

L'analyse des protéines

Le PROTEOME désigne l'ensemble des PROtéines exprimées par le génOME d'une cellule. Ce terme a été proposé en 1995 par Wilkins et fait écho à celui de génome³. C'est un produit du génome mais contrairement à ce dernier qui est unique et statique (en dehors de quelques rares mutations), le protéome évolue selon l'état physiologique de l'organisme et n'est pas identique d'un tissu à un autre d'un même organisme. L'étude du protéome est appelée protéomique. Pendant longtemps, l'analyse des protéines s'est limitée à la détermination de la quantité totale contenue dans un échantillon grâce généralement grâce à des techniques de dosage colorimétriques ou spectrophotométriques⁷. L'identification et la quantification des protéines sont devenues possibles avec l'introduction de l'électrophorèse bidimensionnelle sur gel (gel 2D) et les méthodes d'analyse immunologique ou colorimétrique. La séquence primaire des polypeptides d'intérêt pouvait être déterminée par des méthodes de séquençage comme le séquençage d'Edman⁴. La structure secondaire des protéines purifiées peut être étudiée par spectroscopie des rayons X, Raman ou RMN. L'introduction de la spectrométrie de masse pour l'analyse des protéines s'est faite plus tard avec l'apparition de techniques d'ionisation adaptées aux biomolécules, comme l'ionisation par électrospray (ESI) et l'ionisation laser assistée par une matrice (MALDI) sont celles qui ont connu le plus de succès. La MS est très rapidement devenue une technique de choix pour l'analyse des protéines car elle permet d'accéder à la séquence en acides aminés pour l'identification, la quantification et la caractérisation des modifications post-traductionnelles des protéines qu'elles soient purifiées, enrichies ou dans un échantillon protéique complexe. Par définition, la protéomique est donc l'étude qualitative et quantitative du protéome d'une cellule à un moment donné et dans des conditions données⁵.

1.1.1 Les méthodes spectrophotométriques et colorimétriques

Ces méthodes permettent de déterminer la concentration totale en protéines dans un échantillon. Les méthodes spectrophotométriques sont basées sur la spectroscopie UV et visible. En UV, les protéines peuvent être dosées à 280 et à 205 nm. La méthode à 280 nm est basée sur l'absorption UV des acides aminés aromatiques, la tyrosine et le tryptophane. Elle offre une gamme de mesure de 20 à 3000 µg/mL. Cette méthode a été notamment adaptée dans un appareil appelé Nanodrop1000 (ThermoFisher). Le nanodrop1000, avec sa fonction A280, permet de mesurer les concentrations de protéine purifiée. Par exemple, avec un trajet optique de 0.2 mm, il permet de mesurer la concentration de protéine purifiée sans courbe d'étalonnage avec une gamme de mesure allant de 0.1 à 100 mg/mL. Cet appareil a été utilisé durant cette étude pour déterminer les concentrations de BSA (Bovine Serum Albumin). La concentration en BSA est calculée en utilisant un coefficient d'absorption de 6,7 (BSA, solution de 1% : 10mg/mL, longueur d'onde = 280 nm).

La détermination de la concentration totale en protéine à 205 nm est quant à elle basée sur l'absorption de la lumière à cette longueur d'onde des liaisons peptidiques. Sa limite de quantification (LOQ) s'établit à 1 µg/mL. Ces 2 méthodes présentent l'avantage d'être rapides et simples à mettre en œuvre. La méthode à 280 nm est moins sensible mais elle est toutefois la plus utilisée. Ceci pourrait s'expliquer par le fait qu'elle est aussi moins affectée par les interférents⁶. La détermination des protéines peut aussi se faire en mesurant la fluorescence naturelle des acides aminés aromatiques (tyrosine, tryptophane, phénylalanine). Cette technique permet de déterminer des gammes concentration entre 5 et 50 µg/mL.

Les méthodes colorimétriques sont basées sur l'utilisation d'un composé chimique qui réagit avec les protéines afin de former des complexes colorés. A la fin de la réaction, la mesure de l'intensité de la coloration de la solution permettra de calculer la concentration en protéine. Il existe plusieurs méthodes basées sur ce principe, les plus connues sont les méthodes de Biuret, Lowry, Bradford et BCA (*BiCinchoninic Acid Assay*) (Tableau 1). Ces méthodes ont fait l'objet de publications qui comptent parmi les publications les plus citées. Les méthodes les plus utilisées pour l'analyse d'échantillons protéiques complexes sont les méthodes de Lowry, Bradford et BCA¹¹⁷. Ces méthodes seront décrites plus en détails dans les paragraphes suivants.

Tableau 1: Méthodes colorimétriques pour le dosage des protéines

Méthodes	Réactifs	Longueur d'onde	Gamme d'analyse
Biuret	Cu ²⁺ , pH basique	540	5-160 mg/mL
Lowry	Cu ²⁺ , Folin Ciocalteu, pH basique	750	1-1500 mg/mL
Bradford	Bleu de Coomassie g250, pH basique	595	100-2000 µg/mL
BCA	Cu ²⁺ /BCA/Milieu alcalin	562	20-2000 µg/mL

La méthode de Lowry est basée sur deux réactions différentes. La première est la formation d'un complexe d'ions de cuivre avec des liaisons amides donnant du cuivre réduit dans une solution alcaline. La deuxième est la réduction du réactif de Folin-Ciocalteu (phosphomolybdate et phosphotungstate) par les complexes cuivre-amide mais aussi par les résidus tyrosine et tryptophane. Le réactif de Folin-Ciocalteu une fois réduit devient détectable par mesure spectrophotométrique entre 500 et 750 nm⁷.

La méthode de Bradford est basée sur la liaison entre le bleu de Coomassie (colorant G-250 CBBG) aux résidus arginine, tryptophane, tyrosine, histidine et phénylalanine des protéines. C'est une méthode populaire du fait de sa simplicité, rapidité, de la stabilité du colorant et de son bas coût⁸.

La méthode BCA est similaire avec la méthode de Lowry. Elles sont toutes les 2 basées sur la formation d'un complexe de cuivre. Le réactif BCA (*BiCinchoninic Acid Assay*) remplace le réactif de Folin-Ciocalteu dans la méthode de Lowry. La méthode BCA présente par ailleurs plus de tolérance aux interférents comme les détergents non ioniques¹¹⁸. Le réactif BCA forme un complexe chromophore avec les ions Cu⁺ présentant une force absorbance à 562 nm. La méthode est très utilisée de la stabilité du complexe formé et de sa tolérance aux interférents.

1.1.2 Les méthodes immunologiques

Les tests immunologiques ont connu un grand essor après le développement indépendant et presque simultané de deux méthodes : l'EIA (*Enzyme ImmunoAssay*) en Hollande par Anton Schuurs et Bauke Weemen⁹ et l'ELISA (*Enzyme Linked ImmunoSorbent Assay*) par Peter Perlmann et Eva Engvall à l'université de Stockholm¹⁰. Ces techniques, basées sur le principe de reconnaissance antigène-anticorps, sont très adaptées à l'analyse qualitative et quantitative ciblée de protéine (antigène ou anticorps) dans des échantillons complexes (plasma, urine...). L'anticorps est le composant le plus important dans une méthode immunologique car il apporte la grande spécificité et la sensibilité, envers l'analyte cible, qui caractérisent ces méthodes. La génération et la purification des anticorps constituent par ailleurs une grande étape préalable à l'analyse d'une protéine par des méthodes immunologiques. L'ELISA est la méthode immunologique la plus utilisée. Elle est appliquée à des échantillons biologiques, industriels ou environnementaux. L'utilisation répandue de la méthode ELISA s'explique par la grande sensibilité et spécificité de cette méthode mais aussi par un haut débit d'analyse¹²⁵. Il plusieurs types de méthode ELISA dont le plus simple à mettre en œuvre est appelé ELISA direct (Figure 1)

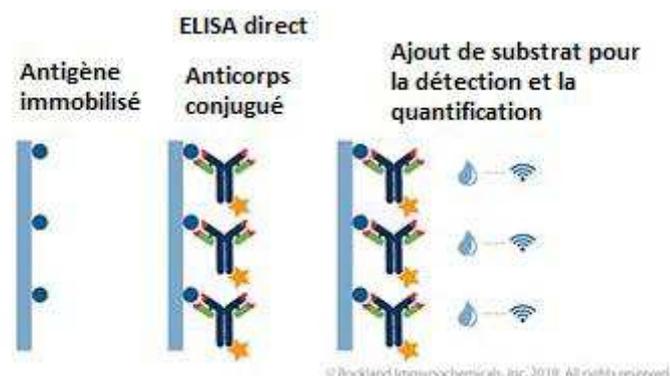


Figure 1 : ELISA direct (Schéma adapté de biocompare.com)

Cependant, pour qu'une protéine puisse être analysée par une méthode immunologique, il faut obligatoirement avoir à disposition un antigène ou anticorps réciproque très spécifique, ce qui limite le nombre de protéines que l'on peut analyser par cette méthode. Par ailleurs, les méthodes immunologiques souffrent d'interférences, de réactivité croisée et de difficultés à différencier les protéoformes¹⁰.

1.1.3 Electrophorèse

L'électrophorèse sur gel est une technique séparative basée sur un support appelé matrice de séparation généralement composée de polyacrylamide, ou acrylamide réticulé. La séparation des protéines peut se faire selon un seul paramètre (généralement la masse moléculaire, électrophorèse monodimensionnelle). Elle peut aussi être basée sur deux paramètres (masse apparente et point isoélectrique, en électrophorèse bidimensionnelle, Figure 2). Publiée en 1975, l'électrophorèse bidimensionnelle sur gel de polyacrylamide permet de séparer les protéines selon leur point isoélectrique par focalisation isoélectrique et le poids moléculaire par électrophorèse en présence de sodium dodecylsulfate (SDS). Une résolution de plus de 1000 protéines est mentionnée et une capacité de séparation pouvant atteindre 5000 protéines¹², ce qui en fait une technique très adaptée aux échantillons complexes.

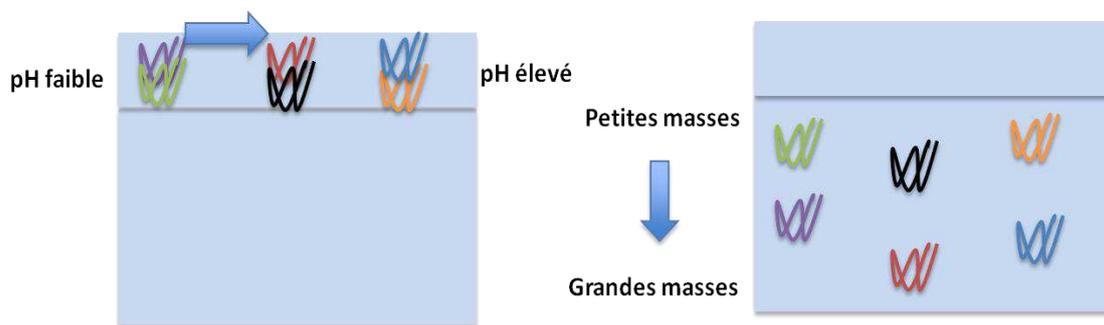


Figure 2 : Gel d'électrophorèse 2D, séparation des protéines selon leur point isoélectrique et leur masse

Après la séparation, les protéines sont fixées et puis visualisées sur les gels grâce à des techniques de coloration dont les plus utilisées sont la coloration au bleu de Coomassie, la coloration négative zinc-Imidazole, la coloration à l'argent¹³ (Tableau 2). Si la technique de coloration utilisée est compatible avec la MS, les spots sont découpés et puis analysés par MS.

Tableau 2: Les principales méthodes de détection sur gel

Technique	Avantages	Inconvénients
Coloration au Bleu de Coomassie	Bas coût, compatibilité avec la MS	Sensibilité insuffisante
Coloration négative (zinc-Imidazole)	Rapidité, simplicité, compatible avec la MS	Gamme dynamique restreinte
Coloration à l'argent	Sensibilité	Reproductibilité, complexe et laborieux

Les points faibles de cette méthode sont caractérisés par l'exigence d'une grande expertise des manipulateurs, pour éviter les risques d'obtenir une faible reproductibilité, une détection difficile des protéines présentes en faible quantité ainsi que les protéines ayant un point isoélectrique trop faible ou très élevé (inférieur à 4 ou supérieur à 12). La

2D PAGE est aussi caractérisée par une faible gamme dynamique. La coloration à l'argent, par exemple, présente une sensibilité de l'ordre du nanogramme et une faible gamme dynamique ($\sim 10^1$)¹²⁶.

1.1.4 La spectrométrie de masse

Un spectromètre de masse est un appareil permettant de déterminer le rapport masse sur charge d'une molécule ionisée à l'état gazeux. Il existe différents types de spectromètres de masse mais tous sont constitués de 3 éléments principaux qui sont la source d'ionisation, le ou les analyseur(s) de masse et le détecteur. Les molécules ionisées en phase gazeuse dans la source d'ionisation sont transférées dans l'analyseur qui les sépare en fonction de leur rapport masse sur charge avant de les envoyer vers le détecteur. Depuis l'introduction des techniques d'ionisation douces à savoir l'ESI (ionisation par *electrospray*) et le MALDI (Ionisation laser assistée par une matrice) capables d'ioniser des biomolécules fragiles en les gardant intactes, la spectrométrie de masse (MS pour *Mass Spectrometry*) est devenue une technique incontournable dans le domaine de la protéomique. En effet, la MS présente entre autres avantages une grande sensibilité et rapidité d'analyse, sa facilité d'automatisation ainsi que sa compatibilité avec des techniques de séparation comme la chromatographie liquide. La MS permet l'identification et la quantification des protéines et la caractérisation des modifications post-traductionnelles (PTM pour *Post-Translational Modification*). Les protéines sont analysées en MS selon 2 approches principales : *top-down* et *bottom-up*. Pour leur caractérisation en MS par l'approche *bottom-up*, les protéines sont identifiées à partir des peptides issus de leur digestion alors qu'en mode *top-down*, c'est la protéine entière qui est ionisée et introduite dans le spectromètre de masse. Ces différentes approches seront d'avantages détaillées dans le Chapitre 2: La protéomique par spectrométrie de masse.

L'étude qualitative et quantitative des PTMs dans un protéome est possible par MS. Par l'approche *bottom-up*, cette étude passe souvent par un marquage différentiel des modifications cibles avec des isotopes stables¹⁴. Cependant, il est nécessaire de passer par une étape d'enrichissement si les PTMs cibles sont à une faible concentration.

1.2 Contexte scientifique

Ces dernières années, les techniques analytiques mises en œuvre dans l'analyse des protéines ont beaucoup évolué. De techniques analytiques ne déterminant que la quantité totale des protéines dans un échantillon, on a aujourd'hui la capacité d'identifier jusqu'à 4000 protéines dans un échantillon complexe en une seule analyse LC-MS/MS voire d'en quantifier une partie¹⁵. Ces avancées offrent de nouveaux outils à la recherche biomédicale notamment. Ces outils pourraient en effet améliorer le diagnostic médical, le suivi d'un traitement et le développement de nouveaux traitements. Cependant, les échantillons biologiques sont complexes et leur analyse aussi. Les protéines y sont présentes avec une grande diversité à des concentrations variant jusqu'à 10 ordres de grandeur¹⁶. Dans le plasma, par exemple, le cytokine interleukine-6 est une protéine impliquée dans la signalisation. Elle est mesurée à des niveaux de concentrations de

l'ordre du pg/mL. Tandis que l'albumine de sérum humain, protéine impliquée dans le transport, est mesurée à des concentrations de l'ordre du dizaine de mg/mL¹²⁷. Il faut aussi noter que les protéines spécifiques à un état physiologique sont généralement disponibles à de faibles concentrations. Par ailleurs, la quantification des protéines est aujourd'hui nécessaire à une meilleure compréhension des processus biologiques. La quantification de ces protéines nécessite une mesure avec un rapport signal-sur-bruit plus élevé que ce qui est utile pour une simple identification. L'analyse de ces protéines, d'une grande importance dans la recherche clinique par exemple¹⁷, nécessite des techniques analytiques très performantes.

La technique la plus utilisée pour l'analyse des protéines sans *a priori* est aujourd'hui la spectrométrie de masse. Elle permet en effet l'identification, la quantification et la caractérisation des PTMs d'un nombre important de protéines en une seule analyse. Cependant, ces précieuses informations obtenues par MS exigent parfois une longue et fastidieuse préparation des échantillons, la nécessité d'utiliser une grande quantité d'échantillon pour des rendements modestes ainsi que l'expertise d'un opérateur pour la réalisation de toutes ces étapes préalables de façon robuste. Dans ce cadre, la miniaturisation présente un certain nombre d'avantages par rapport aux systèmes classiques. Elle permet de réduire la consommation en échantillon et la durée des protocoles d'analyse. Par ailleurs, la miniaturisation peut permettre l'intégration et l'automatisation des étapes de préparation d'échantillon, ce qui favoriserait les approches de protéomique en simplifiant les étapes de préparation et améliorerait encore la démocratisation de la protéomique. La miniaturisation semble donc très utile pour les protocoles d'analyse des protéines par MS et plus encore pour la caractérisation des PTMs des protéines. En effet, l'étude des PTMs par *bottom-up* passe généralement par des étapes de marquage et d'enrichissement qui viennent s'ajouter aux protocoles classiques. C'est le cas des études menées au laboratoire en rédoxomique pour l'étude de l'oxydation des protéines. Pour pouvoir travailler avec moins d'échantillon, exécuter le protocole en un temps plus court et le rendre plus robuste en évitant les oxydations artéfactuelles, il est très prometteur d'avoir recours à la miniaturisation.

La miniaturisation a en effet permis des révolutions sans précédent dans le domaine de l'électronique. Les procédés de fabrication qui ont mené à la miniaturisation des dispositifs électroniques ont ensuite été transposés à la biologie, à la chimie¹⁸. Puis, des techniques de microfabrication spécifiques ont été élaborées pour la miniaturisation dans le domaine des sciences expérimentales. Ce qui a permis le développement de concepts de laboratoire sur puce (*lab-on-chip*)¹⁷, de micro-systèmes d'analyse totale (μ TAS)¹⁹, d'organe sur puce (*Organ-on-a-chip*)²⁰. L'intérêt des sciences expérimentales à la miniaturisation a donné naissance à la microfluidique. Définie comme étant la science qui traite des écoulements de liquide dans des canaux de taille micrométrique, la microfluidique est aussi considérée comme une technologie car elle englobe les techniques de microfabrication héritées de l'industrie électronique mais adaptées aux sciences expérimentales²¹. La miniaturisation a eu un impact considérable sur le séquençage d'ADN²². La protéomique n'échappe pas à cet intérêt vus les avantages que

pourraient avoir la miniaturisation sur l'analyse des protéines. Cependant, la protéomique est toujours à la recherche d'un dispositif qui va révolutionner le séquençage des peptides ou protéines en simplifiant les protocoles de préparation d'échantillon, de séparation et de détection. Il faut noter que le séquençage des protéines est plus complexe que celle de l'ADN ou de l'ARN. Il est en effet difficile d'imaginer, vu le grand nombre d'acide aminé (20) pouvant entrer dans la séquence d'une protéine, un mode de détection intégrée à une micropuce capable de séquencer un millier de protéines sans *a priori*. Cependant, il est tout à fait possible de miniaturiser toute la partie de préparation d'échantillon.

1.3 Présentation du projet de thèse

L'objectif de cette thèse est le développement d'un dispositif microfluidique proposant une miniaturisation de toutes les étapes de la préparation d'échantillon lors de l'analyse des protéines par l'approche *bottom-up*. Dans une première étape le dispositif microfluidique devra permettre la miniaturisation d'un protocole *bottom-up* appliqué à un échantillon complexe pour la caractérisation des modifications post-traductionnelles. Puis le projet vise à la miniaturisation du protocole OcSILAC (Shakir SHAKIR, 2015, Manuscrit de thèse) développé au laboratoire pour la quantification du niveau d'oxydation des cystéines dans le protéome. L'étude de l'oxydation des cystéines est complexe du fait de la grande réactivité des groupements thiols, du caractère labile de certaines modifications de la cystéine (-S-OH, -S-S-, -S-NO) et de la faible concentration des cystéines oxydées. Pour prendre en compte ces difficultés, le protocole OcSILAC inclut plusieurs précipitations acides des protéines dont le but est de contrôler la réactivité des thiols et d'éviter les réactions croisées. Il comprend aussi des étapes de marquage et d'enrichissement des cystéines oxydées de sorte à mieux les caractériser malgré leur faible concentration dans l'échantillon. De plus l'étape de quantification est une étape de quantification relative, c'est-à-dire que celle-ci englobe deux problématiques : tout d'abord comparer l'état d'oxydation de deux échantillons entre eux, puis aussi comparer les abondances des divers états d'oxydation au sein d'un même échantillon. Cette deuxième problématique ne peut se faire par spectrométrie de masse qu'en comparant des espèces moléculaires comparables. C'est pourquoi, la spectrométrie de masse n'étant pas une méthode quantitative *per se*, il faut donc concevoir une stratégie pour pallier à cette limitation. Pour s'affranchir du biais principal de variation de l'efficacité d'ionisation et de détection de deux espèces moléculaires différentes, la stratégie adoptée ici est de résoudre l'ensemble du système à des espèces moléculairement semblables mais avec une composition isotopique différente et maîtrisée. C'est le principe du marquage différentiel avec des isotopes stables qui sera détaillé dans le chapitre suivant. La microfluidique offre aujourd'hui la possibilité d'effectuer des réactions chimique et biochimique en une durée réduite et avec de petites quantités d'échantillon et de réactifs dans des structures de petite taille d'environ 100 μm . Cependant, certaines étapes comme la précipitation acide des protéines ne sont pas compatibles avec la microfluidique. Il ne s'agira pas uniquement d'un changement d'échelle mais aussi d'un changement de stratégie du traitement de l'échantillon. En effet, une précipitation entrainerait un colmatage des canaux

microfluidiques du fait de leur petite taille. La miniaturisation du protocole OcSILAC se fera donc en deux étapes. La première étape consistera à trouver une manière d'éliminer les réactifs en excès et effectuer les changements de tampon sans passer par la précipitation des protéines. Et puis, en deuxième étape nous allons imaginer et fabriquer un dispositif microfluidique permettant de miniaturiser et d'intégrer l'ensemble des étapes du protocole OcSILAC.

Notre approche consiste à utiliser un support solide dans un canal microfluidique ou une chambre de réaction pour la rétention spécifique des protéines et pour le lavage des molécules indésirables. Toutes les réactions de blocage des cystéines, de réduction et de marquage par biotinylation des espèces oxydées seront effectuées directement sur les protéines retenues. Elles seront suivies par la digestion, effectuée aussi directement dans la puce. Les peptides générés seront élués avant de subir un enrichissement de ceux contenant une cystéine biotinylée et enfin une analyse par chromatographie liquide couplée à la spectrométrie de masse. L'utilisation de la rétention des protéines comme alternative à la précipitation acide sera testée dans un premier temps afin de développer un protocole qui peut se prêter à la miniaturisation. La photolithographie douce et la réplique par moulage sont les approches utilisées pour la fabrication de la puce microfluidique.