



Université Batna 2

Faculté des sciences de la nature et de la vie

Département de biologie des organismes

**Cours de
Génomique et protéomique
fonctionnelle
M2 BMC**

2022-2023

Table des matières

Introduction :	1
1. Les approches de la protéomique :	3
a. La protéomique descriptive :	3
b. La protéomique différentielle :	3
c. La protéomique fonctionnelle :	3
d. La protéomique structurale :	3
2. Généralités sur les protéines :	3
a. Structure des protéines :	3
b. Caractéristiques physicochimiques des protéines :	4
3. Principales étapes de la protéomique :	4
a. Extractions des protéines :	5
b. Electrophorèse bidimensionnelle :	6
c. Spectrométrie de masse :	9

Introduction :

Les protéines sont les acteurs principaux de la vie cellulaire, elles ont en charge la structure des tissus, le transport de molécules, la défense de l'organisme, la catalyse de réactions chimiques... Connaître la quantité de chaque protéine présente à un instant donné dans une cellule, et étudier son activité, permet donc de suivre au plus près le fonctionnement cellulaire. C'est ce à quoi s'emploie l'analyse du **protéome** qui, tout comme le transcriptome, évolue au cours du développement cellulaire et dépend de la cellule qui le produit.

Le mot "protéome" est apparu dans la littérature scientifique en 1995. Le protéome est défini comme étant l'ensemble des protéines présentes dans un échantillon biologique donné (liquide physiologique, tissus, organes, etc.), dans une situation physiologique ou pathologique donnée. La **protéomique** a pour but d'identifier et quantifier l'ensemble des protéines synthétisées ou protéome, à un moment donné et dans des conditions données au sein d'un tissu, d'une cellule ou d'un compartiment cellulaire ainsi que les interactions protéine-protéine au sein de cet organisme.

La **diversité** du protéome résulte premièrement de celle du génome à laquelle s'ajoute celle liée aux modifications des ARN (ARNm) messagers et enfin, surtout, aux modifications post-traductionnelles des protéines (clivage, glycosylation, phosphorylation, etc.). Alors que l'homme possède environ 30 000 gènes différents, le protéome est, quant à lui, constitué de plusieurs millions de protéines différentes. Notons que la protéomique présente certains avantages par rapport aux approches génomiques. En particulier, l'information sur les taux d'ARNm n'est pas suffisante pour étudier et analyser la régulation de l'expression d'un gène dans la cellule. En effet, il a été montré qu'il existe parfois une dissociation entre les taux d'ARNm et la présence et/ou l'activité d'une protéine. Compte tenu du "petit" nombre de gènes humains, ce sont plus vraisemblablement les quantités de protéines présentes dans une cellule donnée, ainsi que les modifications post-traductionnelles qui les affectent qui vont être déterminantes. Les données protéomiques sont ainsi souvent plus informatives, car elles permettent d'évaluer le niveau, voire le degré d'activité des protéines qui se trouvent fréquemment modulées et régulées par des modifications post traductionnelles.

Le protéome est extrêmement complexe à plusieurs titres (Fig1) :

- Compte-tenu de l'épissage alternatif des transcrits primaires (plusieurs ARNm pour un gène) et compte-tenu des modifications post-traductionnelles des protéines, on peut estimer à plusieurs dizaines de milliers les formes des protéines synthétisées dans les différents tissus humains par exemple.

- Pour chaque condition environnementale (condition physiologique normale vs. conditions de stress) une cellule est caractérisée par un protéome adapté à cette condition alors qu'elle a toujours le même génome. Le cas des plantes est un exemple flagrant compte-tenu de leur nécessité de s'adapter tant aux variations de la lumière qu'aux effets de stress biotiques ou abiotiques.
- Outre les modifications post-traductionnelles, les protéines subissent des transformations une fois synthétisées : clivage du peptide signal (séquence d'adressage), activation de la forme native à partir d'un précurseur (zymogène), assemblage en complexes oligomériques, association à des cofacteurs.

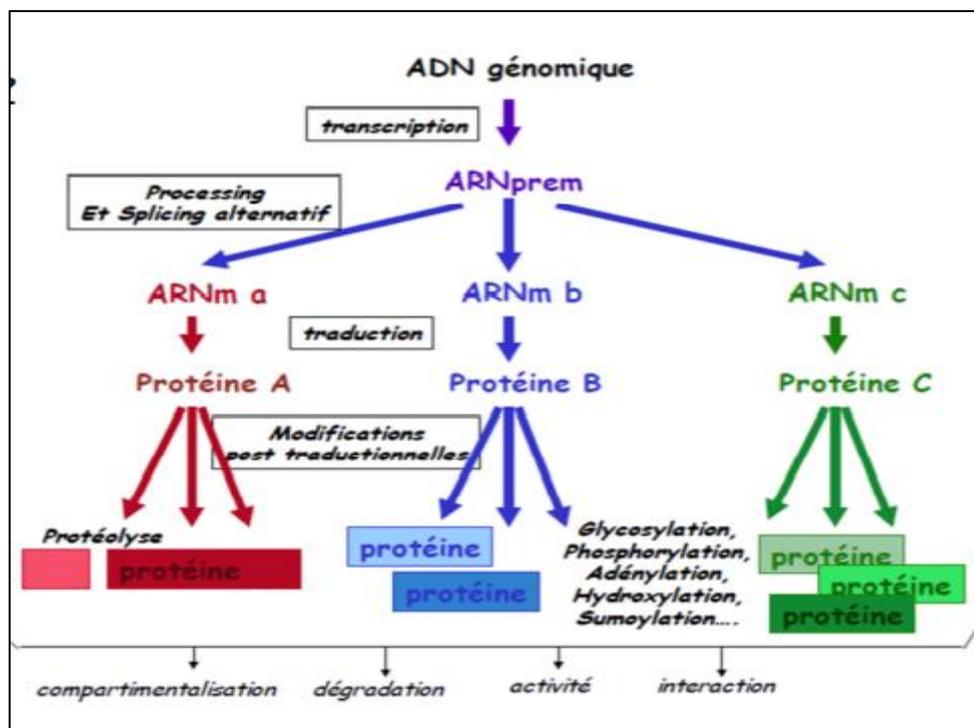


Figure1 : La complexité du protéome

L'exemple de la chenille et du papillon illustre la différence fondamentale qui existe entre les niveaux génomiques et protéomiques (Fig2). Ces deux organismes, apparemment différents, ont exactement le **même génome** mais des protéomes **distincts**. Il est donc important, pour comprendre un organisme, d'étudier son génome mais aussi son **protéome**.

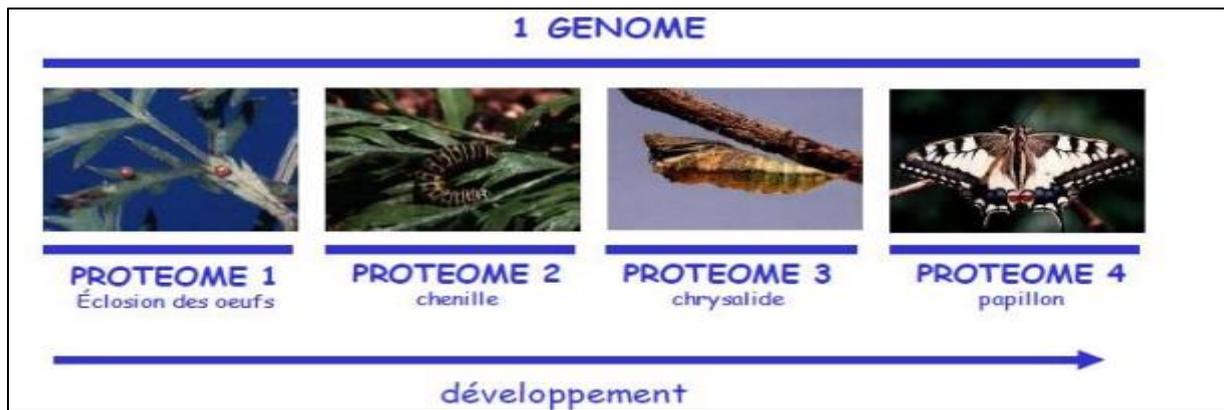


Figure 2 : La différence entre les niveaux génomiques et protéomiques (exemple de la chenille et du papillon).

1. Les approches de la protéomique :

a. La protéomique descriptive :

Le but est de fournir une image de l'ensemble des protéines exprimées par une cellule, un tissu ou un organisme donné (protéomique qualitative, quantitative).

b. La protéomique différentielle :

Il s'agit d'étudier l'expression des protéines de tissus ou d'organismes entre deux conditions différentes et de déterminer les variations d'expression protéique de manière qualitative et quantitative.

c. La protéomique fonctionnelle :

Le but est de pouvoir élucider les interactions entre protéines, pour établir des réseaux d'interaction, de façon à mieux appréhender leurs fonctions et leurs régulations.

d. La protéomique structurale :

Le but est de déterminer la relation entre la structure et la fonction d'une protéine, en particulier par la résolution de sa structure tridimensionnelle.

2. Généralités sur les protéines :

a. Structure des protéines :

Chaque protéine est élaborée à partir de 20 acides aminés. Une protéine est une combinaison, sous la forme d'une chaîne plus au moins longue et orientée, de ces 20 acides aminés (100 à 200 acides aminés) sous différents niveaux d'organisation structurale (Fig3).

- L'enchaînement des acides aminés d'une protéine constitue **la structure primaire** de la protéine.
- Des régions de la protéine peuvent adopter 2 formes particulières : hélices ou feuillets, constitués **la structure secondaire**.

- La structure tridimensionnelle finale qu'adopte la chaîne d'acides aminés, constitue la **structure tertiaire** de la protéine.
- Plusieurs protéines peuvent s'associer pour former des ensembles complexes ; la **structure quaternaire**.

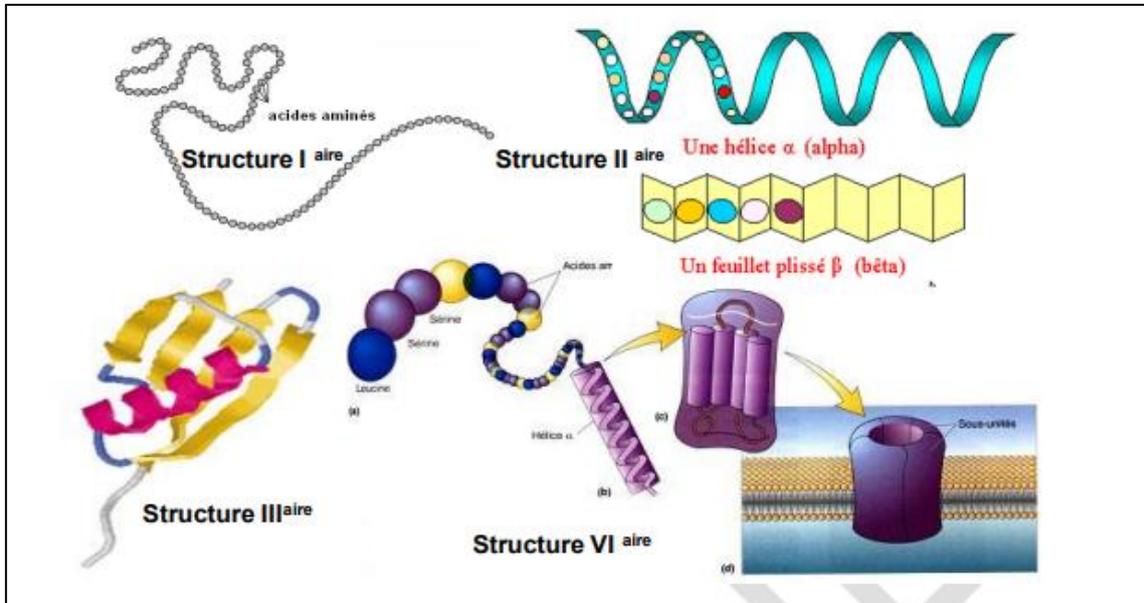


Figure 3 : Différents niveaux d'organisation structurale de protéines

b. Caractéristiques physicochimiques des protéines :

Du fait de leur composition en acides aminés, les protéines diffèrent les unes des autres par leur **masse molaire** et leur **charge nette** à un pH donné. La charge nette est caractérisée par le **point isoélectrique** ou **pI**. Les protéines ont des pI qui s'échelonnent de 2 à 12 et des masses molaires de 5 kDa à 590 kDa.

3. Principales étapes de la protéomique :

L'approche protéomique **différentielle** vise généralement à mettre en évidence les protéines dont l'expression est **modulée** dans une **pathologie** donnée. Cela permet l'établissement d'une vision pertinente des mécanismes physiopathologiques impliqués ainsi que la recherche, l'identification et la validation de nouveaux biomarqueurs. La technologie la plus classiquement utilisée en protéomique repose sur des **gels d'électrophorèse bidimensionnelle (2D)**. Cette approche est fondée sur la séparation des protéines en fonction de deux caractéristiques biochimiques : leur **point isoélectrique** et leur **masse moléculaire**. Après séparation, les protéines d'intérêt sont récupérées des gels, digérées à la trypsine et identifiées par **spectrométrie de masse**. La masse des peptides est mesurée par **spectrométrie**, pour donner un « profil de fragmentation » spécifique de cette protéine. Il faut ensuite retrouver dans les bases de données disponibles à quelle protéine correspond le profil obtenu (Fig4).

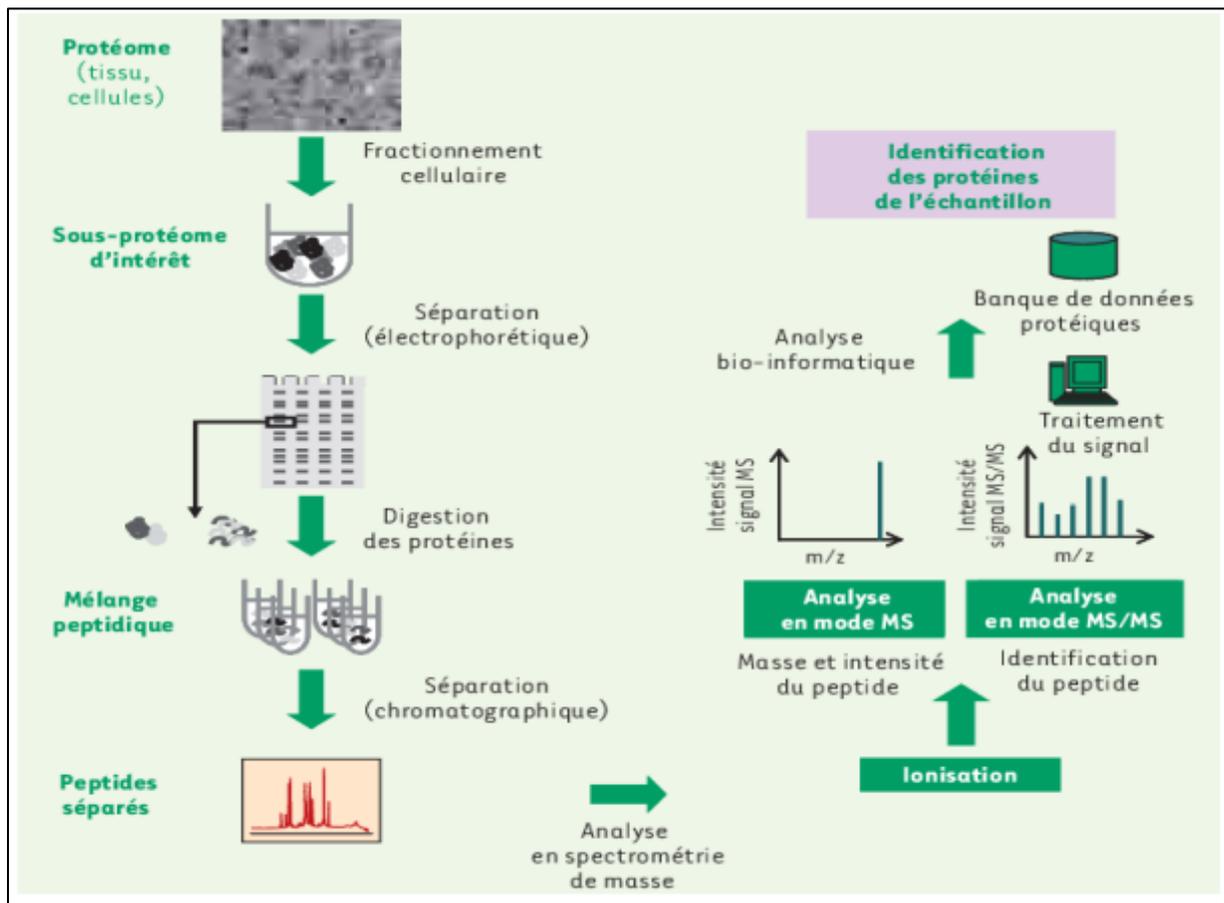
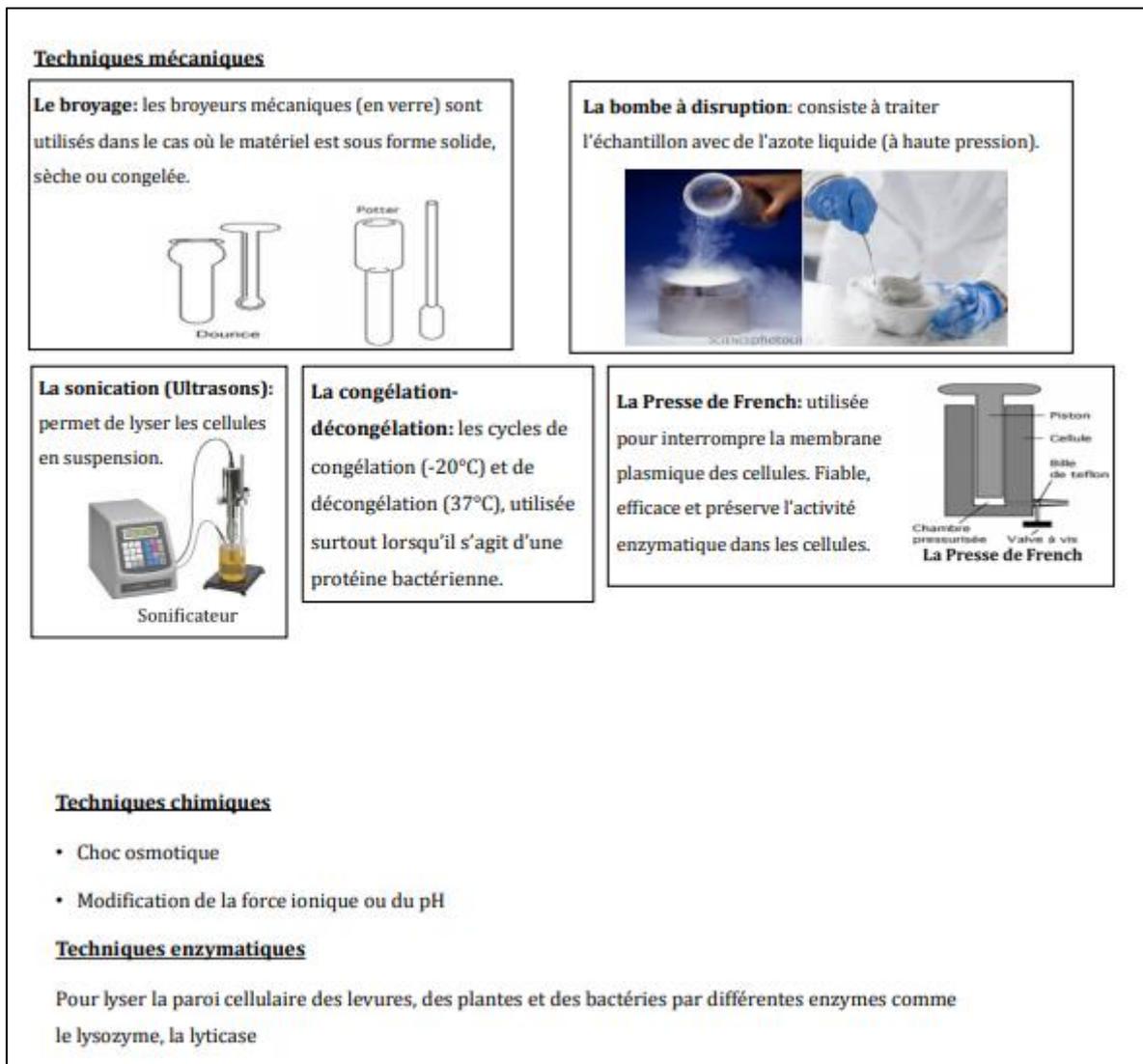


Figure 4 : Principales méthodes d'étude des protéines.

a. Extractions des protéines :

C'est une étape cruciale car elle peut entraîner une dégradation ou une dénaturation irréversible des protéines et compromettre leur identification. Si la protéine se trouve dissoute dans un liquide biologique (ex: plasma, LCR, urines, salive...) aucune méthode d'extraction n'est nécessaire. Si la protéine est à extraire d'un tissu ou d'une culture cellulaire, l'extraction commence par la destruction de l'organisation cellulaire par des méthodes mécaniques, chimiques ou par l'action d'enzymes (Fig5).



Figures 5 : Les méthodes d'extraction des protéines

b.Électrophorèse bidimensionnelle :

Cette technologie permet la séparation des protéines d'un mélange complexe, extraites de tissus, de cellules ou d'autres échantillons biologiques, selon deux propriétés physicochimiques : **le point isoélectrique (pI)** et **la masse moléculaire** des protéines. La combinaison de ces deux principes de séparation électrophorétiques permettant de visualiser un **grand nombre de protéines de façon simultanée** (Fig 6).

Le principe général de l'électrophorèse repose sur la migration de particules chargées sous l'effet d'un champ électrique. Différents supports de migration sont utilisés : électrophorèse en veine liquide, électrophorèse sur papier et électrophorèse sur gels (agarose, polyacrylamide...). Les gels de **polyacrylamide** sont le plus fréquemment utilisés pour la séparation des protéines.

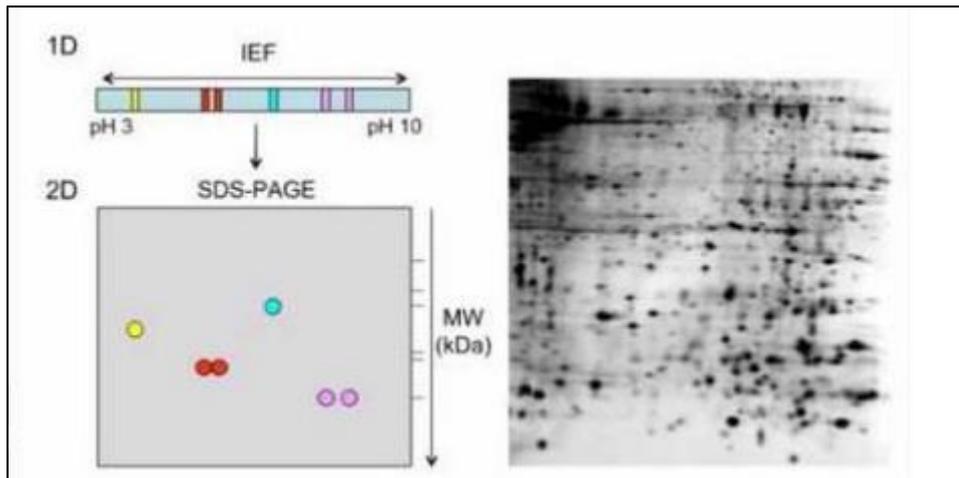


Figure 6 : Schéma simplifié du principe général de l'électrophorèse bidimensionnelle.

Première dimension : séparation selon la charge par Isoélectrofocalisation (IEF)

Au sein d'un même extrait protéique on aura un ensemble des protéines à différentes charges et pHi. Dans la focalisation isoélectrique, les protéines (à l'état natif, non dénaturées) migrent dans un tube étroit de gel de polyacrylamide dans lequel il est établi un gradient de pH. Chaque protéine migre dans le gradient de pH jusqu'à un point qui correspond à son pH isoélectrique et elle se stabilise à ce niveau. L'IEF est une méthode de séparation à haute résolution, elle permet la mesure des pHi de différentes protéines.

Deuxième dimension : séparation selon poids moléculaire (SDS-PAGE)

L'électrophorèse en présence de Sodium Dodecylsulfate (SDS) ou l'électrophorèse de zone est développée pour la première fois par Shapiro et al. en 1967. Le principe de séparation d'un mélange complexe de protéines par l'électrophorèse de zone repose sur deux paramètres critiques à savoir : **Réticulation du gel et l'utilisation du SDS comme agent dénaturant des protéines.**

La séparation est réalisée dans un gel réticulé constitué de polyacrylamide, en présence d'un agent chimique dénaturant, le SDS. La réticulation du gel lui confère la propriété de **tamiser** les molécules et de ce fait, des gels d'acrylamide relativement concentrés 7 à 15 % sont utilisés. Le SDS, ce tensioactif anionique, présente la particularité de se lier aux protéines indépendamment de leur séquence, à la concentration de 1,4 g de SDS/ g de protéine. Cette concentration aboutit à la transformation des protéines en édifices moléculaires composés d'une chaîne polypeptidique unique portant une charge négative importante où l'influence de la charge initiale de la protéine soit négligeable. Les molécules sont séparées ainsi en fonction de leur masse relative.

Détection des spots sur les gels : Il existe aujourd'hui de nombreuses méthodes pour la détection des protéines séparées. Sur le gel, les protéines peuvent être révélées par des techniques de coloration utilisant le bleu de Coomassie, l'argent et les réactifs fluorescents, ou par autoradiographie.

Analyse des gels : La première étape de l'analyse est la digitalisation ou numérisation des gels, c'est à-dire la transformation de l'image expérimentale en une information numérique utilisable par l'ordinateur (Acquisition de l'image). La digitalisation des gels consiste en une transformation de l'image expérimentale en une information numérique utilisable par l'ordinateur. Les gels 2D peuvent être comparés aux gels inscrits dans les bases de données comme celle de **SWISS-2-DPAGE** (<http://expasy.org/swiss-2dpage/>) (Fig 7).

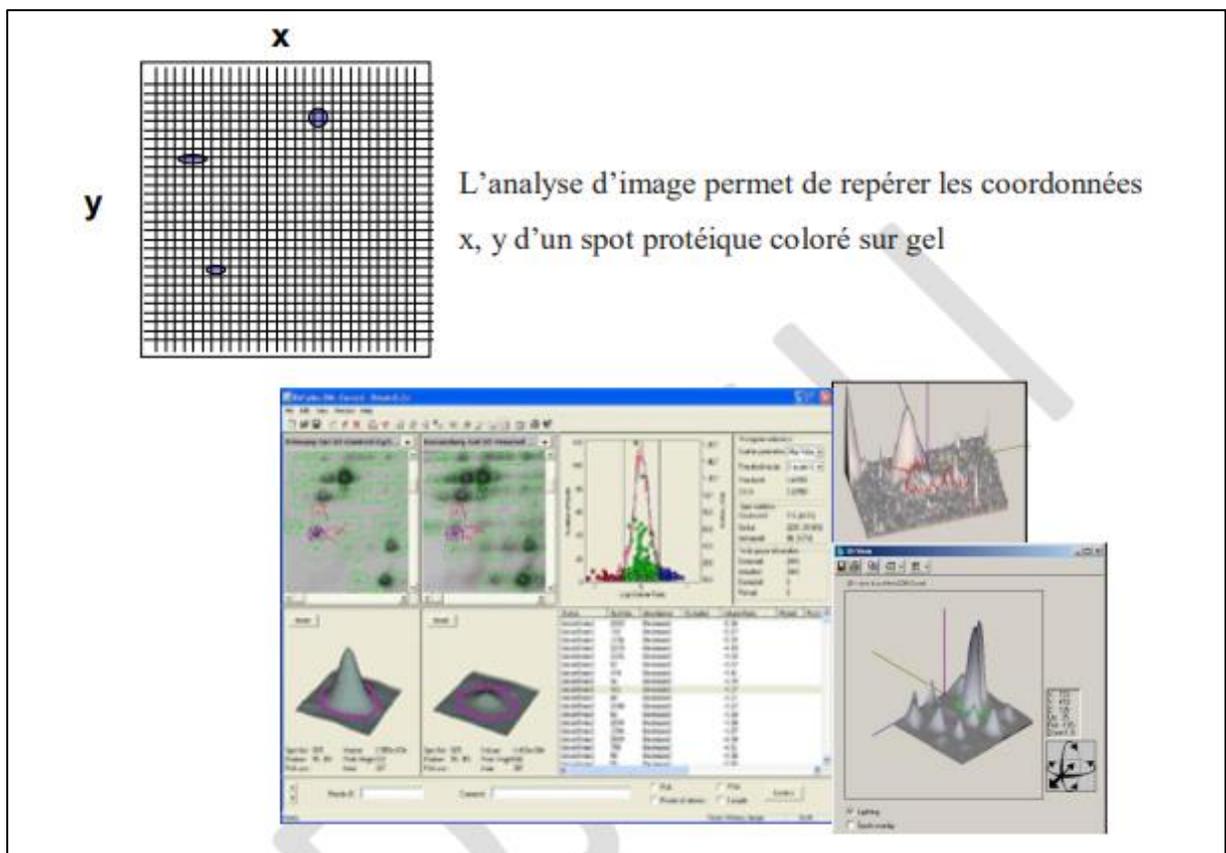


Figure 7 : Schéma représentatif d'un modèle d'acquisition de l'image.

Identification : Après la séparation des protéines par électrophorèse bi-dimensionnelle, les spots sont excisés, digérés (classiquement avec la trypsine). Les étapes allant du prélèvement d'un spot protéique, de sa digestion par une protéase, de l'extraction des peptides de digestion, jusqu'au dépôt sur cible **MALDI** ou en microplaques pour analyse par spectrométrie de masse sont entièrement **automatisables**.

c. Spectrométrie de masse :

La spectrométrie de masse est une technique physique d'analyse permettant de détecter et d'identifier des molécules d'intérêt par mesure de leur masse, et ainsi leur quantification et la caractérisation de leur structure chimique. Le principe de la spectrométrie de masse réside dans la séparation en **phase gazeuse** de molécules chargées (ions) en fonction de leur rapport masse/charge (m/z). Autrement dit, le rapport qu'il existe entre la masse d'une molécule et sa charge électrique. Le spectromètre de masse est un instrument qui comprend plusieurs parties, disposées en série, permettant successivement: Introduction de l'échantillon ; la source d'ionisation permet l'évaporation et l'ionisation des molécules, générer des ions chargés en phase gazeuse (transformation des molécules à l'état naturel en ions à l'état gazeux) ; l'analyseur qui permet de séparer les ions en fonctions de leur rapport m/z ; le détecteur qui permet de détecter et de compter les ions en leur associant leur rapport m/z (c'est-à-dire l'obtention du spectre de masse) et un système informatique permettant l'analyse du signal généré (l'intégration des données) (Fig 8).

Il existe plusieurs types de spectromètres de masse utilisables pour l'analyse protéomique : Maldi-TOF, MS/MS, etc... Un appareil de type MALDI-TOF (pour "Matrix Assisted Laser Desorption/ Ionization Time-of-Flight mass spectrometry") permet d'obtenir une carte peptidique massique représentant la distribution des masses des fragments peptidiques obtenus par digestion trypsique d'une protéine. Le MALDI est une méthode d'ionisation permettant d'analyser des molécules de hautes masses moléculaires (peptides, protéines, oligonucléotides). Dans un spectromètre de masse en mode tandem (ou MS/MS), il est possible d'obtenir des informations de **séquence** en acides aminés.

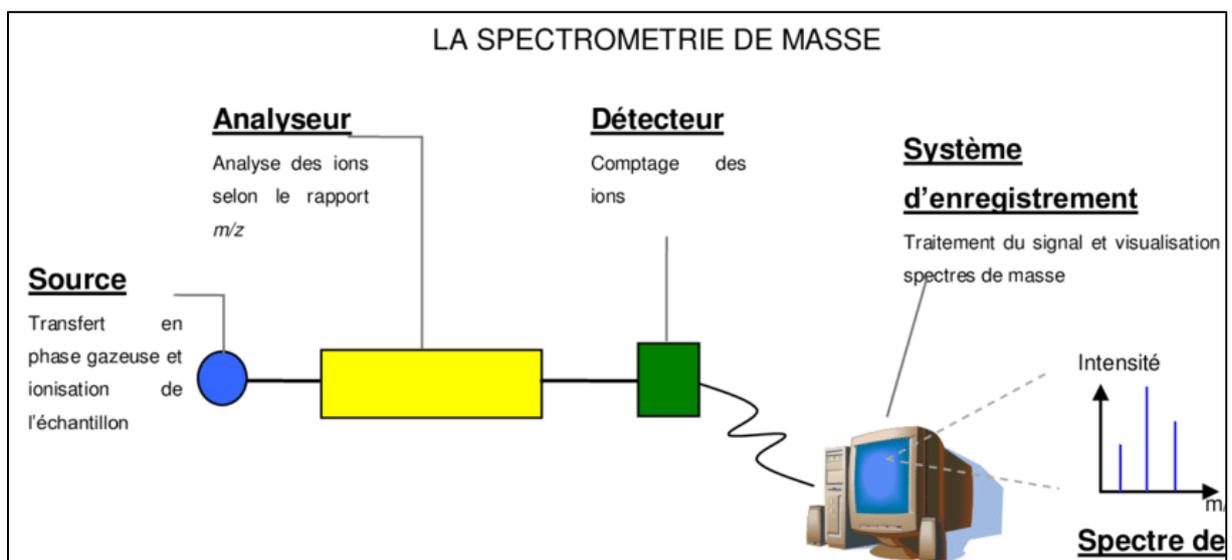


Figure 8 : Principaux éléments d'une spectrométrie de masse