivant la calibration choisie. Ceci donne accès à la mesure du massif isotopique, information très importante pour les composés de haut poids moléculaire permettant d'établir leur masse précise par déconvolution du spectre MS ou de mesurer la masse exacte dans le cas de mélanges complexes.

I.1 Amélioration des performances de la cellule ICR

I.1.1 Traitements du signal

Un traitement des données avant Transformée de Fourier est souvent utile. Ainsi une multiplication du signal par une fonction qui permet de lisser le pic, minimise les artefacts de la transformée de Fourier rapide (FFT, *Fast Fourier Transform*). Cette opération est nommé apodisation (retrait des pieds) [210].

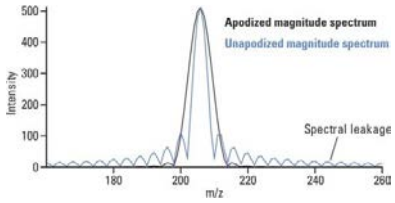


Figure 120: Effet de l'apodisation des spectres FT-ICR. Exemple du molybdène (m/z 94) sans apodisation (spectre bleu) avec apodisation Hanning (spectre noir). L'apodisation consiste à couper le pied du pic afin d'améliorer l'esthétique du pic et éliminer les pics artefacts.[figure de [211]

En amont de la FFT, une digitalisation du signal temporel (transient) est réalisée. Si le signal temporel est d'abord échantillonné avec une période Δt, son spectre obtenu après FFT est limité à une bande de fréquence $1/\Delta t$, donnée par la relation de Nyquist. Cette relation est souvent employée avec un facteur 2 quand seulement la partie réelle du signal est pris en compte. Le signal est alors décrit avec un nombre de point de donnée constant. Un moyen simple d'obtenir plus de point dans le spectre final pour mieux décrire le signal est d'ajouter des zéros à la fin du transient. Cette pratique, appelée *zéro-filling*, permet d'améliorer l'esthétique du spectre en interpolant entre les points et ainsi diminuer l'intervalle de fréquence (effet de lissage du pic)[212].

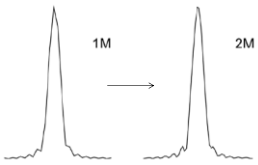


Figure 121: Effet du zéro-filling sur les spectres FT-ICR. Le zéro-filling double le nombre de points et affine les pics. Ici la taille du ficher double de 1méga à 2méga de points. La définition du pic est meilleure, ce qui entraine une amélioration de la mesure de l'intensité. (Ecole thématique CNRS, FTMS 2014, Carlos, 2014)

Un autre moyen simple d'augmenter les performances de la cellule ICR est d'augmenter le temps d'acquisition du transient, ce qui n'est pas chose aisée en couplage chromatographique où le temps est compté.

I.1.2 Améliorations technologiques

Du point de vue instrumental en parallèle de l'augmentation de la résolution par l'augmentation de la fréquence de résonance (et donc du champ magnétique B), et pour répondre à une demande croissante des besoins de la recherche, les cellules ont évolué peu à peu, passant d'une forme cubique à aujourd'hui cylindrique. Dans l'idéal, la cellule doit être constituée d'électrodes hyperboliques pour le piégeage (Figure 122 A) générant mouvement cyclotronique indépendant des autres mouvements naturels des ions (oscillation et magnétron) et indépendant de la position des ions dans le piège, d'électrodes quadripolaires 2D pour l'axialisation des ions (Figure 122 B) et des électrodes planes pour l'excitation dipolaire et la détection des ions (Figure 122 C) mais cela est irréalisable.

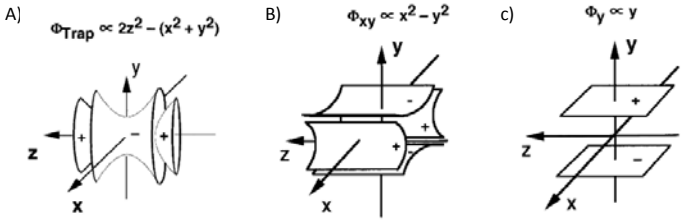


Figure 122: Electrodes idéales. A) Piège 3D de type piège de Penning pour le piégeage des ions B) un quadrupole 2D pour l'axialisation des ions, C) des électrodes planes infinies pour l'excitation et la détection des ions [207].

Du fait de l'application d'un champ électrique dans la cellule pour piéger les ions, le champ magnétique B n'est pas homogène dans la direction axiale magnétique, ce qui conduit à un déphasage des ions et donc une diminution de la résolution. Pour s'approcher de la cellule ICR idéale, une nouvelle cellule a été développée et présentée à l'ASMS en 2010 par l'équipe de Nikolaev [213]. Cette cellule segmentée en plusieurs parties se rapproche d'un piège à ions hyperbolique. Appelée cellule harmonique dynamique, ce concept est aujourd'hui commercialisé par la société Bruker. Cette cellule permet d'atteindre un résolution de plus de 20 millions pour la réserpine (m/z 609) dans un aimant de 7 Tesla [213].

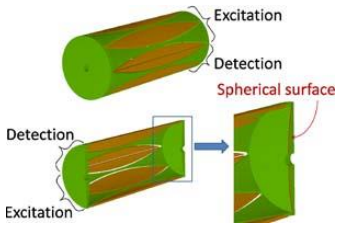


Figure 123: cellules ICR harmonique dynamique[213]

II. Le piège à ions orbitalaire électrostatique (Orbitrap)

L'analyseur à piège à ions orbitalaire électrostatique est une alternative intéressante au FT-ICR puisqu'il va permette l'accès à la très haute résolution sans avoir recours à un instrument à champ magnétique intense couteux et encombrant. Son principe décrit par Makarov [214], est une adaptation du piège à ions développé par Kingdon en 1923 [215]. Il est aujourd'hui commercialisé par la société Thermo Scientific sous le nom d'Orbitrap.

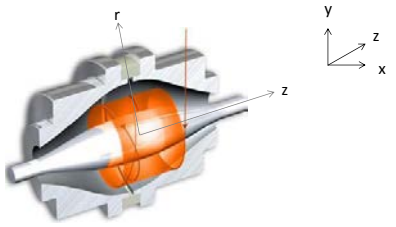


Figure 124: Orbitrap. Les ions sont injectés tangentiellement par la C-Trap à l'Orbitrap. Une fois piégés, les ions oscillent autour de l'électrode centrale et le courant induit est mesuré. La mase est inversement proportionnelle à la fréquence obtenue par transformée de Fourier du transient.

L'Orbitrap se compose d'une électrode creuse, à l'intérieur de laquelle est placée coaxialement une électrode centrale en forme de fuseau (Figure 124). L'application d'une tension continue entre ces deux électrodes de géométrie particulière génèrent un champ électrostatique de type quadro-logarithmique (Equation 47) suivant l'axe z.

$$U_{(r,z)} = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}\,Rm^2 \ln\left(\frac{r}{Rm}\right) + C \qquad \text{(Equation 47)}$$

avec r et z : coordonnées cylindriques, Rm : rayon caractéristique de l'électrode centrale, k : courbure du champ et C : constante.

Les ions sont injectés tangentiellement à l'électrode centrale avec une énergie cinétique de quelques keV. La force électrostatique compense alors les forces centrifuges, piégeant les ions autour de l'électrode centrale, tels des satellites en orbite. Le mouvement des ions se décompose en un mouvement circulaire autour de l'électrode centrale dans le plan (xy) et un mouvement oscillatoire longitudinal de va-et-vient selon l'axe z. Les ions d'un même rapport m/z auront la même trajectoire circulaire et oscillent axialement avec une fréquence ω_z indépendante de la vitesse suivant l'équation (Equation 48):

$$\omega_z = \sqrt{\frac{k}{m/z}} \qquad \text{(Equation 48)}$$

avec k, constante liée à la géométrie du piège.

De la même façon que pour le FT-ICR, à chaque passage devant les électrodes externes, les ions génèrent des courants induits qui permet d'avoir accès au rapport m/z par transformée de Fourier. La résolution étant fonction du temps d'acquisition du courant induit,

de la même façon que pour le FT-ICR, le vide dans le piège est très poussé (10^{-10} mbar) permettant d'augmenter le libre parcours moyen des ions.

Les performances de l'Orbitrap rivalisent avec celles du FT-ICR. En effet, l'erreur de justesse en routine est de l'ordre du ppm et la résolution peut atteindre 140 000 à m/z 200. L'Orbitrap tire son avantage dans l'analyse des hautes masses où la résolution chute plus lentement que dans un système FT-ICR (Figure 125). Cependant, le temps d'analyse nécessaire à résolution égale est en défaveur de l'Orbitrap en basse masse. Un instrument comme l'Orbitrap, présente une gamme dynamique satisfaisante de 4 décades.

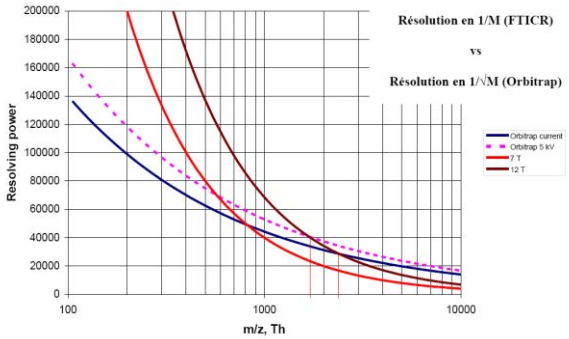


Figure 125: Comparaison de l'évolution de la résolution en fonction du rapport m/z entre l'analyseur Orbitrap et FT-ICR. **Le FT-ICR permet d'atteindre des résolutions plus élevé mais décroit plus vite lorsque le rapport m/z augmente. A haut poids moléculaire, l'emploi d'un orbitrap devient plus judicieux (**Ecole thématique CNRS, FTMS 2014**, Tissen, 2014).**

II.1 Éjection des ions

Pour mesurer la fréquence d'oscillation des ions dans l'Orbitrap, il faut tout d'abord un système capable d'envoyer les ions par paquet vers l'analyseur, via un interstice d'un diamètre de 1mm séparant les deux électrodes externes, à un temps donné. C'est le rôle de la C-Trap, quadrupole en forme de C, fonctionnant en radiofréquence. Les ions sont tout d'abord piégés par paquet dans cette zone, où ils sont refroidis et refocalisés. Ensuite, un abaissement du potentiel RF et en même temps l'application d'une tension continue, provoque l'éjection des ions le long du champ électrique créé entre la C-Trap et le point d'entrée de l'Orbitrap (Figure 126). Contrairement au FT-ICR, il n'y a pas d'étape d'excitation des ions pour les mettre en "orbite", mais une injection pulsée par la C-Trap entrainant peu de déphasage, moyennant que les ions entrent dans l'analyseur à un temps inférieur au 1/10 de leur temps d'oscillation. Le temps zéro est donc le temps d'éjection des ions.

Le refroidissement par ralentissement des ions se fait par collision avec un flux d'azote dans la C-Trap à une pression de 10^{-5}mbar. Cette pression étant 5 décades fois plus élevée que

dans l'analyseur (10^{-10}mbar), un dispositif d'une série de lentilles (Z-Lens) et déflecteurs est nécessaire pour diminuer le flux d'azote avant l'entrée des ions dans l'analyseur (Figure 126).

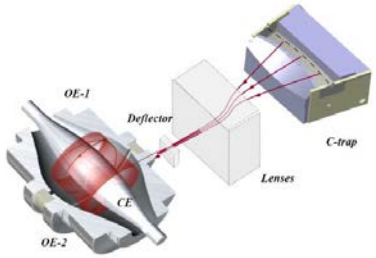


Figure 126: Ejection de la C-Trap vers l'Orbitrap.

II.2 Repousser les limites des performances du piège à ions orbitalaire

La technologie Orbitrap, à l'instar du FT-ICR, s'appuie donc sur la transformée de Fourier pour convertir le signal temporel en fréquence mesurée en un spectre en rapport m/z. En compléments des paramètres cités au paragraphe I.1.1, une autre amélioration du traitement de signal a été développée et appliquée dans les spectromètres de masse contenant un Orbitrap de dernière génération (Q Exactive, Orbitrap Elite et Fusion) nommée eFT™ *(enhanced Fourier Transformation*) [216]. Le calcul de la transformée de Fourier utilise des nombres complexes comportant une partie réelle et une partie imaginaire, qui peuvent être assimilées à l'amplitude et la phase du signal. Etant donné que ce dernier est lié aux paramètres initiaux à l'injection des ions de façon complexe, les spectres FT sont généralement interprétés dans le mode "amplitude" sans considération des informations de phase [217]. Cependant contrairement au FT-ICR, dans le cas de l'Orbitrap, l'excitation par l'éjection fournit une phase initiale indépendante du rapport m/z grâce à la C-Trap rendant l'utilisation de cette composante possible pour créer le spectre. Le processus eFT, utilisant les composantes réelles et imaginaires du signal, améliore la résolution, l'exactitude de la mesure et le profil des pics (Figure 127). La justesse sera optimale si la détection du transient démarre dès la fin de l'injection des ions. Ainsi des modifications techniques ont été apportées pour réduire ce temps de 10ms à une fraction de milliseconde, augmentant la résolution d'un facteur de 1.8 à 2 pour un même transient. Ce gain est plus faible pour les protéines (facteur de 1.4) car le signal décroit par impact avec les molécules de gaz résiduel. Ces traitements du signal se font de façon rapide, de tel sorte que le temps d'un cycle soit déterminé par le temps d'acquisition du transient et temps d'injection des ions.

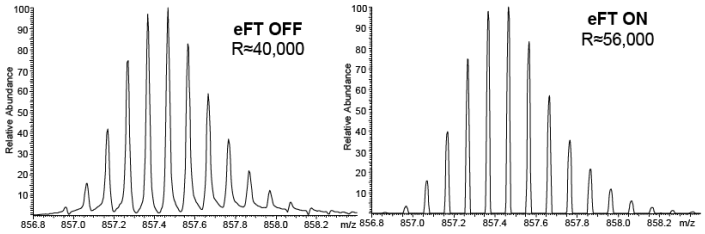


Figure 127: Spectre de la protéine ubiquitine 10 fois chargée (Q Exactive, Thermo Scientific). À gauche le spectre dans traitement eFT et à droite le spectre avec le traitement eFT. LE gain est d'environ 1.4 avec ce traitement [216].

Pour être efficace, l'eFT demande une très bonne synchronisation électronique et reste sensible au délai de transfert d'information du signal. L'exactitude de la mesure reste comparable aux résultats obtenus par FT classique avec le mode "amplitude".

Un autre paramètre, tel que la fréquence d'oscillation, par augmentation le facteur k (Equation 48), spécifique à l'instrument, peut être une solution pour accroitre les performances de l'analyseur. Pour cela, une diminution de la géométrie du piège s'impose (Figure 128). Cette diminution engendre un champ électrostatique plus intense et donc un nombre d'oscillation des ions plus élevé par unité de temps, augmentant le pouvoir résolutif de l'analyseur. Ce nouveau piège a été implémentée dans le spectromètre de masse "Orbitrap Elite" [218].

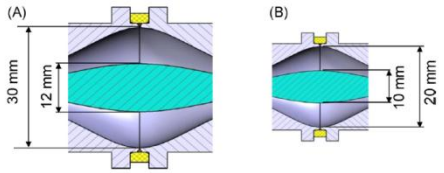


Figure 128: Augmentation des performances de l'Orbitrap par diminution de la géométrie. (A) l'Orbitrap standard. (B) l'Orbitrap de nouvelle génération inclus dans les LTQ Elite. La diminution de l'Orbitrap a pour conséquence l'augmentation du champ électrique à l'intérieur et donc une fréquence d'oscillation plus élevée.

Enfin, l'injection des d'ions de la C-Trap vers l'Orbitrap peut être améliorée en réduisant le flux gazeux ou en remplaçant l'azote par l'hélium à plus faible pression [219], pour diminuer le risque qu'un ion rentre en contact avec des molécules du gaz et fasse décroitre le transient (Figure 119 A).

Annexe 6 : Tableau des transitions choisies pour les expériences tSRM.

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	product ion (m/z)
P01034	Cystatin C	ALDFAVGEYNK	613.8	2	34.1	37.1	24	y7 y6 y5	780.4 709.4 610.3
		ALDFAVGEYNKASNDM[Ox]YHSR	576.8	4	30.6	33.6	28	y8 y4 y5	1025.4 562.3 709.3
		ALDFAVGEYNKASNDMYHSR	572.8	4	30.6	33.6	28	y8 y5 y4	1009.4 693.3 562.3
		LVGGPM[Ox]DASVEEEGVR	830.9	2	24.1	27.1	31	y10 y9 y8	1090.5 975.5 904.4
		LVGGPM[Ox]DASVEEEGVRR	606.3	3	25.5	28.5	30	y7 y6 y9	874.4 745.4 1060.5
		LVGGPMDASVEEEGVR	822.9	2	28.6	31.6	31	y9 y8 y10	975.5 904.4 1090.5
		LVGGPMDASVEEEGVRR	601.0	3	25.0	28.0	30	y7 y6 y9	874.4 745.4 1060.5
		RALDFAVGEYNK	691.9	2	26.1	29.1	27	y7 y6 y8	780.4 709.4 927.5
		TQPNLDNC[Cam]PFHDQPHLK	516.0	4	25.8	28.8	26	y7 y6 y8	874.5 737.4 1021.5
		TQPNLDNC[Cam]PFHDQPHLKR	555.0	4	23.0	26.0	28	y7 y6 y5	893.5 778.5 650.4
P01834	Ig kappa chain C region	DSTYSLSSTLTLSK	751.9	2	35.1	38.1	29	y8 y10 y7	836.5 1036.6 749.4
		HKVYAC[Cam]EVTHQGLSSPVTK	714.4	3	18.8	21.8	35	y8 y9 y6	788.5 916.5 618.3
		SGTASVVC[Cam]LLNNFYPR	899.5	2	46.1	49.1	32	y9 y8 y7	1196.6 1036.6 923.5
		TVAAPSVFIFPPSDEQLK	973.5	2	42.0	45.0	35	y7 y8 y9	816.4 913.5 1060.5
		TVAAPSVFIFPPSDEQLKSGTASVVC[Cam]LLNNFYPR	1242.3	3	55.2	58.2	35	y8 y7 y6	1036.6 923.5 810.4
		VDNALQSGNSQESVTEQDSK	1068.5	2	19.8	22.8	35	y9 y8 y7	1022.5 893.4 806.4
		VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK	1207.2	3	33.6	36.6	35	y8 y10 y11	836.5 1036.6 1199.7
		VQWKVDNALQSGNSQESVTEQDSK	893.1	3	26.9	29.9	35	y9 y8 y6	1022.5 893.4 707.3
		VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK	1041.0	4	35.4	38.4	35	y10 y8 y7	1036.6 836.5 749.4
		VYAC[Cam]EVTHQGLSSPVTK	626.0	3	23.2	26.2	31	y8 y6 y10	788.5 618.3 1053.6

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P01857, P01859, P01860, P01861	Ig gamma-1 chain C , Ig gamma-2 chain C, Ig gamma-3 chain C , Ig gamma-4 chain C	ALPAPIEKTISK	423.3	3	24.1	27.1	22	y6 y7 y8	705.4 818.5 915.6
		FNWYVDGVEVHNAK	559.9	3	31.6	34.6	28	y9 y8 y6	968.5 853.5 697.4
		GPSVFPLAPSSK	593.8	2	30.7	33.7	24	y7 y4 y5	699.4 418.2 489.3
		NQVSLTC[Cam]LVK	581.3	2	29.8	32.8	23	y5 y7 y6	620.3 820.5 733.4
		SC[Cam]DKTHTC[Cam]PPC[Cam]PAPELLGG PSVFLFPPKPK	834.4	4	37.2	40.2	35	y8 y7 y5	973.6 826.5 566.4
		STSGGTAALGC[Cam]LVK	661.3	2	27.5	30.5	26	y7 y8 y6	760.4 831.5 689.4
		TTPPVLDSDGSFFLYSK	937.5	2	41.4	44.4	35	y9 y8 y5	1063.5 948.5 657.4
		VSNKALPAPIEK	422.9	3	18.7	21.7	22	y6 y5 y7	654.4 557.3 767.5
		VVSVLTVLHQDWLNGK	603.3	3	43.0	46.0	30	y8 y9 y7	997.5 1110.6 860.4
		VVSVLTVLHQDWLNGKEYK	557.8	4	39.7	42.7	26	y9 y8 y10	1152.6 1037.5 1280.6
P01871, P04220	Ig mu chain C region , Ig mu heavy chain disease protein	EGKQVGSGVTTDQVQAEAK	644.7	3	17.7	20.7	32	y6 y8 y5	645.4 888.4 546.3
		EGKQVGSGVTTDQVQAEAKESGPTTYK	699.6	4	21.7	24.7	34	y8 y7 y10	882.4 753.4 1081.6
		FTC[Cam]TVTHTDLPSPLK	573.0	3	28.3	31.3	29	y5 y6 y8	541.3 654.4 870.5
		GQPLSPEKYVTSAPMPEPQAPGR	813.1	3	30.7	33.7	35	y10 y9 y8	1079.5 982.5 851.4
		GVALHRPDVYLLPPAR	592.0	3	29.1	32.1	29	y6 y5 y8	666.4 553.3 928.6
		QIQVSWLR	515.3	2	32.9	35.9	21	y4 y5 y6	561.3 660.4 788.4
		QVGSGVTTDQVQAEAK	809.4	2	20.2	23.2	31	y9 y10 y8	989.5 1090.5 888.4
		QVGSGVTTDQVQAEAKESGPTTYK	827.7	3	23.4	26.4	35	y10 y8 y7	1081.6 882.4 753.4
		VFAIPPSFASIFLTK	819.5	2	52.3	55.3	31	y9 y10 y6	1013.6 1110.6 708.4
		YVTSAPMPEPQAPGR	800.9	2	25.4	28.4	31	y8 y6 y10	851.4 625.3 1079.5

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P02647	Apolipoprotein A-I	AHVDALRTHLAPYSDELRQR	587.8	4	24.6	27.6	29	y7 y5 y8	903.5 701.4 1066.5
		ATEHLSTLSEK	608.3	2	15.6	18.6	24	y7 y6 y8	777.4 664.4 914.5
		DLATVYVDVLK	618.3	2	42.3	45.3	24	y7 y6 y8	835.5 736.4 936.5
		LAARLEALKENGGAR	523.6	3	19.5	22.5	26	y6 y7 y5	603.3 731.4 474.2
		LLDNWDSVTSTFSK	806.9	2	40.0	43.0	31	y9 y8 y6	971.5 856.4 670.3
		THLAPYSDELR	651.3	2	22.0	25.0	25	y9 y8 y7	1063.5 950.5 879.4
		THLAPYSDELRQR	529.3	3	19.9	22.9	27	y5 y7 y8	701.4 903.5 1066.5
		VKDLATVYVDVLK	731.9	2	35.8	38.8	28	y8 y7 y6	936.5 835.5 736.4
		VKDLATVYVDVLKDSGR	626.7	3	34.7	37.7	31	y9 y8 y10	988.5 889.5 1151.6
		VSFLSALEEYTK	693.9	2	45.3	48.3	27	y7 y6 y8	853.4 782.4 940.5
P02649	Apolipoprotein E	AATVGSLAGQPLQER	749.4	2	26.3	29.3	29	y7 y8 y5	827.4 898.5 642.4
		AKLEEQAQQIR	657.4	2	16.7	19.7	26	y8 y7 y6	1001.5 872.5 743.4
		AYKSELEEQLTPVAEETR	698.4	3	28.8	31.8	34	y7 y8 y5	801.4 902.5 605.3
		DRLDEVKEQVAEVR	562.6	3	26.5	29.5	28	y6 y4 y7	701.4 474.3 830.4
		LAVYQAGAR	474.8	2	19.4	22.4	19	y6 y5 y7	665.3 502.3 764.4
		LGADM[Ox]EDVC[Cam]GR	619.8	2	18.1	21.1	24	y6 y5 y4	735.3 606.3 491.2
		LGADMEDVC[Cam]GR	611.8	2	23.6	26.6	24	y6 y7 y4	735.3 866.3 491.2
		LGPLVEQGR	484.8	2	22.1	25.1	20	y5 y6 y4	588.3 701.4 489.2
		LQAEAFQAR	517.3	2	21.0	24.0	21	y7 y6 y5	792.4 721.4 592.3
		LSKELQAAQAR	607.8	2	15.1	18.1	24	y7 y8 y6	757.4 886.5 644.3

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P02652	Apolipoprotein A-II	AGTELVNFLSYFVELGTQPATQ	795.7	3	53.7	56.7	35	y9 y8 y7	944.5 815.4 702.3
		EPC[Cam]VESLVSQYFQTVTDYGKDLM[Ox]EK	994.8	3	44.2	47.2	35	y8 y7 y9	999.5 836.4 1114.5
		EPC[Cam]VESLVSQYFQTVTDYGKDLMEK	989.5	3	44.2	47.2	35	y8 y7 y9	983.5 820.4 1098.5
		KAGTELVNFLSYFVELGTQPATQ	838.4	3	49.0	52.0	35	y9 y7 y8	944.5 702.3 815.4
		SKEQLTPLIK	578.9	2	21.6	24.6	23	y8 y7 y4	941.6 812.5 470.3
		SKEQLTPLIKK	642.9	2	17.1	20.1	25	y6 y7 y5	699.5 812.6 598.4
		SPELQAEAK	486.8	2	15.3	18.3	20	y6 y4 y7	659.4 418.2 788.4
		SYFEKSKEQLTPLIK	604.3	3	27.5	30.5	30	y8 y5 y6	941.6 571.4 684.5
		SYFEKSKEQLTPLIKK	647.0	3	26.1	29.1	32	y8 y6 y5	940.6 699.5 598.4
		VKSPELQAEAK	600.3	2	15.1	18.1	24	y8 y6 y7	885.5 659.4 788.4
P02671	Fibrinogen alpha chain	GDFSSANNRDNTYNR	577.6	3	15.3	18.3	29	y5 y4 y6	667.3 553.3 782.3
		GGSTSYGTGSETESPR	786.8	2	15.5	18.5	30	y8 y9 y10	862.4 963.4 1020.5
		GLIDEVNQDFTNR	760.9	2	36.5	39.5	29	y7 y6 y8	894.4 780.4 993.5
		GLIDEVNQDFTNRINK	626.0	3	36.2	39.2	31	y7 y8 y6	892.5 1007.5 745.4
		GSESGIFTNTK	570.8	2	21.0	24.0	23	y5 y6 y7	610.3 723.4 780.4
		M[Ox]ELERPGGNEITR	506.6	3	18.1	21.1	26	y7 y8 y6	746.4 843.4 689.4
		MELERPGGNEITR	501.3	3	21.4	24.4	25	y8 y7 y6	843.4 746.4 689.4
		NSLFEYQK	514.8	2	25.5	28.5	21	y5 y4 y3	714.3 567.3 438.2
		TFPGFFSPM[Ox]LGEFVSETESR	761.0	3	52.0	55.0	35	y7 y6 y8	807.4 708.3 954.5
		TFPGFFSPMLGEFVSETESR	755.7	3	53.7	56.7	35	y6 y7 y8	708.3 807.4 954.5

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P0DJI8, P0DJI9	Serum amyloid A-1, Serum amyloid A-2	AYSDM[Ox]REANYIGSDK	868.4	2	18.8	21.8	33	y6 y5 y4	682.3 519.3 406.2
		AYSDM[Ox]REANYIGSDKYFHAR	603.3	4	23.6	26.6	29	y9 y6 y5	1080.5 821.4 693.3
		AYSDMREANYIGSDK	860.4	2	22.5	25.5	33	y6 y5 y4	682.3 519.3 406.2
		AYSDMREANYIGSDKYFHAR	599.3	4	25.9	28.9	29	y9 y6 y5	1080.5 821.4 693.3
		EANYIGSDKYFHAR	835.9	2	22.6	25.6	32	y6 y5 y8	821.4 693.3 1023.5
		FFGHGAEDSLADQAANEWGR	726.7	3	34.3	37.3	35	y7 y6 y5	803.4 732.3 661.3
		GNYDAAKRGPGGVWAAEAISDAR	778.1	3	28.6	31.6	35	y8 y9 y10	832.4 903.5 1089.5
		GPGGVWAAEAISDAR	728.9	2	35.4	38.4	28	y8 y5 y4	832.4 561.3 448.2
		GPGGVWAAEAISDARENIQR	699.7	3	36.1	39.1	34	y7 y5 y4	886.5 659.3 530.3
		RGPGGVWAAEAISDAR	538.3	3	28.5	31.5	27	y8 y6 y5	832.4 632.3 561.3
		RGPGGVWAAEAISDARENIQR	564.0	4	32.0	35.0	28	y5 y4 y7	659.3 530.3 886.5
		SFFSFLGEAFDGAR	775.9	2	51.7	54.7	30	y8 y9 y6	822.4 935.5 636.3
		SFFSFLGEAFDGARDM[Ox]WR	719.0	3	50.7	53.7	35	y7 y5 y8	907.4 779.4 1022.4
		SFFSFLGEAFDGARDMWR	713.7	3	54.4	57.4	35	y7 y5 y3	891.4 763.4 492.2
		GNYDAAKRGPGGAWAAEVISNAR	583.5	4	29.7	32.7	29	y7 y5 y4	788.4 560.3 447.2
		GPGGAWAAEVISNAR	485.9	3	34.9	37.9	25	y5 y6 y7	560.3 659.4 788.4
		GPGGAWAAEVISNARENIQR	699.4	3	36.9	39.9	34	y6 y5 y4	815.4 659.3 530.3
		RGPGGAWAAEVISNAR	537.9	3	29.5	32.5	27	y9 y6 y5	930.5 659.4 560.3
		RGPGGAWAAEVISNARENIQR	563.8	4	32.0	35.0	28	y5 y4 y7	659.3 530.3 886.5

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P02743	Serum amyloid P-component	AYSLFSYNTQGR	703.8	2	32.1	35.1	27	y7 y8 y6	825.4 972.5 738.4
		AYSLFSYNTQGRDNELLVYKER	667.3	4	34.2	37.2	33	y7 y6 y5	920.6 807.5 694.4
		DNELLVYK	497.3	2	28.6	31.6	20	y3 y4 y5	409.2 522.3 635.4
		DNELLVYKER	426.9	3	24.4	27.4	22	y5 y4 y3	694.4 595.3 432.3
		GLRQGYFVEAQPK	498.3	3	22.7	25.7	25	y5 y6 y4	572.3 671.4 443.3
		GYVIIKPLVWV	643.9	2	46.2	49.2	25	y6 y5 y7	741.5 613.4 854.5
		IVLGQEQDSYGGK	697.4	2	22.8	25.8	27	y8 y6 y5	883.4 626.3 511.3
		IVLGQEQDSYGGKFDR	604.6	3	25.5	28.5	30	y6 y8 y9	679.4 929.4 1044.5
		QGYFVEAQPK	583.8	2	23.3	26.3	23	y6 y5 y4	671.4 572.3 443.3
		VGEYSLYIGR	578.8	2	30.8	33.8	23	y6 y4 y7	708.4 508.3 871.5
P02766	Transthyretin	AADDTWEPFASGK	697.8	2	31.8	34.8	27	y6 y7 y8	606.3 735.4 921.4
		GSPAINVAVHVFR	456.3	3	30.9	33.9	23	y6 y8 y4	728.4 941.5 558.3
		GSPAINVAVHVFRK	499.0	3	27.3	30.3	25	y7 y5 y6	856.5 686.4 785.5
		KAADDTWEPFASGK	508.2	3	26.7	29.7	26	y6 y5 y4	606.3 509.3 362.2
		RYTIAALLSPYSYSTTAVVTNPKE	882.5	3	41.1	44.1	35	y10 y7 y6	1059.6 786.4 687.4
		TSESGELHGLTTEEEFVEGIYK	819.1	3	39.2	42.2	35	y8 y7 y6	984.5 855.5 708.4
		TSESGELHGLTTEEEFVEGIYKVEIDTK	785.9	4	40.9	43.9	35	y8 y7 y6	995.5 832.5 704.4
		VLDAVRGSPAINVAVHVFR	674.1	3	37.8	40.8	33	y8 y7 y6	941.5 827.5 728.4
		VLDAVRGSPAINVAVHVFRK	716.8	3	32.2	35.2	35	y8 y7 y6	955.6 856.5 785.5
		YTIAALLSPYSYSTTAVVTNPKE	830.4	3	46.1	49.1	35	y8 y7 y6	857.5 786.4 687.4

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P02788	Lactotransferrin	C[Cam]LRDGAGDVAFIR	483.9	3	26.4	29.4	25	y4 y5 y6	506.3 605.4 720.4
		FFSASC[Cam]VPGADKGQFPNLC[Cam]R	753.4	3	34.7	37.7	35	y8 y9 y5	991.5 1119.6 659.3
		IDSGLYLGSGYFTAIQNLR	1044.5	2	47.2	50.2	35	y10 y8 y7	1182.6 962.5 815.5
		LC[Cam]AGTGENKC[Cam]AFSSQEPYFSYSGAFK	969.4	3	35.3	38.3	35	y8 y7 y5	906.4 759.4 509.3
		LRPVAAEVYGTER	487.6	3	23.5	26.5	25	y5 y4 y6	625.3 462.2 724.4
		NLLFNDNTEC[Cam]LAR	790.4	2	33.3	36.3	30	y8 y7 y4	978.4 863.4 519.3
		VPPRIDSGLYLGSGYFTAIQNLR	846.5	3	46.6	49.6	35	y6 y7 y8	714.4 815.5 962.5
		VPSHAVVAR	468.3	2	10.5	13.5	19	y5 y6 y7	515.3 652.4 739.4
		VVWC[Cam]AVGEQELR	723.4	2	32.5	35.5	28	y7 y6 y8	830.4 731.4 901.5
		YYGYTGAFR	549.3	2	26.2	29.2	22	y5 y4 y6	551.3 450.2 714.4
P06396	Gelsolin	AGALNSNDAFVLK	660.4	2	30.6	33.6	26	y8 y7 y9	893.5 806.4 1007.5
		AQPVQVAEGSEPDGFWEALGGK	758.0	3	41.9	44.9	35	y10 y9 y7	1079.5 964.5 760.4
		ATEVPVSWESFNNGDC[Cam]FILDLGNNIHQWC[Cam]GSNSNRYER	1122.3	4	46.9	49.9	35	y8 y7 y5	1025.5 938.4 737.4
		EVQGFESATFLGYFK	861.9	2	44.9	47.9	33	y5 y6 y8	627.3 774.4 946.5
		HVVPNEVVVQR	425.9	3	19.5	22.5	22	y5 y4 y6	600.4 501.3 729.4
		IEGSNKVPVDPATYGQFYGGDSYIILYNYR	1133.9	3	43.2	46.2	35	y6 y5 y7	841.5 728.4 954.5
		NWRDPDQTDGLGLSYLSSHIANVER	711.6	4	38.3	41.3	34	y9 y6 y8	1012.5 701.4 925.5
		QTQVSVLPEGGETPLFK	915.5	2	36.9	39.9	34	y10 y8 y4	1074.5 848.5 504.3
		TPSAAYLWVGTGASEAEK	919.5	2	35.0	38.0	33	y11 y9 y10	1134.5 849.4 948.5
		TPSAAYLWVGTGASEAEKTGAQELLR	902.8	3	40.9	43.9	35	y7 y8 y9	786.4 887.5 1015.6

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P06727	Apolipoprotein A-IV	IDQNVEELKGR	434.2	3	18.4	21.4	22	y6 y5 y7	731.4 602.4 830.5
		ISASAEELR	488.3	2	19.2	22.2	20	y6 y5 y7	704.4 617.3 775.4
		ISASAEELRQR	420.6	3	17.1	20.1	22	y6 y5 y8	830.4 701.4 988.5
		KLVPFATELHER	480.6	3	24.6	27.6	24	y7 y8 y6	855.4 1002.5 784.4
		LAPLAEDVR	492.3	2	24.3	27.3	20	y7 y6 y5	799.4 702.4 589.3
		LAPLAEDVRGNLR	475.3	3	26.4	29.4	24	y9 y7 y6	1029.5 829.5 714.4
		LEPYADQLR	552.8	2	24.5	27.5	22	y6 y5 y7	765.4 602.3 862.4
		LVPFATELHER	437.9	3	27.9	30.9	23	y6 y7 y5	784.4 855.4 683.3
		QLTPYAQR	488.8	2	18.3	21.3	20	y4 y5 y6	537.3 634.3 735.4
		TQVNTQAEQLRR	481.9	3	15.6	18.6	25	y6 y5 y8	772.4 701.4 1001.5
P35542	Serum amyloid A-4	AYWDIM[Ox]ISNHQNSNR	932.9	2	30.4	33.4	35	y5 y4 y6	618.3 490.2 755.4
		AYWDIMISNHQNSNR	924.9	2	30.4	33.4	35	y6 y5 y4	755.4 618.3 490.2
		GNYDAAQR	447.7	2	9.4	12.4	19	y6 y5 y4	723.3 560.3 445.3
		GNYDAAQRGPGGVWAAK	573.3	3	25.0	28.0	29	y10 y8 y9	998.6 785.4 842.5
		GPGGVWAAKLISR	656.4	2	29.7	32.7	26	y7 y6 y5	758.5 687.5 616.4
		SFFKEALQGVGDM[Ox]GR	829.4	2	31.6	34.6	32	y9 y10 y7	948.5 1019.5 707.3
		SFFKEALQGVGDMGR	821.4	2	31.4	34.4	31	y10 y9 y7	1003.5 932.5 691.3
		SGKDPDRFRPDGLPK	422.0	4	17.8	20.8	22	y4 y5 y6	414.3 529.3 626.4
		YLYARGNYDAAQR	520.9	3	19.8	22.8	26	y8 y7 y6	894.4 837.4 723.3

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P61626	Lysozyme C	ATNYNAGDRSTDYGIFQINSR	788.4	3	31.1	34.1	35	y9 y6 y8	1097.6 764.4 934.5
		GISLANWM[Ox]C[Cam]LAK	690.3	2	38.0	41.0	27	y7 y5 y8	938.4 638.3 1009.5
		GISLANWMC[Cam]LAK	682.3	2	43.7	46.7	27	y7 y5 y8	922.4 622.3 993.5
		LGM[Ox]DGYR	414.2	2	18.4	21.4	17	y5 y4 y6	657.3 510.2 714.3
		LGMDGYR	406.2	2	17.6	20.6	17	y6 y5 y4	698.3 641.3 510.2
		STDYGIFQINSR	700.8	2	34.2	37.2	27	y8 y7 y6	934.5 877.5 764.4
		TPGAVNAC[Cam]HLSC[Cam]SALLQDNIAD AVAC[Cam]AK	976.5	3	37.2	40.2	35	y8 y7 y6	805.4 734.3 619.3
		TPGAVNAC[Cam]HLSC[Cam]SALLQDNIAD AVAC[Cam]AKR	771.6	4	35.1	38.1	35	y9 y8 y7	961.5 890.5 775.4
		YWC[Cam]NDGKTPGAVNAC[Cam]HLSC[Ca m]SALLQDNIADAVAC[Cam]AK	963.4	4	37.3	40.3	35	y8 y7 y6	805.4 734.3 619.3
		YWC[Cam]NDGKTPGAVNAC[Cam]HLSC[Ca m]SALLQDNIADAVAC[Cam]AKR	802.2	5	35.7	38.7	35	y8 y7 y9	890.5 775.4 961.5
Q08431	Lactadherin	EVTGIITQGAR	572.8	2	22.4	25.4	23	y6 y5 y8	645.4 532.3 815.5
		FELLGCELNGCANPLGLK	631.0	3	41.7	44.7	30	y11 y9 y7	1099.6 872.5 712.4
		IALRLELLGC[Cam]	579.3	2	35.0	38.0	23	y8 y6 y3	973.5 704.4 349.2
		LYPTSC[Cam]HTAC[Cam]TLR	527.3	3	21.6	24.6	26	y8 y6 y4	1018.5 721.4 549.3
		M[Ox]WVTGVVTQGASR	704.4	2	28.0	31.0	27	y8 y10 y6	817.5 975.5 619.3
		MWVTGVVTQGASR	696.4	2	28.0	31.0	27	y10 y8 y6	975.5 817.5 619.3
		NAVHVNLFETPVEAQYVR	696.0	3	36.5	39.5	34	y8 y6 y5	961.5 765.4 636.3
		NPCHNGGLCEEISQEVR	628.9	3	23.8	26.8	30	y9 y7 y5	1092.5 860.4 618.3
		TWGLHLFSWNPSYAR	612.3	3	40.4	43.4	30	y8 y6 y5	980.5 707.3 593.3
		VTFLGLQHWVPELAR	589.3	3	42.4	45.4	29	y7 y5 y6	870.5 585.3 684.4

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
Q15582	Transforming growth factor-beta-induced protein ig-h3	AIISNKDILATNGVIHYIDELLIPDSAK	759.9	4	48.5	51.5	35	y9 y8 y10	985.6 856.5 1100.6
		EGVYTVFAPTNEAFR	850.9	2	38.0	41.0	32	y9 y8 y7	1052.5 905.4 834.4
		GC[Cam]PAALPLSNLYETLGVVGSTTTQLY TDRTEK	1152.6	3	46.8	49.8	35	y7 y8 y6	912.4 1025.5 749.4
		GDELADSALEIFK	704.4	2	46.2	49.2	27	y7 y8 y6	807.5 922.5 720.4
		LTLLAPLNSVFK	658.4	2	44.0	47.0	26	y7 y8 y6	804.5 875.5 707.4
		LTLLAPLNSVFKDGTPPIDAHTR	619.8	4	40.8	43.8	31	y8 y7 y10	906.5 809.4 1064.5
		NHIIKDQLASK	422.9	3	14.3	17.3	22	y6 y7 y4	661.4 789.4 418.3
		QAGLGNHLSGSER	663.3	2	15.1	18.1	26	y9 y8 y6	956.5 899.4 648.3
		YGTLFTM[Ox]DR	560.3	2	30.6	33.6	22	y6 y5 y4	798.4 685.3 538.2
		YGTLFTMDR	552.3	2	31.9	34.9	22	y6 y5 y4	782.4 669.3 522.2
P0CG04, POCG05, P0CG06, P0CF74, A0M8Q6	Ig lambda-1 chain C region, Ig lambda-2 chain C region, Ig lambda-3 chain C region, Ig lambda-6 chain C region, Ig lambda-7 chain C region	ADGSPVKVGVETTKPSK	567.3	3	15.0	18.0	28	y9 y10 y6	946.5 1045.6 661.4
		ATLVC[Cam]LVSDFYPGAVTVAWK	733.1	3	55.1	58.1	35	y10 y8 y6	1091.6 831.5 703.4
		SHRSYSCR	498.2	2	5.2	8.2	20	y5 y7 y3	615.3 908.4 365.2
		SYSC[Cam]RVTHEGSTVEK	580.6	3	12.7	15.7	29	y9 y8 y6	987.5 886.4 620.3
		VGVETTKPSK	523.3	2	7.6	10.6	21	y6 y8 y3	661.4 889.5 331.2
		VGVETTKPSKQSNNK	539.6	3	5.7	8.7	27	y9 y7 y5	1030.6 805.4 590.3
		VTHEGSTVEK	543.8	2	5.9	8.9	22	y6 y8 y4	620.3 886.4 476.3
		VTHEGSTVEKTVAPAEC[Cam]S	476.2	4	15.0	18.0	23	y10 y8 y6	1091.5 834.4 634.2
		ADGSPVNTGVETTTPSK	554.3	3	20.0	23.0	27	y10 y8 y6	1020.5 862.5 634.3

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P0CG04, POCG05, P0CG06, P0CF74, A0M8Q6	Ig lambda-1 chain C region, Ig lambda-2 chain C region, Ig lambda-3 chain C region, Ig lambda-6 chain C region, Ig lambda-7 chain C region	ATLVC[Cam]LISDFYPGAVK	585.3	3	50.0	53.0	28	y10 y8 y6	1096.6 896.5 634.4
		ATLVC[Cam]LISDFYPGAVKVAWK	746.7	3	56.9	59.9	35	y9 y7 y5	955.6 801.5 631.4
		SYSC[Cam]QVTHEGSTVEKTVAPAEC[Cam]S	843.0	3	21.4	24.4	35	y10 y8 y6	1091.5 834.4 634.2
		VAWKADGSPVNTGVETTTPSK	537.0	4	25.3	28.3	27	y10 y8 y6	1020.5 862.5 634.3
		ADGSPVKAGVETTKPSK	558.0	3	13.8	16.8	28	y9 y6 y10	946.5 661.4 1017.6
		AGVETTKPSK	509.3	2	5.9	8.9	21	y6 y3 y8	661.4 331.2 889.5
		ANPTVTLFPPSSEELQANK	681.7	3	37.1	40.1	33	y9 y7 y5	1005.5 831.4 573.3
		ATLVC[Cam]LISDFYPGAVTVAWK	737.7	3	59.2	60.0	35	y5 y8 y9	604.3 831.5 928.5
		GQPKANPTVTLFPPSSEELQANK	614.1	4	33.6	36.6	30	y9 y7 y5	1005.5 831.4 573.3
		QSNNKYAASSYLSLTPEQWK	772.4	3	34.0	37.0	35	y8 y6 y7	988.5 788.4 901.5
		SHRSYSC[Cam]QVTHEGSTVEK	698.0	3	9.7	12.7	33	y10 y8 y6	1086.5 886.4 620.3
		SYSC[Cam]QVTHEGSTVEK	571.3	3	16.0	19.0	28	y8 y9 y6	886.4 987.5 620.3
		SYSC[Cam]QVTHEGSTVEKTVAPTEC[Cam]S	853.0	3	21.5	24.5	35	y5 y9 y7	593.2 992.5 763.3

Accession number	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Collision Energy (eV)	Type product ion	Product ion (m/z)
P0CG04, POCG05, P0CG06, P0CF74, A0M8Q6	Ig lambda-1 chain C region, Ig lambda-2 chain C region, Ig lambda-3 chain C region, Ig lambda-6 chain C region, Ig lambda-7 chain C region	YAASSYLSLTPEQWK	872.4	2	38.4	41.4	33	y8 y6 y5	988.5 788.4 687.3
		YAASSYLSLTPEQWKSHR	708.7	3	33.5	36.5	34	y8 y6 y4	1067.5 841.4 527.3
		SHKSYSC[Cam]QVTHEGSTVEK	688.7	3	10.0	13.0	33	y10 y8 y6	1086.5 886.4 620.3
		YAASSYLSLTPEQWKSHK	699.4	3	33.1	36.1	34	y8 y6 y4	1039.5 813.4 499.3
		AAPSVTLFPPSSEELQANK	662.7	3	35.9	38.9	32	y8 y7 y6	918.5 831.4 702.4
		ADSSPVKAGVETTTPSK	837.9	2	16.9	19.9	32	y9 y7 y10	919.5 763.4 990.5
		AGVETTTPSK	495.8	2	10.0	13.0	20	y7 y6 y8	763.4 634.3 862.5
		AGVETTTPSKQSNNK	781.4	2	8.4	11.4	30	y8 y7 y6	902.5 805.4 718.4
		AGVETTTPSKQSNNKYAASSYLSLTPEQWK	822.4	4	31.6	34.6	35	y8 y6 y5	988.5 788.4 687.3
		GQPKAAPSVTLFPPSSEELQANK	799.4	3	34.1	37.1	35	y8 y6 y7	918.5 702.4 831.4

Annexe 7 : Tableau des transitions choisies pour les expériences tPRM.

Accession number	Accession name	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)	product ion 4 (m/z)	product ion 5 (m/z)
P01034	CYTC_HUMAN	Cystatin C	ALDFAVGEYNK	613.81	2	30.7	34.7	y5,y6,y3,y7,y3	610.28	709.35	424.22	780.39	212.61
			ALDFAVGEYNKASNDM[Ox]YHSR	576.76	4	26.4	30.4	y7,y5,y9,y16,y3	235.35	355.16	1096.45	465.21	399.21
			ALDFAVGEYNKASNDMYHSR	572.77	4	29.5	33.5	y3,y4,y9,y8,y5	399.21	562.27	1080.45	1009.42	693.31
			LVGGPM[Ox]DASVEEEGVR	830.90	2	26.6	30.6	y7,y8,y9,y10,y6	817.40	904.44	975.47	1090.50	718.34
			LVGGPM[Ox]DASVEEEGVRR	606.30	3	21.1	25.1	y4,y5,y6,y7,y8	487.31	616.35	745.39	874.44	973.51
			LVGGPMDASVEEEGVR	822.90	2	30.8	34.8	y3,y8,y9,y10,y6	331.21	904.44	975.47	1090.50	718.34
			LVGGPMDASVEEEGVRR	600.97	3	24.8	28.8	y15,y10,y16,y11,y13	794.87	566.29	844.41	623.80	737.85
			TQPNLDNC[Cam]PFHDQPHLKR	555.02	4	21.1	25.1	y5,y6,y7,y13,y8	650.41	778.47	893.49	416.70	515.78
P01834	IGKC_HUMAN	Ig kappa chain C region	HKVYAC[Cam]EVTHQGLSSPVTK	714.37	3	18.5	22.5	y5,y4,y8,y3,y7	266.16	444.28	788.45	347.23	366.22
			SGTASVVC[Cam]LLNNFYPR	899.45	2	44.9	48.9	y9,y10,y6,y7,y8	1196.59	1295.66	810.39	923.47	1036.56
			TVAAPSVFIFPPSDEQLK	973.52	2	41.1	45.1	y8,y9,y7,y11,y4	913.46	1060.53	816.41	1320.68	517.30
			VDNALQSGNSQESVTEQDSK	712.66	3	19.4	23.4	y6,y8,y9,y5,y7	707.32	893.42	1022.46	606.27	806.39
			VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK	1207.24	3	32.8	36.8	y11,y8,y4,y10,y7	400.56	836.47	448.28	1036.59	749.44
			VQWKVDNALQSGNSQESVTEQDSKDSTYSLSST LTLSK	1041.01	4	34.7	38.7	y8,y4,y10,y7,y20	836.47	448.28	1036.59	749.44	1096.03
			VYAC[Cam]EVTHQGLSSPVTK	625.98	3	22.8	26.8	y4,y8,y8,y6,y10	444.28	788.45	263.49	206.79	1053.57
			VYAC[Cam]EVTHQGLSSPVTKSFNRGEC[Cam]	909.43	3	25.6	29.6	y7,y9,y11,y8,y6	869.36	1098.50	1294.62	997.45	782.32
P01857, P01859,P01860, P01861	IGHG1_HUMAN , IGHG2_HUMAN, IGHG3_HUMAN, IGHG4_HUMAN	Ig gamma-1 chain C, Ig gamma-2 chain C, Ig gamma-3 chain C, Ig gamma-4 chain C	FNWYVDGVEVHNAK	559.94	3	31.2	35.2	y4,y6,y4,y9,y3	469.25	697.36	235.13	968.48	332.19
			GFYPSDIAVEWESNGQPENNYK	848.72	3	39.9	43.9	y6,y4,y10,y3,y6	764.36	538.26	1150.51	424.22	382.68
			GPSVFPLAPSSK	593.83	2	30.1	34.1	y4,y7,y5,y8,y3	418.23	699.40	489.27	846.47	321.18
			SC[Cam]DKTHTC[Cam]PPC[Cam]PAPELLGGPS VFLFPPKPK	667.73	5	36.6	40.6	y5,y6,y8,y13,y14	566.37	713.43	973.59	1370.78	495.29
			SC[Cam]DKTHTC[Cam]PPC[Cam]PAPELLGGPS VFLFPPKPK	834.42	4	36.6	40.6	y5,y6,y8,y13,y25	566.37	713.43	973.59	1370.78	1372.21
			THTC[Cam]PPC[Cam]PAPELLGGPSVFLFPPKPK	711.87	4	38.4	42.4	y5,y13,y25,y6,y8	566.37	1370.78	1372.21	713.43	244.15
			TPEVTC[Cam]VVVDVSHEDPEVK	1070.02	2	31.6	35.6	y4,y7,y12,y9,y11	472.28	427.21	1352.67	1039.50	1253.60
			TTPPVLDSDGSFFLYSK	937.46	2	40.6	44.6	y16,y3,y15,y12,y10	886.94	397.21	836.42	1378.65	1150.54
			ALPAPIEKTISK	423.26	3	23.8	27.8	y8,y6,y7,y8,y4	458.28	705.41	818.50	915.55	448.28
			NQVSLTC[Cam]LVK	581.32	2	29.2	33.2	y7,y5,y8,y5,y4	820.46	620.34	919.53	310.68	519.30
			STSGGTAALGC[Cam]LVK	661.34	2	26.8	30.8	y7,y5,y6,y5,y8	760.44	576.32	689.40	288.66	831.48
			VSNKALPAPIEK	422.92	3	18.5	22.5	y4,y6,y5,y5,y4	486.29	654.38	557.33	186.45	243.65
			VVSVLTVLHQDWLNGK	603.34	3	42.1	46.1	y3,y3,y8,y7,y9	159.59	318.18	997.48	860.43	1110.57
			VVSVLTVLHQDWLNGKEYK	743.41	3	38.9	42.9	y11,y9,y8,y6,y10	1417.69	1152.57	1037.54	738.38	1280.63

Accession number	Accession name	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)	product ion 4 (m/z)	product ion 5 (m/z)
P01871, P04220	IGHM_HUMAN, MUCB_HUMAN	Ig mu chain C region , Ig mu heavy chain disease protein	EGKQVGSGVTTDQVQAEAK	644.66	3	17.6	21.6	y6,y8,y5,y4,y9	645.36	888.44	546.29	418.23	989.49
			EGKQVGSGVTTDQVQAEAKESGPTTYK	699.60	4	21.3	25.3	y12,y8,y7,y10,y11	1281.63	882.42	753.38	1081.55	1210.59
			GVALHRPDVYLLPPAR	592.01	3	28.8	32.8	y4,y14,y5,y10,y12	440.26	809.46	553.35	1140.64	717.40
			NVPLPVIAELPPK	693.92	2	38.3	42.3	y3,y4,y9,y7,y3	341.22	227.65	963.59	767.47	171.11
			NVPLPVIAELPPKVSVFVPPR	756.79	3	43.9	47.9	y11,y11,y3,y10,y12	1222.73	611.87	369.22	1125.68	1335.81
			QIQVSWLR	515.30	2	32.6	36.6	y4,y6,y5,y3,y7	561.31	788.44	660.38	474.28	901.52
			QVGSGVTTDQVQAEAK	809.41	2	19.8	23.8	y10,y5,y4,y14,y8	1090.54	546.29	418.23	1390.68	888.44
			QVGSGVTTDQVQAEAKESGPTTYK	827.74	3	23.2	27.2	y12,y10,y7,y8,y13	1281.63	1081.55	753.38	882.42	1409.69
			FTC[Cam]TVTHTDLPSPLK	572.96	3	28.0	32.0	y3,y5,y8,y6,y9	357.25	541.33	870.49	654.42	1007.55
			GVALHRPDVYLLPPAR	592.01	3	29.0	33.0	y4,y14,y5,y10,y12	440.26	809.46	553.35	1140.64	717.40
			LIC[Cam]QATGFSPR	625.32	2	24.9	28.9	y7,y9,y6,y8,y5	735.38	1023.47	664.34	863.44	563.29
			SKLIC[Cam]QATGFSPR	732.89	2	26.8	30.8	y5,y6,y8,y9,y10	563.29	664.34	863.44	1023.47	1136.55
			VFAIPPSFASIFLTK	819.47	2	51.4	55.4	y11,y11,y10,y9,y4	1207.67	604.34	1110.62	1013.57	508.31
			VSVFVPPRDGFFGNPR	597.65	3	34.5	38.5	y11,y13,y12,y6,y15	630.32	753.39	679.85	246.46	846.44
P02647	APOA1_HUMAN	Apolipoprotein A-I	AHVDALRTHLAPYSDELRQR	587.81	4	24.5	28.5	y18,y16,y17,y4,y5	714.05	642.68	681.02	191.46	351.21
			ATEHLSTLSEK	608.31	2	15.8	19.8	y6,y7,y3,y8,y8	664.35	777.43	363.19	457.75	914.49
			LAARLEALKENGGAR	523.63	3	19.7	23.7	y6,y5,y4,y4,y5	603.28	474.24	180.60	360.20	158.75
			LLDNWDSVTSTFSK	806.90	2	39.2	43.2	y6,y9,y8,y5,y10	670.34	971.47	856.44	569.29	1157.55
			THLAPYSDELR	651.33	2	22.1	26.1	y7,y8,y9,y3,y9	879.42	950.46	1063.54	417.25	532.27
			THLAPYSDELRQR	793.41	2	20.0	24.0	y5,y9,y5,y11,y10	351.21	1163.58	701.40	674.35	1234.62
			VKDLATVYVDVLK	488.29	3	34.2	38.2	y6,y4,y5,y3,y6	736.42	474.29	573.36	359.26	246.15
			VKDLATVYVDVLKDSGR	626.68	3	34.5	38.5	y10,y8,y3,y9,y6	1151.61	889.47	319.17	988.54	675.38
P02649	APOE_HUMAN	Apolipoprotein E	AATVGSLAGQPLQER	749.40	2	25.9	29.9	y7,y8,y11,y5,y10	827.44	898.47	1155.61	642.36	1098.59
			AKLEEQAQQIR	438.58	3	16.6	20.6	y3,y4,y5,y4,y9	416.26	544.32	615.36	182.11	372.20
			LGADM[Ox]EDVC[Cam]GR	619.76	2	18.0	22.0	y4,y5,y6,y3,y7	491.24	606.27	735.31	392.17	882.34
			LGADMEDVC[Cam]GR	611.76	2	23.3	27.3	y4,y5,y6,y7,y10	491.24	606.27	735.31	866.35	1109.43
			LGPLVEQGR	484.78	2	21.9	25.9	y5,y4,y6,y3,y8	588.31	489.24	701.39	360.20	855.47
			LGPLVEQGRVR	612.36	2	20.5	24.5	y10,y9,y7,y6,y5	555.82	527.31	843.48	744.41	615.37
			LQAEAFQAR	517.27	2	20.9	24.9	y7,y5,y6,y4,y3	792.40	592.32	721.36	521.28	374.21
			LSKELQAAQAR	405.57	3	15.1	19.1	y5,y4,y3,y7,y6	516.29	445.25	374.21	757.43	215.45
P02652	APOA2_HUMAN	Apolipoprotein A-II	AGTELVNFLSYFVELGTQPATQ	795.74	3	51.8	55.8	y6,y8,y7,y10,y9	645.32	815.43	702.34	1043.54	944.47
			EPC[Cam]VESLVSQYFQTVTDYGK	784.04	3	41.9	45.9	y8,y7,y5,y3,y9	911.45	783.39	583.27	367.20	1058.51
			KAGTELVNFLSYFVELGTQPATQ	838.44	3	47.3	51.3	y5,y6,y7,y8,y9	544.27	645.32	702.34	815.43	944.47
			QAKEPC[Cam]VESLVSQYFQTVTDYGK	893.10	3	38.1	42.1	y8,y9,y10,y11,y12	911.45	1058.51	1221.58	1349.64	1436.67
			SKEQLTPLIK	578.85	2	21.5	25.5	y4,y5,y8,y8,y5	470.33	571.38	941.57	471.29	286.19
			SKEQLTPLIKK	428.93	3	17.1	21.1	y5,y3,y10,y5,y4	598.43	388.29	599.38	299.72	501.38
			SPELQAEAK	486.75	2	14.2	18.2	y4,y6,y7,y5,y8	418.23	659.37	788.41	546.29	885.47
			VKSPELQAEAK	600.34	2	15.2	19.2	y8,y7,y6,y5,y9	885.47	788.41	659.37	546.29	972.50

Accession number	Accession name	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)	product ion 4 (m/z)	product ion 5 (m/z)
P02671	FIBA_HUMAN	Fibrinogen alpha chain	GDFSSANNRDNTYNR	577.59	3	15.2	19.2	y12,y11,y10,y5,y13	706.32	662.80	619.29	667.32	779.85
			GGSTSYGTGSETESPR	786.84	2	15.5	19.5	y10,y11,y8,y9,y12	1020.46	1183.52	862.39	963.44	1270.55
			GLIDEVNQDFTNR	760.87	2	35.7	39.7	y7,y10,y4,y8,y9	894.41	1237.54	537.28	993.47	1122.52
			GLIDEVNQDFTNRINK	625.99	3	35.4	39.4	y7,y8,y10,y6,y9	892.50	1007.53	1249.63	745.43	1135.58
			M[Ox]ELERPGGNEITR	506.58	3	17.9	21.9	y8,y11,y10,y9,y3	843.43	621.33	564.79	500.27	389.25
			MELERPGGNEITR	501.25	3	21.2	25.2	y11,y8,y12,y4,y9	621.33	843.43	685.85	259.65	500.27
			TFPGFFSPM[Ox]LGEFVSETESR	761.02	3	51.0	55.0	y10,y8,y11,y6,y4	1140.52	954.45	1253.60	708.32	492.24
			TFPGFFSPMLGEFVSETESR	755.69	3	52.7	56.7	y6,y7,y10,y8,y4	708.32	807.38	1140.52	954.45	492.24
P02743	SAMP_HUMAN	Serum amyloid P-component	AYSLFSYNTQGR	703.84	2	31.8	35.8	y7,y8,y10,y6,y5	825.38	972.45	1172.57	738.35	575.29
			AYSLFSYNTQGRDNELLVYKER	667.34	4	33.7	37.7	y17,y5,y4,y15,y18	1043.02	694.39	595.32	917.97	1116.56
			DNELLVYK	497.27	2	28.2	32.2	y4,y3,y5,y6,y6	522.33	409.24	635.41	764.45	382.73
			DNELLVYKER	639.84	2	24.1	28.1	y6,y5,y4,y7,y5	807.47	694.39	595.32	920.56	347.70
			IVLGQEQDSYGGK	697.35	2	22.5	26.5	y4,y5,y10,y3,y6	212.61	511.25	1068.46	261.16	626.28
			IVLGQEQDSYGGKFDR	604.64	3	25.1	29.1	y6,y8,y9,y10,y7	679.35	929.45	1044.47	1172.53	842.42
			QGYFVEAQPK	583.80	2	23.1	27.1	y5,y6,y3,y4,y7	572.30	671.37	186.62	443.26	818.44
			VGEYSLYIGR	578.80	2	30.4	34.4	y6,y9,y7,y4,y5	708.40	1057.53	871.47	508.29	621.37
P02766	TTHY_HUMAN	Transthyretin	AADDTWEPFASGK	697.81	2	31.6	35.6	y6,y3,y4,y7,y8	606.32	291.17	362.20	735.37	921.45
			ALGISPFHEHAEVVFTANDSGPR	613.56	4	31.6	35.6	y6,y4,y7,y5,y8	645.29	416.22	716.33	531.25	817.38
			ALGISPFHEHAEVVFTANDSGPRR	652.58	4	29.0	33.0	y15,y5,y16,y17,y20	828.42	286.67	892.94	961.47	751.70
			GSPAINVAVHVFR	456.26	3	30.5	34.5	y6,y4,y8,y5,y4	728.42	558.31	941.53	657.38	279.66
			GSPAINVAVHVFRK	498.96	3	26.5	30.5	y5,y7,y4,y6,y5	686.41	856.51	549.35	785.48	343.71
			KAADDTWEPFASGK	508.24	3	26.6	30.6	y6,y3,y4,y5,y7	606.32	291.17	362.20	509.27	245.79
			RYTIAALLSPYSYSTTAVVTNPKE	882.47	3	40.7	44.7	y3,y6,y7,y4,y5	373.21	687.37	786.44	487.25	588.30
			YTIAALLSPYSYSTTAVVTNPKE	830.43	3	45.2	49.2	y3,y7,y6,y4,y5	373.21	786.44	687.37	487.25	588.30
P02788	TRFL_HUMAN	Lactotransferrin	DGAGDVAFIR	510.76	2	28.7	32.7	y5,y7,y4,y6,y3	605.38	777.42	506.31	720.40	435.27
			FFSASC[Cam]VPGADKGQFPNLC[Cam]R	753.35	3	33.3	37.3	y13,y5,y13,y8,y9	1459.71	659.33	730.36	991.48	1119.57
			IDSGLYLGSGYFTAIQNLR	1044.54	2	46.5	50.5	y12,y13,y7,y11,y8	1326.68	1439.76	815.47	1269.66	962.54
			NLLFNDNTEC[Cam]LAR	790.38	2	33.4	37.4	y10,y9,y7,y11,y8	1239.54	1092.47	863.40	1352.63	978.43
			SNLC[Cam]ALC[Cam]IGDEQGENKC[Cam]VPN SNER	922.07	3	26.5	30.5	y6,y8,y7,y12,y10	716.33	975.43	815.40	1403.63	1217.57
			VPPRIDSGLYLGSGYFTAIQNLR	846.46	3	43.5	47.5	y8,y7,y12,y6,y5	962.54	815.47	1326.68	714.43	643.39
			VVWC[Cam]AVGEQELR	723.36	2	31.7	35.7	y4,y6,y8,y9,y7	273.16	731.37	901.47	1061.50	830.44
			YYGYTGAFR	549.26	2	27.2	31.2	y7,y5,y8,y4,y6	771.38	551.29	934.44	450.25	714.36
P06727	APOA4_HUMAN	Apolipoprotein A-IV	IDQNVEELKGR	434.23	3	18.3	22.3	y6,y6,y5,y8,y4	731.40	244.47	602.36	944.52	473.32
			ISASAEELRQR	630.34	2	17.3	21.3	y10,y5,y9,y8,y4	573.80	351.21	530.28	494.76	286.68
			LAPLAEDVR	492.28	2	24.1	28.1	y5,y6,y7,y7,y4	589.29	702.38	799.43	400.22	518.26
			LAPLAEDVRGNLR	475.27	3	26.3	30.3	y6,y10,y9,y11,y4	357.72	571.82	515.28	620.34	459.27
			LEPYADQLR	552.79	2	24.3	28.3	y5,y6,y7,y3,y4	602.33	765.39	862.44	416.26	531.29
			LVPFATELHER	437.91	3	28.0	32.0	y4,y7,y6,y4,y3	185.44	855.43	784.39	554.30	441.22
			QLTPYAQR	488.76	2	18.3	22.3	y5,y4,y6,y3,y5	634.33	537.28	735.38	374.21	317.67
			TQVNTQAEQLRR	481.93	3	15.6	19.6	y10,y9,y8,y3,y7	607.83	558.30	501.28	222.66	450.75

Accession number	Accession name	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)	product ion 4 (m/z)	product ion 5 (m/z)
P0CG04, P0CG05, P0CG06, P0CF74, A0M8Q6	LAC1_HUMAN, LAC2_HUMAN, LAC3_HUMAN, LAC6_HUMAN, LAC7_HUMAN	Ig lambda-1 chain C region, Ig lambda-2 chain C region, Ig lambda-3 chain C region, Ig lambda-6 chain C region, Ig lambda-7 chain C region	AAPSVTLFPPSSEELQANK	662.68	3	34.5	38.5	y10,y11,y10,y9,y5	1102.54	600.30	368.18	1005.48	573.34
			ADGSPVKAGVETTKPSK	418.73	4	13.6	17.6	y6,y6,y5,y11,y10	331.20	661.39	187.45	573.33	1017.56
			ADGSPVKVGVETTKPSK	567.31	3	15.4	19.4	y6,y7,y8,y9,y10	661.39	790.43	889.50	946.52	1045.59
			ADGSPVNTGVETTTPSK	554.27	3	20.0	24.0	y6,y7,y8,y9,y10	634.34	763.38	862.45	919.47	1020.52
			ADSSPAKAGVETTTPSK	549.61	3	15.9	19.9	y6,y7,y8,y9,y10	634.34	763.38	862.45	919.47	990.51
			ADSSPVKAGVETTTPSK	558.96	3	16.5	20.5	y3,y10,y6,y9,y11	331.20	990.51	634.34	919.47	559.81
			AGVETTKPSK	509.28	2	6.8	10.8	y3,y4,y5,y6,y7	331.20	459.29	560.34	661.39	790.43
			AGVETTTPSK	495.76	2	10.7	14.7	y3,y6,y7,y5,y8	331.20	634.34	763.38	533.29	862.45
			ANPTVTLFPPSSEELQANK	681.68	3	36.1	40.1	y6,y8,y10,y11,y7	702.38	918.45	1102.54	1199.59	831.42
			ATLVC[Cam]LISDFYPGAVK	877.46	2	48.2	52.2	y5,y6,y7,y8,y9	471.29	634.36	781.42	896.45	983.48
			ATLVC[Cam]LISDFYPGAVKVAWK	746.74	3	54.8	58.8	y6,y7,y8,y9,y10	730.46	801.50	858.52	955.57	1118.64
			ATLVC[Cam]LISDFYPGAVTVAWK	737.72	3	56.9	60.0	y9,y8,y7,y6,y10	928.52	831.47	774.45	703.41	1091.59
			ATLVC[Cam]LVSDFYPGAVTVAWK	733.05	3	54.2	58.2	y6,y7,y8,y9,y10	703.41	774.45	831.47	928.52	1091.59
			QSNNKYAASSYLSLTPEQWK	772.38	3	33.5	37.5	y5,y6,y8,y7,y9	687.35	788.39	988.51	901.48	1101.59
			SHKSYSC[Cam]QVTHEGSTVEK	688.65	3	10.7	14.7	y6,y7,y8,y9,y10	620.32	749.37	886.43	987.47	1086.54
			SHRSYSC[Cam]QVTHEGSTVEK	697.99	3	10.4	14.4	y5,y6,y7,y8,y9	563.30	620.32	749.37	886.43	987.47
			SYSC[Cam]QVTHEGSTVEK	571.26	3	16.0	20.0	y14,y9,y6,y7,y7	542.25	987.47	620.32	749.37	250.46
			SYSC[Cam]QVTHEGSTVEKTVAPAEC[Cam]S	843.04	3	23.7	27.7	y5,y10,y14,y6,y9	563.21	1091.50	1435.67	634.25	962.46
			SYSC[Cam]QVTHEGSTVEKTVAPTEC[Cam]S	853.04	3	23.7	27.7	y5,y4,y5,y6,y22	593.22	248.59	198.41	664.26	1235.55
			SYSC[Cam]RVTHEGSTVEK	580.61	3	13.1	17.1	y5,y6,y7,y8,y9	563.30	620.32	749.37	886.43	987.47
			SYSC[Cam]RVTHEGSTVEKTVAPAEC[Cam]S	852.39	3	18.2	22.2	y6,y7,y8,y9,y10	634.25	733.32	834.37	962.46	1091.50
			VGVETTKPSK	523.30	2	8.4	12.4	y4,y5,y6,y7,y8	459.29	560.34	661.39	790.43	889.50
			VTHEGSTVEK	543.77	2	6.8	10.8	y4,y5,y6,y7,y8	476.27	563.30	620.32	749.37	886.43
			YAASSYLSLTPEQWK	872.43	2	37.8	41.8	y5,y8,y6,y9,y7	687.35	988.51	788.39	1101.59	901.48
P0DJI8, P0DJI9	SAA1_HUMAN, SAA2_HUMAN	Serum amyloid A-1, Serum amyloid A-2	AYSDM[Ox]REANYIGSDK	868.39	2	18.6	22.6	y4,y7,y10,y3,y3	406.19	796.38	1152.56	175.09	349.17
			AYSDM[Ox]REANYIGSDKYFHAR	603.28	4	23.3	27.3	y6,y9,y5,y4,y6	206.12	1080.52	693.35	530.28	821.44
			AYSDMREANYIGSDK	573.93	3	22.2	26.2	y4,y11,y13,y3,y5	406.19	642.31	743.34	349.17	519.28
			AYSDMREANYIGSDKYFHAR	599.28	4	25.8	29.8	y6,y9,y5,y4,y6	206.12	1080.52	693.35	530.28	821.44
			EANYIGSDKYFHAR	557.60	3	22.5	26.5	y5,y4,y9,y9,y3	693.35	530.28	1080.52	540.76	383.21
			EANYIGSDKYFHARGNYDAAK	797.38	3	22.4	26.4	y16,y3,y19,y11,y10	900.43	289.19	730.68	1249.61	1102.54
			SFFSFLGEAFDGARDM[Ox]WR	718.99	3	49.8	53.8	y7,y12,y4,y13,y15	454.21	713.81	208.42	770.35	887.40
			SFFSFLGEAFDGARDMWR	713.66	3	53.4	57.4	y7,y12,y3,y15,y13	446.22	705.81	492.24	879.41	762.36
			GNYDAAKRGPGGAWAAEVISNAR	583.55	4	29.4	33.4	y4,y5,y6,y7,y3	447.23	560.31	659.38	788.43	360.20
			GNYDAAKRGPGGVWAAEAISDAR	583.79	4	30.0	34.0	y5,y4,y3,y14,y9	561.30	448.21	361.18	700.35	226.62
			GPGGAWAAEVISNAR	485.92	3	34.2	38.2	y4,y5,y6,y7,y3	447.23	560.31	659.38	788.43	360.20
			GPGGVWAAEAISDAR	728.86	2	33.0	37.0	y10,y9,y8,y6,y7	1089.53	903.45	832.42	632.34	761.38
			RGPGGAWAAEVISNAR	537.95	3	29.2	33.2	y5,y4,y6,y7,y8	560.31	447.23	659.38	788.43	859.46
			RGPGGVWAAEAISDAR	538.28	3	28.2	32.2	y6,y4,y5,y7,y8	632.34	448.21	561.30	761.38	832.42
P35542	SAA4_HUMAN	Serum amyloid A-4	AYWDIM[Ox]ISNHQNSNR	932.92	2	29.9	33.9	y9,y8,y7,y6,y10	1069.51	956.43	869.40	755.35	1216.55
			AYWDIMISNHQNSNR	924.93	2	29.9	33.9	y9,y10,y8,y7,y6	1069.51	1200.55	956.43	869.40	755.35
			GNYDAAQR	447.71	2	10.1	14.1	y3,y4,y5,y6,y7	374.21	445.25	560.28	723.34	837.38
			GNYDAAQRGPGGVWAAK	859.42	2	24.8	28.8	y11,y10,y9,y8,y4	1126.61	998.55	842.45	785.43	475.27
			GPGGVWAAK	421.73	2	20.7	24.7	y4,y5,y6,y7,y8	475.27	574.33	631.36	688.38	785.43
			SFFKEALQGVGDM[Ox]GR	829.40	2	30.8	34.8	y11,y10,y9,y7,y6	1148.54	1019.49	948.46	707.31	650.29
			SFFKEALQGVGDMGR	821.41	2	30.8	34.8	y10,y11,y9,y7,y3	1003.50	1132.54	932.46	691.32	363.18
			SGKDPDRFRPDGLPK	421.97	4	18.2	22.2	y13,y12,y10,y11,y12	514.28	471.58	600.83	433.24	706.87

Accession number	Accession name	Protein name	Peptides	Precurseur ion (m/z)	Charge State	Start RT (min)	End RT (min)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)	product ion 4 (m/z)	product ion 5 (m/z)
P61626	LYSC_HUMAN	Lysozyme C	ATNYNAGDRSTDYGIFQINSR	788.37	3	30.6	34.6	y6,y9,y17,y5,y8	764.40	1097.57	957.46	617.34	934.51
			C[Cam]QNRDVRQYVQGC[Cam]GV	613.62	3	22.7	26.7	y5,y13,y3,y3,y6	520.22	775.88	168.07	335.14	619.29
			DVRQYVQGC[Cam]GV	640.81	2	27.9	31.9	y4,y3,y3,y4,y5	196.58	335.14	168.07	392.16	520.22
			GISLANWM[Ox]C[Cam]LAK	690.34	2	37.5	41.5	y4,y8,y5,y7,y6	491.26	1009.46	638.30	938.42	824.38
			GISLANWMC[Cam]LAK	682.35	2	42.9	46.9	y4,y5,y8,y6,y3	491.26	622.30	993.46	808.38	331.23
			STDYGIFQINSR	700.84	2	33.4	37.4	y6,y8,y5,y9,y7	764.40	934.51	617.34	1097.57	877.49
			TPGAVNAC[Cam]HLSC[Cam]SALLQDNIADAV										
			AC[Cam]AK	976.47	3	36.7	40.7	y8,y4,y3,y6,y11	805.39	449.22	378.18	619.32	1147.54
			VVRDPQGIR	520.30	2	13.6	17.6	y5,y8,y3,y6,y7	570.34	470.77	345.22	685.36	421.24
P61769	B2MG_HUMAN	Beta-2-microglobulin	DWSFYLLYYTEFTPTEK	735.01	3	51.1	55.1	y6,y7,y8,y4,y5	722.37	851.41	952.46	474.26	575.30
			DWSFYLLYYTEFTPTEKDEYAC[Cam]R	999.78	3	50.3	54.3	y8,y7,y6,y5,y9	1070.46	941.41	813.32	698.29	1171.50
			IEKVEHSDLSFSK	759.89	2	13.7	17.7	y10,y8,y9,y7,y5	1148.56	920.45	1049.49	783.39	581.33
			IQVYSRHPAENGK	500.26	3	13.3	17.3	y6,y3,y3,y9,y11	615.31	318.18	159.59	995.50	629.32
			SNFLNC[Cam]YVSGFHPSDIEVDLLK	852.08	3	45.0	49.0	y10,y3,y11,y4,y18	1128.61	373.28	1265.67	488.31	1047.00
			SNFLNC[Cam]YVSGFHPSDIEVDLLKNGER	753.36	4	43.5	47.5	y12,y10,y8,y6,y9	1400.74	1172.63	944.52	716.40	1043.58
			VEHSDLSFSK	574.78	2	17.1	21.1	y4,y5,y6,y7,y8	468.24	581.33	696.36	783.39	920.45
			VNHVTLSQPK	561.82	2	15.0	19.0	y6,y7,y8,y4,y3	673.39	772.46	455.26	459.26	186.62
Q08431	MFGM_HUMAN	Lactadherin	EVTGIITQGAR	572.82	2	22.3	26.3	y5,y6,y7,y8,y9	532.28	645.37	758.45	815.47	916.52
			FELLGC[Cam]ELNGC[Cam]ANPLGLK	1003.00	2	40.4	44.4	y11,y10,y8,y5,y9	1156.61	1043.53	872.47	527.35	929.49
			M[Ox]WVTGVVTQGASR	704.36	2	26.0	30.0	y6,y7,y8,y9,y10	619.32	718.38	817.45	874.47	975.52
			MWVTGVVTQGASR	696.36	2	26.1	30.1	y8,y9,y10,y7,y6	817.45	874.47	975.52	718.38	619.32
			NAVHVNLFETPVEAQYVR	696.03	3	32.1	36.1	y8,y7,y6,y5,y4	961.51	864.46	765.39	636.35	565.31
			NLFETPILAR	587.33	2	33.9	37.9	y5,y6,y7,y8,y4	569.38	670.42	799.47	946.54	472.32
			TWGLHLFSWNPSYAR	917.95	2	39.3	43.3	y9,y8,y7,y10,y5	1127.53	980.46	893.43	1240.61	593.30
			VTFLGLQHWVPELAR	589.33	3	42.6	46.6	y5,y7,y6,y8,y3	585.34	870.48	684.40	1007.54	359.24
Q15582	BGH3_HUMAN	Transforming growth factor-beta-induced protein ig-h3	AIISNKDILATNGVIHYIDELLIPDSAK	759.92	4	51.0	55.0	y5,y19,y6,y27,y7	517.26	518.03	630.35	742.16	743.43
			EGVYTVFAPTNEAFR	850.92	2	37.4	41.4	y9,y7,y8,y10,y11	1052.52	834.41	905.45	1151.58	1252.63
			GDELADSALEIFK	704.35	2	42.0	46.0	y7,y8,y7,y9,y4	404.23	922.49	807.46	993.52	536.31
			LLGDAKELANILK	466.62	3	36.2	40.2	y5,y7,y6,y5,y9	186.79	800.49	671.44	558.36	500.31
			LTLLAPLNSVFK	658.40	2	45.0	49.0	y7,y8,y9,y6,y10	804.46	875.50	988.58	707.41	1101.67
			LTLLAPLNSVFKDGTPPIDAHTR	619.84	4	40.1	44.1	y8,y7,y8,y4,y5	906.48	809.43	453.74	484.26	599.29
			NHIIKDQLASK	422.91	3	14.5	18.5	y3,y6,y4,y7,y7	305.18	661.35	418.27	789.45	395.23
			QAGLGNHLSGSER	442.56	3	15.2	19.2	y5,y6,y10,y4,y7	535.25	648.33	535.27	448.21	393.20
P06396	GELS_HUMAN	Gelsolin	AGALNSNDAFVLK	660.35	2	30.0	34.0	y9,y8,y4,y3,y7	1007.52	893.47	506.33	359.26	403.72
			AQPVQVAEGSEPDGFWEALGGK	758.03	3	41.0	45.0	y11,y3,y7,y5,y4	1176.57	261.16	760.40	445.28	187.62
			EVQGFESATFLGYFK	861.92	2	44.1	48.1	y4,y5,y5,y9,y7	514.27	314.18	627.35	1033.53	875.47
			HVVPNEVVVQR	425.91	3	19.3	23.3	y4,y3,y5,y9,y9	501.31	402.25	600.38	520.30	347.20
			NWRDPDQTDGLGLSYLSSHIANVER	711.60	4	37.8	41.8	y9,y5,y10,y6,y8	1012.52	588.31	1125.60	701.39	925.48
			QTQVSVLPEGGETPLFK	915.49	2	36.2	40.2	y10,y4,y11,y13,y5	1074.55	504.32	1187.63	1373.73	605.37
			TPSAAYLWVGTGASEAEK	919.45	2	34.3	38.3	y9,y10,y11,y5,y3	849.39	948.46	1134.54	563.27	347.19
			TPSAAYLWVGTGASEAEKTGAQELLR	902.80	3	40.0	44.0	y8,y7,y11,y17,y9	887.49	786.45	1215.67	859.44	1015.59

Annexe 8 : Tableau des transitions choisies pour les expériences tPRM pour la comparaison des instruments Q Exactive et Q Exactive HF (Thermo Scientific)

Accession nu	Accession name	Protein name	Peptides	Precursor ion (m/z)	Charge State	Start RT Qex (min)	End RT Qex (min)	Start RT QexHF (min)	End RT QexHF (min)	Type product ions	product ion 1 (m/z)	product ion 2 (m/z)	product ion 3 (m/z)	product ion 4 (m/z)	product ion 5 (m/z)	product ion 6 (m/z)
P02647	APOA1_HUMAN	Apolipoprotein A-I	ATEHLSTLSEK	608.314	2	12.8	16.8	22.2	26.2	y3,y6,y7,y8,b3,y8	363.187	664.351	777.435	914.494	439.193	457.751
			THLAPYSDELR	651.328	2	20.9	24.9	31.1	35.1	y3,y7,y8,y9,b2,b3	417.245	879.420	950.457	1063.541	352.197	423.235
P02652	APOA2_HUMAN	Apolipoprotein A-II	SKEQLTPLIK	578.850	2	20.8	24.8	31.0	35.0	y3,y4,y5,y6,y7,y8	373.280	470.333	571.381	684.465	812.523	941.566
			SPELQAEAK	486.753	2	12.1	16.1	21.3	25.3	y6,y7,b2,y4,y8,b2	659.372	788.414	314.134	209.618	443.237	157.571
			VKSPELQAEAK	400.559	3	12.3	16.3	22.0	26.0	y4,y5,y6,b4,b5,b7	418.229	546.288	659.372	541.298	218.799	285.164
P06727	APOA4_HUMAN	Apolipoprotein A-IV	LAPLAEDVR	492.280	2	23.7	27.7	34.1	38.1	y4,y5,y6,y7,b2,y7	518.256	589.294	702.378	799.430	282.181	400.219
			LEPYADQLR	552.788	2	23.7	27.7	33.9	37.9	y3,y4,y5,y6,y7,y7	416.261	531.288	602.325	765.388	862.441	431.724
			LVPFATELHER	656.356	2	29.7	33.7	39.5	43.5	y4,y6,y7,y8,y9,y9	554.304	784.394	855.431	1002.500	1099.553	550.280
P02649	APOE_HUMAN	Apolipoprotein E	AATVGSLAGQPLQER	749.405	2	26.3	30.3	36.3	40.3	y5,y7,y8,y10,y11,b5	642.356	827.437	898.474	1098.590	1155.611	244.129
			LGPLVEQGR	484.780	2	20.6	24.6	31.3	35.3	y4,y5,y6,y7,y8,y7	489.241	588.309	701.394	798.446	855.468	399.727
			LQAEAFQAR	517.275	2	19.1	23.1	29.2	33.2	y3,y4,y5,y6,y7,y8	374.214	521.283	592.320	721.362	792.399	920.458
P61769	B2MG_HUMAN	Beta-2-microglobulin	SNFLNC[Carboxyamidomethyl]YVSGFHPSDIEVDLLKN	753.365	4	48.6	52.6	58.5	62.5	y14,y18,y19,y5,y9,b22	792.915	1007.005	1056.539	201.778	348.533	884.092
			VNHVTLSQPK	561.817	2	12.3	16.3	21.8	25.8	y6,y7,b2,y8,b3,b5	673.387	772.456	351.177	455.261	225.626	332.692
Q15582	BGH3_HUMAN	Transforming growth factor-beta-induced protein ig-h3	LTLLAPLNSVFK	658.403	2	52.5	56.5	62.3	66.3	y7,y8,y9,b2,y3,b4	804.461	875.498	988.582	328.223	197.128	256.676
			QAGLGNHLSGSER	663.331	2	11.9	15.9	21.5	25.5	y4,y5,y6,y7,y9,y11	448.215	535.247	648.331	785.390	956.454	1126.559
P02671	FIBA_HUMAN	Fibrinogen alpha chain	GGSTSYGTGSETESPR	786.842	2	12.3	16.3	21.7	25.7	y3,y8,y9,y10,y11,y12	359.203	862.390	963.437	1020.459	1183.522	1270.554
			GLIDEVNQDFTNRINK	938.482	2	38.9	42.9	48.8	52.8	y7,y8,y9,y10,y11,b10	892.499	1007.526	1135.585	1249.628	1348.696	1232.579
P06396	GELS_HUMAN	Gelsolin	AGALNSNDAFVLK	660.351	2	31.6	35.6	41.8	45.8	y3,y5,y7,y8,y9,b2	359.265	577.370	806.440	893.472	1007.515	200.102
			HVVPNEVVVQR	638.362	2	18.0	22.0	28.1	32.1	y8,y9,y10,b2,y7,y10	940.521	1039.589	1138.657	336.203	422.238	569.832
			QTQVSVLPEGGETPLFK	915.486	2	40.4	44.4	50.3	54.3	y4,y8,y10,y11,y13,b5	504.318	848.451	1074.546	1187.630	1373.731	322.174
P01857	IGHG1_HUMAN	Ig gamma-1 chain C	FNWYVDGVEVHNAK	839.405	2	33.4	37.4	43.5	47.5	y4,y8,y9,y10,y11,b2	469.251	853.452	968.479	1067.548	1230.611	448.197
P01859	IGHG2_HUMAN	Ig gamma-2 chain C	GPSVFPLAPSSK	593.827	2	32.0	36.0	42.7	46.7	y4,y5,y7,y8,b2,y10	418.229	489.266	699.403	846.471	242.113	516.790
P01859	IGHG2_HUMAN	Ig gamma-2 chain C	SC[Carboxyamidomethyl]DKTHTC[Carboxyamidome	1112.219	3	42.2	46.2	51.8	55.8	y2,y4,y5,y8,y13,y17	244.165	469.313	566.366	973.586	1370.783	1823.046
P01860	IGHG3_HUMAN	Ig gamma-3 chain C	TTPPVLDSDGSFFLYSK	937.465	2	46.1	50.1	55.9	59.9	y3,y8,y10,y11,y12,y15	397.208	948.482	1150.541	1265.568	1378.652	836.417
P01861	IGHG4_HUMAN	Ig gamma-4 chain C	ALPAPIEKTISK	634.385	2	24.4	28.4	34.5	38.5	y6,y8,y10,y8,y9,y10	705.414	915.550	1083.640	458.279	493.797	542.324
			VSNKALPAPIEK	633.874	2	17.3	21.3	27.1	31.1	y4,y6,y7,y8,b4,b7	486.292	654.382	767.466	838.503	500.282	781.456
			VVSVLTVLHQDWLNGKEYK	743.407	3	44.7	48.7	54.4	58.4	y11,b2,y11,y14,y15,y17	1417.685	286.176	709.346	865.946	922.488	1015.539
P01871	IGHM_HUMAN	Ig mu chain C region	GVALHRPDVVLLPPAR	887.510	2	30.9	34.9	41.1	45.1	y3,y4,y5,y8,y9,y10	343.208	440.261	553.345	928.561	1043.588	1140.641
P04220	MUCB_HUMAN	Ig mu heavy chain disease protein	NVPLPVIAELPPK	693.921	2	43.6	47.6	54.2	58.2	y3,y7,y9,y11,y4,y9	341.218	767.466	963.587	1173.724	227.655	482.297
			NVPLPVIAELPPKVSVFVPPR	756.787	3	52.8	56.8	61.4	65.4	y3,y7,y9,y11,y13,y12	369.224	801.461	1028.625	1222.730	1464.857	1335.814
			QIQVSWLR	515.296	2	34.4	38.4	44.9	48.9	y3,y4,y5,y6,y7,y7	474.282	561.314	660.382	788.441	901.525	451.266
			GVALHRPDVYLPPAR	887.512	2	30.8	34.8	41.1	45.1	y4,y8,y10,b7,b11,b7	440.261	928.561	1140.641	846.458	1334.757	423.732
			LIC[Carboxyamidomethyl]QATGFSPR	625.322	2	24.8	28.8	35.0	39.0	y3,y5,y6,y7,y8,y9	359.203	563.293	664.341	735.378	863.437	1023.467
			VFAIPPSFASIFLTK	819.469	2	60.8	62.8	70.0	74.0	y9,y10,y11,b2,y10,y11	1013.566	1110.619	1207.672	318.181	555.813	604.339
P01834	IGKC_HUMAN	Ig kappa chain C region	SGTASVVC[Carboxyamidomethyl]LLNNFYPR	899.451	2	52.2	56.2	61.7	65.7	y4,y5,y6,y7,y8,y9	582.303	696.346	810.389	923.473	1036.557	1196.588
			TVAAPSVFIFPPSDEQLK	973.517	2	47.3	51.3	57.4	61.4	y8,y9,y10,y11,b2,b11	913.462	1060.530	1173.615	1320.683	272.160	614.342
			VYAC[Carboxyamidomethyl]EVTHQGLSSPVTK	938.467	2	22.9	26.9	32.8	36.8	y4,y5,y8,y10,y11,y6	444.281	531.313	788.451	1053.568	1154.616	618.345
P0CG04	LAC1_HUMAN	Ig lambda-1 chain C region	ADGSPVKAGVETTKPSK	557.969	3	10.8	14.8	20.2	24.2	y9,y10,y6,y11,y13,y5	946.520	1017.557	331.197	573.330	671.390	187.451
P0CG05	LAC2_HUMAN	Ig lambda-2 chain C region	ADSSPVKAGVETTTPSK	837.931	2	14.6	18.6	24.3	28.3	y3,y9,y10,y11,y13,y15	331.197	919.473	990.510	1118.605	657.867	744.899
P0CG06	LAC3_HUMAN	Ig lambda-3 chain C region	AGVETTTPSK	495.759	2	5.8	9.8	17.6	21.6	y3,y4,y6,y7,b2,y9	331.197	432.245	634.340	763.383	228.134	460.240
P0CF74	LAC6_HUMAN	Ig lambda-6 chain C region	SHKSYSC[Carboxyamidomethyl]QVTHEGSTVEK	516.742	4	5.7	9.7	18.0	22.0	y6,y7,y8,y4,y8,b12	620.325	749.367	886.426	238.639	443.717	501.221
P61626	LYSC_HUMAN	Lysozyme C	ATNYNAGDRSTDYGIFQINSR	788.372	3	32.5	36.5	42.8	46.8	y3,y4,y5,y7,y8,y9	376.193	489.277	617.336	877.489	934.510	1097.573
			C[Carboxyamidomethyl]QNRDVRQYVQGC[Carboxya	919.923	2	13.7	17.7	32.2	36.2	y3,y5,y8,y9,y10,y11	335.138	520.218	910.408	1066.509	1165.578	1280.605
			STDYGIFQINSR	700.844	2	35.9	39.9	46.4	50.4	y5,y6,y7,y8,y9,b2	617.336	764.404	877.489	934.510	1097.573	304.113
P0DJI8	SAA1_HUMAN	Serum amyloid A-1	AYSDM[Oxid]REANYIGSDK	868.383	2	16.4	20.4	26.2	30.2	y3,y4,y5,b6,y3,y11	349.171	406.193	519.277	869.345	175.089	650.303
P0DJI9	SAA2_HUMAN	Serum amyloid A-2	AYSDM[Oxid]REANYIGSDKYFHAR	1205.548	2	22.9	26.9	32.9	36.9	y5,y6,y9,y12,y15,y16	693.346	821.441	1080.522	1470.712	913.950	987.468
			AYSDMREANYIGSDK	860.386	2	21.2	25.2	31.2	35.2	y3,y4,y11,b8,b9,b10	349.171	406.193	1283.604	1038.430	1201.494	1314.578
			AYSDMREANYIGSDKYFHAR	1197.550	2	26.4	30.4	36.3	40.3	y4,y6,y9,b5,y16,b17	530.283	821.441	1080.522	724.308	979.470	1074.976
			EANYIGSDKYFHAR	835.900	2	21.5	25.5	31.6	35.6	y6,y9,y10,b2,b3,b4	821.441	1080.522	1193.606	315.129	478.193	296.142
			EANYIGSDKYFHARGNYDAAK	797.377	3	21.7	25.7	31.8	35.8	y3,b2,y13,y16,y17,b4	289.187	315.129	770.886	900.426	956.968	296.142
			GPGGVWAAEAISDAR	728.863	2	37.8	41.8	47.8	51.8	y5,y11,y13,b2,b3,b6	561.299	1188.600	1302.643	212.102	269.124	313.158
			RGPGGAWAAEVISNAR	806.421	2	30.6	34.6	40.7	44.7	y15,b9,b10,b11,b3,b4	1455.733	953.458	1052.527	1165.611	184.605	213.116
			RGPGGVWAAEAISDAR	806.913	2	29.3	33.3	39.5	43.5	y6,y10,y15,b9,b10,b13	632.336	1089.532	1456.717	981.490	1052.527	1367.670
P02743	SAMP_HUMAN	Serum amyloid P-component	AYSLFSYNTQGR	703.839	2	33.2	37.2	43.5	47.5	y6,y7,y8,y9,y10,b2	738.352	825.384	972.453	1085.537	1172.569	322.139
			DNELLVYK	497.266	2	28.4	32.4	38.8	42.8	y3,y4,y5,y6,b2,b5	409.244	522.328	635.412	764.455	359.119	342.682
			IVLGQEQDSYGGK	697.352	2	21.3	25.3	31.3	35.3	y5,y6,y8,y10,y11,y4	511.251	626.278	883.379	1068.459	1181.543	212.613
			QGYFVEAQPK	583.796	2	21.9	25.9	32.6	36.6	y4,y5,y6,y7,b2,b2	443.261	572.303	671.372	818.440	349.150	175.079
			VGEYSLYIGR	578.803	2	31.4	35.4	41.6	45.6	y4,y5,y6,y7,y9,b2	508.287	621.371	708.403	871.467	1057.531	286.139
P02788	TRFL_HUMAN	Lactotransferrin	DGAGDVAFIR	510.759	2	27.5	31.5	39.6	43.6	y3,y4,y5,y6,y7,b2	435.271	506.308	605.376	720.403	777.425	244.092
			VVWC[Carboxyamidomethyl]AVGEQELR	723.364	2	34.5	38.5	44.0	48.0	y5,y6,y7,y8,y9,y10	674.346	731.368	830.436	901.473	1061.504	1247.583
			YYGYTGAFR	549.256	2	26.4	30.4	37.3	41.3	y4,y5,y6,y7,y8,b2	450.245	551.293	714.356	771.378	934.441	384.155

Références

1. Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev. 1996;13:19‑50.
2. Consortium IHGS. Finishing the euchromatic sequence of the human genome. Nature. 21 oct 2004;431(7011):931‑45.
3. Smith LM, Kelleher NL, Proteomics TC for TD. Proteoform: a single term describing protein complexity. Nat Methods. mars 2013;10(3):186‑7.
4. Legrain P, Aebersold R, Archakov A, Bairoch A, Bala K, Beretta L, et al. The Human Proteome Project: Current State and Future Direction. Mol Cell Proteomics. 7 janv 2011;10(7):M111.009993.
5. HUPO (Human Proteome Organization) 1st World Congress. 21-24 November 2002, Versailles, France. Abstracts. Mol Cell Proteomics MCP. sept 2002;1(9):651‑752.
6. Paik Y-K, Jeong S-K, Omenn GS, Uhlen M, Hanash S, Cho SY, et al. The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome. Nat Biotechnol. mars 2012;30(3):221‑3.
7. Aebersold R, Bader GD, Edwards AM, van Eyk JE, Kussmann M, Qin J, et al. The Biology/Disease-driven Human Proteome Project (B/D-HPP): Enabling Protein Research for the Life Sciences Community. J Proteome Res. 4 janv 2013;12(1):23‑7.
8. Anderson NL, Anderson NG. Proteome and proteomics: new technologies, new concepts, and new words. Electrophoresis. août 1998;19(11):1853‑61.
9. Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez J-C. The dynamic range of protein expression: A challenge for proteomic research. ELECTROPHORESIS. 1 avr 2000;21(6):1104‑15.
10. Lachmann HJ, Booth DR, Booth SE, Bybee A, Gilbertson JA, Gillmore JD, et al. Misdiagnosis of Hereditary Amyloidosis as AL (Primary) Amyloidosis. N Engl J Med. 6 juin 2002;346(23):1786‑91.
11. Solomon A, Murphy CL, Westermark P. Misclassification of amyloidosis is unwarranted. Blood. 15 juill 2006;108(2):776‑7.
12. Comenzo RL, Zhou P, Fleisher M, Clark B, Teruya-Feldstein J. Seeking confidence in the diagnosis of systemic AL (Ig light-chain) amyloidosis: patients can have both monoclonal gammopathies and hereditary amyloid proteins. Blood. 1 mai 2006;107(9):3489‑91.
13. Linke RP. On Typing Amyloidosis Using Immunohistochemistry. Detailled Illustrations, Review and a Note on Mass Spectrometry. Prog Histochem Cytochem. août 2012;47(2):61‑132.
14. Theis JD, Dasari S, Vrana JA, Kurtin PJ, Dogan A. Shotgun-proteomics-based clinical testing for diagnosis and classification of amyloidosis. J Mass Spectrom. 1 oct 2013;48(10):1067‑77.
15. Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat Methods. juin 2012;9(6):555‑66.
16. Picotti P, Bodenmiller B, Aebersold R. Proteomics meets the scientific method. Nat Methods. 2013;10(1):24‑7.
17. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted

proteomics. Mol Cell Proteomics [Internet]. 3 août 2012 [cité 19 oct 2012]; Disponible sur: http://www.mcponline.org/content/early/2012/08/03/mcp.O112.020131
18. Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, Domon B. Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. Mol Cell Proteomics. 12 janv 2012;11(12):1709‑23.
19. Virchow R. Ueber eine im Gehirn und Rückenmark des Menschen aufgefundene Substanz mit der chemischen Reaction der Cellulose. Arch Für Pathol Anat Physiol Für Klin Med. 1 mars 1854;6(1):135‑8.
20. Bennhold H. Eine spezifische Amyloidfärbung mit Kongorot. Münch Med Wochenschr. 1922;69:1537‑8.
21. Cohen AS, Calkins E. Electron Microscopic Observations on a Fibrous Component in Amyloid of Diverse Origins. Nature. 25 avr 1959;183(4669):1202‑3.
22. Pepys MB, Dyck RF, de Beer FC, Skinner M, Cohen AS. Binding of serum amyloid P-component (SAP) by amyloid fibrils. Clin Exp Immunol. nov 1979;38(2):284‑93.
23. Grateau G, Benson MD, Delpech M. Les amyloses. Médecine Sciences Publications; 2000. 580 p.
24. Jaccard A, Fermand J-P. Amyloses. EMC - Hématologie. juin 2004;1(2):46‑58.
25. Iranchahr N. Amylose et nécrose linguale : présentation d'un premier cas? [Genève]: Université de Genève; 2007.
26. Eanes ED, Glenner GG. X-Ray Diffraction Studies on Amyloid Filaments. J Histochem Cytochem. 11 janv 1968;16(11):673‑7.
27. Lebrazi H, Hachulla E, Saïle R. Mécanismes de l'amylose et protéines impliquées. Rev Médecine Interne. janv 2000;21(1):35‑49.
28. Grateau G, Verine J, Delpech M, Ries M. Les amyloses, un modèle de maladie du repliement des protéines. MS Médecine Sci. 2005;21(6-7):627‑33.
29. Hardy J, Selkoe DJ. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. Science. 19 juill 2002;297(5580):353‑6.
30. Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, Cotman CW, et al. Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis. Science. 18 avr 2003;300(5618):486‑9.
31. Lee H-G, Casadesus G, Zhu X, Takeda A, Perry G, Smith MA. Challenging the Amyloid Cascade Hypothesis: Senile Plaques and Amyloid-β as Protective Adaptations to Alzheimer Disease. Ann N Y Acad Sci. 1 juin 2004;1019(1):1‑4.
32. Grateau G. Physiopathologie des amyloses. Rev Rhum. avr 2000;67(3):189‑96.
33. Sipe JD, Benson MD, Buxbaum JN, Ikeda S, Merlini G, Saraiva MJM, et al. Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis. Amyloid. déc 2012;19(4):167‑70.
34. Ren R, Hong Z, Gong H, Laporte K, Skinner M, Seldin DC, et al. Role of Glycosaminoglycan Sulfation in the Formation of Immunoglobulin Light Chain Amyloid Oligomers and Fibrils. J Biol Chem. 26 nov 2010;285(48):37672‑82.
35. Emsley J, White HE, O'Hara BP, Oliva G, Srinivasan N, Tickle IJ, et al. Structure of pentameric human serum amyloid P component. Nature. 27 janv 1994;367(6461):338‑45.
36. Coria F, Castaño E, Prelli F, Larrondo-Lillo M, van Duinen S, Shelanski ML, et al. Isolation and characterization of amyloid P component from Alzheimer's disease and other types of cerebral amyloidosis. Lab Investig J Tech Methods Pathol. avr 1988;58(4):454‑8.

37. Hamazaki H. Ca2+-mediated association of human serum amyloid P component with heparan sulfate and dermatan sulfate. J Biol Chem. 2 mai 1987;262(4):1456‑60.
38. Hamazaki H. Calcium-dependent polymerization of human serum amyloid P component is inhibited by heparin and dextran sulfate. Biochim Biophys Acta. 19 oct 1989;998(3):231‑5.
39. Pepys MB, Booth DR, Hutchinson WL, Gallimore JR, Collins IM, Hohenester E. Amyloid P component. A critical review. Amyloid. 1 janv 1997;4(4):274‑95.
40. Pepys MB, Rademacher TW, Amatayakul-Chantler S, Williams P, Noble GE, Hutchinson WL, et al. Human serum amyloid P component is an invariant constituent of amyloid deposits and has a uniquely homogeneous glycostructure. Proc Natl Acad Sci U S A. 7 juin 1994;91(12):5602‑6.
41. Tennent GA, Lovat LB, Pepys MB. Serum amyloid P component prevents proteolysis of the amyloid fibrils of Alzheimer disease and systemic amyloidosis. Proc Natl Acad Sci U S A. 9 mai 1995;92(10):4299‑303.
42. Colombat M, Stern M. Maladies pulmonaires et dépôts de chaînes légères non amyloïdes. Rev Mal Respir. sept 2013;30(7):524‑6.
43. Colombat M, Holifanjaniaina S, Guillonneau F, Mal H, Hirschi S, Reynaud-Gaubert M, et al. Mass Spectrometry–based Proteomic Analysis: A Good Diagnostic Tool for Cystic Lung Light Chain Deposition Disease. Am J Respir Crit Care Med. 1 août 2013;188(3):404‑5.
44. Magy N. Amylose AA : données actuelles sur la physiopathologie. Rev Médecine Interne. févr 2004;25(2):129‑34.
45. Hamidi Asl L, Liepnieks JJ, Hamidi Asl K, Uemichi T, Moulin G, Desjoyaux E, et al. Hereditary Amyloid Cardiomyopathy Caused by a Variant Apolipoprotein A1. Am J Pathol. janv 1999;154(1):221‑7.
46. Westermark P, Davey E, Lindbom K, Enqvist S. Subcutaneous fat tissue for diagnosis and studies of systemic amyloidosis. Acta Histochem. 1 août 2006;108(3):209‑13.
47. Gameren IIV, Hazenberg BPC, Bijzet J, Rijswijk MHV. Diagnostic accuracy of subcutaneous abdominal fat tissue aspiration for detecting systemic amyloidosis and its utility in clinical practice. Arthritis Rheum. 1 juin 2006;54(6):2015‑21.
48. Vrana JA, Theis JD, Dasari S, Mereuta OM, Dispenzieri A, Zeldenrust SR, et al. Clinical diagnosis and typing of systemic amyloidosis in subcutaneous fat aspirates by mass spectrometry-based proteomics. Haematologica. 1 juill 2014;99(7):1239‑47.
49. Rodriguez FJ, Gamez JD, Vrana JA, Theis JD, Giannini C, Scheithauer BW, et al. Immunoglobulin derived depositions in the nervous system: novel mass spectrometry application for protein characterization in formalin-fixed tissues. Lab Invest. 18 août 2008;88(10):1024‑37.
50. Murphy CL, Eulitz M, Hrncic R, Sletten K, Westermark P, Williams T, et al. Chemical Typing of Amyloid Protein Contained in Formalin-Fixed Paraffin-Embedded Biopsy Specimens. Am J Clin Pathol. 7 janv 2001;116(1):135‑42.
51. Vrana JA, Gamez JD, Madden BJ, Theis JD, Bergen HR, Dogan A. Classification of Amyloidosis by Laser Microdissection and Mass Spectrometry–based Proteomic Analysis in Clinical Biopsy Specimens. Blood. 12 mars 2009;114(24):4957‑9.
52. Sethi S, Theis JD, Leung N, Dispenzieri A, Nasr SH, Fidler ME, et al. Mass Spectrometry-Based Proteomic Diagnosis of Renal Immunoglobulin Heavy Chain Amyloidosis. Clin J Am Soc Nephrol CJASN. déc 2010;5(12):2180‑7.
53. Lavatelli F, Vrana JA. Proteomic typing of amyloid deposits in systemic amyloidoses. Amyloid. déc 2011;18(4):177‑82.

54. Sethi S, Vrana JA, Theis JD, Leung N, Sethi A, Nasr SH, et al. Laser microdissection and mass spectrometry–based proteomics aids the diagnosis and typing of renal amyloidosis. Kidney Int. juill 2012;82(2):226‑34.
55. Vrana JA, Theis JD, Dasari S, Mereuta OM, Dispenzieri A, Zeldenrust SR, et al. Clinical diagnosis and typing of systemic amyloidosis in subcutaneous fat aspirates by mass spectrometry-based proteomics. Haematologica. 1 juill 2014;99(7):1239‑47.
56. Rowczenio D, Dogan A, Theis JD, Vrana JA, Lachmann HJ, Wechalekar AD, et al. Amyloidogenicity and Clinical Phenotype Associated with Five Novel Mutations in Apolipoprotein A-I. Am J Pathol. oct 2011;179(4):1978‑87.
57. Loo D, Mollee PN, Renaut P, Hill MM. Proteomics in Molecular Diagnosis: Typing of Amyloidosis. J Biomed Biotechnol. 2011;2011.
58. Mollee P, Renaut P, Gottlieb D, Goodman H. How to diagnose amyloidosis. Intern Med J. 1 janv 2014;44(1):7‑17.
59. Wiśniewski JR, Duś K, Mann M. Proteomic workflow for analysis of archival formalin-fixed and paraffin-embedded clinical samples to a depth of 10 000 proteins. PROTEOMICS – Clin Appl. 1 avr 2013;7(3-4):225‑33.
60. Shi S-R, Liu C, Taylor CR. Standardization of Immunohistochemistry for Formalin-fixed, Paraffin-embedded Tissue Sections Based on the Antigen-retrieval Technique: From Experiments to Hypothesis. J Histochem Cytochem. 1 févr 2007;55(2):105‑9.
61. Bose AK, Manhas MS, Banik BK, Robb EW. Microwave-induced organic reaction enhancement (more) chemistry: Techniques for rapid, safe and inexpensive synthesis. Res Chem Intermed. 1 janv 1994;20(1):1‑11.
62. Sun W, Gao S, Wang L, Chen Y, Wu S, Wang X, et al. Microwave-assisted Protein Preparation and Enzymatic Digestion in Proteomics. Mol Cell Proteomics. 4 janv 2006;5(4):769‑76.
63. López-Ferrer D, Capelo JL, Vázquez J. Ultra Fast Trypsin Digestion of Proteins by High Intensity Focused Ultrasound. J Proteome Res. 2005;4(5):1569‑74.
64. López-Ferrer D, Cañas B, Vázquez J, Lodeiro C, Rial-Otero R, Moura I, et al. Sample treatment for protein identification by mass spectrometry-based techniques. TrAC Trends Anal Chem. nov 2006;25(10):996‑1005.
65. Santos HM, Rial-Otero R, Fernandes L, Vale G, Rivas MG, Moura I, et al. Improving Sample Treatment for In-Solution Protein Identification by Peptide Mass Fingerprint Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. J Proteome Res. 1 sept 2007;6(9):3393‑9.
66. Rial-Otero R, Carreira RJ, Cordeiro FM, Moro AJ, Santos HM, Vale G, et al. Ultrasonic assisted protein enzymatic digestion for fast protein identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry: Sonoreactor versus ultrasonic probe. J Chromatogr A. 28 sept 2007;1166(1–2):101‑7.
67. Carrera M, Canas B, Lopez-Ferrer D, Pineiro C, Vazquez J, Gallardo JM. Fast Monitoring of Species-Specific Peptide Biomarkers Using High-Intensity-Focused-Ultrasound-Assisted Tryptic Digestion and Selected MS/MS Ion Monitoring. Anal Chem. 15 juill 2011;83(14):5688‑95.
68. Suslick KS, Fang MM, Hyeon T, Mdleleni MM. Applications of sonochemistry to materials synthesis. Sonochemistry and sonoluminescence [Internet]. Springer; 1999 [cité 17 août 2014]. p. 291 - 320. Disponible sur: http://link.springer.com/chapter/10.1007/978-94-015-9215-4_24
69. Liu W, Huang D, Yi P, Li Y. Status and Prospects of Concentrated Organic Wastewater Degradation. In: Show K-Y, éditeur. Industrial Waste [Internet]. InTech. Kuan-Yeow Show and Xinxin Gu; 2012 [cité 17 août 2014]. p. 284 pages. Disponible

sur: http://www.intechopen.com/books/industrial-waste/status-and-prospects-of-concentrated-organic-wastewater-degradation

70. López-Ferrer D, Heibeck TH, Petritis K, Hixson KK, Qian W, Monroe ME, et al. Rapid Sample Processing for LC-MS-Based Quantitative Proteomics Using High Intensity Focused Ultrasound. J Proteome Res. 5 sept 2008;7(9):3860‑7.
71. Chait BT. Mass Spectrometry: Bottom-Up or Top-Down? Science. 10 juin 2006;314(5796):65‑6.
72. Zubarev RA, Kelleher NL, McLafferty FW. Electron capture dissociation of multiply charged protein cations. A nonergodic process. J Am Chem Soc. 1998;120(13):3265‑6.
73. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A. 29 juin 2004;101(26):9528‑33.
74. Madsen JA, Gardner MW, Smith SI, Ledvina AR, Coon JJ, Schwartz JC, et al. Top-Down Protein Fragmentation by Infrared Multiphoton Dissociation in a Dual Pressure Linear Ion Trap. Anal Chem. 1 nov 2009;81(21):8677‑86.
75. Savaryn JP, Catherman AD, Thomas PM, Abecassis MM, Kelleher NL. The emergence of top-down proteomics in clinical research. Genome Med. 27 juin 2013;5(6):53.
76. Hesse A-M, Ndiaye S, Vinh J. Reversed-Phase HPLC and Hyphenated Analytical Strategies for Peptidomics. In: Merighi A, éditeur. Neuropeptides [Internet]. Humana Press; 2011 [cité 10 oct 2014]. p. 203 ‑ 21. Disponible sur: http://link.springer.com/protocol/10.1007/978-1-61779-310-3_13
77. Nesvizhskii AI. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. J Proteomics. 10 oct 2010;73(11):2092‑123.
78. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. ELECTROPHORESIS. 8 déc 1999;20(18):3551‑67.
79. Eng JK, McCormack AL, Yates III JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom. nov 1994;5(11):976‑89.
80. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. J Proteome Res. 1 avr 2011;10(4):1794‑805.
81. Jimenez CR, Huang L, Qiu Y, Burlingame AL. Searching Sequence Databases Over the Internet: Protein Identification Using MS-Tag. In: Coligan JE, Dunn BM, Speicher DW, Wingfield PT, éditeurs. Current Protocols in Protein Science [Internet]. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2001 [cité 15 oct 2014]. Disponible sur: http://www.currentprotocols.com/WileyCDA/CPUnit/refId-ps1606.html
82. Zhang N, Aebersold R, Schwikowski B. ProbID: a probabilistic algorithm to identify peptides through sequence database searching using tandem mass spectral data. Proteomics. oct 2002;2(10):1406‑12.
83. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. Bioinformatics. 6 déc 2004;20(9):1466‑7.
84. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, et al. Open Mass Spectrometry Search Algorithm. J Proteome Res. 1 oct 2004;3(5):958‑64.
85. Zamdborg L, LeDuc RD, Glowacz KJ, Kim Y-B, Viswanathan V, Spaulding IT, et al. ProSight PTM 2.0: improved protein identification and characterization for top down mass spectrometry. Nucleic Acids Res. juill 2007;35(Web Server issue):W701‑6.

86. Tanner S, Shu H, Frank A, Wang L-C, Zandi E, Mumby M, et al. InsPecT: Identification of Posttranslationally Modified Peptides from Tandem Mass Spectra. Anal Chem. 1 juill 2005;77(14):4626‑39.
87. Poullet P, Carpentier S, Barillot E. myProMS, a web server for management and validation of mass spectrometry-based proteomic data. PROTEOMICS. 1 août 2007;7(15):2553‑6.
88. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, et al. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun Mass Spectrom. 30 oct 2003;17(20):2337‑42.
89. Bairoch A, Boeckmann B. The SWISS-PROT protein sequence data bank. Nucleic Acids Res. 25 avr 1991;19(suppl):2247‑9.
90. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. mars 2007;4(3):207‑14.
91. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. Mol Cell Proteomics. 5 janv 2002;1(5):376‑86.
92. Bantscheff M, Lemeer S, Savitski MM, Kuster B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal Bioanal Chem. 1 sept 2012;404(4):939‑65.
93. Ong S-E, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc. janv 2007;1(6):2650‑60.
94. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M. Super-SILAC mix for quantitative proteomics of human tumor tissue. Nat Methods. mai 2010;7(5):383‑5.
95. Rose K, Simona MG, Offord RE, Prior CP, Otto B, Thatcher DR. A new mass-spectrometric C-terminal sequencing technique finds a similarity between gamma-interferon and alpha 2-interferon and identifies a proteolytically clipped gamma-interferon that retains full antiviral activity. Biochem J. 1 nov 1983;215(2):273‑7.
96. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. Proteolytic 18O Labeling for Comparative Proteomics: Model Studies with Two Serotypes of Adenovirus. Anal Chem. 1 juill 2001;73(13):2836‑42.
97. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol. oct 1999;17(10):994‑9.
98. Hansen KC, Schmitt-Ulms G, Chalkley RJ, Hirsch J, Baldwin MA, Burlingame AL. Mass Spectrometric Analysis of Protein Mixtures at Low Levels Using Cleavable 13C-Isotope-coded Affinity Tag and Multidimensional Chromatography. Mol Cell Proteomics. 5 janv 2003;2(5):299‑314.
99. Murray CI, Uhrigshardt H, O'Meally RN, Cole RN, Eyk JEV. Identification and Quantification of S-Nitrosylation by Cysteine Reactive Tandem Mass Tag Switch Assay. Mol Cell Proteomics. 2 janv 2012;11(2):M111.013441.
100. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed Protein Quantitation in Saccharomyces cerevisiae Using Amine-reactive Isobaric Tagging Reagents. Mol Cell Proteomics. 12 janv 2004;3(12):1154‑69.
101. Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem. 15 avr 2003;75(8):1895‑904.
102. Brun V, Masselon C, Garin J, Dupuis A. Isotope dilution strategies for absolute quantitative proteomics. J Proteomics. 21 juill 2009;72(5):740‑9.

103. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods. août 2005;2(8):587‑9.
104. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci. 6 oct 2003;100(12):6940‑5.
105. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, et al. Less label, more free: Approaches in label-free quantitative mass spectrometry. PROTEOMICS. 1 févr 2011;11(4):535‑53.
106. Liu H, Sadygov RG, Yates JR. A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. Anal Chem. 1 juill 2004;76(14):4193‑201.
107. Rappsilber J, Ryder U, Lamond AI, Mann M. Large-Scale Proteomic Analysis of the Human Spliceosome. Genome Res. 8 janv 2002;12(8):1231‑45.
108. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides Per Protein. Mol Cell Proteomics. 9 janv 2005;4(9):1265‑72.
109. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol. janv 2007;25(1):117‑24.
110. Griffin NM, Yu J, Long F, Oh P, Shore S, Li Y, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomics analysis. Nat Biotechnol. janv 2010;28(1):83‑9.
111. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP. Statistical Analysis of Membrane Proteome Expression Changes in Saccharomyces cerevisiae. J Proteome Res. 1 sept 2006;5(9):2339‑47.
112. Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ. Absolute Quantification of Proteins by LCMSE A Virtue of Parallel ms Acquisition. Mol Cell Proteomics. 1 janv 2006;5(1):144‑56.
113. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li G-Z, et al. Quantitative Proteomic Analysis by Accurate Mass Retention Time Pairs. Anal Chem. 1 avr 2005;77(7):2187‑200.
114. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. Mol Cell Proteomics. 6 janv 2012;11(6):O111.016717.
115. Thompson AJ, Abu M, Hanger DP. Key issues in the acquisition and analysis of qualitative and quantitative mass spectrometry data for peptide-centric proteomic experiments. Amino Acids. sept 2012;43(3):1075‑85.
116. Chevallet M, Santoni V, Poinas A, Rouquié D, Fuchs A, Kieffer S, et al. New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. Electrophoresis. août 1998;19(11):1901‑9.
117. Rabilloud T. Solubilization of Proteins in 2DE: An Outline. In: Tyther R, Sheehan D, éditeurs. Two-Dimensional Electrophoresis Protocols [Internet]. Humana Press; 2009 [cité 20 mai 2014]. p. 19 - 30. Disponible sur: http://link.springer.com/protocol/10.1007/978-1-59745-281-6_2
118. Fey SJ, Larsen PM. 2D or not 2D. Curr Opin Chem Biol. 1 févr 2001;5(1):26‑33.

119. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 10 juin 1989;246(4926):64‑71.
120. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T, et al. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom. 1 août 1988;2(8):151‑3.
121. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem. 1 août 2007;389(4):1017‑31.
122. Kelly RT, Page JS, Zhao R, Qian W-J, Mottaz HM, Tang K, et al. Capillary-Based Multi-Nanoelectrospray Emitters. Anal Chem. 1 janv 2008;80(1):143‑9.
123. Gale DC, Smith RD. Small volume and low flow-rate electrospray lonization mass spectrometry of aqueous samples. Rapid Commun Mass Spectrom. 1 nov 1993;7(11):1017‑21.
124. Emmett MR, Caprioli RM. Micro-electrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins. J Am Soc Mass Spectrom. juill 1994;5(7):605‑13.
125. Wilm M, Mann M. Analytical Properties of the Nanoelectrospray Ion Source. Anal Chem. 1 janv 1996;68(1):1‑8.
126. Valaskovic GA, Kelleher NL, Little DP, Aaserud DJ, McLafferty FW. Attomole-sensitivity electrospray source for large-molecule mass spectrometry. Anal Chem. 15 oct 1995;67(20):3802‑5.
127. Tang K, Lin Y, Matson DW, Kim T, Smith RD. Generation of multiple electrosprays using microfabricated emitter arrays for improved mass spectrometric sensitivity. Anal Chem. 15 avr 2001;73(8):1658‑63.
128. Zhang S, Van Pelt CK, Henion JD. Automated chip-based nanoelectrospray-mass spectrometry for rapid identification of proteins separated by two-dimensional gel electrophoresis. ELECTROPHORESIS. 1 nov 2003;24(21):3620‑32.
129. Zubarev R, Mann M. On the Proper Use of Mass Accuracy in Proteomics. Mol Cell Proteomics. 3 janv 2007;6(3):377‑81.
130. Zubarev RA, Håkansson P, Sundqvist B. Accuracy Requirements for Peptide Characterization by Monoisotopic Molecular Mass Measurements. Anal Chem. 1 janv 1996;68(22):4060‑3.
131. Olsen JV, Godoy LMF de, Li G, Macek B, Mortensen P, Pesch R, et al. Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap. Mol Cell Proteomics. 12 janv 2005;4(12):2010‑21.
132. Anderson NL, Anderson NG. The Human Plasma Proteome History, Character, and Diagnostic Prospects. Mol Cell Proteomics. 11 janv 2002;1(11):845‑67.
133. Issaq HJ, Xiao Z, Veenstra TD. Serum and Plasma Proteomics. Chem Rev. 1 août 2007;107(8):3601‑20.
134. Whiteaker JR, Zhao L, Zhang HY, Feng L-C, Piening BD, Anderson L, et al. Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers. Anal Biochem. 1 mars 2007;362(1):44‑54.
135. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom. nov 1984;11(11):601.
136. Biemann K. Appendix 5. Nomenclature for peptide fragment ions (positive ions). In: James A. McCloskey, éditeur. Methods in Enzymology [Internet]. Academic Press; 1990 [cité 16 juill 2014]. p. 886 - 7. Disponible sur: http://www.sciencedirect.com/science/article/pii/0076687990934603
137. Johnson RS, Martin SA, Biemann K, Stults JT, Watson JT. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass

spectrometer: differentiation of leucine and isoleucine. Anal Chem. 1 nov 1987;59(21):2621‑5.
138. Johnson RS, Martin SA, Biemann K. Collision-induced fragmentation of (M + H)+ ions of peptides. Side chain specific sequence ions. Int J Mass Spectrom Ion Process. 29 déc 1988;86:137‑54.
139. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods. sept 2007;4(9):709‑12.
140. Giddings JC. Maximum number of components resolvable by gel filtration and other elution chromatographic methods. Anal Chem. 1 juill 1967;39(8):1027‑8.
141. Ettre LS. MS Tswett and the. LC• GC Eur. 2003;2.
142. Garcıa MC, Hogenboom AC, Zappey H, Irth H. Effect of the mobile phase composition on the separation and detection of intact proteins by reversed-phase liquid chromatography–electrospray mass spectrometry. J Chromatogr A. 31 mai 2002;957(2):187‑99.
143. Chervet JP, Ursem M, Salzmann JP. Instrumental Requirements for Nanoscale Liquid Chromatography. Anal Chem. 1 mai 1996;68(9):1507‑12.
144. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol. 14 oct 2008;4:222.
145. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, et al. The PeptideAtlas project. Nucleic Acids Res. 1 janv 2006;34(Database issue):D655‑8.
146. Jones P, Cote RG, Martens L, Quinn AF, Taylor CF, Derache W, et al. PRIDE: a public repository of protein and peptide identifications for the proteomics community. Nucleic Acids Res. 1 janv 2006;34(Database issue):D659‑63.
147. Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, et al. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol Cell Proteomics MCP. oct 2007;6(10):1809‑17.
148. Method of the Year 2012. Nat Methods. janv 2013;10(1):1‑1.
149. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. Full Dynamic Range Proteome Analysis of S. cerevisiae by Targeted Proteomics. Cell. 21 août 2009;138(4):795‑806.
150. Kiyonami R, Schoen A, Prakash A, Peterman S, Zabrouskov V, Picotti P, et al. Increased Selectivity, Analytical Precision, and Throughput in Targeted Proteomics. Mol Cell Proteomics MCP [Internet]. févr 2011 [cité 9 avr 2014];10(2). Disponible sur: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3033677/
151. Michalski A, Damoc E, Hauschild J-P, Lange O, Wieghaus A, Makarov A, et al. Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. Mol Cell Proteomics [Internet]. 9 janv 2011 [cité 9 oct 2012];10(9). Disponible sur: http://www.mcponline.org/content/10/9/M111.011015
152. Gallien S, Duriez E, Demeure K, Domon B. Selectivity of LC-MS/MS analysis: Implication for proteomics experiments. J Proteomics. 9 avr 2013;81:148‑58.
153. Gallien S, Bourmaud A, Kim SY, Domon B. Technical considerations for large-scale parallel reaction monitoring analysis. J Proteomics [Internet]. [cité 28 févr 2014]; Disponible sur: http://www.sciencedirect.com/science/article/pii/S1874391913005447
154. Hauschild J-P, Denisov E, Peterson AC, Lange O, Makarov A, Damoc E, et al. New Developments of a Bench-top Quadrupole-Orbitrap Mass Spectrometer. 2014.

155. Grove P, Neale PH, Peck M, Schiller B, Haas M. Monoclonal immunoglobulin G1-kappa fibrillary glomerulonephritis. Mod Pathol Off J U S Can Acad Pathol Inc. janv 1998;11(1):103‑9.
156. Akinbi HT, Epaud R, Bhatt H, Weaver TE. Bacterial Killing Is Enhanced by Expression of Lysozyme in the Lungs of Transgenic Mice. J Immunol. 15 nov 2000;165(10):5760‑6.
157. Dajani R, Zhang Y, Taft PJ, Travis SM, Starner TD, Olsen A, et al. Lysozyme Secretion by Submucosal Glands Protects the Airway from Bacterial Infection. Am J Respir Cell Mol Biol. juin 2005;32(6):548‑52.
158. Tufte ER. Beautiful Evidence. Graphics Press; 2006. 213 p.
159. Farnaud S, Evans RW. Lactoferrin—a multifunctional protein with antimicrobial properties. Mol Immunol. nov 2003;40(7):395‑405.
160. Kumar V, Hassan MI, Tomar AK, Kashav T, Nautiyal J, Singh S, et al. Proteomic analysis of heparin-binding proteins from human seminal plasma: a step towards identification of molecular markers of male fertility. J Biosci. déc 2009;34(6):899‑908.
161. Thaler CJ, Vanderpuye OA, McIntyre JA, Faulk WP. Lactoferrin binding molecules in human seminal plasma. Biol Reprod. 10 janv 1990;43(4):712‑7.
162. Murthy KR, Goel R, Subbannayya Y, Jacob HK, Murthy PR, Manda SS, et al. Proteomic analysis of human vitreous humor. Clin Proteomics. 14 juill 2014;11(1):29.
163. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367‑72.
164. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. Mol Cell Proteomics. 17 juin 2014;mcp.M113.031591.
165. Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nat Biotechnol. juill 2010;28(7):710‑21.
166. Deutsch EW, Lam H, Aebersold R. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. EMBO Rep. mai 2008;9(5):429‑34.
167. Siepen JA, Keevil E-J, Knight D, Hubbard SJ. Prediction of Missed Cleavage Sites in Tryptic Peptides Aids Protein Identification in Proteomics. J Proteome Res. 1 janv 2007;6(1):399‑408.
168. Gershon PD. Cleaved and Missed Sites for Trypsin, Lys-C, and Lys-N Can Be Predicted with High Confidence on the Basis of Sequence Context. J Proteome Res. 7 févr 2014;13(2):702‑9.
169. Krokhin OV, Spicer V. Peptide Retention Standards and Hydrophobicity Indexes in Reversed-Phase High-Performance Liquid Chromatography of Peptides. Anal Chem. 15 nov 2009;81(22):9522‑30.
170. Krokhin OV, Craig R, Spicer V, Ens W, Standing KG, Beavis RC, et al. An Improved Model for Prediction of Retention Times of Tryptic Peptides in Ion Pair Reversed-phase HPLC Its Application to Protein Peptide Mapping by Off-Line HPLC-MALDI MS. Mol Cell Proteomics. 9 janv 2004;3(9):908‑19.
171. Jones P, Côté RG, Cho SY, Klie S, Martens L, Quinn AF, et al. PRIDE: new developments and new datasets. Nucleic Acids Res. 1 janv 2008;36(suppl 1):D878‑83.
172. Craig R, Cortens JP, Beavis RC. Open Source System for Analyzing, Validating, and Storing Protein Identification Data. J Proteome Res. 1 déc 2004;3(6):1234‑42.
173. Picotti P, Clément-Ziza M, Lam H, Campbell DS, Schmidt A, Deutsch EW, et al. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. Nature [Internet]. 2013 [cité 28 janv 2013]; Disponible sur:

http://www.nature.com.gate1.inist.fr/nature/journal/vaop/ncurrent/full/nature11835.html

174. Gallien S, Duriez E, Demeure K, Domon B. Selectivity of LC-MS/MS analysis: Implication for proteomics experiments. J Proteomics [Internet]. [cité 3 oct 2012];(0). Disponible sur: http://www.sciencedirect.com/science/article/pii/S1874391912007452
175. Gallien S, Peterman S, Kiyonami R, Souady J, Duriez E, Schoen A, et al. Highly multiplexed targeted proteomics using precise control of peptide retention time. Proteomics. avr 2012;12(8):1122‑33.
176. Krokhin OV, Spicer V. Predicting Peptide Retention Times for Proteomics. Current Protocols in Bioinformatics [Internet]. John Wiley & Sons, Inc.; 2002 [cité 29 août 2014]. Disponible sur: http://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi1314s31/abstract
177. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 1 avr 2010;26(7):966‑8.
178. Rakotomalala R. TANAGRA: une plate-forme d'expérimentation pour la fouille de données. Rev Modul. 2005;32:70‑85.
179. Lam H, Deutsch EW, Eddes JS, Eng JK, Stein SE, Aebersold R. Building Consensus Spectral Libraries for Peptide Identification in Proteomics. Nat Methods. oct 2008;5(10):873‑5.
180. Blades AT, Ikonomou MG, Kebarle P. Mechanism of electrospray mass spectrometry. Electrospray as an electrolysis cell. Anal Chem. 1 oct 1991;63(19):2109‑14.
181. Cole RB. Some tenets pertaining to electrospray ionization mass spectrometry. J Mass Spectrom. 1 juill 2000;35(7):763‑72.
182. Ochran RA, Konermann L. Effects of ground loop currents on signal intensities in electrospray mass spectrometry. J Am Soc Mass Spectrom. 1 déc 2004;15(12):1748‑54.
183. Mora JF de la, Van Berkel GJ, Enke CG, Cole RB, Martinez-Sanchez M, Fenn JB. Electrochemical processes in electrospray ionization mass spectrometry. J Mass Spectrom. août 2000;35(8):939‑52.
184. Kebarle P, Tang L. From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry. Anal Chem. 1 nov 1993;65(22):972A ‑ 986A.
185. Gaskell SJ. Electrospray: Principles and Practice. J Mass Spectrom. 1 juill 1997;32(7):677‑88.
186. Wilm MS, Mann M. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? Int J Mass Spectrom Ion Process. 22 sept 1994;136(2–3):167‑80.
187. Allen MH, Vestal ML. Design and performance of a novel electrospray interface. J Am Soc Mass Spectrom. janv 1992;3(1):18‑26.
188. Rayleigh, Lord. XX. On the equilibrium of liquid conducting masses charged with electricity. Philos Mag Ser 5. 1882;14(87):184‑6.
189. Gomez A, Tang K. Charge and fission of droplets in electrostatic sprays. Phys Fluids 1994-Present. 1 janv 1994;6(1):404‑14.
190. Dole M, Mack LL, Hines RL, Mobley RC, Ferguson LD, Alice MB. Molecular Beams of Macroions. J Chem Phys. 1 sept 1968;49(5):2240‑9.
191. Iribarne JV, Thomson BA. On the evaporation of small ions from charged droplets. J Chem Phys. 15 mars 1976;64(6):2287‑94.

192. Konermann L, Rodriguez AD, Liu J. On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. Anal Chem. 7 août 2012;84(15):6798‑804.
193. Konermann L, Ahadi E, Rodriguez AD, Vahidi S. Unraveling the Mechanism of Electrospray Ionization. Anal Chem. 2 janv 2013;85(1):2‑9.
194. Fernandez de la Mora J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. Anal Chim Acta. 1 févr 2000;406(1):93‑104.
195. Banerjee S, Mazumdar S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int J Anal Chem. 12 mars 2012;2012:e282574.
196. Iavarone AT, Jurchen JC, Williams ER. Supercharged Protein and Peptide Ions Formed by Electrospray Ionization. Anal Chem. 1 avr 2001;73(7):1455‑60.
197. Hesse A-M, Marcelo P, Rossier J, Vinh J. Simple and universal tool to remove on-line impurities in mono- or two-dimensional liquid chromatography–mass spectrometry analysis. J Chromatogr A. 2 mai 2008;1189(1–2):175‑82.
198. Karlsson KE, Novotny M. Separation efficiency of slurry-packed liquid chromatography microcolumns with very small inner diameters. Anal Chem. 1 sept 1988;60(17):1662‑5.
199. Kennedy RT, Jorgenson JW. Preparation and evaluation of packed capillary liquid chromatography columns with inner diameters from 20 to 50 micrometers. Anal Chem. 1 mai 1989;61(10):1128‑35.
200. Paul W, Raether M. Das elektrische Massenfilter. Z Für Phys. 1 mai 1955;140(3):262‑73.
201. Paul W. Electromagnetic traps for charged and neutral particles. Rev Mod Phys. 1 juill 1990;62(3):531‑40.
202. Bouchoux G. Spectrométrie de masse [Internet]. Ed. Techniques Ingénieur; [cité 30 juill 2014]. Disponible sur: http://books.google.com/books?hl=en&lr=&id=-uxOQjFPQesC&oi=fnd&pg=PA1&dq=%22en+savoir+plus%22+%22en+revue.+Le+fonctionnement+des+divers+analyseurs%22+%22la+formation+initiale+d%E2%80%99ions,+%C3%A0+l%E2%80%99%C3%A9tat+gazeux,+%C3%A0+partir%22+%22que+soit+leur+%C3%A9tat+physique+:+gazeux,+liquide+ou+solide.+Bien%22+&ots=xm_06DAsYj&sig=GsiKS3xOE2wAyj49yOeS605PZ6c
203. Paul W. Apparatus for separating charged particles of different specific charges. US 2,939,952, 1960.
204. Schwartz JC, Senko MW, Syka JEP. A two-dimensional quadrupole ion trap mass spectrometer. J Am Soc Mass Spectrom. juin 2002;13(6):659‑69.
205. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, et al. A Dual Pressure Linear Ion Trap Orbitrap Instrument with Very High Sequencing Speed. Mol Cell Proteomics. 12 janv 2009;8(12):2759‑69.
206. Guan S, Marshall AG. Ion traps for Fourier transform ion cyclotron resonance mass spectrometry: principles and design of geometric and electric configurations. Int J Mass Spectrom Ion Process. 31 août 1995;146–147:261‑96.
207. Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry: A primer. Mass Spectrom Rev. 1 janv 1998;17(1):1‑35.
208. Penning FM. Die glimmentladung bei niedrigem druck zwischen koaxialen zylindern in einem axialen magnetfeld. Physica. nov 1936;3(9):873‑94.
209. Dehmelt H. Experiments with an isolated subatomic particle at rest. Rev Mod Phys. 1 juill 1990;62(3):525‑30.

210. Aarstol M, Comisarow MB. Apodization of FT-ICR spectra. Int J Mass Spectrom Ion Process. 22 juin 1987;76(3):287‑97.
211. Scigelova M, Hornshaw M, Giannakopulos A, Makarov A. Fourier Transform Mass Spectrometry. Mol Cell Proteomics MCP [Internet]. juill 2011 [cité 30 juill 2014];10(7). Disponible sur: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3134075/
212. Comisarow MB, Melka JD. Error estimates for finite zero-filling in Fourier transform spectrometry. Anal Chem. 1 nov 1979;51(13):2198‑203.
213. Boldin IA, Nikolaev EN. Fourier transform ion cyclotron resonance cell with dynamic harmonization of the electric field in the whole volume by shaping of the excitation and detection electrode assembly. Rapid Commun Mass Spectrom. 15 janv 2011;25(1):122‑6.
214. Makarov A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. Anal Chem. mars 2000;72(6):1156‑62.
215. Kingdon KH. A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. Phys Rev. 1 avr 1923;21(4):408‑18.
216. Lange O, Damoc E, Wieghaus A, Makarov A. Enhanced FT for Orbitrap Mass Spectrometry. pES. 1:1.
217. Marshall AG, Hendrickson CL. Fourier transform ion cyclotron resonance detection: principles and experimental configurations. Int J Mass Spectrom. 2002;215(1):59‑75.
218. Michalski A, Damoc E, Lange O, Denisov E, Nolting D, Muller M, et al. Ultra High Resolution Linear Ion Trap Orbitrap Mass Spectrometer (Orbitrap Elite) Facilitates Top Down LC MS/MS and Versatile Peptide Fragmentation Modes. Mol Cell Proteomics MCP [Internet]. mars 2012 [cité 21 juill 2014];11(3). Disponible sur: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3316736/
219. Shaw JB, Brodbelt JS. Extending the Isotopically Resolved Mass Range of Orbitrap Mass Spectrometers. Anal Chem. 3 sept 2013;85(17):8313‑8.

Fiche envoi des échantillons pour une analyse par spectrométrie de masse

Où adresser sa demande ?

Référents : Dr. Joëlle VINH/ Emmanuelle DEMEY

Adresse : ESPCI ParisTech / SMBP
Bâtiment F, 4ème étage, pièce F4.23
10 rue Vauquelin
75231 Paris cedex 05
Téléphone : 01.40.79.51.78 / 47.65
Fax : 01.40.79.52.18

Mail :joelle.vinh@espci.fr

Quel type d'échantillon ?

Les différents échantillons sont acceptés:

- tissus fixés et/ou colorés, coupés ou non par microtomie
- tissus micro-disséqués par capture laser (LCM)

Mode opératoire

- Dans un minimum de tampon (50mM bicarbonate d'ammonium, pH 7,5) **dans le tube fourni par nos soins**.
  En général, 200µl suffisent (100µl pour les coupes au LCM).

Stockage pour transport

- au congélateur à -20°C avant envoi
- envoi sous carboglace

Informations à fournir : (* : information obligatoire)

	Echantillon 1	Echantillon 2	Echantillon 3	Echantillon 4
Analyse urgente?				
Code échantillon *				
Origine tissu *				
Diganostic (IHC)				
Type fixateur *				
Type colorant				
Surface (mm²) *				
Epaisseur (mm) *				
% dépôt amylose *				
Volume tampon rajouté (µl) *				

ANALYSE PROTEOMIQUE CIBLEE A HAUTE RESOLUTION : UN OUTIL PUISSANT POUR LE DIAGNOSTIC CLINIQUE A PARTIR DE PRELEVEMENTS DE TISSUS AMYLOÏDES BRUTS

par **Sophie LIUU**

Résumé

L'amylose une maladie rare, caractérisée par des agrégats insolubles de protéines extracellulaires. Trente types d'amyloses sont répertoriés et différenciés par la protéine amyloïdogène associée. Leur identification repose sur l'analyse immunohistochimique du tissu malade. Cependant, l'interprétation peut être ambigüe. La micro-dissection laser (LCM) et la spectrométrie de masse sont une alternative prometteuse.

Nous avons développé une approche protéomique ciblée à partir de tissus traités par ultrasons. Nous avons montré que les biopsies prélevées chez des patients atteints de différents types d'amyloses peuvent être protéolysées sous ultrasons, ce qui augmente l'efficacité et réduit le temps de la digestion (90s au lieu de 15h), tout en minimisant la quantité de matériel nécessaire.

Après une première étape de protéomique sans *a priori* pour le diagnostic clinique des patients présentant une amylose dans différents organes/tissus (reins, poumons, glandes salivaires, testicule, humeur vitrée et rate), nous avons optimisé une méthode de criblage dédiée à un transfert technologique vers les départements cliniques français. Pour cela, deux méthodes montrent des résultats très encourageants : la SRM (*selected reaction monitoring*) sur un spectromètre de masse à basse résolution, disponible dans certains laboratoires cliniques et la PRM (*parallel reaction monitoring*) accessible sur les instruments de dernière génération à très haute résolution, qui fournit rapidement des résultats.

En conclusion, nous proposons ici un protocole spécifique pour le diagnostic robuste et la classification rapide des amyloïdoses.

Mots clés : protéomique, spectrométrie de masse, diagnostic, amylose, biopsie, analyse ciblée, SRM, PRM.

Abstract

Amyloidosis is rare disease that appears as an insoluble deposition of specific extracellular proteins. Thirty classes of amyloidosis have been reported and their diagnosis relies on the identification of the associated proteins by immunohistochemistry analysis of the affected tissues. However, its interpretation is highly dependent on the pathologist and can be ambiguous. Laser capture micro-dissection (LCM) and mass spectrometry are a promising alternative.

We have developed a targeted proteomics approach from raw tissues treated with ultrasound. We showed that the biopsies taken from patients presenting different classes of amyloidosis can be proteolysed by ultrasonic treatment. This technique increases the efficiency and reduces the time of digestion (90s instead of 15h), while minimizing the amount of material needed.

After a first unbiased proteomics study for clinical diagnosis of patients with amyloidosis in various organs/tissues (kidney, lung, salivary glands, testicle, vitreous humor and spleen), we have optimized a dedicated screening method for a technological transfer to the French clinical departments. For this, two methods show very promising results: SRM (selected reaction monitoring) on a low-resolution mass spectrometer, available in some clinical laboratories, and PRM (parallel reaction monitoring) available on the latest generation of high-resolution instruments which provide fast results.

In conclusion, we propose here a specific protocol for a robust diagnosis and rapid classification of amyloidoses.

Keywords : proteomics, mass spectrometry, diagnosis, amyloidosis, biopsy, targeted analysis, SRM, PRM.