



HAL
open science

Apport du fractionnement peptidique dans l'identification et la quantification par spectrométrie de masse des protéines plasmatiques : application sur du plasma de rats transgéniques

Lorie Spano, Christophe C. Chambon

► To cite this version:

Lorie Spano, Christophe C. Chambon. Apport du fractionnement peptidique dans l'identification et la quantification par spectrométrie de masse des protéines plasmatiques : application sur du plasma de rats transgéniques. Sciences du Vivant [q-bio]. 2022. hal-03696497

HAL Id: hal-03696497

<https://hal.inrae.fr/hal-03696497>

Submitted on 16 Jun 2022

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

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

SPANO Lorie

Du 04 Avril 2022 au 10 Juin 2022

Université Clermont Auvergne
Institut Universitaire de Technologie, Clermont Auvergne
Campus universitaire Simone Veil, Aurillac
Département Génie Biologique
DUT Génie Biologique
2ème année option biologie - informatique



Centre de recherche de Clermont-Ferrand, site de Theix
Unité de recherche QuaPA, Qualité des Produits Animaux
PFEM, Plateforme d'Exploration du Métabolisme,
composante protéomique



RAPPORT DE STAGE

APPORT DU FRACTIONNEMENT PEPTIDIQUE DANS L'IDENTIFICATION ET
LA QUANTIFICATION PAR SPECTROMÉTRIE DE MASSE DES PROTÉINES
PLASMATIQUES : APPLICATION SUR DU PLASMA DE RATS
TRANSGÉNIQUES

Maître de stage : CHAMBON Christophe

Enseignant référent : VEISSEIRE Philippe

Année Universitaire 2021-2022

SOMMAIRE

Remerciements

Résumé

Abstract

Introduction.....	1
1. Présentation de l'entreprise.....	2
1.1. INRAE, l'institut en quelques mots.....	2
1.2. INRAE, unité de recherche QuaPA.....	2
1.3. INRAE, la plateforme PFEM, composante protéomique.....	3
2. Présentation des problématiques technique et scientifique.....	3
3. Matériels et Méthodes.....	4
3.1. Matériels.....	4
3.1.1. Direct Detect™ (Merck)	4
3.1.2. Électrophorèse avec gel 1D-SDS-PAGE.....	5
3.1.3. Chromatographie Liquide de Haute Performance (HPLC).....	6
3.1.4. Spectrométrie de masse.....	6
3.1.5. Logiciels d'analyses.....	7
3.2. Mode opératoire.....	9
3.2.1. Préparation et quantification protéiques des pools d'échantillons..	9
3.2.2. Immobilisation des échantillons sur gel 1D-SDS-PAGE.....	10
3.2.3. Hydrolyse des protéines immobilisées dans le gel 1D-SDSPAGE.....	11
3.2.4. Fractionnement des peptides à pH élevé en phase inverse.....	12
3.2.5. Analyse LC-MS/MS.....	13
3.2.6. Identification et quantification des protéines.....	13
4. Résultats et Discussion.....	14
4.1. Caractérisation physicochimique des peptides tryptiques.....	14
4.2. Identification des protéines et des peptides.....	15
4.3 Apport du fractionnement sur l'identification des peptides	15
4.4. Apport des répétitions techniques LC-MS/MS sur l'identification des peptides.....	16

4.5. Influence de la méthode (Fractionnée/Non Fractionnée) sur l'étude protéomique de l'effet d'un traitement à la dexaméthasone de rats transgéniques avec une pathologie oculaire.....	17
4.6. Bioanalyse des résultats protéomiques obtenus.....	18
Conclusion.....	20
Bibliographie	
Table des sigles	
Glossaire/Lexique	
Table des figures	

REMERCIEMENTS

Je remercie le centre de recherche INRAE de Theix et plus particulièrement la composante protéomique de la PFEM de m'avoir accueillie et permis le bon déroulement de mon stage.

Je remercie Mme Véronique SANTE-LHOUTELLIER, Directrice de l'unité Qualité des Produits Animaux et M. Michel HEBRAUD, Responsable Scientifique de la plateforme protéomique pour leur accueil.

Je remercie mon maître de stage M. Christophe CHAMBON, Ingénieur Analyste, pour ses explications sur le fonctionnement des machines et pour sa confiance. Je le remercie aussi pour l'autonomie qu'il m'a offert pendant ce stage.

Je remercie M. Didier VIALA, Ingénieur, pour son aide précieuse lors de certaines manipulations ainsi que pour l'initiation à certaines techniques de laboratoire mais aussi pour sa patience pour répondre à toutes mes questions.

Je remercie Mme Claude FERREIRA, Technicienne de laboratoire, de m'avoir accompagné et suivi durant toute la durée de mon stage. Je la remercie pour son encadrement de qualité, sa patience et ses conseils qui m'ont permis de réaliser mon stage dans de bonnes conditions.

Merci à l'ensemble de la plateforme protéomique de m'avoir accueillie avec bienveillance et intérêt.

RÉSUMÉ

Les **protéines** sont des macromolécules biologiques présentes dans toutes les cellules vivantes. Elles assurent une multitude de fonctions au sein des cellules mais aussi des tissus [1]. Au laboratoire, il est possible de **séparer les protéines** des autres constituants cellulaire à l'aide de diverses techniques notamment l'**électrophorèse** et la **chromatographie**. En protéomique, le **système LC-MS/MS*** est capable de détecter et **identifier des quantités infimes de peptides***. Cependant, cette identification peut être plus problématique dans certaines matrices biologiques, comme par exemple le **plasma**. En effet, la matrice plasmatique possède un grand nombre de protéines majoritaires comme l'albumine, la protéine la plus représentée parmi ces protéines surabondantes dans le plasma.

Ainsi, des solutions de **prépurification** peuvent être utilisées pour essayer de simplifier la composition des échantillons et de réduire la présence de ces protéines majoritaires. Une méthode de **fractionnement peptidique** a été testée au cours de ce stage. Pour cela, une migration partielle et un traitement des bandes sont réalisés. Ensuite, les **échantillons sont divisés en deux** : une partie **non fractionnée** et une partie **fractionnée**. La partie fractionnée est passée sur colonne* grâce au **kit de fractionnement des peptides à pH élevé en phase inverse Thermo Scientific™ Pierce™**. Chaque échantillon est alors divisé en **huit fractions** qui sont analysés par le système LC-MS/MS. Pour la partie non fractionnée, les échantillons sont directement analysés par le système LC-MS/MS avec quatre **répétitions techniques** par échantillons afin de pouvoir les comparer aux échantillons fractionnés.

Les **résultats obtenus** indiquent que le fractionnement permet d'**identifier** un nombre plus important de protéines et de peptides comparé à la méthode sans fractionnement. Au niveau de la **découverte des protéines différentielles**, les deux méthodes ont montré des différences. Néanmoins les **études bio-informatiques** replaçant les protéines d'intérêts dans la cellule et/ou au sein des réseaux restent cohérentes entre les méthodes. De plus, il est possible de faire un lien sur le plan clinique avec l'**effet du traitement anti-inflammatoire** sur les rats et nos résultats. Il est possible de dire qu'un design avec des échantillons poolés et un fractionnement permet d'obtenir des résultats pertinents en **limitant le nombre de runs LC-MS/MS** et l'**effet de saturation des protéines majoritaires** du plasma.

ABSTRACT

Proteins are biological macromolecules present in all living cells. They perform a multitude of functions within cells and tissues [1]. In laboratories, it is possible to **separate proteins** from other cellular components using various techniques, notably **electrophoresis** and **chromatography**. In proteomics, the **LC-MS/MS* system** is able to detect and **identify minute quantities of peptides***. However, this identification can be more problematic in certain biological matrices, such as plasma. Indeed, the plasma matrix has a large number of major proteins such as albumin, which is the most represented protein among the overabundant proteins in **plasma**.

Thus, **pre-purification** solutions can be used to try to simplify the composition of the samples and to reduce the presence of these major proteins. A **peptide fractionation** method was tested during this internship. For this, a partial migration and processing of the strips is performed. Then, the **samples are divided into two parts**: an **unfractionated** part and a **fractionated** part. The fractionated portion is run on column* using **the Thermo Scientific™ Pierce™ Reverse Phase High pH Peptide Fractionation Kit**. Each sample is then split into **eight fractions** that are analysed using the LC-MS/MS system. For the unfractionated portion, the samples are analysed directly using the LC-MS/MS system. However, four **technical repeats** per sample are performed in order to compare them to the fractionated samples.

The **results obtained** indicate that the fractionation allows the **identification** of a greater number of proteins and peptides compared to the non-fractionated method. At the level of **differential protein discovery**, the two methods showed differences. However, **bioinformatics studies** locating the proteins of interest within the cell and/or within the networks remain consistent between the methods. Furthermore, it is possible to make a clinical link with **the effect of the anti-inflammatory treatment** on the rats and our results. It is possible to say that a design with pooled samples and fractionation could provide relevant results by **limiting the number of LC-MS/MS runs** and the **saturation effect of the majority proteins in plasma**.

INTRODUCTION

INRAE, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, possède une plateforme PFEM (Plateforme d'Exploration du Métabolisme) avec une composante protéomique. Outre l'activité analytique de routine, la composante développe aussi des méthodes pour les échantillons complexes qui sont analysés à la PFEM. La matrice plasma fait partie de ces échantillons complexes, son étude est toujours un défi analytique.

En effet, lors de l'étude de plasma, les protéines majoritaires empêchent une analyse exhaustive de cette matrice car ces dernières « masquent » les protéines minoritaires. Une solution souvent employée dans l'étude du plasma est de dépler les protéines majoritaires avec des kits d'immunoaffinité. Néanmoins, la déplétion* des protéines n'est pas toujours efficace, c'est pourquoi, un fractionnement peptidique est testé ici dans le but de simplifier l'échantillon pour permettre une meilleure « profondeur » d'analyse. Ce rapport de stage détaillera le matériel et les modes opératoires utilisés pour effectuer ce projet, puis les résultats seront exposés et discutés et enfin une conclusion sur la méthode de fractionnement sera faite.

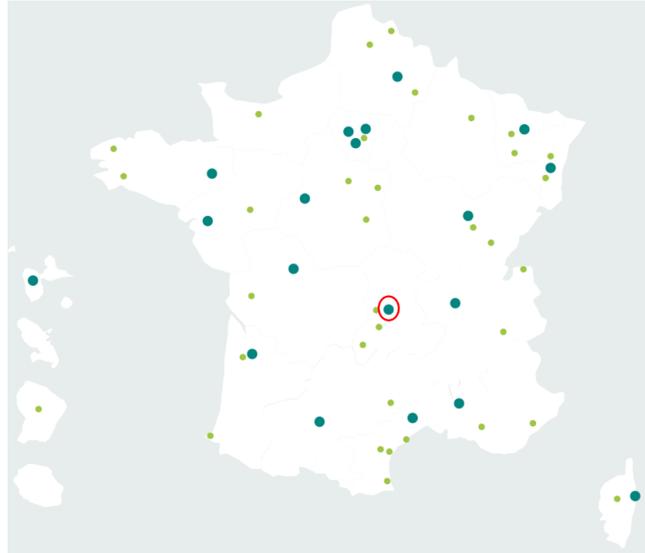


Figure 1. Carte représentant l'implantation des centres INRAE en France

[Liste des centres INRAE | INRAE INSTIT](#)

- Centres de grande taille ● Centres de taille moyenne à petite
- Centre INRAE Clermont-Theix

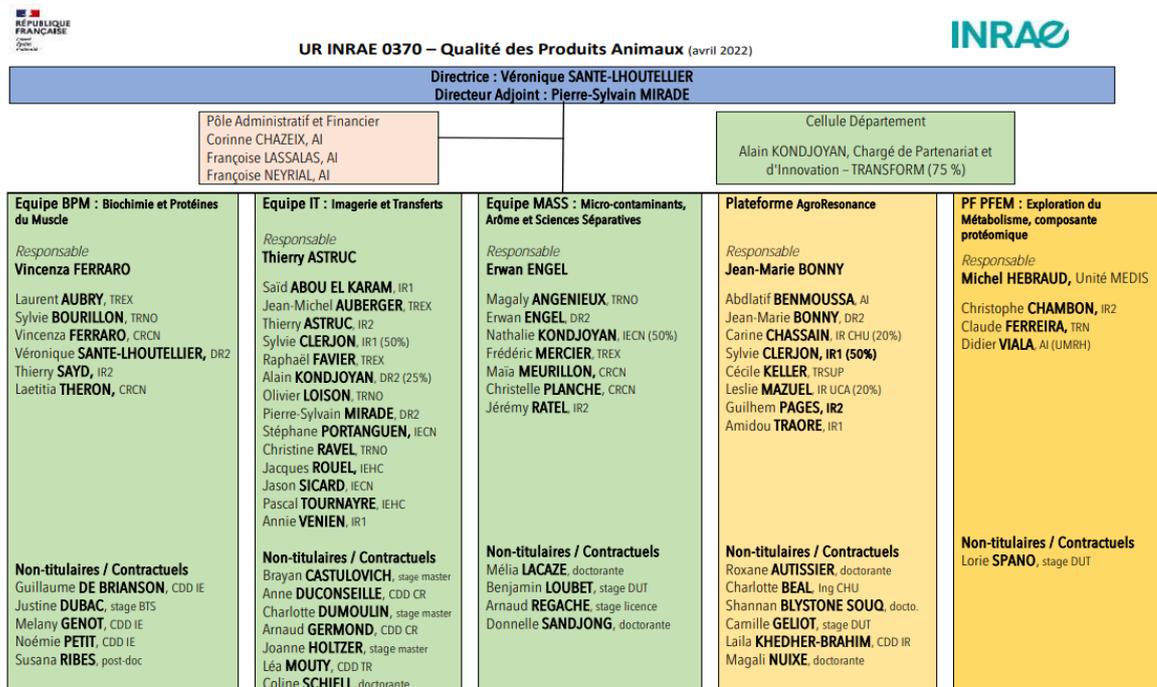


Figure 2. Organigramme de l'unité QuaPA

1. Présentation de l'entreprise

1.1. INRAE, l'institut en quelques mots

INRAE, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, est un acteur majeur de la recherche et de l'innovation créé le 1^{er} janvier 2020. Institut de recherche finalisée issu de la fusion entre l'INRA et IRSTEA. INRAE rassemble une communauté de 12 000 personnes, avec 14 départements divisés en 267 unités de recherche et expérimentales implantées dans 18 centres de recherche sur toute la France. L'institut se positionne parmi les tout premiers organismes de recherche au monde en sciences agricoles et alimentaires, en sciences du végétal et de l'animal, et en écologie-environnement. Il est le premier organisme de recherche mondial spécialisé sur l'ensemble « agriculture-alimentation-environnement ».

INRAE a pour ambition d'être un acteur clé des transitions nécessaires pour répondre aux grands enjeux mondiaux. Face à l'augmentation de la population, au changement climatique, à la raréfaction des ressources et au déclin de la biodiversité, l'institut construit des solutions pour des agricultures multi-performantes, une alimentation de qualité et une gestion durable des ressources et des écosystèmes.

1.2. INRAE, unité de recherche QuaPA

L'unité de recherche QuaPA (Qualité des Produits Animaux) appartient au département TRANSFORM : Aliments, produits biosourcés et déchets. Elle rassemble plusieurs équipes de recherche : l'équipe BPM (Biochimie et Protéines du Muscle), l'équipe IT (Imagerie et Transfert), l'équipe MASS (Micro-contaminants, Arômes et Sciences Séparatives), mais aussi de plateformes comme la plateforme AgroResonance ou encore la plateforme PFEM (Plateforme d'Exploration du Métabolisme, composante protéomique).

Cette unité de recherche est dirigée par Véronique SANTE-LHOUELLIER et chaque équipe et plateforme a un responsable.

1.3. INRAE, la plateforme PFEM, composante protéomique

La Plateforme d'Exploration du Métabolisme (PFEM) est une plateforme analytique composée de la composante protéomique et de la composante métabolomique. Celle-ci travaille sur des domaines touchants l'analyse des molécules par spectrométrie de masse et l'analyse des données générées. En effet, cette composante est spécialisée dans l'analyse des macromolécules que sont les protéines et les peptides. Ces molécules peuvent provenir de différentes origines que ce soit humaines, animales, végétales ou encore de micro-organismes.

Pour réaliser ces analyses, la composante protéomique utilise des spectromètres de masse couplés à des HPLC et des logiciels d'interrogation de données et de quantification. Les résultats permettent d'identifier et quantifier les protéines d'une matrice biologique. Les projets soumis à la plateforme protéomique ont pour but de réaliser des comparaisons qualitatives et quantitatives des protéines dans des échantillons appartenant à des designs expérimentaux différents (exemple traités/non traités)

2. Présentation des problématiques technique et scientifique

Le plasma* est une matrice biologique complexe à étudier en protéomique du fait de l'extrême abondance de certaines protéines majoritaires comme l'albumine*. En effet, ces protéines représentent 95% de la quantité totale des protéines dans le plasma [1].

Cette surexpression est problématique pour l'identification et la quantification d'échantillons plasmatiques car les peptides issus de ces protéines majoritaires masquent par compétition ionique la détection par spectrométrie de masse des peptides issus des protéines minoritaires.

Pour diminuer cette compétition ionique, il est possible d'avoir une action sur le nombre de protéines présentes dans l'échantillon (passage dans des colonnes de déplétion par exemple) et/ou une action sur le nombre de peptides qui rentrent à un instant t dans l'appareil de masse par un fractionnement peptidique.

Ainsi, dans cette étude, il a été choisi de tester ce fractionnement peptidique sur l'identification et la quantification des protéines plasmatiques. Les échantillons de plasma utilisés proviennent d'un projet scientifique soumis à la plateforme.

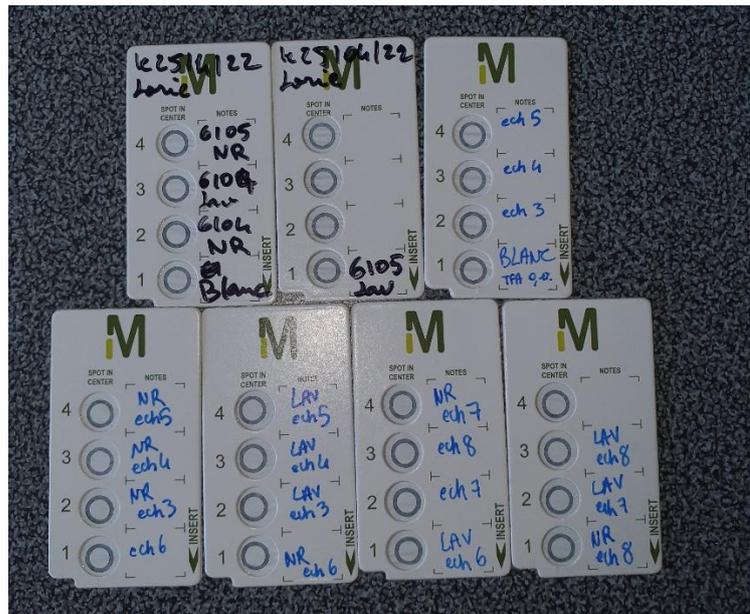


Figure 3. Plaque de Direct Detect™ avec les échantillons des pools après passage sur colonne

[Source personnelle](#)

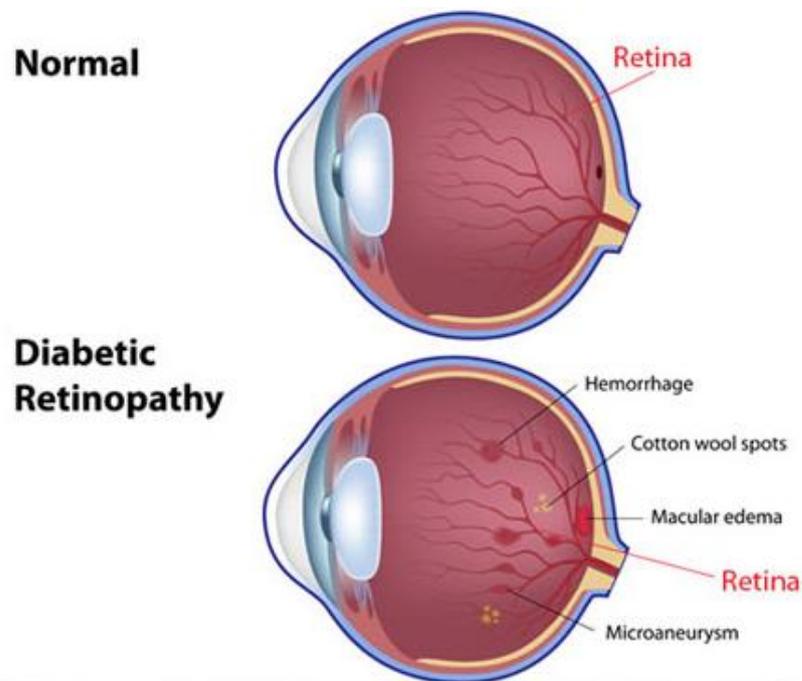


Figure 4. Image d'un œil atteint de rétinopathie diabétique

[odeme maculaire - Bing images](#)

Le but du projet est d'apporter des informations sur les mécanismes de formation des œdèmes chez les patients avec une pathologie oculaire.

L'œdème maculaire est une augmentation de volume de la macula, zone centrale de la rétine, responsable de l'acuité visuelle. Des symptômes visuels handicapent la vie de millions de patients atteints d'œdème maculaire secondaire à une maladie chronique et parfois aiguë de la rétine. Les protéines qui neutralisent la voie du facteur de croissance de l'endothélium vasculaire (VEGF) ou les glucocorticoïdes*, au prix d'injections intraoculaires répétées pendant des années, limitent les symptômes visuels. Mieux comprendre pourquoi et comment l'œdème se forme et comment les molécules thérapeutiques exercent un effet anti-œdémateux permettra de mieux prévenir la survenue de cette complication rétinienne handicapante [7].

Ainsi, deux populations de rats (sauvages (WT) ou mutés (TG1) pour un récepteur minéralocorticoïdes (sur expression)) ont été traités par une injection de dexaméthasone, un glucocorticoïde puissant. Le plasma de ces rats a été prélevé avant (J0, rats non traités (NT)) et après injection (J7, rats traités (T)), puis l'étude protéomique sur des pools* a été réalisée avec les deux workflows: fractionnement (F) ou sans fractionnement (NF). Ainsi, l'apport du fractionnement peptidique sera analysé sur la base du nombre de peptides et de protéines identifiées mais aussi sur les résultats de l'analyse de quantification label free entre les protéines des différents designs pouvant être étudiés (T vs NT, par exemple).

3. Matériels et Méthodes

3.1. Matériels

3.1.1. Direct Detect™ (Merck)

L'utilisation de Direct Detect™ permet de réaliser un dosage des protéines sans analyse biochimique. C'est un système de spectroscopie basée sur l'infrarouge (IR) avec lecture directe. En effet, cette technologie utilise une membrane en polytétrafluoroéthylène (PTFE) hydrophile conçue pour être transparente et permet l'application de solutions de biomolécules directement sur la membrane. Le système Direct Detect™ a été optimisé pour la détection et la quantification des protéines.

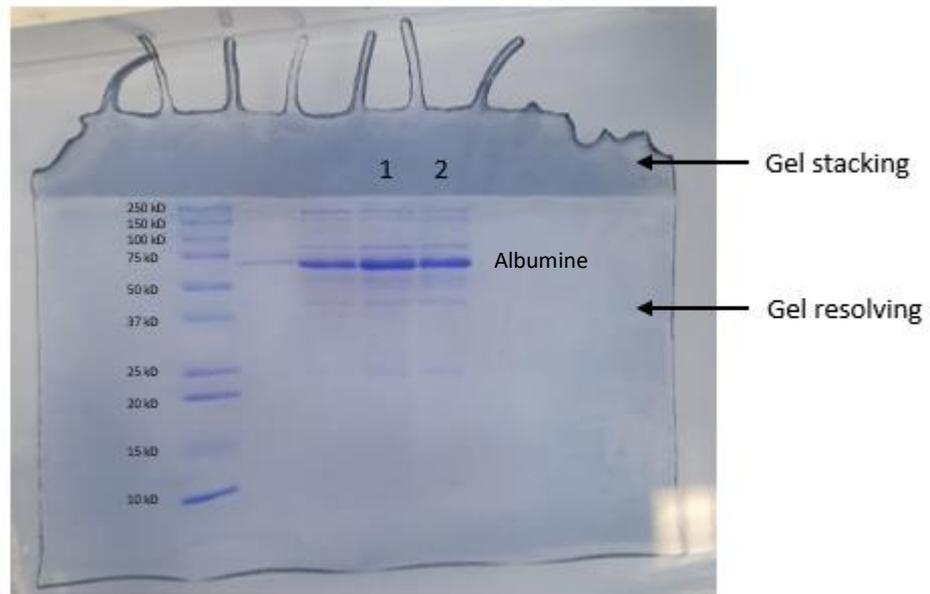


Figure 5. *Photo d'un gel 1D-SDS-PAGE après migration et coloration*

[Source personnelle](#)

1 : NT-TG1-Mâle 2 : NT-WT-Mâle

Exemple de séparation (migration totale) des protéines de rats transgéniques (TG1) et sauvages (WT). Chaque bande correspond à une classe de protéines ayant une masse moléculaire proches

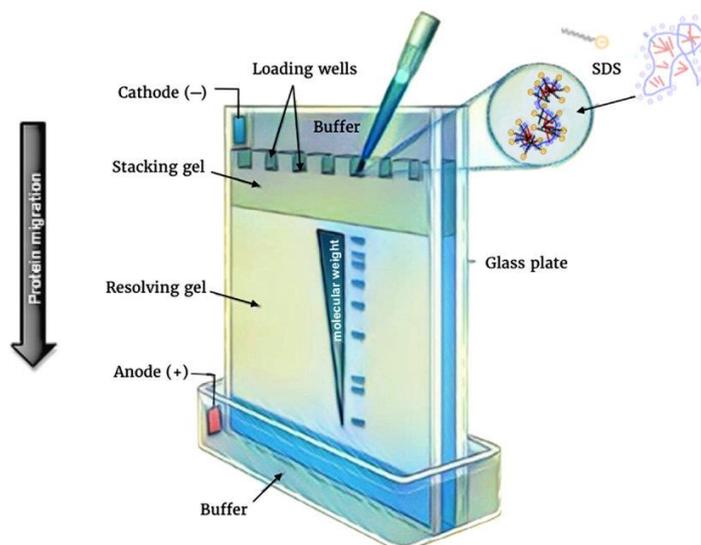


Figure 6. *Montage d'un gel 1D-SDS-PAGE*

[gel sds-page - Bing images](#)

Chaque liaison amide est constituée de sous-structures chimiques qui absorbent le rayonnement IR. Par conséquent, le Direct Detect™ permet de mesurer des concentrations de protéines de 0,2 mg/mL à 5 mg/mL en quelques minutes, sans aucune coloration bio- ou immuno-chimique, directement à partir d'échantillons, y compris de solutions tamponnées.[14].

3.1.2. Électrophorèse avec gel 1D-SDS-PAGE*

L'Électrophorèse 1D est une technique séparative des protéines selon leurs poids moléculaires. En effet, elle se fait en présence d'agents dénaturants tel que le SDS*, surfactant polyanionique*. En présence de ce surfactant, les protéines se comportent toutes comme des anions (chargé négativement). Ainsi, le seul critère de séparation est le poids moléculaire [11].

La technique du gel 1D-SDS-PAGE repose sur ce principe. Il est composé de deux gels distincts : un gel stacking* (comportant les puits de dépôts des échantillons) qui est faiblement concentré en acrylamide* afin de concentrer les protéines avant qu'elles passent dans le deuxième gel, qui est un gel resolving* plus concentré en acrylamide qui permet de séparer les protéines en fonction de leur poids moléculaire. Il est donc possible de faire plusieurs types de migrations : partielles ou totales. L'utilisation d'une migration partielle signifie qu'il faut arrêter celle-ci avant que les protéines passent dans le deuxième gel. Le but est de concentrer les protéines. L'utilisation d'une migration totale signifie que les protéines passent dans les deux gels afin que ces macromolécules soient séparées en classe sur la base de poids moléculaires similaires. En effet, la taille des pores des gels de polyacrylamide est liée par le rapport bisacrylamide/acrylamide, plus ce rapport est élevé, plus la taille des pores diminue. Ainsi, en faisant varier ce rapport entre deux gels, il est possible de réaliser deux séparations différentes [11].

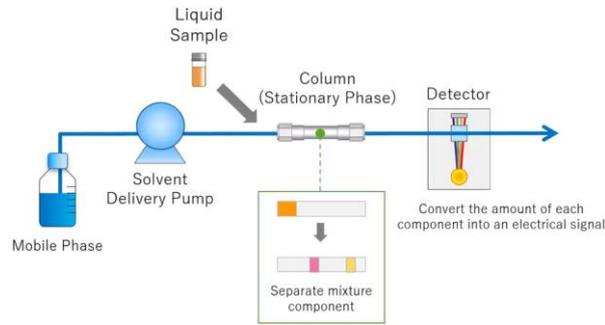


Figure 7. *Schéma simplifié de la vue d'ensemble du fonctionnement d'une HPLC*

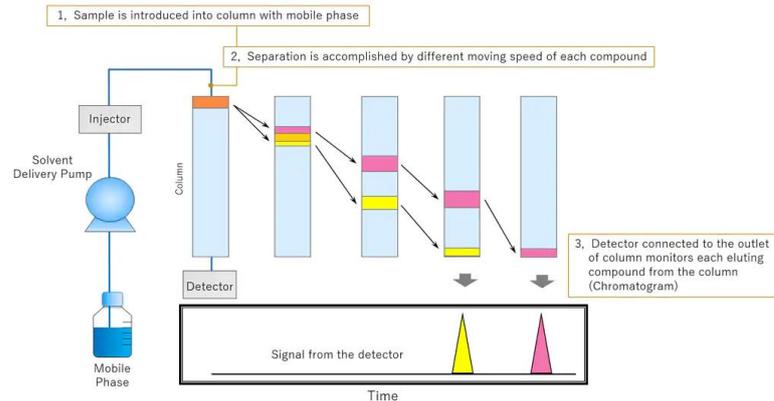


Figure 8. *Schéma simplifié expliquant la séparation par HPLC*

[What is HPLC \(High Performance Liquid Chromatography\) ? : SHIMADZU \(Shimadzu Corporation\)](#)

Ce schéma montre un exemple dans lequel le composant jaune a une forte affinité avec la phase mobile et se déplace rapidement dans la colonne, tandis que le composant rose a une forte affinité avec la phase stationnaire et se déplace lentement. La vitesse d'éluion dans la colonne dépend de l'affinité entre le composé et la phase stationnaire.

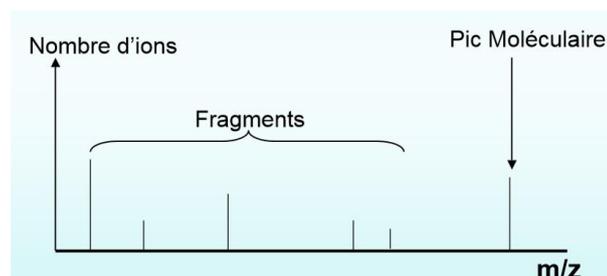


Figure 9. *Spectre de masse*

[Cours protéomique-1502-PowerPoint](#)

Le spectre de masse d'une molécule ionisée (spectre MS) et/ou de ses fragments (spectre MS/MS) se définit en abscisse par un rapport masse (m) sur charge (z) et en ordonnée comme une intensité (lié aux nombres d'ions détectés pour cette m/z).

3.1.3. Chromatographie liquide de Haute Performance (HPLC)

HPLC est l'abréviation de « High Performance Liquid Chromatography ». C'est une technique d'analyse permettant la séparation des constituants d'un mélange, celui-ci devant être dissous dans des solvants* car la HPLC ne peut analyser que des échantillons liquide. Le principe de cette technique est basé sur la distribution, dans une colonne analytique, des constituants de l'échantillon entre une phase mobile (liquide qui dissout le composé cible) distribuée par une pompe, et une phase stationnaire (liquide ou solide qui interagit avec le composé cible), possédant une large surface de contact. Chaque constituant est caractérisé par un temps de rétention (temps mis par les composés pour être élué de la colonne) et par le signal obtenu dans le détecteur placé en aval (détecteur uv et/ou spectromètre de masse). [11][13][2].

L'injection de l'échantillon est réalisée grâce à un passeur d'échantillon qui prélève dans une fiole HPLC et l'échantillon est ensuite acheminé à la colonne grâce au débit de solvant de la pompe HPLC. En protéomique, la pompe est une pompe nanodébit (200 à 400 nl/min), les solvants sont l'eau et l'acétonitrile en milieu acide, la colonne est une phase inverse (C18). Dans ce type de colonne, plus l'affinité (basée sur des liaisons faibles (liaisons hydrogène, force de van der Waals)) entre les peptides et la phase stationnaire est forte, plus ils se déplacent lentement dans la colonne, et donc plus leur temps de rétention (rt) est élevé [11][13][2]. Pour un même peptide, dans les mêmes conditions HPLC, le temps de rétention varie très peu entre deux analyses HPLC (quelques secondes).

3.1.4. Spectromètres de masse

La spectrométrie de masse est une méthode analytique mesurant la masse (m) d'une molécule en relation avec sa charge (z). En effet, le fait d'ioniser les molécules à analyser en leur donnant une charge électrique permet de mesurer leurs réponses à des forces électriques et magnétiques. Le spectromètre de masse est un appareil de mesure très performant dont la précision est de l'ordre de la partie par millions (ppm, 3 à 4 chiffres après la masse unitaire). De plus outre la précision de mesure, il faut aussi une résolution [11] très importante pour pouvoir séparer deux pics de masse très proches.

De façon générale, un spectromètre de masse est composé de trois parties : la source, l'analyseur et le détecteur.

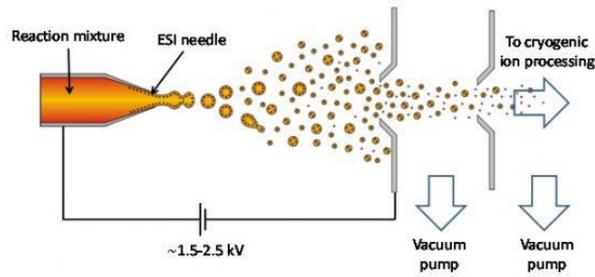


Figure 10. *Schéma expliquant la formation de nébulisats dans une source ESI*

[source ESI - Bing images](#)

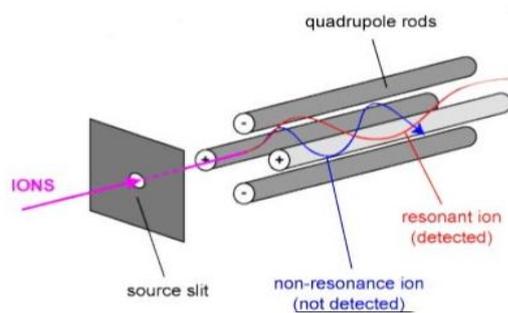


Figure 11. *Schéma d'un quadripôle*

[Mass part 2 2 \(slideshare.net\)](#)

Seuls les ions de trajectoires stables passent au travers du quadripôle, les autres sont déviés sur les barres dû aux différences de potentiels et à une tension alternative de celles-ci appliqués sur les barres du quadripôle.

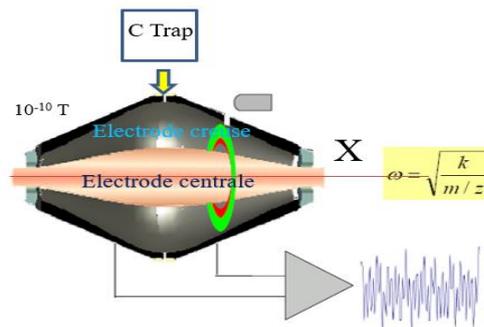


Figure 12. *Schéma d'un OrbiTrap*

[Cours protéomique-1502-PowerPoint](#)

La source est la partie où les ions sont formés. L'analyseur permet la séparation des ions en fonction du rapport masse sur charge. Enfin, le détecteur est la partie où les ions collectés viennent activer une cascade d'émission d'électrons aboutissant au final à l'amplification très importante du signal ionique. Ensuite ce signal est digitalisé et l'ordinateur de pilotage construit un spectre [11]. Il est possible de réaliser des analyses de spectrométrie de masses à deux échelles : MS ou MS/MS. La MS correspond à l'analyse m/z des peptides non fragmentés alors que la MS/MS correspond à l'analyse des ions fragments de chaque peptide détecté en MS et fragmentés. Cette double empreinte obtenue en MS-MS/MS est utilisée lors de l'analyse d'un mélange complexe, où l'information de la masse des peptides n'est pas suffisante pour l'identification.

Sur la composante protéomique, le spectromètre de masse est un ESI-Q-OrbiTrap. ESI correspond à la source utilisée (Electrospray Ionisation). Ce type de source provoque une accumulation de charges à la surface du liquide en sortie du capillaire, ce qui va former des fines gouttelettes hautement chargées appelées nébulisats. La source est reliée à un quadripôle. Il sert de filtre pour envoyer dans le deuxième analyseur un ion unique de masse donnée. Le deuxième analyseur est un OrbiTrap. Les ions vont d'abord être piégés dans une Trappe ionique (C trap) puis injectés en paquet dans l'Orbitrap. Au sein, de cette « cage à champ électronique » les ions ne vont pas percuter les parois mais vont se mettre en orbite autour de l'électrode centrale. L'obtention de la masse de l'ion est une mesure indirecte basée sur la fréquence d'oscillation axiale des ions autour de l'électrode centrale [11][12].

3.1.5. Logiciels d'analyses

Il existe plusieurs logiciels permettant de traiter et d'analyser les données protéomiques issues d'un spectromètre de masse. Au laboratoire, les plus utilisés sont le logiciel de quantification Progenesis QI (Nonlinear) et le moteur de recherche Mascot (Matrix Science) [21].

ProQI est un logiciel d'analyse de données LC-MS/MS de nouvelle génération. Il permet la quantification des pics de masse détectés par le spectromètre de masse. Le Workflow* s'organise en plusieurs étapes, dont les principales sont l'alignement automatique des temps de rétention de toutes les cartes ioniques des échantillons passés en LC-MS/MS,

la quantification des signaux ioniques communs et l'identification via Mascot des peptides (étape externalisée). Une fois identifiés, les abondances des peptides de chaque protéine sont utilisées pour calculer l'abondance des protéines présentes dans chaque échantillon. Une étude statistique détermine, suivant le design de comparaison choisi, les protéines ayant une ANOVA significativement différente et donc faisant partie des protéines dont l'abondance est modifiée (protéines différentielles) dans le design étudié (exemple Echantillons Traités vs Non Traités) [16].

Le logiciel d'identification Mascot est très utilisé en protéomique. Il est basé sur un algorithme de comparaison des spectres observés contre des spectres théoriques (banques de données de séquences de protéines). En effet, les données sont exportées vers ce logiciel et les banques de données sont interrogées avec des paramètres de recherche choisis par l'analyste. Mascot compare donc les listes de masses MS puis MS/MS expérimentales produites à partir de l'échantillon lors de l'analyse LC-MS/MS avec celles théoriques de la base de données utilisée. Spectre MS/MS par spectre MS/MS, le logiciel calcule alors un score basé sur la probabilité que les spectres MS/MS observés/théoriques soient identiques et donne des paramètres expérimentaux (séquences des peptides, annotations des spectres, écarts en masse entre mesure et théorie) pour permettre la validation de la protéine dont sont issus les peptides analysés [5][8].

Ainsi, grâce à ces deux logiciels, il est possible d'identifier et de quantifier les protéines étudiées. A la fin d'un Workflow LC-MS/MS_ProQI-MASCOT, la liste de protéines obtenues peut être analysée par des outils de bioanalyse permettant de caractériser les protéines dans leur environnement biologique. Deux outils en open source sont utilisés au laboratoire : String (Search Tool for the Retrieval of Interacting Genes/Proteins) et Panther (Protein ANalysis THrough Evolutionary Relationships).

L'outil String est une base de données biologiques et une ressource web d'interactions protéine-protéine connues et prédites. La base de données String contient des informations provenant de nombreuses sources, notamment des données expérimentales, des méthodes de prédiction computationnelles et des collections de textes publics. Les réseaux d'interaction protéine-protéine sont un élément important pour la compréhension des processus cellulaires. Ces réseaux peuvent être utilisés pour filtrer et évaluer les données de génomique fonctionnelle et fournir une plateforme intuitive pour annoter les

propriétés structurelles, fonctionnelles et évolutives. Ils permettent aussi de fournir des prédictions inter-espèces pour une cartographie efficace des interactions [17].

Le logiciel Panther est un système complet qui combine la fonction des gènes, l'ontologie, les voies d'accès et les outils d'analyse statistique permettant aux biologistes d'analyser des données à grande échelle sur le génome entier provenant d'expériences de séquençage, de protéomique ou d'expression génétique. Ainsi, avec Panther, il est possible d'obtenir des informations sur un gène d'intérêt mais aussi d'explorer les familles de protéines. De plus, il peut informer sur les fonctions moléculaires, les processus biologiques, les composants cellulaires et les voies d'accès. Il permet de générer des listes de gènes appartenant à une famille ou sous-famille de protéines données. Panther fournit deux tests statistiques : le test de surreprésentation (compare une liste de gènes de test à une liste de gènes de référence, et détermine si une classe particulière de gènes est surreprésentée ou sous-représentée) et le test d'enrichissement (détermine si les valeurs numériques des gènes associés à une classe d'ontologie ou à une voie particulière ont été tirées au hasard de la distribution globale des valeurs. Le test U de Mann-Whitney (test de Wilcoxon Rank-Sum) est utilisé pour déterminer la valeur P) [18].

3.2. Mode opératoire

3.2.1. Préparation et quantification protéiques des pools d'échantillons

Les échantillons de plasma de rats utilisés font partie d'une analyse à plus grande échelle destinée à déterminer les biomarqueurs de pathologie oculaire. Dans ce cadre-là, 75 échantillons sont reçus et seront analysés individuellement dans une analyse en juin prochain sur la plateforme. Dans le cadre de notre analyse, nécessitant de réaliser des fractionnements (8 fractions par échantillon) et des répétitions techniques, il n'est pas possible de réaliser 900 runs LC-MS/MS (75x8 échantillons (Fractionné) + 75 x 4 échantillons (Non Fractionné)). Ainsi, la première préparation consiste à former des pools regroupant tous les échantillons similaires. Ainsi, 8 pools ont été créés : WT-J0 femelle, WT-J7 femelle, TG1-J0 femelle, TG1- J7 femelle, WT-J0 mâle, WT-J7 mâle, TG1-J0 mâle, TG1- J7 mâle. Pour cela, 10µl de chacun des 75 échantillons sont prélevés et poolés en 8 échantillons selon la répartition donnée plus haut.

10.1.1 Préparation du gel resolving

Le rapport Acrylamide/Bis-acrylamide est de 37,5/1

	Resolving 8%	Resolving 10%	Resolving 12%	Resolving 15%
Acrylamide 40%	5 ml	6.25 ml	7.5 ml	9.38 ml
Tampon Tris-HCl 1.5 M pH 8.8	6.25 ml	6.25 ml	6.25 ml	6.25 ml
S.D.S 10 %	0.25 ml	0.25 ml	0.25 ml	0.25 ml
H2O	13.36 ml	12.11 ml	10.86 ml	8.98 ml
Ajouter les 2 composés suivant avant de couler les gels				
APS 10 % (0.1g / 1ml H2O)	0.125 ml	0.125 ml	0.125 ml	0.125 ml
Temed	12.5 µl	12.5 µl	12.5 µl	12.5 µl
volume final	25 ml	25 ml	25 ml	25 ml

Une fois le gel coulé ajouter de l'eau plus alcool (1/1) à la surface des gels.
Attendre 30mn pour une polymérisation complète.

10.1.2 Préparation du gel stacking à 5%

Le rapport Acrylamide/Bis-acrylamide est de 37,5/1

	Stackink 5%
Acrylamide 40%	1.25ml
Tampon Tris-HCl 0,5 M pH 6.8	2.5 ml
S.D.S 10 %	0.1 ml
H2O	6.1 ml
Éliminer l'eau+alcool en surface du resolving puis ajouter les 2 composés suivant avant de couler vos gels	
APS 10 %	50 µl
Temed	5 µl
volume final	10 ml

Une fois le gel coulé, insérer les peignes au format des spacers choisis.
Attendre 1 heure pour polymérisation complète avant l'étape suivante.

Figure 13. Protocoles de préparation des gels pour une électrophorèse 1D-SDS-PAGE

Source interne à la PFEM

Reagent	Molecular weight	Concentration (M or %)			Add for 50 ml of 2X	Add for 50 ml of 4X
		1X	2X	4X		
Tris base ¹	121.14	0.0625 M	0.125 M	0.250 M	0.747 g	1.514 g
SDS ²	288.37	0.07 M (2%)	0.14 M (4%)	0.28 M (8%)	2 g	4 g
glycerol	92.09	10%	20%	40%	10 ml	20 ml
2-mercapto-ethanol	78.13	5%	10%	20%	5 ml	10 ml
Bromphenol blue ³	691.94	-	-	-	100 mg	200 mg

¹Tris base is tris (hydroxymethyl) aminomethane. You can avoid using crystalline Tris by using Tris buffer, adjusted with HCl to 6.8.

²SDS is sodium dodecyl sulfate.

³Bromphenol blue is available as sodium salt or solution. Some modern protocols are using higher concentration (0.005 % Bio-rad, 0.002 % Sigma-Aldrich) to obtain bright colour.

Figure 14. Composition du tampon dénaturant : Laemmli

Laemmli buffer: Preparation (1x,2x & 4x) and principle - Sharebiology

Pour assurer la traçabilité des échantillons selon la procédure qualité ISO 9001 et NFX50-900 en vigueur au laboratoire, il faut les identifier. Il existe un logiciel dédié : Labcollector. Il est nécessaire de réaliser un tableau Excel contenant toutes les informations du demandeur (noms d'échantillons, nombre d'échantillons etc.). Ce tableau est exporté dans Labcollector et le logiciel va créer des numéros uniques pour chaque échantillon. Cela permet de les suivre tout au long de la manipulation et de les retrouver dans les stocks. Une fois les pools terminés, il faut faire un dosage pour savoir la quantité de protéines dans chaque échantillon. Un dosage au Direct Detect™ est réalisé. Il faut déposer 2µl de l'échantillon à doser sur la membrane de la plaque Direct Detect™. Une fois cela réalisé, la plaque est laissée à l'air libre toute la nuit pour sécher. Ensuite, elle passe dans la machine et une lecture infra-rouge est réalisée. Les résultats obtenus donnent la quantité de protéines en µg/µl dans chaque échantillon. A partir de ces résultats, il est possible de calculer le volume à prélever dans le pool pour déposer dans les puits des gels la quantité désirée de protéines pour la migration des échantillons.

3.2.2. Immobilisation des échantillons sur bande 1D-SDS-PAGE

Les volumes déposés dans les puits formés dans le gel stacking (5% de bisacrylamide), sont ajustés pour obtenir une concentration de 40µg de protéines pour chaque pool. Une migration partielle est réalisée dans le but de concentrer les échantillons. Au préalable les échantillons sont dénaturés à 100°C au bain marie pendant cinq minutes avec un tampon Laemmli de dénaturation (voir figure 14). Puis la migration dans le gel de concentration est lancée avec un voltage de 80 V pendant 18 minutes. Une fois les échantillons au bout du premier gel, il faut stopper la migration avant que ceux-ci ne passent dans le deuxième gel resolving (12% de bisacrylamide). Dès que la migration est stoppée, il s'en suit plusieurs étapes successives d'incubation dans des bains de fixation des protéines, de coloration au bleu de Coomassie et de décoloration. Vient alors le prélèvement des bandes. Pour cela, il faut déposer le gel sur une plaque de verre et l'humidifier avec de l'eau milliQ. Ensuite, il convient de se munir d'un scalpel et de découper une à une les bandes souhaitées qui sont ensuite placées dans des tubes Eppendorf de 0.5ml pour être traitées.

3.2.3. Traitement et hydrolyse des protéines immobilisées dans le gel 1D-SDSPAGE

Le traitement des bandes s'effectue en plusieurs étapes : la réduction-alkylation, la décoloration, la déshydratation, l'hydrolyse* à la trypsine et l'extraction des peptides.

Les protéines sont réduites et alkylées par des incubations successives dans différentes solutions : d'abord avec 100µl de tampon ABC* 50mM*-50% éthanol pendant 20 minutes à température ambiante. Il faut éliminer le tampon avant d'ajouter 100 µl d'éthanol pur laissé 15 minutes. Les ponts disulfures sont ensuite réduits avec 200 µl de tampon DTT* 10mM. Les tubes sont mis à incuber à 56°C pendant 1 heure.

Après cette étape, les protéines sont alkylées avec 200µl de tampon IAA* -ABC 50mM. Il faut laisser incuber 30 min à l'abri de la lumière, puis la solution d'alkylation est éliminée. Les bandes sont alors lavées avec 200µl de tampon ABC 50mM pendant 15 minutes.

Vient alors l'étape de décoloration avec premièrement un tampon ABC 25mM-5% acétonitrile qu'il faut laisser incuber pendant 15 min puis, avec un deuxième tampon ABC 25mM-50% acétonitrile qui doit incuber 30 min sous agitation. Cette dernière étape est réalisée deux fois. Cependant, pour des échantillons dont la coloration est encore très soutenue après les deux lavages, il est possible de faire un troisième bain en laissant un temps suffisant pour une décoloration plus complète. Néanmoins, il n'est pas rare que le gel reste coloré même après les trois bains.

Les gels sont ensuite déshydratés par 5 à 10 min d'incubation dans 200µl d'acétonitrile pur. Le temps nécessaire pour déshydrater est fonction de la dimension du gel. Il faut alors éliminer cette solution et passer les tubes Eppendorf contenant les bandes traitées au SpeedVac* pour une élimination complète du solvant.

Par la suite, les protéines contenues dans les prélèvements de gels sont hydrolysées en peptides par la trypsine*. Pour cela, une solution de trypsine à 10 ng/µl dans un tampon ABC 25mM est préparée. La trypsine est alors déposée directement sur le gel déshydraté. Puis, les échantillons sont incubés toute la nuit à 37°C [15].

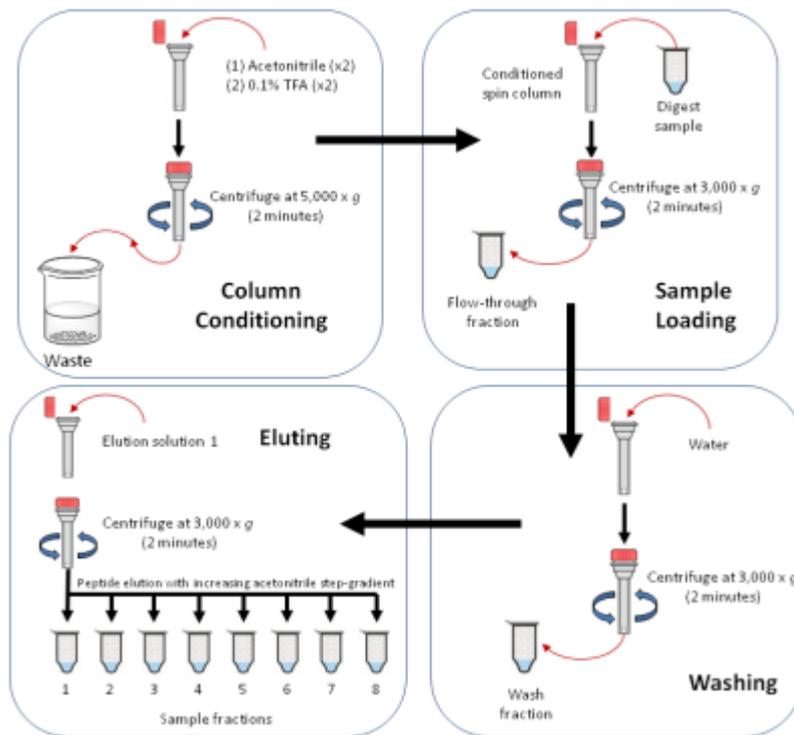


Figure 15. *Workflow pour le fractionnement des échantillons*

[84868_highph_rp_peptidfrac_PIERCE.pdf](#)

Enfin, les peptides sont extraits du gel par de l'acétonitrile acidifié (0.1% acide formique). Le volume de celui-ci correspond à 80% du volume d'incubation dans la trypsine. Une fois passer au vortex* puis au bain à ultrasons* pendant 10 minutes, les échantillons sont transférés dans des tubes Eppendorf, puis séchés au SpeedVac pour éliminer l'acétonitrile. Les échantillons sont repris avec 25µl d'une solution H₂O/TFA* 0,05% puis après sonication 10 min, ils sont transférés dans un Vial HPLC pour être analysés dans le spectromètre de masse ESI-QOrbiTrap.

3.2.4. Fractionnement des peptides en milieu basique sur colonne de phase inverse

Le kit de fractionnement des peptides de Thermo Scientific™ Pierce™ fournit un protocole optimisé pour séparer les peptides plasmatiques en plusieurs fractions et ainsi décomplexifier l'échantillon en vue d'une analyse LC-MS/MS.

La chromatographie en phase inverse en condition basique est une méthode robuste de fractionnement des peptides qui permet de séparer les peptides en fonction de leur hydrophobicité* et fournit une excellente orthogonalité aux gradients LC-MS/MS en phase inverse en condition acide. Le kit comprend une solution à pH élevé (0,1% de triéthylamine*) et 12 colonnes à centrifuger contenant une résine en phase inverse résistant donc aux conditions basiques. Chaque colonne de fractionnement en phase inverse permet de traiter 10 à 100 µg d'échantillon de peptides.

Ainsi, le protocole se déroule en deux étapes : le conditionnement des colonnes et le fractionnement des échantillons. Tout d'abord, il faut placer la colonne dans un tube de 2ml pour que le passage dans une centrifugeuse soit possible. Le premier passage dans la centrifugeuse dure 2 minutes à 5000g. Cela va permettre d'éliminer la solution de stockage de la colonne et de tasser la résine. Puis, viennent les étapes de préparation de la colonne au fractionnement. Un volume de 300µl d'une solution d'acétonitrile 100% est ajouté et éliminé par centrifugation. Cette étape est renouvelée 1 fois, puis une équilibration de la colonne a lieu de la même manière en remplaçant l'acétonitrile par du TFA 0,1%. Ainsi la colonne est prête à l'emploi, équilibrée dans le même tampon que les échantillons.

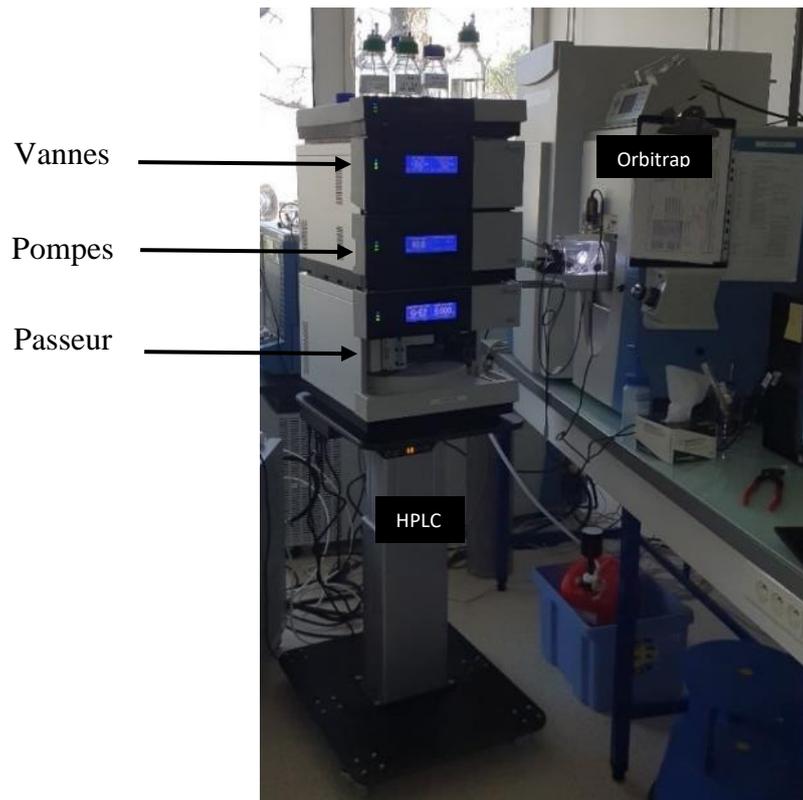


Figure 16. Couplage HPLC - spectromètre de masse

[Source personnelle](#)

Ici est représenté le matériel utilisé pour l'analyse des peptides. C'est une nanoHPLC (High Performance Liquid Chromatography), à gauche, reliée à un appareil de masse qui est ici, un Orbitrap (HFX), à droite.

Pour les étapes de fractionnement, la procédure est la même que l'étape de préparation de la colonne. Il est toutefois nécessaire de préparer au préalable les solutions d'élution de concentrations croissantes en acétonitrile (de 5% (F1) à 50% (F8)). Ainsi, après avoir ajouté 300µl de TFA à 0,1% dans chaque échantillon, ils sont individuellement fractionnés sur la colonne, en suivant les étapes décrites dans le protocole Pierce (voir figure 13) [9].

3.2.5. Analyses LC-MS/MS

Pour réaliser une analyse MS/MS, il faut tout d'abord créer une séquence afin de conserver la traçabilité des échantillons et paramétrer la méthode d'analyse à utiliser. La méthode LC comporte l'ensemble de la programmation des blocs (pompes, passeur, vannes) composant la LC. Cette dernière est couplée à l'appareil de masse de type Q-Orbitrap HFX (ThermoScientific) par une source nanoESI. La méthode masse MS/MS employée en protéomique est une méthode dite « Data Dependant Analysis » (DDA). Dans ce type de méthode, le spectromètre de masse alterne une détection en full scan MS (ions parents) et une détection en MS/MS après fragmentation dans une cellule de collision, des ions présents dans le full scan (ions fragments)). La résolution du mode MS est fixée à 60000 alors que pour le mode MS/MS, elle est fixée à 15000 [15].

Les hydrolysats peptidiques avec (F) ou sans fractionnement (NF) sont injectés (6µl) successivement dans la nano-HPLC. Ils passent alors sur une colonne de chargement pour les concentrer et éliminer les contaminants avant de passer sur une colonne analytique C18 (Acclaim PepMap 100, 75µm x 25 cm, nanoViper, C18, 3µm, 100 Å) alimentée par un débit 400nl/min, afin de séparer les peptides. Cette séparation est réalisée en milieu acide, grâce à un gradient d'acétonitrile dont la concentration augmente de 4% à 25% en 60 minutes. Cela permet l'élution des peptides selon leur hydrophobicité en condition acide.

3.2.6. Identification des peptides et Quantification des protéines

Les spectres obtenus lors des événements de fragmentation sont associés à une séquence peptidique et accompagnés d'un score Mascot. Plus celui-ci est haut, plus la probabilité de corrélation est bonne. Cependant, plus le nombre de séquences présentes dans une base de données est important, plus le risque de trouver un match « par hasard » (ou faux positif) devient important.

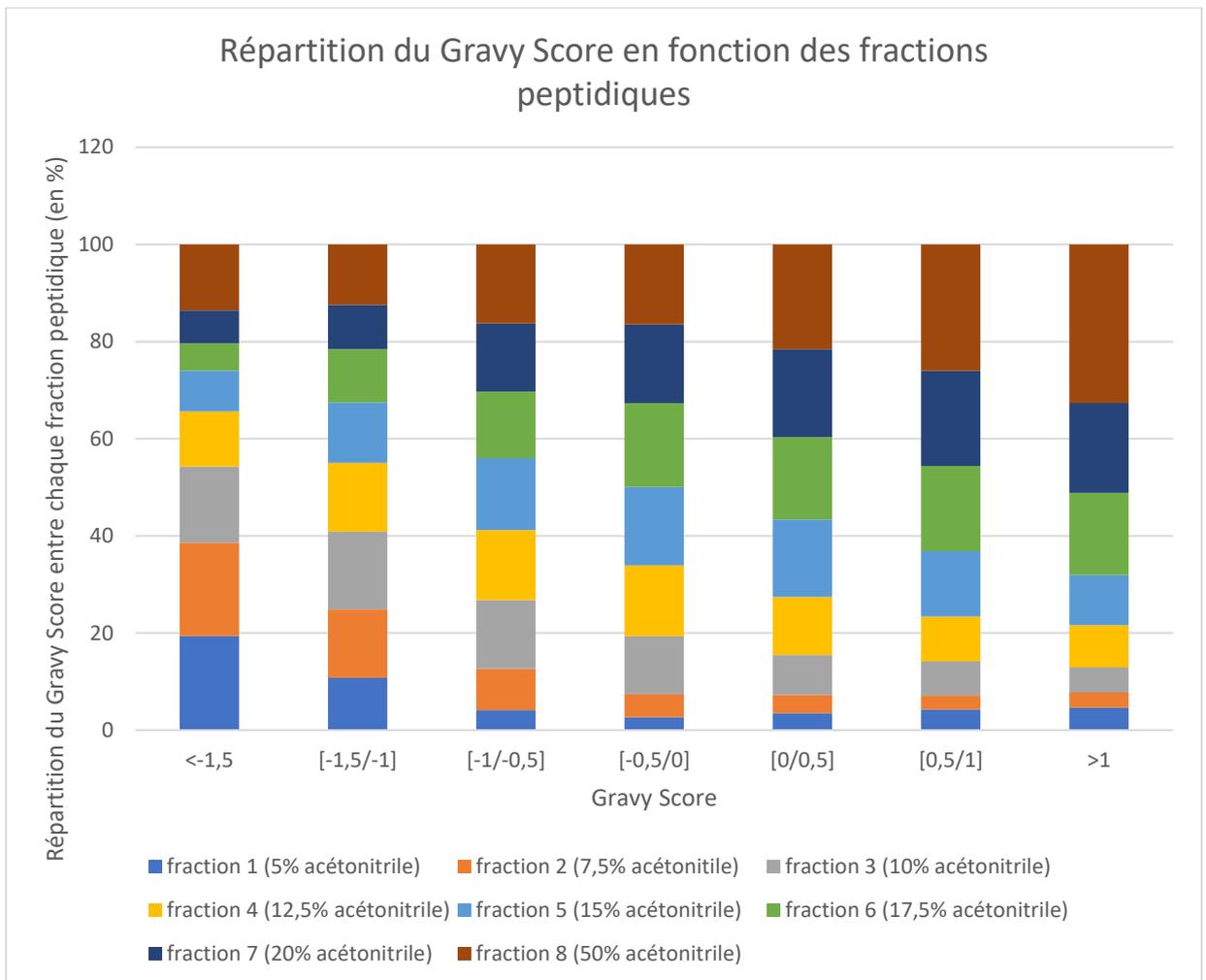


Figure 17. Répartition du Gravy Score en fonction des fractions peptidiques

Les paramètres de recherche dans la banque de rats pour notre étude ont été fixés à 10 ppm pour les ions parents (MS) et fragments (MS/MS), les modifications post traductionnelles ont été réduites à Oxydation* des méthionines (M), carbamidomethylation des Cytéines (C) et déamidation possible des Asparangine (N) et Glutamine (N) et les protéines validées avec 2 peptides uniques.

Pour la validation, une séquence est validée pour un peptide quand le score attribué par Mascot à cette interprétation a un taux de faux positif (FDR) inférieur à 1% [5][8][16].

4. Résultats et Discussion

4.1. Caractérisation physicochimique des peptides tryptique

Un calculateur de la valeur d'hydropathie (Gravy) [19] a été utilisé pour déterminer la répartition des scores d'hydrophobie des peptides en fonction du fractionnement [4]. Un score Gravy plus négatif indique un peptide aux propriétés plus hydrophiles, tandis qu'un score Gravy plus positif indique un peptide avec des propriétés plus hydrophobes. Les résultats (figure 17) montrent que les fractions éluées avec au moins 90% d'eau (F1 à F3), contiennent effectivement les peptides de Gravy score les plus faibles. Plus de 55% des peptides ayant des Gravy score inférieur à -1.5 appartiennent aux fractions F1, F2 et F3. C'est aussi vrai pour les peptides de Gravy Score -1.5 à -1, le pourcentage étant quand même plus faible avec un peu plus de 40%. Ainsi, un pourcentage élevé de peptides hydrophiles a été élué à l'aide d'un solvant avec un fort pourcentage d'eau. A l'inverse, les fractions contenant un fort pourcentage d'acétonitrile, supérieur ou égal à 30%, éluent le plus de peptides à Gravy score supérieur à 0. Ainsi, il y a 55 à 65% de ces peptides qui sont élués dans les dernières fractions 6, 7, et 8. En effet, les molécules hydrophobes sortent en dernier de la colonne. Il est nécessaire d'augmenter le gradient d'acétonitrile pour que les molécules les plus accrochées à la colonne puissent être éluées.

Par conséquent, la colonne de fractionnement en milieu basique (pH élevé) est efficace pour séparer les peptides en fonction de leur hydrophobicité puisque chaque fraction n'a pas la même composition sur la base des Gravy score. Cette rétention différentielle permet donc de réduire la complexité de l'échantillon et augmenter théoriquement la capacité d'analyse du spectromètre de masse [3].

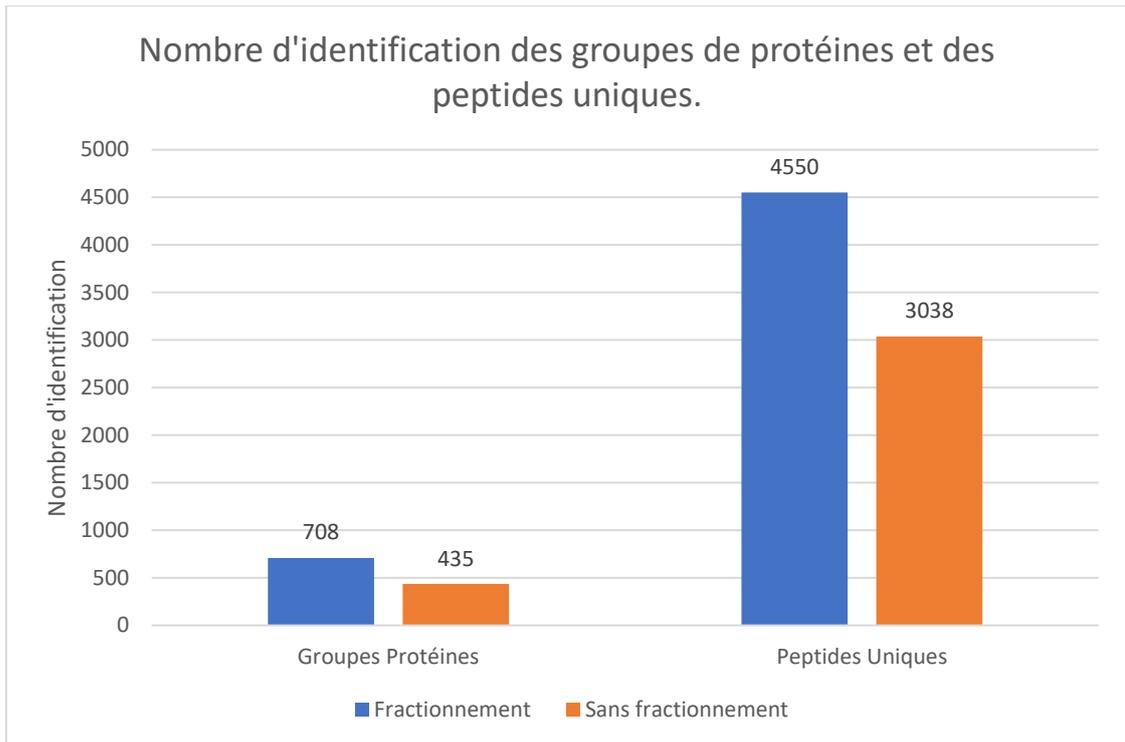


Figure 18. Nombre d'identification des groupes de protéines et des peptides uniques

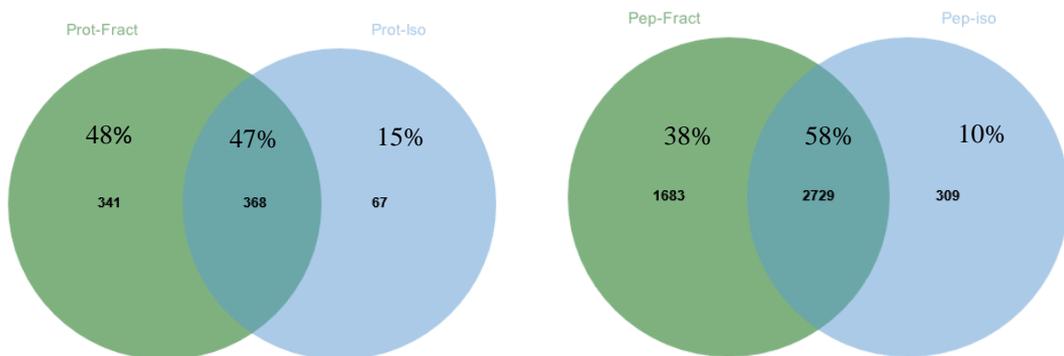


Figure 19 Diagramme de Venn montrant la distribution de l'identification des protéines et des peptides

4.2. Identification des protéines et des peptides

La figure 18 détaille le nombre de peptides et de protéines identifiés selon la méthode de préparation des échantillons. Les résultats montrent l'identification de 708 protéines et 4550 peptides avec le fractionnement contre 435 et 3038 sans fractionnement. En combinant les résultats ci-dessus, un total de 776 groupes de protéines et 4721 peptides uniques ont été identifiés en enlevant les séquences communes. Le fractionnement des peptides en milieu basique sur phase inverse a permis d'augmenter de 67% le nombre de protéines et d'environ 55 % le nombre de groupes de protéines et de peptides uniques identifiés, respectivement, par rapport aux résultats obtenus sans fractionnement. Ces résultats montrent donc que la méthode avec fractionnement permet une augmentation très significative de l'identification des protéines et des peptides avec la moitié des identifications de protéines obtenues grâce à cette méthode [3].

De plus, le diagramme de Venn (figure 19) montre que les groupes de protéines (distribution en pourcentage) et les peptides (distribution en pourcentage) possèdent des séquences communes : 368 (47 %) et de 2729 (58 %) respectivement. Aussi, il est notable qu'un nombre supérieur de protéines et de peptides spécifiques soient identifiés avec la méthode avec fractionnement (341 pour les protéines et 1683 pour les peptides).

La combinaison des résultats d'identification générés par les fractions (F1 à F8) améliore donc le nombre de peptides et de protéines identifiés par rapport aux échantillons non fractionnés.

4.3. Influence du fractionnement sur l'identification de séquences peptidiques différentes

Ainsi, la capacité à séparer les peptides de séquences différentes par la colonne Pierce a été évaluée en fonction des chevauchements des séquences des peptides uniques identifiés entre les fractions.

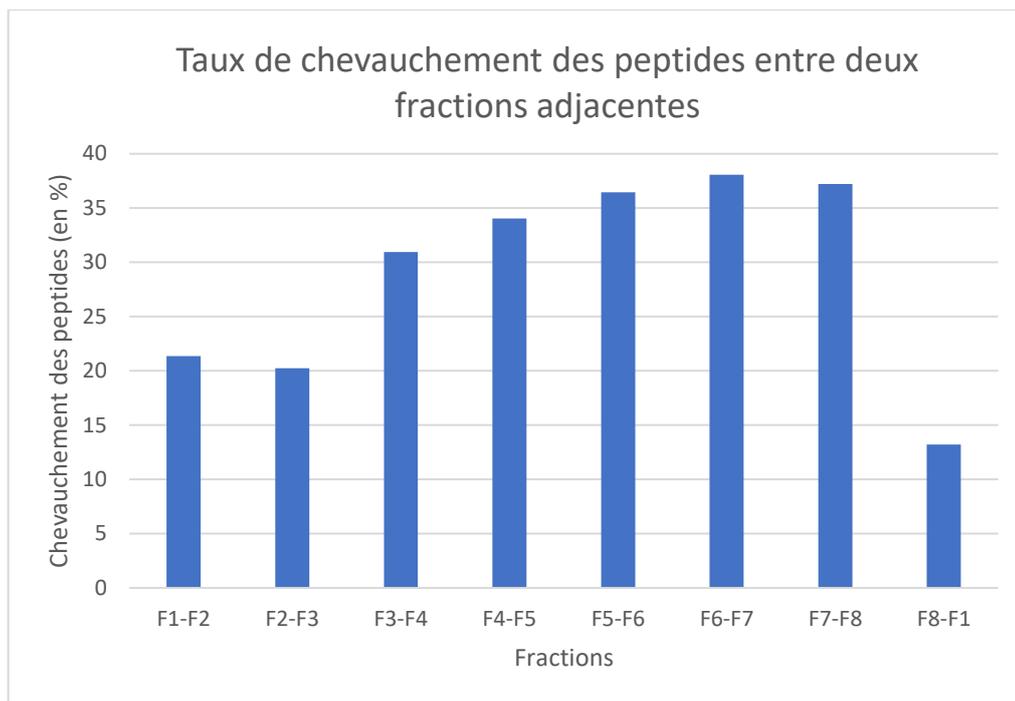


Figure 20. Taux de chevauchement des peptides identiques entre deux fractions adjacentes

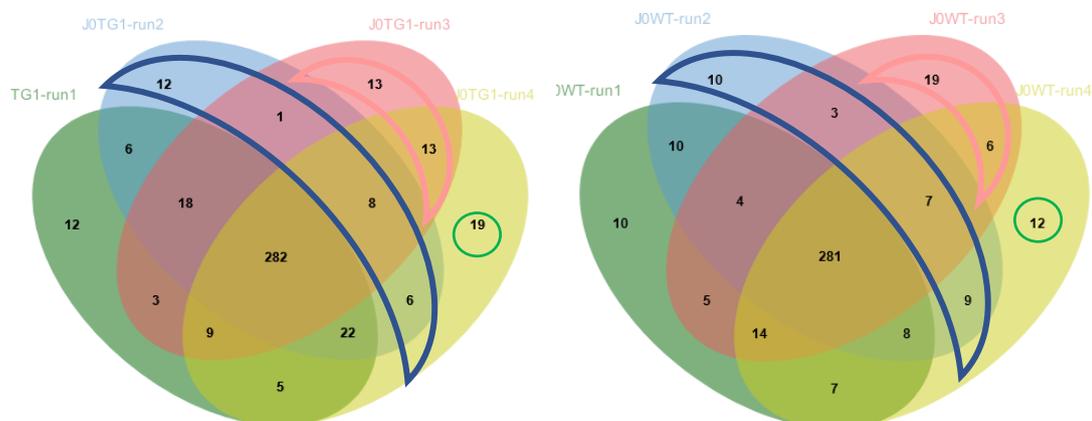


Figure 21. Diagrammes de Venn montrant le nombre de protéines dans la méthode sans fractionnement (4 répétitions techniques de LC-MS/MS)

Les protéines entourées correspondent aux protéines identifiées lors de chaque run :

-  Protéines identifiées lors du run-2
-  Protéines identifiées lors du run-3
-  Protéines identifiées lors du run-4

La figure 20 montre le taux de chevauchement des peptides entre deux fractions adjacentes. Il est possible de voir qu'entre les trois premières fractions, plus de 75% des peptides élués sont différentes entre les fractions. En effet, le taux de chevauchement n'étant que d'environ 20%. Ces trois fractions contiennent les peptides les plus hydrophiles (voir figure 17). Pour les peptides moins hydrophiles et même hydrophobes (F4 à F8), le chevauchement est plus important mais ne dépasse pas 40%. L'objectif de dé-complexification est bien atteint par la méthode. Il reste néanmoins un peu plus de 10% des peptides élués dans la fraction 8 qui étaient présents dans la fraction 1, montrant que d'autres paramètres que les caractéristiques physicochimiques distinctes des peptides seraient aussi à évaluer (une adaptation du protocole d'élution pourrait aussi permettre de réduire ce taux de chevauchement F1F8). En effet, dans la fraction 8, le pourcentage d'acétonitrile est de 50% alors que dans la fraction 1, il est de 5%. Sur la base des interactions hydrophobes pures, il ne devrait pas avoir de peptides de séquences identiques chevauchantes entre ces deux fractions [3]. Ainsi, il serait peut-être plus adapté de passer de 5% d'acétonitrile (élution pour F1) à 30% (élution pour F8) au lieu des 50% (élution pour F8) indiqué dans le kit Pierce. Cela permettrait peut-être une efficacité de fractionnement encore meilleure.

4.4. Influence des répétitions techniques LC-MS/MS sur l'identification des protéines

Chaque pool non fractionné a été analysé avec 4 répétitions techniques afin de pouvoir être comparé au même pool fractionné en 8 éluions. (8 répétitions auraient été parfaitement identiques aux 8 runs des fractions (F1 à F8), néanmoins les contraintes de temps machine ne l'ont pas permis).

Il est connu que les répétitions techniques permettent d'obtenir beaucoup plus de « peptide spectrum mass » (PSM) et donc d'identification (du fait du caractère aléatoire des déclenchements des MS/MS au cours d'une analyse) qu'une simple analyse sans répétition. Pour connaître l'influence des répétitions sur l'identification des protéines, les résultats d'identification obtenus sur les répétitions J0WT et J0TG1 ont été analysés (figure 21).

Designs	Fractionné (Tf*)	Commun (%)	Non fractionné (Tnf**)	Total
D0-TG1(T-NT,M-F) vs WT(T-NT,M-F)	6 (7)	1 (3%)	26 (27)	33
D1-T(TG1(M)-WT(M)) vs NT(TG1(M)-WT(M))	24 (58)	34 (14%)	140 (174)	198
D2-T(TG1(F)-WT(F)) vs NT(TG1(F)-WT(F))	15 (21)	6 (3%)	172 (178)	193
D3-T(TG1(M,F)-WT(M,F)) vs NT(TG1(M,F)-WT(M,F))	45 (108)	63 (28%)	115 (178)	223
D4-T(TG1(M,F)) vs T(WT(M,F))	15 (16)	1 (3%)	34 (35)	50
Total (%)	105 (14%)	105 (14%)	487 (72%)	697

*Total protéines fractionnées

**Total protéines non fractionnées

Tableau 1. Nombre de protéines quantifiées différentielles selon chaque design

D0-TG1(T-NT,M-F) vs WT(T-NT,M-F) : Design 0 comparant les rats sauvages et transgéniques, traités/ non traités et mâles/femelles tout confondus.

D1-T(TG1(M)-WT(M)) vs NT(TG1(M)-WT(M)) : Design 1 comparant les rats traités et non traités, transgéniques/sauvages mâles.

D2-T(TG1(F)-WT(F)) vs NT(TG1(F)-WT(F)) : Design 1 comparant les rats traités et non traités, transgéniques/sauvages femelles.

D3-T(TG1(M,F)-WT(M,F)) vs NT(TG1(M,F)-WT(M,F)) : Design 3 comparant les rats traités et non traités, transgéniques/sauvages mâles et femelles confondus.

D4-T(TG1(M,F)) vs T(WT(M,F)) : Design 4 comparant les rats traités, transgéniques/sauvages mâles et femelles confondus.

Nombre total de protéines identifiées : 697.

Le nombre moyen d'identification est de 355 pour les identifications J0TG1 et 338 pour les J0WT, avec 282 (TG1) et 281(WT) communes à tous les runs. A chaque répétition technique, le nombre de nouvelles protéines identifiées augmente. Dans le détail, entre les run-1 et les run-2, 27 (TG1) et 29 (WT) protéines supplémentaires, puis 26 et 25 entre run-2- et run-3 et 19 et 12 entre le run-3 et le run-4 (voir figure 19, sous-ensembles).

Ainsi, après 4 répétitions techniques, le pourcentage de protéines supplémentaires identifiées est donc d'environ +20 % dans les J0TG1 et +17% pour J0WT correspondant à 72 et 66 protéines supplémentaires.

4.5. Influence de la méthode (Fractionnée/Non Fractionnée) sur l'étude protéomique de l'effet d'un traitement à la dexaméthasone de rats transgéniques

Outre l'identification des peptides et des protéines, la quantification des protéines identifiées a été analysée afin de mettre en évidence les protéines d'intérêts, c'est-à-dire les protéines dont l'abondance est statistiquement différentielle (Anova $p < 0.05$) dans le design de comparaison choisi. Dans cette étude, 5 designs ont été choisis (tableau 1), D0 comparant les rats transgéniques (TG1) aux rats sauvages (WT), D1 comparant les rats Mâles (M) traités (T (j7)) aux rat non traités (NT (j0)), D2 idem pour les Femelles (F), D3 même comparaison mais mâle et femelles confondus et enfin D4 comparant les rats TG1 T aux rats WT T.

Les résultats obtenus sont importants pour la comparaison des deux méthodes de préparation des échantillons de plasma mais aussi pour l'étude clinique sur la pathologie oculaire étudiée. Ils montrent un nombre important de protéines différentielles pour les deux techniques utilisées (tableau 1). Ce nombre varie de 7 (D0) à 108 (D3) pour la méthode Fractionnée et de 27 (D0) à 178 ((D2) et (D3)) pour la méthode Non Fractionnée. Selon les designs utilisés lors des comparaisons, il existe plus ou moins de protéines différentielles communes. Ce nombre de protéines communes est dans tous les cas minoritaire. Par exemple pour D1, étudiant l'effet du traitement à la Dexaméthasone chez les mâles (TG1 et WT), parmi les 198 protéines (pour les deux méthodes) ayant

PANTHER GO-Slim Cellular Component	Client Text Box Input (expected)	Client Text Box Input (over/under)	Client Text Box Input (fold Enrichment)	Client Text Box Input (raw P-value)	Client Text Box Input (FDR)
extracellular matrix (GO:0031012)	0,68	+	10,26	6,84E-06	1,16E-03
external encapsulating structure (GO:0030312)	0,68	+	10,26	6,84E-06	8,72E-04
extracellular space (GO:0005615)	4,18	+	6,21	2,88E-14	1,47E-11
extracellular region (GO:0005576)	4,34	+	5,99	6,74E-14	1,72E-11
nucleus (GO:0005634)	11,23	-	0,09	1,38E-04	1,41E-02

Tableau 2. Résultats de l'analyse Panther avec les protéines différentielles du design D3 fractionné

PANTHER GO-Slim Cellular Component	Client Text Box Input (expected)	Client Text Box Input (over/under)	Client Text Box Input (fold Enrichment)	Client Text Box Input (raw P-value)	Client Text Box Input (FDR)
high-density lipoprotein particle (GO:0034364)	0,04	+	48,56	1,19E-03	4,67E-02
collagen-containing extracellular matrix (GO:0062023)	0,5	+	12	1,58E-05	1,61E-03
extracellular matrix (GO:0031012)	1,08	+	7,39	1,77E-05	1,50E-03
external encapsulating structure (GO:0030312)	1,08	+	7,39	1,77E-05	1,29E-03
extracellular space (GO:0005615)	6,64	+	7,08	3,10E-27	1,58E-24
extracellular region (GO:0005576)	6,9	+	6,82	1,47E-26	3,75E-24
organelle (GO:0043226)	37,2	-	0,56	1,17E-03	4,98E-02
membrane (GO:0016020)	46,26	-	0,41	1,31E-07	2,22E-05
intracellular membrane-bounded organelle (GO:0043231)	29,82	-	0,4	8,14E-05	4,61E-03
membrane-bounded organelle (GO:0043227)	31,33	-	0,38	1,84E-05	1,17E-03
nucleus (GO:0005634)	17,83	-	0,17	1,48E-05	1,89E-03
intrinsic component of membrane (GO:0031224)	7,61	-	< 0.01	9,32E-04	4,32E-02
integral component of membrane (GO:0016021)	7,4	-	< 0.01	8,82E-04	4,50E-02

Tableau 3. Résultats de l'analyse Panther avec les protéines différentielles du design D3 non fractionné

PANTHER Go-Slim Cellular Component : composant cellulaire des résultats obtenus et attendus.

Client Text Box Input (expected) : résultats attendus selon la banque de donnée du rat du logiciel Panther.

Client Text Box Input (over/under) : sur ou sous expressions des résultats obtenus.

Client Text Box Input (fold Enrichissement) : résultats obtenus avec les données utilisées.

Client Text Box Input (raw P-value) : statistiques de la P-value obtenue.

Client Text Box Input (FDR) : statistiques sur le taux de faux positifs obtenus.

pour ces deux populations (TG1vsWT) des abondances différentielles ($p < 0.05$) seulement 34 protéines sont communes. Cela représente seulement 14% des protéines différentielles quantifiées.

Pour les autres designs, les résultats sont similaires, seule la comparaison D3, comparant la population (M et F) traitée à la même population non traitée montrent environ 30% de protéines différentielles communes avec les deux méthodes fractionnées et non fractionnées. D'autre part, on constate qu'il y a plus de protéines différentielles détectées au cours du Workflow d'analyse sans Fractionnement (487 vs 105). C'est vrai quel que soit le design (voir tableau 1). Néanmoins, dans la méthode sans fractionnement, les répétitions techniques (4 runs par pool) sont interprétées comme des répétitions biologiques dans le test statistique (pas de possibilités dans le logiciel de les interpréter différemment) et donc cela rend « artificiellement » plus robuste l'analyse et donc peut augmenter le nombre de protéines différentielles. De plus, le nombre de protéines quantifiées est très différent : 438 pour la méthode Fractionnée et seulement 263 protéines pour la méthode Non Fractionnée. Cela peut aussi jouer un rôle dans le résultat du test statistique au niveau du calcul des faux positifs (FDR).

Indépendamment du nombre de protéines différentielles, il est intéressant pour l'interprétation biologique que doit faire le clinicien de placer les protéines d'intérêt dans le contexte de la cellule et son métabolisme.

4.6. Bioanalyse des résultats protéomiques obtenus

Le système Panther a été conçu pour classer les protéines en famille et/ou processus biologiques. Il compare donc une liste de protéines test, ici la liste de protéines différentielles dans un design donné par exemple, à une liste de référence. Ainsi, il est possible de déterminer si une classe particulière de protéines est surreprésentée ou sous-représentée [18]. Pour l'étude, plusieurs tests ont été réalisés. Le tableau 2 montre les résultats pour le design D3 fractionné pour un test Gene Ontologie basé sur la localisation cellulaire des protéines. Les résultats montrent une surreprésentation de 4 classes GO, toutes regroupant des protéines localisées à l'extérieur de la cellule.

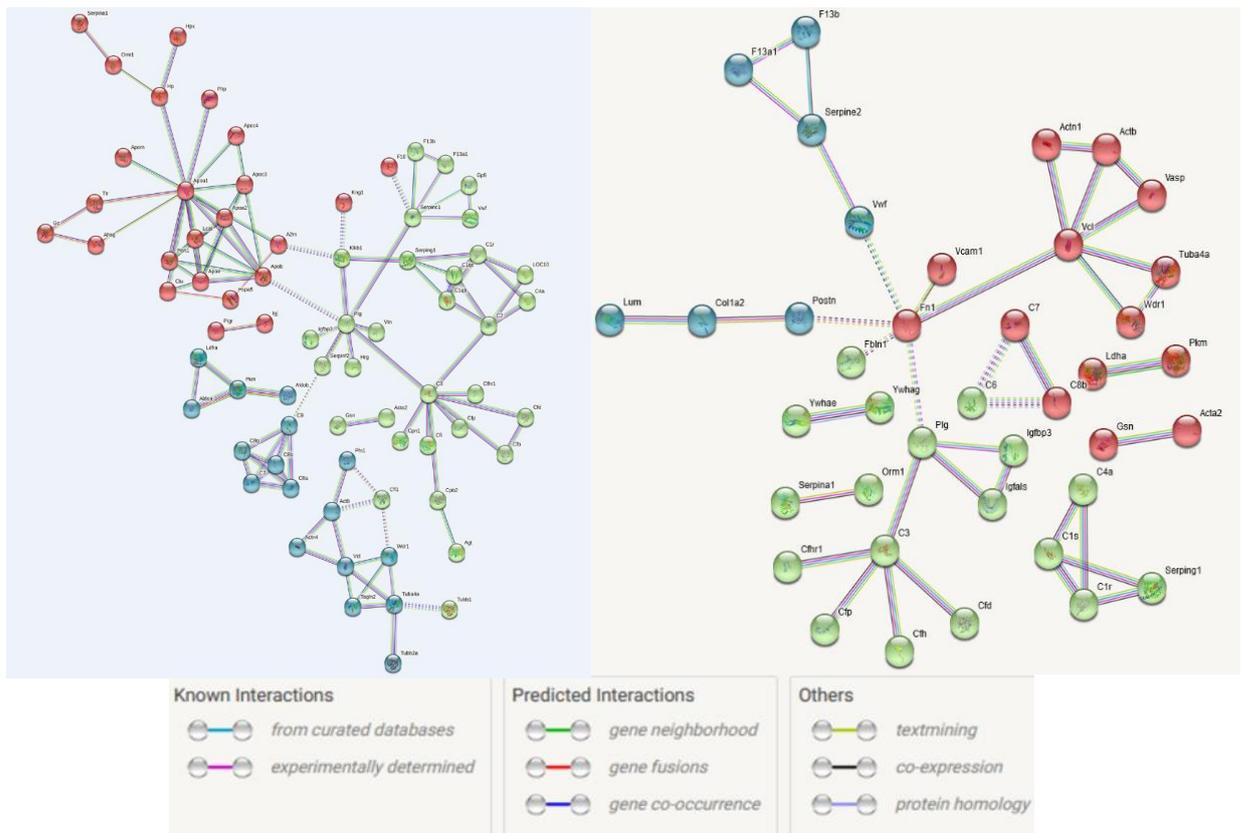


Figure 22. Réseau issu du logiciel String (Design D3-non fractionné et D3-fractionné)

https://string-b.org/cgi/input?sessionId=bu0vwYQ7GDaZ&input_page_show_search=on

Les taux d'enrichissement sont très significatifs avec des tests statistiques et des FDR très faibles. Pour le design D3 non fractionné (tableau 3), il est possible de voir la même chose. En effet, 6 classes GO sont surreprésentées et les protéines qui les composent sont des protéines extracellulaires. Aussi, il est notable que certaines classes sont communes entre les deux designs. En regardant les valeurs attendues et les valeurs de nos analyses, un enrichissement fort est aussi présent. De plus, les tests statistiques sont très significatifs et les FDR sont faibles.

Ainsi, malgré des méthodes de préparation utilisées différentes, il est possible de retrouver des protéines d'intérêts communes montrant que les résultats semblent bien refléter des phénomènes biologiques spécifiques au Design étudié. C'est aussi une façon de « valider » l'approche expérimentale, par la validation biologique des résultats. Pour rappel, dans le cadre du D3, la comparaison est faite sur les animaux T (J7) versus les animaux NT(J0). L'injection d'anti-inflammatoire semble donc augmenter la quantité de protéines extracellulaires dans le plasma des rats traités. Ce phénomène est en accord avec la bibliographie. Il est connu en effet que l'injection d'anti-inflammatoire augmente le drainage des organes et de la rétine en particulier [20] et pourrait donc être à l'origine de l'enrichissement des protéines extracellulaire dans le plasma des animaux traités (T) par rapport au non traités (NT). Ce drainage permettant la diminution de l'œdème dans le cas de problème oculaire.

Les résultats de quantification sont aussi analysés avec l'outil STRING qui permet de mettre en évidence les interactions existantes entre des protéines d'intérêts. Comme dans le cas de l'analyse PANTHER, ce sont les résultats de quantification du D3 TvsNT qui sont analysés. Les résultats montrent qu'il existe des interactions fortes (connues ou prédites entre les protéines différentielles découvertes (D3 Fractionné ou Non Fractionné) (Figure 22).

De plus, les protéines peuvent être regroupées en cluster avec des caractéristiques et fonctions communes [17] ajoutant de la cohérence aux réseaux établis. On peut visualiser 3 clusters, dont certains communs aux deux réseaux D3F et D3NF. De plus, l'analyse de la bibliographie montre que plusieurs articles citent des protéines des 2 réseaux impliquées dans des problèmes oculaires. C'est le cas de Lumican, Periostin, Von Willebrand Factor, Alpha-1-antiprotéinase, Plasma Protease C1 inhibito, Gelsolin* par exemple (*données non montrées).

CONCLUSION

Afin d'obtenir une analyse plus exhaustive des échantillons complexes comme le plasma, une nouvelle technique a été testée. Le plasma ne contient que très peu de protéines (3000 théoriques) mais les meilleurs résultats sur le plasma sans fractionnement identifient moins de 300 protéines. C'est peu par rapport aux possibilités d'identification. Cette difficulté est due à la présence de protéines majoritaires qui masquent par compétition ionique, la détection par spectrométrie de masse des peptides issus des protéines minoritaires.

Dans cette étude, un fractionnement peptidique en amont de l'utilisation de la spectrométrie de masse a été réalisé. Les résultats obtenus indiquent que ce fractionnement sur colonne induit un nombre plus conséquent d'identification de peptides et de protéines. Le fait de fractionner engendre une simplification des échantillons afin de réduire le nombre de peptides issus des protéines majoritaires rentrant à l'instant t dans le spectromètre de masse. La quantification des protéines et la mise en évidence des protéines différentielles ont montré des différences en nombre de protéines d'intérêts. Néanmoins, le traitement bioinformatique des résultats a permis d'observer de nombreuses similitudes entre les protéines d'intérêts. En effet, des réseaux communs entre les protéines ont été mis en évidence avec certaines protéines déjà citées dans la bibliographie relative aux problèmes oculaires. De plus, la surreprésentation des protéines extracellulaires dans le plasma des rats traités à la dexaméthasone (D3), vient confirmer l'effet de drainage des anti-inflammatoires. L'ensemble de ces données confirme la possibilité de travailler avec des Pools d'échantillons limitant les runs LC-MS/MS. L'apport du fractionnement de peptides est indéniable en termes d'identification mais il reste à confirmer en termes de quantification label-free, avec l'observation d'un nombre de protéines différentielles moins importants que sans fractionnement. Il n'est néanmoins pas exclu qu'un nombre plus restreint de protéines identifiées comme différentielles soit plus pertinent qu'une liste plus conséquente mais moins spécifique (FDR plus élevé). Les résultats qui seront obtenus dans les prochaines semaines sur les échantillons de plasma individuels (sans pools) viendront apporter des informations supplémentaires sur le workflow le plus pertinent dans ce type d'étude sur les protéines plasmatiques.

BIBLIOGRAPHIE

- [1] « Biochimie_Clinique_Chapitre_7.doc ». Consulté le : 12 mai 2022 [En ligne]. Disponible sur : <https://view.officeapps.live.co> .
- [2] « Chromatographie en phase liquide à haute performance », *Wikipédia*. 27 avril 2022. Consulté le : 13 mai 2022. [En ligne]. Disponible sur : https://fr.wikipedia.org/w/index.php?title=Chromatographie_en_phase_liquide_%C3%A0_haute_performance&oldid=193217094
- [3] Yu-Ching Liu · Yu-Ting Huang · Chao-Jung Chen, « Development of a high-pH reversed-phase well plate for peptide fractionation and deep proteome analysis of cells and exosomes ». Springer-Verlag GmbH Germany, part of Springer Nature 2022, 31 janvier 2022. Consulté le : 18 avril 2022. [En ligne]. Disponible sur : file:///Z:/proteo-data/DA/Fiches%20projets/2022/DUT_RAMI/Dossier-Biblio/Liu2022_Article_DevelopmentOfAHigh-pHReversed-.pdf
- [4] « Hydropathie : définition et explications », *AquaPortail*. Consulté le : 19 mai 2022. [En ligne]. Disponible sur : <https://www.aquaportail.com/definition-8863-hydropathie.html>
- [5] « Identifications de protéines par moteur MASCOT | Plateforme Protéomique (3P5) ». Consulté le : 10 mai 2022. [En ligne]. Disponible sur : <https://3p5.recherche.parisdescartes.fr/prestations/mascot/>
- [6] P. Bardou, J. Mariette, F. Escudié, C. Djemiel, et C. Klopp, « jvenn: an interactive Venn diagram viewer », *BMC Bioinformatics*, , déc. 2014, Consulté le : 9 mai 2022. [En ligne]. Disponible sur : <http://jvenn.toulouse.inra.fr/app/index.html>
- [7] F. Behar-Cohen *et al.*, « Les œdèmes maculaires », *médecine/sciences*, août 2020. Consulté le : 11 avril 2022. [En ligne]. Disponible sur : <https://hal-archives-ouvertes.fr/hal-02919330>
- [8] « Mascotte (logiciel) - fr.wikichamsoc.com », *Wikichamsoc*. Consulté le : 10 mai 2022. [En ligne]. Disponible sur : <https://fr.wikichamsoc.com/497789-mascot-software-DWAFSH>.
- [9] Thermo Scientific, « Pierce High pH Reversed-Phase Peptide Fractionation Kit ». 2016. Consulté le: 28 avril 2022. [Protocole]. < Pub. No. MAN0015701>
- [10] « Protéine », *Wikipédia*. Consulté le : 20 mai 2022. [En ligne]. Disponible sur : <https://fr.wikipedia.org/w/index.php?title=Prot%C3%A9ine&oldid=193217101>
- [11] M. Hebraud, « PROTEOME ET ANALYSES PROTEOMIQUES ». [Cours]

- [12] Adrabovich, *Q Exactive*, (5 mai 2015). Consulté le: 20 mai 2022. [En ligne Vidéo]. Disponible sur: <https://www.youtube.com/watch?v=K1VSYjuw6os>
- [13] « What is HPLC (High Performance Liquid Chromatography) », *Shimadzu*. Consulté le : 19 mai 2022. [En ligne]. Disponible sur : https://www.shimadzu.com/an/service-support/technical-support/analysis-basics/basic/what_is_hplc.html
- [14] Merck Millipore Bioscience, « The Direct Detect™ Biomolecular Quantitation System Exploits Membrane Technology to Enable Accurate, Infrared-based Detection. » Consulté le : 12 mai 2022. [En ligne]. Disponible sur : [1327531832_DirectDetectApplicationNote.MM%20.pdf](https://www.merckmillipore.com/~/media/1327531832_DirectDetectApplicationNote.MM%20.pdf)
- [15] C Ferreira, « Identification d'une protéine par spectrométrie de masse », *INRAE*. Consulté le : 16 mai 2022. [Mode Opérateur]. <MO-131>
- [16] C Chambon, « Quantification des protéines en label-free : LCProgenesis », *INRAE*, 26 juin 2019. Consulté le : 25 avril 2022. [Mode Opérateur]. <I-MAT-097>
- [17] « STRING: functional protein association networks ». Consulté le : 16 mai 2022. [En ligne]. Disponible sur : https://string-db.org/cgi/input?sessionId=bu0vwYQ7GDaZ&input_page_show_search=on
- [18] « Panther Protein ». Consulté le : 16 mai 2022. [En ligne]. Disponible sur : <http://pantherdb.org/>
- [19] « Peptide Analyzing Tool - FR ». Consulté le : 1 juin 2022. [En ligne]. Disponible sur : <https://www.thermofisher.com/fr/fr/home/life-science/protein-biology/peptides-proteins/custom-peptide-synthesis-services/peptide-analyzing-tool.html>
- [20] V. S. Dedania et S. J. Bakri, « Novel Pharmacotherapies in Diabetic Retinopathy », *Middle East Afr J Ophthalmol*. Consulté le : 31 juin. [En ligne]. Disponible sur : <https://pubmed.ncbi.nlm.nih.gov/25949073/>
- [21] « Mascot search engine | Protein identification software for mass spec data ». Consulté le : 12 mai 2022. [En ligne]. Disponible sur : <https://www.matrixscience.com/>

TABLE DES SIGLES

LC-MS/MS Liquid Chromatography – Mass Spectrometry

SDS Sodium Dodecyl Sulfate

SDS-PAGE Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

HPLC High Performance Liquid Chromatography

ABC Ammonium Bicarbonate

mM milli Mole

DTT DL-Dithiothreitol

IAA IodoAcetAmide

TFA TriFluoroacetic Acid

ACN Acetonitrile

GLOSSAIRE/LEXIQUE

Acrylamide/ Bis-acrylamide Composé organique utilisé comme monomère participant à la formation de polymères réticulés.

Albumine Principale protéine qui circule dans le sang.

Bain à ultrasons Procédure accélérée de dissolution de produits par l'effet mécanique d'ondes ultrasonores.

Centrifugeuse Machine qui utilise la force centrifuge pour séparer les différents composants d'un fluide.

Colonne Tube dans lequel se déroule des échanges de matière permettant une séparation des composés utilisés.

Déplétion Diminution d'un liquide organique normalement présent dans une cavité en quantité.

Elution Procédé permettant de mettre en solution (dite éluee) un composé adsorbé à l'aide d'un solvant nommé l'éluant.

Gel resolving Gel très concentré en bis-acrylamide, assez pour permettre la séparation des protéines selon leur poids moléculaire.

Gel stacking Gel très faiblement concentré en bis-acrylamide, pas suffisamment pour séparer les protéines. Elles migrent toutes en même temps, tout poids moléculaires confondus.

Glucocorticoïdes Molécules exerçant une action sur le métabolisme protidique et glucidique.

Hydrolyse Réaction chimique et enzymatique dans laquelle une liaison covalente est rompue.

Hydrophobie Caractère de ce qui est hydrophobe, qui repousse l'eau.

Label-free Quantification sans marquages.

Oxydation Réaction chimique au cours de laquelle se produit un transfert d'électrons.

Peptides Polymère d'acides aminés reliés entre eux par des liaisons peptidiques.

Plasma Partie du sang représentant composée d'eau, de lipides, d'hormones, de facteurs de coagulation et de protéines.

Polymérisation Réaction chimique ou le procédé par lesquels des petites molécules réagissent entre elles pour former des molécules de masses molaires plus élevées. Ici, elle représente la prise en masse du gel.

Pools Ensemble des éléments cellulaires ou des molécules biologiques de même structure.

Solvants substance qui a la propriété de dissoudre, de diluer ou d'extraire d'autres substances sans les modifier chimiquement et sans lui-même se modifier.

SpeedVac Evaporateur utilisé pour concentrer des solutions par apport d'une force centrifuge sous vide avec un système de chauffage

Surfactant polyanionique Composé qui modifie la tension superficielle entre deux surfaces.

Trypsine Peptidase participant à la digestion des protéines. Il s'agit d'une protéase à sérine qui hydrolyse les liaisons peptidiques situées côté C d'un résidu de lysine (K) ou d'arginine (R), qui sont des acides aminés basiques.

Vortex Matériel utilisé en biologie moléculaire pour mélanger des solutions, notamment dans des microtubes.

Workflow Représentation d'un ensemble de tâches nécessaires à la réalisation d'un processus opérationnel

TABLE DES FIGURES

Figure 1. Carte représentant l'implantation des centres INRAE en France

Figure 2. Organigramme de l'unité QuaPA

Figure 3. Plaque de Direct Detect™ avec les échantillons des pools après passage sur colonne

Figure 4. Image d'un œil atteint de rétinopathie diabétique

Figure 5. Photo d'un gel 1D-SDS-PAGE après migration et coloration

Figure 6. Montage d'un gel 1D-SDS-PAGE

Figure 7. Schéma simplifié de la vue d'ensemble du fonctionnement d'une HPLC

Figure 8. Schéma simplifié expliquant la séparation par HPLC

Figure 9. Spectre de masse

Figure 10. Schéma expliquant la formation de nébulisats dans une source ESI

Figure 11. Schéma d'un quadripôle

Figure 12. Schéma d'une OrbiTrap

Figure 13. Protocoles de préparation des gels pour une électrophorèse 1D-SDS-PAGE

Figure 14. Composition du tampon dénaturant : Leampli

Figure 15. Workflow pour le fractionnement des échantillons

Figure 16. Spectromètre de masse

Figure 17. Répartition du Gravy Score en fonction des fractions peptidiques

Figure 18. Nombre d'identification des groupes de protéines et des peptides uniques

Figure 19. Diagramme de Venn montrant la distribution de l'identification des protéines et des peptides

Figure 20. Taux de chevauchement des peptides identiques entre deux fractions adjacentes

Figure 21. Diagrammes de Venn montrant le nombre de protéines dans la méthode sans fractionnement (4 répétitions techniques de LC-MS/MS)

Figure 22. Réseau issu du logiciel String (Design D3-non fractionné et D3-fractionné)

Tableau 1. Nombre de protéines quantifiées différentielles selon chaque design

Tableau 2. Résultats de l'analyse Panther avec les protéines différentielles du design D3 fractionné

Tableau 3. Résultats de l'analyse Panther avec les protéines différentielles du design D3 non fractionné