

CAHIER DE

# Formation

Biologie médicale

N°46

2010

---

**Protéomique,  
spectrométrie de masse  
et analyses multiples**



**BIOFORMA**

FORMATION CONTINUE DES BIOLOGISTES

---



*Chère Consœur, Cher Confrère,*

*Nous avons le plaisir de vous présenter un nouveau cahier Bioforma. Depuis les années 60 de très grands progrès ont été accomplis dans les méthodes d'analyses des phénomènes biologiques.*

*Au fil des ans, la mise à disposition des chercheurs d'équipements de haute technologie de plus en plus pointus, et de plus en plus performants, à permis l'étude intime du protéome, conduisant à la mise en évidence de tous les peptides et protéines présents à un instant donné et dans un contexte physiopathologique précis à l'intérieur d'un échantillon de tissu ou dans un prélèvement biologique relevant d'une pathologie particulière. L'électrophorèse bi-dimensionnelle, les chromatographes en phase liquide ou gazeuse liés à des appareils de spectrométrie de masse apportent aux biopathologistes des éléments de réponses précises aux questions qu'ils se posent sur telles ou telles pathologies.*

*Ce cahier de formation n°46, préparé par une équipe de chercheurs et de cliniciens de grande qualité, sous la coordination du Professeur Bruno BAUDIN, vous apporte un ensemble d'informations essentielles qui relèveront demain de la pratique courante des laboratoires de biologie médicale.*

*Déjà, dans le domaine de l'immuno analyse, des avancées significatives produisent des informations nouvelles offrant aux biologistes un terrain de collaboration avec les cliniciens tant dans le domaine de la détermination de la pathologie constatée que dans le suivi du traitement. Nous espérons que ce nouveau numéro des cahiers Bioforma puisse être un outil d'amélioration de vos connaissances et de votre formation continue.*

*Nous vous souhaitons, chère Consœur, cher Confrère, une bonne réception, une lecture attentive, avec nos fraternelles et cordiales salutations.*

**Adrien BEDOSSA**  
Président

86 rue du Cherche-Midi  
75006 Paris

Tél. 01.56.54.39.39

Fax : 01.56.54.39.30

site internet : [www.bioforma.net](http://www.bioforma.net)

E-mail : [bioforma@wanadoo.fr](mailto:bioforma@wanadoo.fr)

Association régie par la loi de 1901

siret : 391 155 744 00025

code APE : 8040

# **Protéomique, spectrométrie de masse et analyses multiples**

Ouvrage réalisé sous la direction du Professeur Bruno Baudin

# PRÉFACE

---

Un numéro spécial sur les nouvelles technologies, quelle gageure et quels défis à relever ! Et, finalement, quels en sont les intérêts pour le biologiste des laboratoires, publics ou privés, de biologie médicale (LBM) ?

La gageure fut le choix des technologies à vous présenter, à vous chers collègues de ces laboratoires. Par mon expérience et mes connaissances dans le domaine de l'analyse protéomique, de la spectrométrie de masse et des méthodes d'immuno-analyse, mon choix a naturellement porté sur ces thématiques, laissant volontairement le champ de la génétique et de la biologie moléculaire à d'autres confrères ou spécialistes ; il y aurait tout autant à dire et présenter. Et finalement je me suis résolu à rester au plus près de ma discipline : la Biochimie, quelle soit fondamentale, métabolique ou clinique comme disaient mes maîtres, et sans désirer les opposer l'une à l'autre. Autant le dire d'emblée, nous ferons la part belle aux protéines et aux métabolites après une ère, de près de 20 ans, du tout ADN.

Même en se restreignant aux technologies choisies, le défi fut ensuite d'organiser les chapitres en une suite cohérente et pédagogique tout en distillant les applications, déjà actuelles et plus encore certainement du proche futur, de ces nouvelles techniques. Sans avoir la prétention d'y être totalement parvenu, la somme des informations données dans ce numéro devrait permettre au biologiste, néophyte dans ces domaines ou juste un peu éclairé par ses souvenirs d'études universitaires, d'appréhender les principes et les aspects techniques de, en particulier, l'électrophorèse bidimensionnelle, les divers spectromètres de masse et leurs couplages à la chromatographie, les nouvelles méthodes de production de fragments d'anticorps ou encore les analyses multiples par anticorps.

Et donc quel intérêt pour le biologiste, médecin ou pharmacien, praticien du public ou du privé ? Tout d'abord, l'analyse protéomique est actuellement la méthode de choix pour identifier de nouvelles protéines, marqueurs potentiels de maladies ou d'états pathologiques particuliers. Cette analyse ne se conçoit pas sans la spectrométrie de masse ; certains spectromètres de masse y sont particulièrement bien adaptés. Ensuite, à n'en pas douter ces technologies connaissent un réel engouement dans de nombreux domaines de la recherche, mais aussi pour des applications médicales, pharmacologiques, toxicologiques et agroalimentaires. La spectrométrie de masse s'implante à grande vitesse dans nos laboratoires de Biochimie, de Pharmacologie médicale, et même de Microbiologie. C'est dans les hôpitaux à forte activité pédiatrique qu'elle est d'abord apparue ; elle a déjà révolutionné le diagnostic et la prise en charge thérapeutique des maladies héréditaires, en particulier les maladies métaboliques avec le dosage des métabolites déficients, excédentaires ou anormaux. Elle commence à s'appliquer à l'étude et la caractérisation des maladies de l'adulte, avec encore le dosage de petites molécules comme les acides et sels biliaires, les acides aminés, les hormones stéroïdiennes et bien d'autres potentiellement. Et de façon inattendue la spectrométrie de masse vient d'entrer dans les laboratoires de Bactériologie avec des spectromètres de masse à ionisation douce permettant l'identification des microorganismes en routine, tout simplement en caractérisant leur contenu en protéines et bien mieux qu'une galerie API même en technologie robotisée. Ses performances, sa simplicité d'utilisation et son faible coût en réactifs assurent d'ores et déjà une amélioration nette du rendu des résultats ; nous aurions dû lui réserver un

chapitre dans ce numéro, mais tout va tellement si vite en ce domaine qu'on se voit rapidement dépassé. A l'aube de l'accréditation généralisée des LBM tant publics que privés, ajoutons que la spectrométrie de masse, en particulier avec dilution isotopique, est devenue la méthode de référence pour caractériser les étalons nécessaires à l'étude de la justesse des mesures. Autant dire qu'il est urgent que nous, biologistes, prenions ce train en marche pour ne pas laisser d'autres, technocrates et financiers, nous devancer comme cela aurait pu être le cas avec l'ouverture du capital des LBM privés aux capitaux d'origine non professionnelle.

Une autre partie de ce numéro est réservée aux analyses multiples de protéines avec quelques applications et perspectives. La dernière partie, complémentaire des deux autres, présente les nouvelles méthodes de production d'anticorps, des anticorps monoclonaux aux anticorps recombinants, à l'origine d'un renouveau de l'immuno-analyse, ces nouveaux anticorps constituant des traceurs uniques, spécifiques d'un antigène donné, d'affinité élevée et produits toujours à l'identique.

Ainsi, la protéomique permet la découverte de nouveaux marqueurs, l'ingénierie des anticorps vise la mise en place d'immuno-dosages performants, et si l'on possède un panel de marqueurs d'un état physiologique ou pathologique particulier on peut tous les doser dans une méthode de multiplexage, à laquelle il faudra joindre une analyse bio-informatique devant le nombre élevé de résultats que cette étude peut générer.

Ainsi, dans le **premier Chapitre**, après un bref rappel de connaissances « universitaires », j'expose les principes et méthodes de l'analyse protéomique, ainsi que les applications actuelles ou prévisibles de l'électrophorèse bidimensionnelle.

L'équipe de **Joëlle Vinh**, du CNRS à l'ESPCI, a été chargée dans le **deuxième Chapitre** d'expliquer le principe des appareils de spectrométrie de masse dévolus à l'analyse protéomique ; elle a développé en particulier le MALDI (TOF et TOF/TOF), le couplage LC-MS/MS et la FT-ICR, avec une brève description de ces appareils. L'exemple de notre chaîne d'analyse protéomique à la Faculté de Pharmacie de Châtenay-Malabry est donné **Chapitre 3** par Florence Poirier et Céline Boursier, respectivement chercheur INSERM et ingénieur à la plateforme « Transcriptome-Protéome ».

J'ai confié à mon collègue le **Pr Sylvain Lehmann** du CHU de Montpellier (**Chapitre 4**) le soin de nous présenter la manière d'appréhender des milliers de données par l'analyse bio-informatique et de nous rappeler les bonnes pratiques en terme de pré- et de post-analytique, en particulier dans le cadre de l'analyse protéomique clinique, donc appliquée aux cohortes de patients.

Nous glissons maintenant vers l'étude des petites molécules par spectrométrie de masse. Dans le **cinquième Chapitre**, **Marie-Claude Menet**, Maître de Conférences à la Faculté de Pharmacie de Paris et analyste reconnu dans ce domaine, nous donne les définitions des termes physiques ou chimiques concernant la spectrométrie de masse et explique les principes de fractionnement nécessaires à bien comprendre pour étudier les petites molécules organiques. Dans le **Chapitre 6**, **Jean-François Benoist**, biologiste des hôpitaux, donne un exemple actuel d'utilisation de couplages de spectromètres de masse à des appareils de chromatographie pour explorer les maladies métaboliques touchant les acides-aminés et les acides organiques.

Quittons la spectrométrie de masse pour découvrir l'analyse multiple des protéines ; l'équipe du **Pr Sylvain Lehmann** décrit dans le **Chapitre 7** les méthodes de multiplexage en milieu solide (puces à protéines) ou liquide, en donnant quelques applications en biologie médicale, mais aussi en montrant que les panels de protéines ou cytokines proposés actuellement ne semblent pas tout à fait convenir à notre pratique de biochimie clinique.

J'ai confié dans le **huitième et dernier chapitre** à mon collègue le **Pr Philippe Billiald** le soin de traiter des méthodes de production des anticorps, et tout particulièrement des anticorps recombinants, déjà très utilisés en thérapeutique, et dont pourrait bénéficier à court terme l'immuno-analyse en biologie médicale.

La science et la technique sont en perpétuel mouvement, l'une alimentant l'autre ; la biologie clinique s'y abreuve depuis longtemps et notre début de 21<sup>ème</sup> siècle n'est qu'à l'aube d'une nouvelle ère technologique. C'est à nous, biologistes des hôpitaux, des LBM et autres structures de santé, d'œuvrer vers un futur certes encore plus technologique, mais pas à tout prix et certainement pas au détriment du patient et de la Santé Publique ; rappelons-nous que l'humain doit rester au centre de notre réflexion et de notre travail.

*Professeur Bruno Baudin*

# Liste des auteurs

---

- **Pr BAUDIN Bruno**  
*Service de Biochimie A - Pôle Imagerie-Biologie, Hôpital Saint-Antoine, APHP  
184 Rue du faubourg Saint-Antoine, 75571 Paris cedex  
Laboratoire de Biochimie et Biologie cellulaire - Faculté de Pharmacie, Université Paris-Sud 11  
5 Rue Jean-Baptiste Clément, 92296 Châtenay-Malabry cedex*
  
- **Dr HESSE Anne-Marie et Dr VINH Joëlle**  
*Laboratoire de Spectrométrie de Masse et Protéomique (SMBP)  
CNRS USR3149, ESPCI Paris Tech, 10 Rue Vauquelin, 75005 Paris*
  
- **Dr PFLIEGER Delphine**  
*Laboratoire d'Analyse et de Modélisation pour la Biologie et l'Environnement (LAMBE)  
CNRS UMR8587, Université d'Evry Val d'Essonne, Boulevard François Mitterrand, 91025 Evry cedex*
  
- **Dr POIRIER Florence et BOURSIER Céline**  
*Plate-forme Transcriptome-Protéome de l'IFR-141, Faculté de Pharmacie, Université Paris-Sud 11  
5 Rue Jean-Baptiste Clément, 92296 Châtenay-Malabry cedex*
  
- **Pr LEHMANN Sylvain et Dr ROCHE Stéphane**  
*Service de Biochimie, Plate-forme de Protéomique Clinique - Institut de Génétique Humaine  
CNRS UPR1142, 141 Rue de la Cardonille, 34396 Montpellier cedex 5  
Laboratoire de Biochimie, Hôpital Saint-Eloi, CHRU de Montpellier  
80 Avenue Augustin Fliche, 34295 Montpellier cedex 5*
  
- **Dr DUPUY Anne-Marie**  
*Laboratoire de Biochimie, Hôpital Lapeyronie, CHRU de Montpellier  
371 Avenue du Doyen Gaston Giraud, 34295 Montpellier cedex 5*
  
- **Dr MENET Marie-Claude**  
*Laboratoire de Chimie Analytique - UFR de Pharmacie, Université René Descartes, Paris V  
4 Avenue de l'Observatoire, 75270 Paris cedex*
  
- **Dr BENOIST Jean-François**  
*Service de Biochimie-Hormonologie, Hôpital Robert Debré, APHP  
48 Boulevard Sérurier, 75935 Paris cedex*
  
- **Pr BILLIALD Philippe**  
*Laboratoire de Biochimie et Biologie cellulaire - Faculté de Pharmacie, Université Paris-Sud 11  
5 Rue Jean-Baptiste Clément, 92296 Châtenay-Malabry cedex*

# SOMMAIRE

## CHAPITRE I

**Bases de l'analyse protéomique - Applications cliniques de l'électrophorèse bidimensionnelle et ses alternatives .....09**

## CHAPITRE II

**Spectrométrie de masse et analyse protéomique .....33**

## CHAPITRE III

**Exemple d'une chaîne d'analyse protéomique :  
de l'électrophorèse bidimensionnelle à la LC-MS/MS .....55**

## CHAPITRE IV

**Pré- et post-analytique en protéomique clinique .....71**

## CHAPITRE V

**La spectrométrie de masse et applications aux petites molécules .....81**

## CHAPITRE VI

**Apport de la spectrométrie de masse dans l'exploration des maladies  
héréditaires du métabolisme .....103**

## CHAPITRE VII

**Analyses multiples des protéines : multiplexage en phase solide (puces à protéines)  
ou liquide : applications en biologie médicale et perspectives .....119**

## CHAPITRE VIII

**Les anticorps recombinants : un répertoire élargi source de nouvelles  
applications en biologie médicale .....129**

**Bases de l'analyse  
protéomique -  
Applications cliniques  
de l'électrophorèse  
bidimensionnelle  
et ses alternatives**

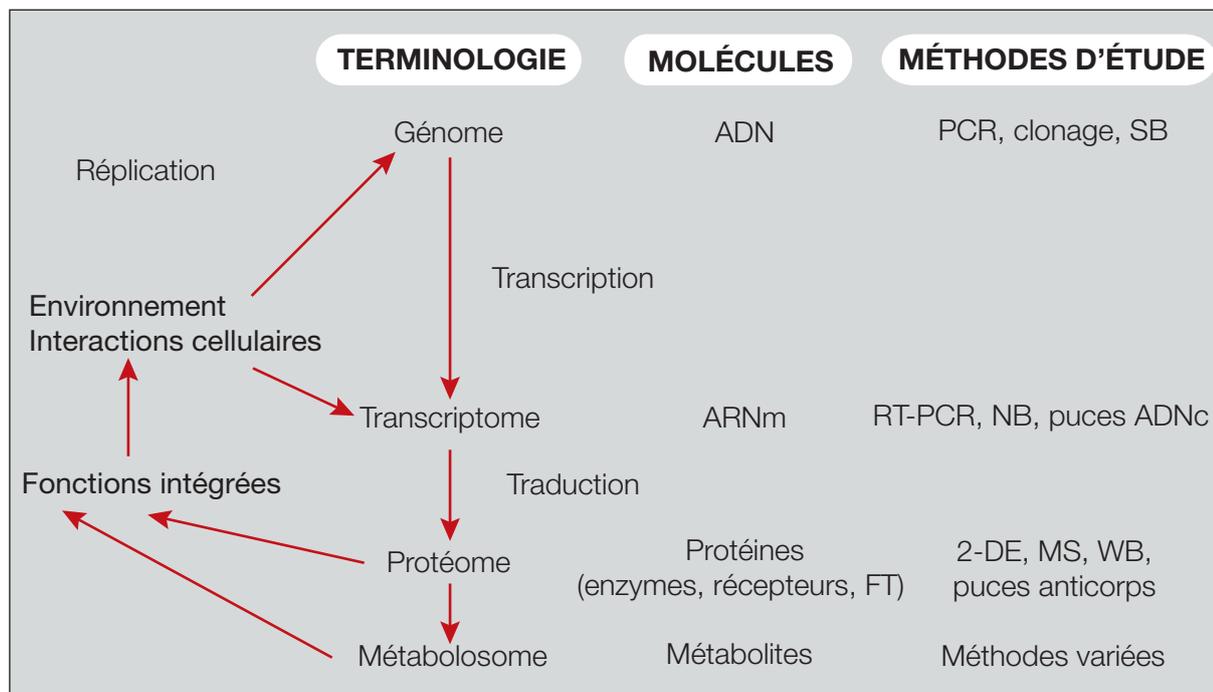
**Pr Bruno Baudin**

CHAPITRE I

# Introduction

L'ère de la protéomique a commencé en 1994 en définissant le protéome, mot hybride entre PROTÉine et génOME, comme l'ensemble des protéines codées par un génome mettant un trait d'union entre le monde de l'ADN et celui des acteurs du vivant, les protéines. Depuis, on a étendu l'étude du protéome à l'analyse systématique des protéines à l'échelle d'un organisme, d'un organe, d'un tissu, d'une cellule ou d'un organite. L'intérêt pour la protéomique humaine a été renforcé par le décryptage du génome humain et par la disparition du paradigme « un gène code pour une protéine qui possède une fonction », beaucoup trop schématique pour apprécier l'ensemble des mécanismes cellulaires et des états pathologiques ; un autre paradigme s'est imposé : « tout processus biologique est multifactoriel » [1]. Des méthodes d'analyses multiples ont progressivement vu le jour, tant pour l'étude des séquences d'ADN ou d'ARN que pour celles des protéines (figure 1).

Figure 1 : Du gène à la protéine et leurs méthodes d'analyse.



PCR (Polymerase Chain reaction)

RT-PCR (Reverse-Transcriptase - PCR)

SB, NB et WB (Southern, Northern et Western-blot, respectivement)

2-DE (Electrophorèse bidimensionnelle)

MS (Spectrométrie de Masse)

FT (Facteur de Transcription).

En recherche clinique, l'ensemble de ces méthodes est basé pour l'essentiel sur les profils d'expression génomique ou protéomique de cohortes de patients présentant une pathologie donnée, celle-ci étant parfois déjà subdivisée à partir de données cliniques ou biologiques existantes [2,3]. L'analyse protéomique nécessite en particulier les techniques d'électrophorèse bidimensionnelle et de spectrométrie de masse, même si ces dernières permettent aussi l'analyse

d'autres molécules comme les métabolites, les hormones, les médicaments et les toxiques, ou encore les acides nucléiques. Avant d'entrer dans le vif du sujet, nous nous proposons de rappeler quelques notions fondamentales sur les gènes et les protéines, notions nécessaires à la compréhension éclairée des parties qui vont suivre et des autres chapitres de ce dossier.

## Des gènes aux protéines

Les chromosomes sont d'énormes molécules d'ADN qui peuvent être transmises de façon assez stable d'une génération de cellules en division à une autre (réplication de l'ADN au cours de la mitose et de la méiose). Chez les eucaryotes supérieurs, les gènes ne représentent qu'une petite partie de l'ADN chromosomique ; la partie de l'ADN ne transmettant pas l'information est appelé ADN non codant. Dans l'espèce humaine, les gènes qui codent pour les protéines comportent en moyenne 28000 paires de bases (pb), huit exons (séquences codantes) contenant en moyenne 145 pb séparés par des introns (séquences non codantes, éliminées au cours de l'épissage) dont la longueur moyenne est de 3000 pb. Le nombre de gènes dans l'espèce humaine serait de seulement 30000 à 35000 alors que des estimations récentes en prévoient près de 100000. Le séquençage des 22 autosomes humains et des deux chromosomes sexuels X et Y est en cours d'achèvement ; la taille du génome haploïde serait de  $3,3 \cdot 10^9$  pb (« Human Genome Project »).

La transcription de l'ARN est la première étape de récupération de l'information codée dans l'ADN chromosomique. L'ARN-polymérase I transcrit les gènes de classe I codant pour les ARN ribosomiques. L'ARN-polymérase II transcrit les gènes de classe II codant pour les ARN messagers (ARNm) et un certain nombre de petites molécules d'ARN intervenant dans la maturation des ARNm. L'ARN-polymérase III transcrit les gènes de classe III codant pour les ARN de transfert ainsi que les plus petits ARN ribosomiques. La transcription est régulée au niveau de la zone promotrice du gène (site de liaison de l'ARN-polymérase) et de régions de régulation beaucoup plus en amont ou en aval du gène. Les produits initiaux de la transcription font l'objet de nombreuses modifications dites post-transcriptionnelles. En particulier, les ARNm subiront une maturation dans le noyau des cellules avec clivage des introns et aboutement des exons (épissage), adjonction d'une coiffe sur l'extrémité 5' et polyadénylation de l'extrémité 3'. Ces deux dernières modifications ont pour but de protéger l'ARNm de l'action des exonucléases, de permettre le passage des pores nucléaires pour se diriger vers le cytoplasme et le réticulum endoplasmique rugueux (porteur de ribosomes), enfin de reconnaître ces ribosomes, sièges de la traduction des séquences nucléotidiques de l'ARNm en séquences peptidiques des chaînes polypeptidiques des protéines en utilisant le code génétique. Les niveaux d'expression protéique ne sont pas un simple reflet de ceux des ARNm, en effet à un niveau constant en ARNm, le niveau de la protéine peut varier d'un facteur 10 à 25, et inversement, à niveau constant en protéine, le niveau de l'ARNm peut varier de 10 à 30 [4].

Les protéines sont donc issues de la traduction d'ARNm matures et stables ; puis elles vont subir des modifications, dites post-traductionnelles (MPTs), pour parvenir à leur forme biologiquement active. Cette maturation des protéines a lieu d'abord dans la cavité du réticulum puis l'appareil de Golgi avec le repliement tridimensionnel (co-traductionnel), des protéolyses limitées, par exemple

pour éliminer la séquence signal ayant permis le passage dans la cavité du réticulum, ou encore des modifications covalentes, comme l'adjonction de chaînes glycaniques (glycosylations), l'assemblage quaternaire covalent par des ponts disulfures (ou non covalent par des liaisons faibles). D'autres MPTs pourront survenir après la sécrétion ou l'adressage de la protéine dans le compartiment où elle doit manifester ses fonctions ; il s'agit encore de modifications covalentes comme les phosphorylations (intervenant souvent dans des régulations métaboliques), les modifications chimiques des extrémités N et C terminales de la chaîne polypeptidique ou encore des chaînes latérales de résidus d'acides aminés (AA) (figure 1).

La séquence (succession des résidus d'AA du N-terminus au C-terminus) définit la structure primaire des chaînes polypeptidiques. La structure secondaire provient de repliements réguliers, issus de rotations permises autour des C $\alpha$  porteurs des chaînes latérales des résidus d'AA stabilisés par des liaisons hydrogènes, formant des motifs structuraux retrouvés dans de nombreuses protéines, les hélices  $\alpha$  et les feuillets  $\beta$  principalement. Seuls les polypeptides de moins de 5 kDa ne présentent pas de structures secondaires. La structure tertiaire regroupe tous les paramètres tridimensionnels des polypeptides, tant des zones repliées en structures secondaires que les fragments de séquence repliés sans structure secondaire, et qui peuvent porter des MPTs comme des phosphates et des glycanes. La stabilité de l'édifice tridimensionnel (minimum entropique de la conformation la plus stable) est assurée par des liaisons de covalence (ponts disulfures intrachaînes) et des liaisons faibles (hydrogènes, ioniques, hydrophobes...). La structure tertiaire rend compte de la plupart des propriétés fonctionnelles des protéines, qui sont soit des protéines de structure, soit des protéines douées de fonctions particulières : enzymes, récepteurs, transporteurs, facteurs de transcription... En particulier, les sites actifs des enzymes, les sites de liaison des ligands des récepteurs et de nombreux sites de reconnaissance moléculaire sont formés du regroupement dans l'espace d'atomes de résidus d'AA qui se sont rapprochés dans l'espace suite au repliement conformationnel. Les anticorps sont dirigés soit contre une séquence d'AA (épitopes séquentiels), soit contre des résidus d'AA regroupés dans l'espace par repliement conformationnel (épitopes conformationnels). Certaines protéines sont formées de plusieurs chaînes polypeptidiques reliées soit par des ponts disulfures interchaînes (comme dans les immunoglobulines), soit par des liaisons faibles (comme dans l'hémoglobine) ; c'est une structure quaternaire, souvent symétrique et formée d'un nombre pair (2, 4, 8..) de chaînes, qui peuvent être identiques ou différentes (codées par des gènes différents) comme les immunoglobulines formées de deux chaînes lourdes identiques et de deux chaînes légères identiques (codées par des gènes différents « lourd » ou « légère »), ou encore l'hémoglobine formée de deux chaînes dites  $\alpha$  (codées par le gène  $\alpha$ ) et de deux chaînes  $\beta$  (codées par le gène  $\beta$ ). L'existence d'une structure quaternaire rend généralement compte de propriétés particulières comme la liaison spécifique de l'antigène et la formation du réseau idiotypique pour les immunoglobulines, et la liaison du dioxygène pour l'hémoglobine avec régulation allostérique, autotrope par le dioxygène lui-même, et hétérotrope exercée par le CO<sub>2</sub> et le bis-phosphoglycérate. De nombreuses enzymes présentant une structure quaternaire subissent des régulations allostériques ; ce serait aussi le cas de récepteurs et de facteurs de transcription.

## Buts et promesses de la protéomique

L'établissement du profil d'expression globale des protéines dans un type cellulaire ou un liquide biologique est appelé protéomique descriptive. Si le génome fournit le répertoire de chaque cellule d'un organisme, la réalité cellulaire va dépendre de l'emploi de ce répertoire en termes de choix des gènes exprimés, du niveau d'expression de ces gènes et des modulations des fonctionnalités des protéines, par exemple par les MPTs. Dans une cellule, 80 % des protéines ne lui sont pas spécifiques ; c'est le cas des « protéines de ménage » souvent fortement exprimées, car elles portent les fonctions essentielles à la vie de toute cellule. Ainsi seulement 500 à 1000 protéines sont caractéristiques d'un type cellulaire ; or si l'on considère qu'il existe environ 250 cellules hautement différenciées dans le corps humain, 100 000 protéines seraient potentiellement intéressantes. L'étude de celles-ci, et plus particulièrement de leurs différences d'expression en fonction de l'environnement et des conditions de vie, est appelée protéomique fonctionnelle ou d'expression ; ce qui donne une vision plus dynamique de la vie cellulaire. Les modifications de l'environnement induisent des variations d'expression spécifiques à la cellule (soit 1 à 2 % des protéines d'une cellule différenciée), ou communes à différents types cellulaires. Néanmoins l'établissement du protéome constitutif (« cartographie ») reste la ligne de base des études fonctionnelles. Quant aux MPTs, elles induisent des changements dans la masse des polypeptides analysables par spectrométrie de masse, et sont visibles par les méthodes bidimensionnelles, en particulier par l'électrophorèse bidimensionnelle (2-DE). Finalement la plus grande différence entre génomique et protéomique réside dans le fait qu'à un génome fixé par l'espèce va correspondre une infinité de protéomes par organisme et en évolution constante. Plus encore, les protéines ne fonctionnent pas de façon isolée dans la cellule, mais sous forme de complexes multimoléculaires ; l'étude de l'ensemble des interactions entre protéines dans une cellule est appelée « interactomique » et celle des complexes multiprotéiques la « complexomique ». L'analyse protéomique reste un défi analytique en raison de la diversité des protéines (pI allant de 2 à 14 et Mr de 5 à plus de 500 kDa) et de la gamme de concentrations auxquelles elles sont exprimées (par exemple dans une quelconque cellule humaine, l'actine est présente à  $10^8$  molécules par cellule alors que les facteurs de transcription ne le sont qu'à quelques centaines d'exemplaires) [2,3].

Quels sont les buts et les promesses de la protéomique ? En biologie fondamentale, elle améliore déjà notre compréhension de la physiologie des cellules et de leurs régulations sous l'influence de stimuli (approche sans a priori). L'un des premiers challenges sera d'acquérir de nouvelles données reliées à la croissance, à la différenciation, à la sénescence, aux modifications de l'environnement et aux manipulations génétiques. Un grand espoir se porte aussi sur la puissance de la protéomique pour identifier de nouvelles cibles d'interventions thérapeutiques sur les maladies, étant donné que près de 80 % des cibles des médicaments sont des protéines. En biologie clinique, la protéomique devrait permettre d'identifier des nouveaux marqueurs spécifiques et sensibles de maladies, qui pourraient s'avérer utiles aux diagnostic et pronostic des maladies humaines comme le cancer et les maladies cardiovasculaires, neurodégénératives ou encore infectieuses ou inflammatoires. L'analyse protéomique peut également être utilisée avec profit dans les études toxicologiques. Avant de

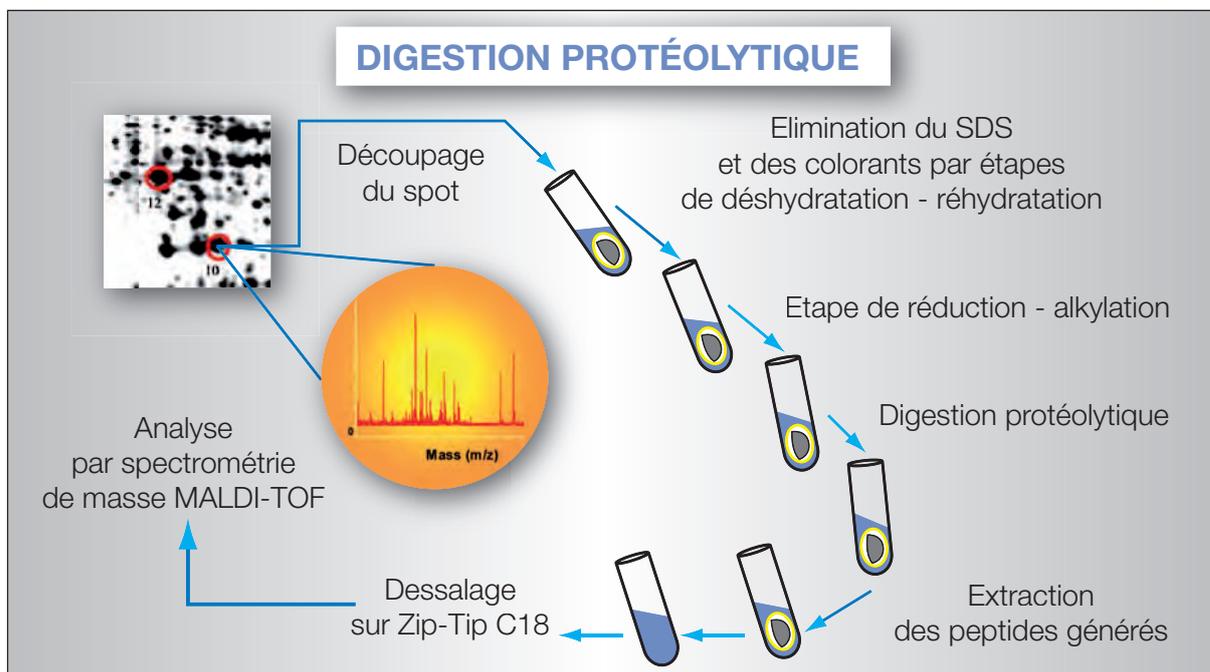
développer les acquis et de dessiner les promesses de la protéomique en biologie clinique, nous nous proposons de décrire les différentes approches techniques actuelles avec, en particulier dans ce chapitre, la 2-DE appliquée à l'analyse protéomique.

## **Electrophorèse bidimensionnelle et analyse protéomique**

### **Principes de l'analyse protéomique alliant l'électrophorèse bidimensionnelle en gel (2-DE) et la spectrométrie de masse (MS)**

La 2-DE est le principal outil permettant de séparer des milliers de protéines ; cette technique se situe donc au premier plan des approches dites « post-génomiques ». L'approche protéomique classique associe la 2-DE à l'analyse des taches (ou spots) d'intérêt par spectrométrie de masse (MS pour « Mass Spectrometry ») en particulier en utilisant le principe du « Peptide Mass Fingerprinting » (PMF), c'est-à-dire l'hydrolyse des protéines par une protéase, généralement la trypsine, puis la détermination précise de la masse des peptides obtenus, ou encore la fragmentation des peptides parents par un deuxième spectromètre de masse en peptides plus petits permettant de retrouver la séquence primaire de la protéine (MS/MS) (figure 2). Dans les deux cas, les masses des peptides (parents ou fragments) sont enregistrées dans des logiciels qui interrogent sur Internet des banques de données protéiques. L'identité des protéines peut être retrouvée par la connaissance des sites de coupure par la protéase utilisée (MS) ou par les masses des principaux peptides fragments retrouvés dans les polypeptides (MS/MS) (voir Chapitre 2). La 2-DE associe la séparation des protéines contenues dans l'échantillon biologique d'abord par isoélectrofocalisation (IEF) en fonction de leur point isoélectrique (pI), c'est la première dimension, puis par électrophorèse dénaturante en gel de polyacrylamide (PAGE-SDS) en fonction de la masse moléculaire (Mr), c'est la deuxième dimension. Un grand soin doit être apporté à la préparation des échantillons, toutes les protéines n'étant pas facilement solubles ou extractibles ; classiquement le culot de précipitation des protéines totales (par l'acétone par exemple) est repris dans le tampon de dépôt sur le « strip » d'IEF, tampon dénaturant contenant de l'urée, de la thiourée, du Chaps (détergent zwitterionique), du Nonidet P40 (détergent neutre), du dithiothréitol (DTT, agent réducteur) et des ampholytes (pour établir le gradient de pH). Les méthodes actuelles permettent de travailler sur quelques dizaines de microgrammes de protéines et de réaliser des électrophorèses multiples en parallèle (IEF et PAGE-SDS). Si la 2-DE a fait ses preuves, elle a aussi montré ses limites en particulier pour étudier les protéines minoritaires (faiblement exprimées) et les protéines membranaires ou très hydrophobes qui pénètrent mal dans les gels [5-7].

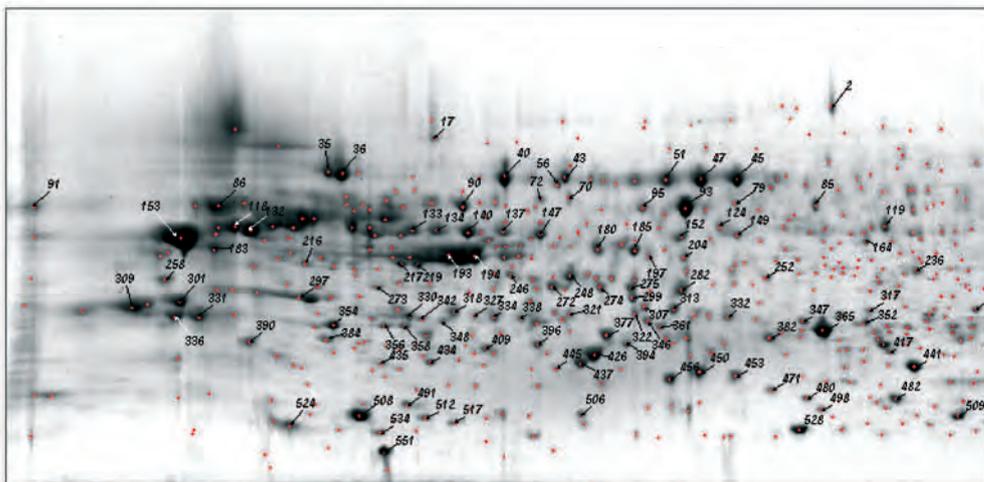
**Figure 2 : Principe du Peptide Mass Fingerprinting (PMF) en analyse protéomique : de la 2-DE pour la séparation des protéines à la MS/MALDI-TOF pour leur identification (Protocole de l'ESPCI-CNRS-USR3149).**



L'IEF sur gel est devenue aisée et reproductible grâce aux gels de polyacrylamide à gradient de pH immobilisé (IPG) utilisant essentiellement des Immobilines® (GE Healthcare), dont il existe de nombreuses gammes (larges : 3-10, 3-11, ou plus étroites : 4-7 « zone acide », 9-12 « zone alcaline », voire très étroites : 4,35-5,55...) sous plusieurs dimensions (les strips de 11, 18 et 24 cm sont courants, mais des gels d'IEF bien plus grands ont aussi été proposés). De même qu'il existe des méthodes d'IEF « off gel » donc en milieu liquide. Les conditions d'IEF en gel sont classiquement des incréments de voltage en paliers pour atteindre 30000 V/h, sous 1 mA et à moins de 20 °C (thermorégulation nécessaire).

Pour la 2<sup>ème</sup> dimension, après libération des thiols par un agent réducteur et blocage de ces groupements cystéinyles par l'iodoacétamide (alkylant des thiols), les protéines sont séparées par PAGE-SDS dans des gels de dimensions correspondantes aux strips d'IEF et soit dans un gel de polyacrylamide d'une concentration fixe (13 % par exemple), soit dans un gradient de polyacrylamide (8-18 % par exemple) ; dans les deux cas la séparation se fait par tamisage moléculaire. Le tampon de migration est de type Tris-HCl avec urée, SDS, DTT et glycérol. Une fois les gels 2D réalisés, les protéines sont révélées par des méthodes de colorations classiques comme l'emploi du bleu de Coomassie colloïdal (ou CBB pour « Coomassie Brilliant Blue ») ou du nitrate d'argent ( $\text{AgNO}_3$ ), 20 à 100 fois plus sensible, mais moins reproductible et donnant une moins bonne réponse dynamique, aussi moins compatible avec la MS que le CBB. Ainsi, sur un gel classique de 18 x 20 cm on révèle quelques centaines de spots au Bleu et plusieurs milliers au nitrate d'argent (figure 3).

**Figure 3 : Exemple d'un profil de 2-DE avec étiquetage des taches contenant les protéines identifiées par MS (communication personnelle et résultats disponibles sur <http://www.huvec.com>).**



On peut aussi utiliser la radioactivité après marquage sélectif, par exemple l'incorporation de  $^{35}\text{S}$ -méthionine en cours de traduction ou le marquage par le  $^{32}\text{P}$  pour l'étude des phosphorylations (phosphoprotéome). L'emploi de marqueurs fluorescents est en plein essor ; le marquage peut être réalisé avant la 2-DE, ou après l'étape d'IEF ou encore après la 2-DE, par exemple avec le SyproRuby®. On peut obtenir un marquage différentiel si on utilise des marqueurs fluorescents émettant à des longueurs d'onde différentes, par exemple des cyanines fluorescentes dans la méthode DIGE® (« Differential Gel Electrophoresis »). Toutes ces méthodes de détection très sensibles sont compatibles avec la MS mais elles nécessitent un appareillage (en particulier pour la détection de radioactivité ou de fluorescence) et des réactifs coûteux [5-7].

L'analyse des gels par un logiciel dédié va consister en l'acquisition des données 2D, puis à diminuer le bruit de fond, éliminer des défauts sur les gels, dénombrer les spots, les comparer par superposition de fichiers, enfin effectuer une quantification relative par des mesures d'intensité et de volume des spots, travail souvent long et nécessaire, suivi par des études statistiques. Une analyse protéomique complète va permettre d'établir une référence qui pourra se présenter sous forme d'une banque de données éventuellement mise à disposition sur Internet [8]. Les spots d'intérêt (plus ou moins intenses par rapport à la référence) sont découpés des gels, soit manuellement au scalpel, soit de façon automatisée avec un « spot-picker » qui travaille sur des plaques à 96 puits. Chaque spot est ensuite traité individuellement, généralement par la trypsine, en microtubes ou de façon automatisée en plaques 96 puits dans un « digester » ; des protocoles précis d'hydrolyse trypsique ont été décrits (voir Chapitre 3). La masse des peptides peut être obtenue aussi bien par MS de type MALDI-MS que ESI-MS, cette dernière étant généralement couplée à la chromatographie liquide et utilisée en tandem MS/MS (voir Chapitres 2 et 3). Tous les MALDI-MS actuels permettent une acquisition automatique des données et des bombardements programmables ; ils sont tous équipés du mode de détection en Temps de Vol (TOF pour « Time of Flight ») et d'un réflectron qui améliore la précision sur la masse et la résolution. Ce système de

détection nécessite un étalonnage par des peptides de masse connue à  $10^{-4}$  Da, souvent des angiotensines et des neuropeptides ; on utilise aussi des peptides d'autodigestion de la trypsine pour un étalonnage interne [9,10]. L'automatisation de toutes ces étapes, de la 2-DE à l'analyse en particulier par MALDI-MS (« Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry »), correspond à ce qu'on appelle la protéomique à haut débit, menant à identifier plusieurs centaines de protéines dans une même série (voir Chapitre 3). La dernière étape, même si elle utilise des logiciels, reste longue et pas toujours couronnée de succès. L'interrogation des banques de données par différents logiciels (Mascot®, ProteomeDiscoverer®,...), sur la base des masses des peptides déterminées par MS, demande des renseignements complémentaires tels l'emploi de réducteur et d'alkylant dans l'étape de 2-DE, la précision de détermination de masse souhaitée, les gammes de pI et Mr probables (figure 4A). La masse d'un seul peptide ne permet pas d'identifier la protéine dont il est issu ; il en faut au moins cinq identiques aux séquences trouvées dans la protéine, ce qui doit représenter au moins 20 % de la séquence. Le logiciel donne, dans l'ordre statistique décroissant, la probabilité d'identité de la protéine ; il faut alors vérifier la cohérence de ses pI et Mr avec ceux mesurés dans le gel et si les fonctions et localisations sub-cellulaires de la protéine identifiée correspondent à l'hypothèse de travail formulée (figure 4B). Il est conseillé de vérifier le résultat dans plusieurs banques de données différentes (SwissProt®, Prosite®, la banque de données de la NCBI...), qui sont constamment enrichies en séquences provenant de coupures in silico des produits d'expression des banques génomiques (figure 2).

**Figure 4 : Identification de protéines dans des banques de données par un logiciel dédié.**

A- Tableau de recherche de ProteomeDiscoverer® indiquant la banque testée (ici celle de la NCBI), l'enzyme utilisée pour le PMF (ici la trypsine), les modifications covalentes effectuées (ici par l'iodoacétamide pour blocage des ponts disulfures), et les masses monoisotopiques des ions  $MH^+$  provenant des pics sélectionnés sur le spectrogramme. B- Résultat de la recherche : la meilleure statistique est pour l'énolase non neuronale.

**A**

**B**

ProFound - Search Result Summary						Version 4.10.5 The Rockefeller University Edition	
Protein Candidates for search BFBCD655-0494-5768FB77 [168990 sequences searched]							
Rank	Probability	Est'd Z	Protein Information and Sequence Analyse Tools (T)	%	pl	kDa	Ⓜ
+1	1.0e+000	1.72	T gi 13637776 sp P17182 ENOA_MOUSE ALPHA ENOLASE (2-PHOSPHO-D-GLYCERATE HYDRO-LYASE) (NON-NEURAL ENOLASE) (NNE)	51	6.4	47.47	Ⓜ
2	1.9e-005	0.17	T gi 13644877 ref XP_011304.2  transcription factor CA150 [Homo sapiens]	36	9.2	70.10	Ⓜ
3	9.4e-006	0.07	T gi 2498896 sp Q60563 SCP1_MESAU SYNAPTONEMAL COMPLEX PROTEIN 1 (SCP1 PROTEIN) (MEIOTIC CHROMOSOME SYNAPTIC PROTEIN)	30	5.8	99.95	Ⓜ
+4	1.2e-008	-	T gi 14741868 ref XP_011908.3  tumor necrosis factor alpha-inducible cellular protein containin [Homo sapiens]	33	5.2	66.30	Ⓜ
5	2.7e-009	-	T gi 14150056 ref NP_115676.1  hypothetical protein MGC10854 [Homo sapiens]	40	6.2	61.17	Ⓜ

Dans un spot d'un gel 2D, plusieurs protéines peuvent coexister et le MALDI-MS est capable de les découvrir en même temps, le logiciel triant les peptides appartenant aux unes et aux autres protéines. On peut vérifier la présence d'une protéine dans un spot en réalisant un western-blot sur le gel 2D ou un gel 1D après transfert, bien sûr si l'anticorps est disponible.

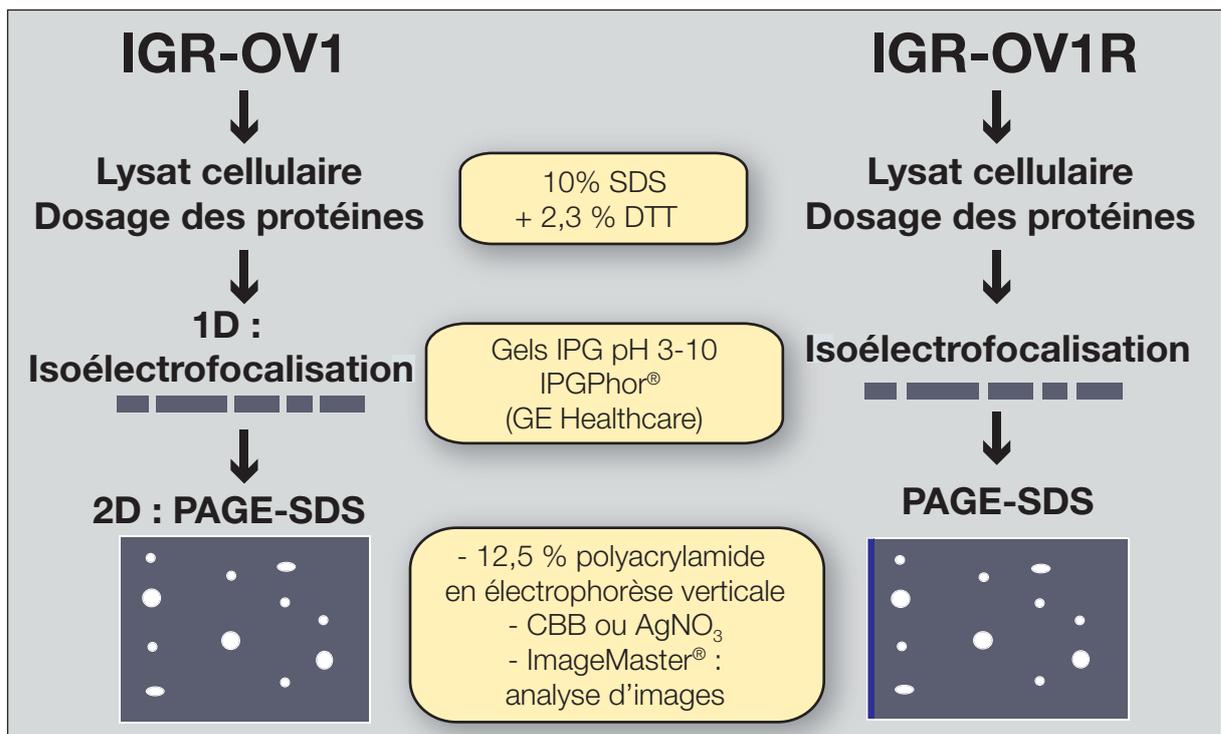
L'ESI-MS (« Electro-Spray Ionization Mass Spectrometry ») est une méthode de spectrométrie de masse aussi bien adaptée aux petites molécules qu'aux macromolécules, en particulier aux protéines entières. En protéomique, l'ESI-MS présente un intérêt double ; le premier réside dans son couplage MS/MS (dit en tandem) pour l'analyse de peptides séparés par LC (« Liquid Chromatography ») à phase inversée, la seconde MS fragmentant le peptide sélectionné (ion parent) en dérivés peptidiques ionisés [11-13]. La mesure des masses de ces fragments, analysés un à un en interrogeant d'autres banques de données, permet de reconstruire la séquence du peptide grâce à des logiciels adaptés. On peut ainsi vérifier des identités douteuses en MALDI-MS et même établir la séquence d'une protéine inconnue ; cette double approche en MS augmente souvent les réussites d'identification. Des modèles évolués couplant nanoLC et nanoSpray permettent le séquençage à partir d'un microlitre d'échantillon (voir [Chapitres 2 et 3](#)). Le second grand intérêt de l'ESI-MS réside dans sa haute précision à identifier des MPTs, en particulier les groupements phosphate (+ 39), acétate, méthyle et bien d'autres, même si les possibilités d'étude dépendent beaucoup du type d'analyseur utilisé [14]. Des grands progrès sur les performances de ces appareils sont présentés chaque année. Gageons que de nouveaux spectromètres bientôt associeront toutes les qualités pour l'identification rapide et sûre des protéines et la quantification directe des variations d'expression observées.

Les applications de la protéomique en biologie cellulaire sont multiples, tant en biologie fondamentale qu'en biologie appliquée, et aussi bien dans le règne animal que végétal ou encore chez les micro-organismes ; plusieurs revues leurs ont été consacrées [7,15,16].

## Applications cliniques de la protéomique classique

En médecine, pharmacie et biologie clinique, une approche protéomique classique concerne, d'une part, la conception de médicaments par la mise en évidence de nouvelles cibles protéiques et, d'autre part, la découverte de nouveaux marqueurs de maladies (bio-marqueurs), potentiellement utiles à leurs diagnostic, pronostic, suivi évolutif et thérapeutique [7]. Dans ce cadre, le fait que l'analyse protéomique ne suppose pas d'hypothèses mécanistiques, elle permet une approche sans a priori ; son intérêt dans la recherche contre le cancer a tout de suite été évident (oncoprotéomique) [17]. Deux approches sont intéressantes en oncoprotéomique, l'une consiste à analyser un fluide biologique, comme le plasma, dans le but de découvrir un bio-marqueur mesurable dans ce milieu. L'autre approche, plus exhaustive, passe par l'étude de la tumeur elle-même, ce qui nécessite plus de préparation des échantillons comme la microdissection des tissus cancéreux ou des séparations cellulaires. Des études sont aussi menées sur des lignées cancéreuses, en particulier comme dans la recherche de marqueurs de résistance aux chimiothérapies ; un exemple de protocole est donné [figure 5](#) [18,19].

Figure 5 : Exemple d'un protocole de protéomique différentielle. Réf. Karen Le Moguen et coll. [18,19].



## Approche tissulaire et cellulaire

L'analyse protéomique est souvent capable de fournir une description moléculaire phénotypique des maladies. Néanmoins de nombreuses maladies font intervenir des protéines faiblement abondantes, souvent difficiles à mettre en évidence y compris par l'analyse protéomique classique couplant 2-DE et MS. Certains résultats récents sont cependant prometteurs.

Le cancer le plus étudié par cette approche protéomique est certainement le cancer colorectal, caractérisé par son évolution à bas bruit et son incidence élevée. Une étude a mis en évidence la

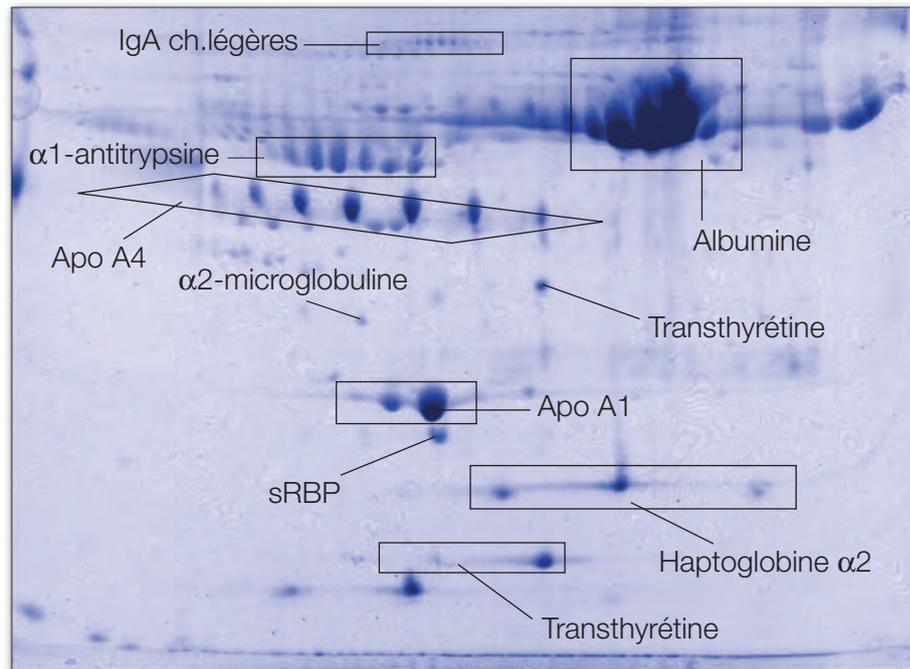
diminution d'expression dans le tissu cancéreux colorectal de la « Selenium-Binding Protein 1 », un facteur de survie [20]. Diehl *et al.* [21] ont établi le secrétome (ensemble des protéines secrétées) de cellules de cancer du côlon par analyse 2-DE différentielle en fluorescence ; ils ont mis en évidence l'augmentation d'expression de la phosphoglycérate-kinase, d'une protéine-disulfure isomérase et de la protéine CSJ-3. Chen *et al.* [22] ont proposé la protéine CLIC-1 comme marqueur pronostique des tumeurs gastriques en la validant par immunohistochimie ; ce canal chlore intracellulaire régule l'acidification de l'estomac, en relation avec la gastrine. Après microdissection laser de tumeurs hépatiques primitives, des marqueurs de ce carcinome ont été proposés par Ai *et al.* [23], comme la peroxyrédoxine-2 (enzyme anti-oxydante) diminuée dans ces tumeurs. Sarto *et al.* [24] ont montré la diminution de l'expression de l'ubiquinol-cytochrome C-réductase dans des cellules de carcinome rénal, faisant de cette protéine de la chaîne respiratoire un marqueur potentiel de ce cancer souvent diagnostiqué trop tardivement. Dans le cancer du poumon, plusieurs marqueurs ont été proposés à partir d'une étude de microdissection du tissu bronchique cancéreux en comparaison du tissu épithélial sain adjacent, en particulier mdm-2 et c-jun, deux produits d'oncogènes, et l'EGFR (« Epidermal Growth Factor Receptor »), récepteur membranaire d'un facteur de croissance souvent incriminé dans les phénomènes de cancérisation des tissus épithéliaux ; ce dernier représente une cible thérapeutique extrêmement prometteuse [25]. En dernier exemple, Wozny *et al.* [26] ont montré la sur-expression de l'annexine-III dans des tumeurs de la prostate.

Dans le cadre de la compréhension moléculaire des relations hôte/pathogène et de la réponse immunitaire induite, la protéomique semble particulièrement attractive. L'étude comparative d'agents infectieux, pour corrélérer expression protéique et virulence, permet de définir des cibles thérapeutiques. L'étude de la réponse immunitaire induite au cours des infections ouvre les champs de nouvelles stratégies vaccinales.

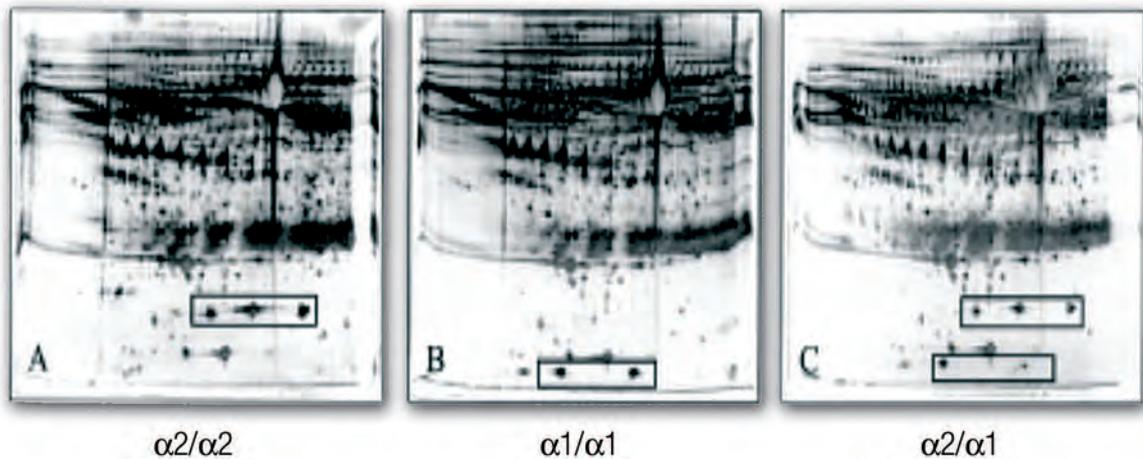
## **Approche directe sur les fluides biologiques**

Dans le plasma humain, déjà plus de 300 protéines ont été identifiées par 2-DE et MS [27-29]. Seules quelques dizaines d'entre elles sont utilisées actuellement comme marqueurs pathologiques (figure 6). On retrouve les isoformes de protéines déjà connues, comme celles de l' $\alpha_1$ -antitrypsine ; d'autres représentent probablement un intérêt pour l'avenir mais devront être validées. Dans notre laboratoire, nous avons mis au point une méthode de phénotypage de l'haptoglobine par 2-DE plus fiable que la méthode classique de PAGE-SDS 1D (figure 7). Bruneel *et al.* [30] ont appliqué la 2-DE et la MS à l'étude des glycosylations dans le cadre des maladies congénitales avec des anomalies portant sur les O- et/ou N-glycosylations, en particulier pour le CDG (« Carbohydrate-Deficient Glycoprotein Syndrome ») de type II sur les glycanes de l'apoC-III et de la transferrine. Le protéome plasmatique est complexe et hétérogène d'un individu à l'autre ; les marqueurs susceptibles d'être intéressants se trouvent probablement dans le plasma à des concentrations très inférieures (< 10 ng/ml) aux seuils de détection des approches actuelles ( $\mu\text{g/ml}$ ).

**Figure 6 : Profil 2-DE d'un sérum normal après coloration au Bleu de Coomassie (CBB) et l'emplacement des protéines majoritaires (Résultats personnels).**



**Figure 7 : Phénotypage de l'haptoglobine par 2-DE (sérum humain), coloration à l'AgNO<sub>3</sub> (Résultats personnels).**



A- Phénotype 2/2 ; B- Phénotype 1/1 ; C- Phénotype 2/1.

Des techniques de séparation ont été proposées pour éliminer les protéines majoritaires du plasma, mais elles sont difficiles d'emploi en particulier pour de grandes cohortes de patients, peu reproductibles et on n'est jamais sûr de pas avoir aussi éliminé des protéines intéressantes. La collection des échantillons est une étape essentielle de l'analyse des liquides biologiques ; la « Human Proteome Organization » (HUPO) a publié des recommandations pour le recueil du plasma et du sérum en vue d'une standardisation des résultats [31,32].

On connaît maintenant le protéome, et même le peptidome, des urines humaines et dans certaines pathologies rénales [33], le protéome du liquide céphalo-rachidien (LCR) [34], celui du liquide de

lavage broncho-alvéolaire, par exemple après rejet de greffe du poumon [35], le protéome du liquide amniotique normal [36] ou celui d'un fœtus atteint de mongolisme [37], ou encore les protéomes du sperme [38] et de la salive [39]. Certaines de ces études ont utilisé directement la MS, généralement en tandem (MS/MS). Des études de conservation de ces échantillons biologiques ont aussi été menées, comme pour le plasma et les urines, même s'il est encore difficile de trouver des consensus pour chacun d'eux [40].

Dans le LCR, la protéine 14-3-3  $\sigma$  a été mise en évidence comme marqueur de la maladie de Creutzfeldt-Jacob ; la présence en grande quantité dans le LCR de cette protéine signe une destruction cérébrale particulière et caractéristique de cette maladie mais aussi de l'encéphalite herpétique [41]. Par ailleurs, une diminution de son expression a été montrée dans des tumeurs du sein, cette protéine étant le produit d'un gène suppresseur de tumeur [42].

## Etude des complexes et des protéines membranaires

### Par 2-DE

Pour l'étude des complexes membranaires (et autres complexomes), on peut utiliser l'association de deux types d'électrophorèses : en 1<sup>ère</sup> dimension le Blue Native-PAGE et en 2<sup>ème</sup> dimension le PAGE-SDS. Une pré-purification des protéines est souhaitable, soit par « Free Flow Electrophoresis » (IEF en veine liquide) soit par chromatographie d'exclusion stérique ; la préparation de microsomes (vésicules contenant du cytosol) doit être purifiée par éclatement du culot membranaire en détergent puis recentrifugation, opérations répétées plusieurs fois si nécessaire. Les protéines constitutives du complexe sont solubilisées en détergent neutre (digitonine ou dodécyl- $\beta$ D-maltoside) puis traitées par le Bleu de Coomassie G-250 qui se lie aux résidus basiques ou aromatiques ; dans ces conditions les protéines se séparent assez bien dans le gel de polyacrylamide en gradient (4-18 % ou 7-12 %) et dans un tampon pH 7 contenant de l'acide amino-6-caproïque (Blue Native-PAGE). Puis, comme dans la 2-DE classique, les protéines subissent le tamisage moléculaire en PAGE-SDS. On peut colorer le gel 2D par le CBB, le nitrate d'argent ou encore utiliser un anticorps [7].

### Par TAP-TAG (« Tandem Affinity Purification »)

L'étude des interactions entre protéines (formant ou non des complexes) a connu un énorme développement ces dernières années (interactomique). Rappelons le succès du « double hybride » pour l'étude d'interactions binaires (deux à deux), mais qui ne fait pas appel à la MS. Nous développerons un peu plus le TAP-TAG qui lui y fait appel. La première étape consiste à créer une protéine de fusion (dans la levure par exemple) entre une protéine appât et, d'une part, la protéine A qui présente la propriété de se lier aux IgG par leur fragment Fc et, d'autre part, la « Calmodulin-Binding-Protein » (CBP) en prenant soin de placer un site de coupure par une protéase entre ces deux séquences du Tag (« l'étiquette »). En purifiant en conditions non dénaturantes la protéine appât par affinité d'abord sur colonne d'IgG, avec élution en présence de la protéase, puis sur

colonne de calmoduline, avec élution en EGTA (complexant le calcium nécessaire à l'interaction), on va obtenir la protéine appât qui lie ses protéines partenaires si elles forment un complexe stable dans ces conditions. On pourra ensuite identifier les protéines du complexe par MS par exemple après digestion trypsique (PMF). Il existe beaucoup de variantes de cette méthode avec l'affinité avidine/biotine, ou le multi-Tag, ou encore le pontage chimique avant l'étape d'affinité permettant de mieux étudier les protéines membranaires. Parmi les succès du TAP-TAG on peut citer la caractérisation de la plupart des interactions protéiques (interactomes) dans la levure (*Saccharomyces cerevisiae*) et maintenant dans le modèle procaryote *Escherichia coli*.

## Par spectrométrie de masse

L'analyse des complexes macromoléculaires par MS est difficile car ces derniers sont souvent instables voire transitoires et le MALDI les dénature par co-cristallisation avec la matrice. Des progrès ont récemment été réalisés avec l'ESI-MS (souvent il s'agit d'un Q-TOF ou hybride quadripôle), en solvant volatil non dénaturant comme l'acétate ou le bicarbonate d'ammonium. Le nombre de protonations est faible ce qui augmente le rapport m/z, avec la formation d'adduits non spécifiques élargissant les pics. Il faut travailler sur des protéines pures, fraîchement préparées et correctement dessalées. On peut jouer sur la pression et la ddp pour trouver les conditions optimales d'analyse. Les premiers progrès ont été obtenus par l'étude d'états d'oligomérisation de protéines multimériques, parfois de la dynamique d'assemblage de complexes, ou encore de l'interaction antigène/anticorps avec la détermination de la stoechiométrie.

Pour étudier l'interaction entre deux protéines, il a été proposé d'utiliser l'échange isotopique H/D (hydrogène/deutérium) basé sur le principe que les zones cachées par l'interaction ne sont pas accessibles au solvant deutéré ; il faut travailler dans ce solvant vers 0 °C, à pH acide (2,5) et rapidement (1 min) ; l'échange H/D se fait au niveau des ponts peptidiques (H amide accessible et échangeable). L'hydrolyse par la pepsine produit des peptides deutérés et non deutérés identifiables par ESI-MS. Par cette méthode il est aussi possible de mettre en évidence les zones ancrées dans la membrane des protéines membranaires intrinsèques, puisque là encore cet enfouissement interdit l'accès au solvant deutéré. En mode cinétique, ce type d'analyse par échange permet l'étude du repliement conformationnel ; en particulier l'enfouissement des hélices  $\alpha$  les rend moins accessibles au solvant. Sur les mêmes principes on peut étudier les zones accessibles au réactif, par exemple de protéolyse ménagée, d'oxydation de surface (par photochimie ou radiolyse de l'eau), ou encore de marquage des chaînes latérales des résidus d'AA, en particulier en réalisant des pontages chimiques bloquant les interactions entre protéines partenaires.

## La protéomique sans gel

Il existe des méthodes qui paraissent plus simples que celles décrites précédemment pour l'analyse des protéomes. Cependant, toutes ces méthodes hautement résolutes souffrent d'un défaut de capacité suggérant la nécessité d'un pré-fractionnement des échantillons et le développement d'outils bio-informatiques plus performants.

## Les couplages chromatographiques

Les méthodes de couplage CLHP/MS présentent l'avantage de garder l'échantillon en milieu liquide, donc de pouvoir l'analyser secondairement, voire de l'aliquoter pour effectuer de nouvelles analyses et de réaliser d'autres couplages plus ou moins directs [43]. Parmi ces méthodes le MudPIT® (« Multidimensional Protein Identification Technology ») combine la CLHP bidimensionnelle (échangeur cationique fort et phase inversée) à l'ESI-MS à trappe d'ions. Cette méthode a permis de produire la plus grande analyse protéomique à ce jour avec 1484 protéines identifiées dans une lignée de levure, mais encore sans applications cliniques à notre connaissance [44]. Selon des principes proches, la technologie ProteomeLab® (Beckman-Coulter) couple deux chromatographies sur colonne (chromatofocalisation en 1<sup>ère</sup> dimension et phase inversée en 2<sup>ème</sup>) avec élution en continu des protéines et peptides de l'échantillon en fonction de leurs propriétés physico-chimiques (pI et hydrophobie). Les molécules recueillies peuvent aussi être analysées par le classique PMF en MS, ou encore par électrophorèse capillaire étant donné que des couplages avec ces appareils ont été prévus [45, 46]. Comme une analyse différentielle est possible, des résultats en clinique sont prévisibles à court terme ; on peut déjà citer une étude sur le cancer de l'ovaire différenciant deux sous-types de tumeurs [47].

Un système basé sur la séparation des protéines et peptides sur des billes magnétiques recouvertes de phases chromatographiques des mêmes types que précédemment est également disponible pour établir des profils protéiques. Une nouvelle technologie vient de paraître avec l'analyse directe des tissus par MALDI-TOF ; elle permet de suivre l'expression d'un peptide ou d'une protéine en 3D donc d'en faire la cartographie tissulaire et même sub-cellulaire [48].

## La protéomique quantitative utilisant des marquages

Plusieurs méthodes de quantification s'appuyant sur un marquage isotopique différentiel des protéomes ont été développées. Les échantillons sont marqués par des isotopes stables et analysés de façon simultanée. Parmi les plus anciennes, la technologie ICAT® (« Isotope-Coated Affinity Tag ») est basée sur le marquage différentiel des protéines de deux séries d'échantillons par conjugaison de la biotine aux thiols cystéiniques par l'intermédiaire d'un groupe léger ou lourd (enrichi en deutérium). Les protéines différemment marquées sont mélangées et traitées par la trypsine ; les peptides générés sont ensuite séparés sur colonne d'avidine (affinité en CLHP) et analysés par MS ; on retrouve sur le spectre de masse le massif isotopique de chaque peptide sous forme légère et lourde, dont on va déterminer le rapport des hauteurs de pics pour une quantification relative [49, 50]. Citons l'exemple d'une étude utilisant cette méthode sur le liquide pancréatique provenant de tumeurs du pancréas [51]. Sur des principes proches on peut encore citer la technologie iTRAQ® (« isobaric Tags for Relative and Absolute Quantitation ») qui remporte actuellement le plus grand succès. Ici le marquage est isobarique et se fait sur les NH<sub>2</sub> (N-ter libre et chaîne latérale de la lysine) ; cette méthode présente l'avantage de permettre la réalisation d'analyses multiples (jusqu'à 8) et sans compliquer le spectre MS (tag isobarique) [52]. Par exemple, cette approche a été utilisée pour étudier la progression du cancer de la prostate [53]. L'approche SILAC® (« Stable Isotope Labelling by Amino acids in Cell culture ») est utilisée pour comparer les protéomes

de deux types cellulaires amplifiables en culture. Un des types cellulaires est cultivé en présence d'une forme isotopiquement alourdie de certains AA (généralement  $^{13}\text{C}_6$ -arginine et  $^{13}\text{C}_6$ -lysine) qui vont être incorporés dans les chaînes polypeptidiques au cours de la traduction ; en MS, chaque peptide (trypsique par exemple) va apparaître sous la forme d'une paire de massifs isotopiques qui est utilisée là encore pour calculer le rapport donnant la proportion relative du peptide dans chacune des conditions, marquée ou non marquée [54]. Cette méthode n'est pas applicable aux tissus et aux échantillons biologiques comme le plasma.

Ces méthodes d'analyse différentielle par marquage isotopique chimique nécessitent l'emploi de logiciels d'analyse bio-informatique performants, indispensables à l'intégration de l'ensemble des données, tout en diminuant les risques de mauvaise identification ou quantification, qui n'est souvent que relative. Une quantification absolue (quantité exacte) est pourtant requise en biologie clinique pour l'évaluation de candidats bio-marqueurs et en pharmacie pour le contrôle de qualité des médicaments. Suivant le principe de dilution isotopique, une quantification exacte d'un composé peut être obtenue en MS en ajoutant dans l'échantillon à analyser une quantité connue d'un étalon interne chimiquement identique au composé à doser, mais alourdi par l'incorporation d'isotopes stables. Plusieurs méthodes de quantification exacte de peptides et protéines par IDMS (« Isotopique Dilution Mass Spectrometry ») ont été développées : AQUA, QconCAT, ou encore PSAQ pour la quantification du PSA (« Prostate Specific Antigen ») dans le sérum. De nouvelles approches de protéomique quantitative émergent constamment, en particulier sans marquage covalent (approches « label free ») et même directement par MS, par exemple avec les spectromètres de masse de type Orbitrap® ou LTQ-FT couplé à la nanoLC.

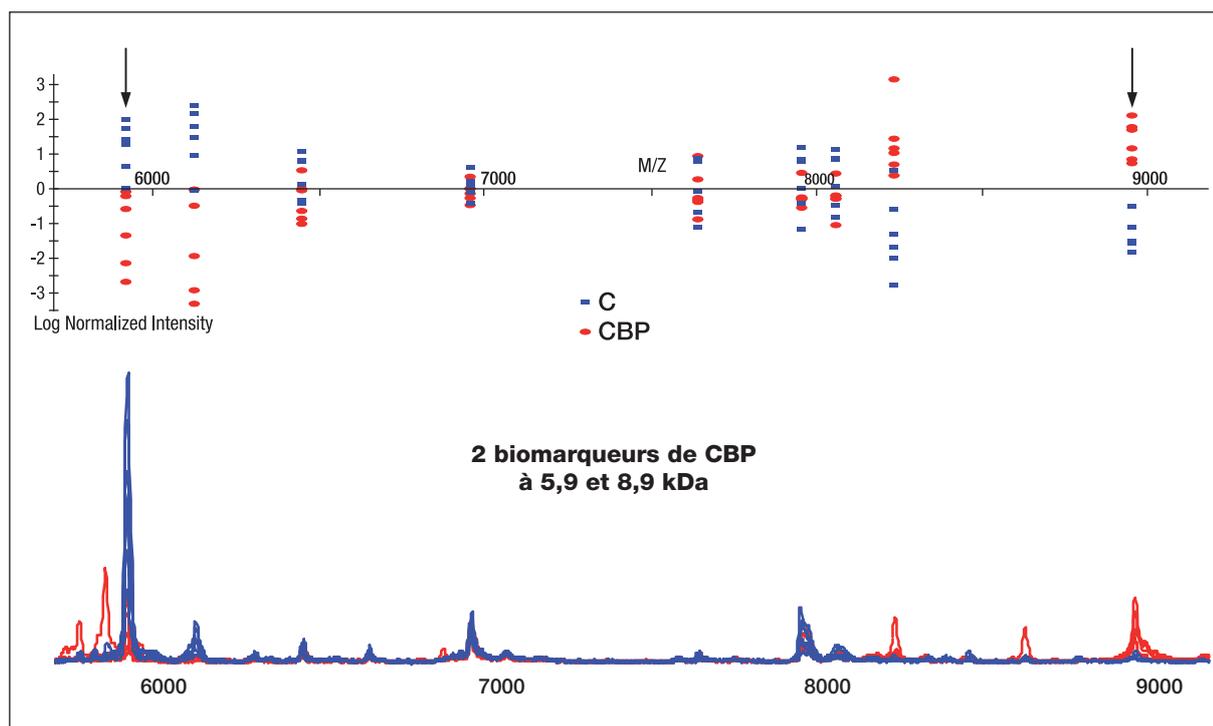
## **L'approche SELDI**

### *Principes*

La technologie SELDI-ProteinChip® associe le couplage du SELDI (« Surface-Enhanced Laser Desorption/Ionization Mass Spectrometry »), cousin du MALDI en détection TOF, avec un système de séparation des protéines et peptides par chromatographie sur surface plane (ProteinChip® Array de chez Ciphergen, maintenant chez BioRad) [55]. Plusieurs types de barrettes « array » sont disponibles ; pratiquement on peut effectuer de l'échange d'ions, anionique ou cationique, de la chromatographie d'adsorption en phase inversée et de l'affinité métallique. Le principe consiste à déposer un échantillon complexe sur la surface chromatographique puis à l'éluer de façon à ne laisser sur cette surface qu'un petit nombre de protéines et de peptides de l'échantillon. Ceux-ci seront directement analysés par le SELDI. Ce système est particulièrement intéressant en analyse comparative, de type patient/témoin, pour mettre en évidence des marqueurs présents dans un échantillon par rapport à un autre. En variant le type de barrette et le schéma d'élution (choix du pH et de la concentration ionique en échange d'ions, du taux d'acétonitrile en phase inversée, du ligand métallique en affinité sur métal), de nombreux marqueurs différentiels ont été proposés pour améliorer le diagnostic, le typage ou le pronostic de diverses maladies. Un logiciel construit des profils protéomiques à partir des données de MS et permet des superpositions de profils mettant en évidence les différences entre échantillons. Si l'on a analysé suffisamment d'échantillons, le logiciel

pourra directement établir une statistique pour proposer les meilleurs marqueurs potentiels et rendre des diagrammes et autres représentations attractives déjà décrites pour les puces à ADN (figure 8). Le ou les marqueurs qui semblent les plus intéressants peuvent être purifiés de l'échantillon de départ en mettant en œuvre les principes chromatographiques qui les ont mis en évidence ; une fois purifiés, ils pourront être identifiés par le SELDI après digestion protéolytique. Divers fluides biologiques ont été étudiés par cette technologie, ainsi que des extraits cellulaires.

**Figure 8 : Approche protéomique par SELDI-ProteinChip® CIPHERGEN. Profil SELDI sur IMAC-40 chargé avec du Zn<sup>2+</sup> ; deux représentations montrant la découverte de marqueurs sériques potentiels de la Cirrhose Biliaire Primitive (CBP), comparaison à du sérum normal (C), échelles en Da (daltons).**



### Applications cliniques

Les premiers succès de cette méthode, appliquée au plasma, ont été dans la détection des cancers de l'ovaire et leur classification pronostique [56] ; mais ces résultats n'ont pas été reproduits [57]. En effet, on sait maintenant que les différences de profils observées étaient essentiellement dues à la nature des tubes de prélèvement du sang différant entre les populations de malades et la population témoin. D'autres études ont permis l'identification de marqueurs du cancer de l'ovaire, comme la protéine SAA (« Serum Amyloid A ») pour laquelle une grande étude prospective par comparaison au marqueur utilisé actuellement, le CA-125, est en cours [58]. Notons que cette protéine de l'inflammation à la phase aiguë différencie aussi l'ostéosarcome de l'ostéochondrome bénin [59]. Une autre étude a identifié l'ubiquitine et la ferritine (chaîne légère), dont les expressions sériques, vérifiées par immunohistochimie dans les tumeurs, sont apparues en miroir l'une de l'autre permettant d'établir un pronostic du cancer du sein [60]. Par ailleurs, un « profiling » SELDI sur plasma a proposé des profils différentiels entre nodule bénin de la thyroïde et cancer du corps

thyroïdien [61]. D'autres résultats semblent en faveur de profils distinctifs entre cancer de la prostate et hyperplasie bénigne [62], ou pour la détection du carcinome hépato-cellulaire [63] en particulier après hépatite C [64, 65]. L'analyse des protéines urinaires par SELDI devrait permettre d'améliorer le diagnostic précoce du cancer de la vessie en comparaison de l'analyse cytologique traditionnelle des urines [66]. Le principal défaut de la technologie SELDI réside dans sa faible reproductibilité ; en effet, plusieurs équipes travaillant sur la même pathologie ont proposé des marqueurs ou des profils protéiques différents [67-70].

## Conclusion et perspectives

---

En conclusion, l'avènement de nouvelles technologies permettant l'étude des transcriptomes, des protéomes et maintenant des métabolomes, lipidomes et glycomes constitue un challenge de la biologie médicale pour le début du 21<sup>ème</sup> siècle, tant dans la compréhension de la physiopathologie que dans le diagnostic et le pronostic de maladies, qu'elles soient rares ou fréquentes, monogéniques ou polygéniques.

Elles permettront, et en particulier la protéomique :

- d'affiner le classement de pathologies d'histologie et/ou de clinique similaires (ontologie)
- de pronostiquer leur progression
- de proposer de nouvelles cibles thérapeutiques et d'individualiser les traitements (« théranostique »)
- d'approfondir notre connaissance des mécanismes physiopathologiques.

## Bibliographie

---

1. Edelman A. Physiopathologie à l'ère de la protéomique contemporaine. *Médecine Sciences* 2005 ; 21 (8-9) : 675-676.
2. Baudin B., Bruneel A. L'analyse protéomique : intérêts et perspectives en biologie clinique. *Immuno-analyse et Biologie spécialisée* 2004 ; 19 : 313-322.
3. Lehmann S., Dupuy A., Peoc'h K., Roche S., Baudin B., Quillard M., Berger F., Briand G., Chwetzoff S., Dine G., Gonzalo P., Dastugue B., Sève M., Siest G., Beaudeau JL. Présent et futur de la protéomique clinique. *Ann Biol Clin* 2007 ; 65 : 463-471.
4. Gygi SP., Rochon Y., Franza BR., Aebersold R. Correlation between protein and mRNA abundance in yeast. *Mol Cell Proteomics* 1999 ; 19 : 1720-1730.
5. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J. Biol Chem* 1975 ; 250 : 4007-4021.
6. Görg A., Obermaier C., Boguth G., Harder A., Scheibe B., Wildgruber R., Weiss W. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 2000 ; 21 : 1037-1053.

7. Baudin B., Bruneel A. Introduction to proteomics : goals, technical aspects and applications to fundamental biology, drug discovery and clinical chemistry. *Recent Res Devel Biophys Biochem* 2003 ; 3 : 977-993.
8. Pernet P., Bruneel A., Baudin B., Vaubourdolle M. PHProteomicDB : A module for two-dimensional gel electrophoresis database creation on personal Web sites. *Genomics, Proteomics and Bioinformatics* 2006 ; 4 : 124-136.
9. Aebersold R., Mann M. Mass spectrometry-based proteomics. *Nature* 2003 ; 422 : 198-207.
10. Shevchenko A., Chernushevich I., Wilm M., Mann M. De novo peptide sequencing by nanoelectrospray tandem mass spectrometry using triple quadrupole and quadrupole/time-of-flight instruments. *Methods Mol Biol* 2000 ; 146 : 1-16.
11. Dancik V., Addona TA., Clauser KR., Vath JE., Pevzner PA. De novo peptide sequencing via tandem mass spectrometry. *J. Comput Biol* 1999 ; 6 : 327-342.
12. Jensen ON., Wilm M., Shevchenko A., Mann M. Peptide sequencing of 2DE gel-isolated proteins by nanoelectrospray tandem mass spectrometry. *Methods Mol Biol* 1999 ; 112 : 571-588.
13. Shevchenko A., Loboda A., Shevchenko A., Ens W., Standing KG. MALDI quadrupole time-of-flight mass spectrometry : a powerful tool for proteomic research. *Anal Chem* 2000 ; 72 : 2132-2141.
14. Mann M., Jensen ON. Proteomic analysis of post-translational modifications. *Nat Biotechnol* 2003 ; 21 : 255-261.
15. Hanash SM., Teichroew D. Mining the human proteome: experience with the human lymphoid protein database. *Electrophoresis* 1998 ; 19 : 2004-2009.
16. Celis JE., Ostergaard M., Jensen N., Gromova I., Rasmussen HH., Gromov P. Human and mouse databases : novel resources in the protein universe. *FEBS Lett* 1998 ; 430 : 64-72.
17. Srinivas PR., Srivastava S., Hanash S., Wright GL. Proteomics in early detection of cancer. *Clin Chem* 2001 ; 47 : 1901-1911.
18. Le Moguen K., Lincet H., Deslandes E., Hubert-Roux M., Lange C., Poulain L., Gauduchon P., Baudin B. Comparative proteomic analysis of cisplatin sensitive IGROV1 ovarian carcinoma cell line and its resistant counterpart IGROV1-10. *Proteomics* 2006 ; 6 : 5183-5192.
19. Le Moguen K., Lincet H., Marcelo P., Lemoisson E., Heutte N., Duval M., Poulain L., Vinh J., Gauduchon P., Baudin B. A proteomic kinetic analysis of IGROV1 ovarian carcinoma cell line response to cisplatin treatment. *Proteomics* 2007 ; 7 : 4090-4101.
20. Kim H., Kang HJ., You KT., Kim SH., Lee KY., Kim TK., Kim C., Song SY., Kim HJ., Lee C., Kim H. Suppression of human selenium-binding protein 1 is a late event in colorectal carcinogenesis and is associated with poor survival. *Proteomics* 2006 ; 6 : 3466-3476.
21. Diehl HC., Stühler K., Klein-Scory S., Volmer MW., Schöneke A., Bieling C., Schmiegel W., Meyer HE., Schwarte-Waldhoff I. A catalogue of proteins released by colorectal cancer cells in vitro as an alternative source for biomarker discovery. *Proteomics Clin Appl* 2007 ; 1 : 47-61.
22. Chen CD., Wang CS., Huang YH., Chien KY., Liang Y., Chen WJ., Lin KH. Overexpression of CLIC1 in human gastric carcinoma and its clinicopathological significance. *Proteomics* 2007 ; 7 : 155-167.
23. Ai J., Tan Y., Ying W., Hong Y., Liu S., Wu M., Qian X., Wang H. Proteome analysis of hepatocellular carcinoma by laser capture microdissection. *Proteomics* 2006 ; 6 : 538-546.
24. Sarto C., Marocchi A., Sanchez JC., Giannone D., Frutiger S., Golaz O., Wilkins ML., Doro G., Cappellano F., Hughes G., Hochstrasser DF., Mocarelli P. Renal cell carcinoma and normal kidney protein expression. *Electrophoresis* 1997 ; 18 : 599-604.
25. Li C., Xiao Z., Chen Z., Zhang X., Li J., Wu X., Li X., Yi H., Li M., Zhu G., Liang S. Proteome analysis of human lung squamous carcinoma. *Proteomics* 2006 ; 6 : 547-558.
26. Wozny W., Schroer K., Schwall GP., Poznanovic S., Stegmann W., Dietz K., Rogatsch H., Schaefer G., Huebl H., Klocker H., Schrattenholz A., Cahill MA. Differential radioactive quantification of protein abundance ratios between benign and malignant prostate tissues: cancer association of annexin A3. *Proteomics* 2007 ; 7 : 313-322.

27. Anderson NL., Anderson NG. The human plasma proteome : history, character, and diagnosis prospects. *Mol Cell Proteomics* 2002 ; 1 : 845-867.
28. Hortin GL. The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. *Clin Chem* 2006 ; 52 : 1223-1237.
29. Hu S., Loo JA., Wong DT. Human body fluid proteome analysis. *Proteomics* 2006 ; 6 : 6326-6353.
30. Bruneel A., Morelle W., Carre Y., Harabou F., Dupont D., Hesbert A., Durand G., Michalski JC., Drouin-Garraud V., Seta N. Two dimensional gel electrophoresis of apolipoprotein C-III and MALDI-TOF MS are complementary techniques for the study of combined defects in N- and mucin type O-glycan biosynthesis. *Proteomics Clin Appl* 2008 ; 2 : 1670-1674.
31. Rai AJ., Gelfand CA., Haywood BC., Warunek DJ., Yi J., Schuchard MD., Mehig R., Cockrill SL., Scott GB., Tammen H., Schulz-Knappe P., Speicher DW., Vitzthum F., Haab BB., Siest G., Chan DW. HUPO Plasma Proteome Project specimen collection and handling : towards the standardization of parameters for plasma proteome samples. *Proteomics* 2005 ; 5 : 3262-3277.
32. Rai AJ., Vitzthum F. Effects of preanalytical variables on peptide and protein measurements in human serum and plasma : implications for clinical proteomics. *Expert Rev Proteomics* 2006 ; 3 : 409-26.
33. Zerefos PG., Vougas K., Dimitraki P., Kossida S., Petrolekas A., Stravodimos K., Giannopoulos A., Fountoulakis M., Vlahou A. Characterization of the human urine proteome by preparative electrophoresis in combination with 2-DE. *Proteomics* 2006 ; 6 : 4346-4355.
34. Liu XD., Zeng BF., Xu JG., Zhu HB., Xia QC. Proteomic analysis of the cerebrospinal fluid of patients with lumbar disk herniation. *Proteomics* 2006 ; 6 : 1019-1028.
35. Zhang Y., Wroblewski M., Hertz MI., Wendt CH., Cervenka TM., Nelsestuen GL. Analysis of chronic lung transplantation rejection by MALDI-TOF profiles of bronchoalveolar lavage fluid. *Proteomics* 2006 ; 6 : 1001-1010.
36. Park SJ., Yoon WG., Song JS., Kim CJ., Oh SY., Yoon BH., Jung G., Kim HJ., Nirasawa T. Proteome analysis of human amnion and amniotic fluid by two-dimensional electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proteomics* 2006 ; 6 : 349-363.
37. Tsangaris GT., Karamessinis P., KoliAlexi A., Garbis SD., Antsaklis A., Mavrou A., Fountoulakis M. Proteomic analysis of amniotic fluid in pregnancies with Down syndrome. *Proteomics* 2006 ; 6 : 4410-4419.
38. Martinez-Heredia J., Estanyol JM., Balleca JL., Oliva R. Proteomic identification of human sperm proteins. *Proteomics* 2006 ; 6 : 4356-4369.
39. Walz A., Stühler K., Wattenberg A., Hawranke E., Meyer HE., Schmalz G., Blüggel M., Ruhl S. Proteome analysis of glandular parotid and submandibular-sublingual saliva in comparison to whole human saliva by two-dimensional gel electrophoresis. *Proteomics* 2006 ; 6 : 1631-1639.
40. Traum AZ., Wells MP., Aivado M., Libermann TA., Ramoni MF., Schachter AD. SELDI-TOF MS of quadruplicate urine and serum samples to evaluate changes related to storage conditions. *Proteomics* 2006 ; 6 : 1676-1680.
41. Hsich G., Kenney K., Gibbs CJ., Lee KH., Harrington MG. The 14.3.3  $\sigma$  brain protein in cerebrospinal fluid as a marker for transmissible spongiform encephalopathies. *N Engl J. Med* 1996 ; 335 : 924-930.
42. Vercoutter-Edouart AS., Lemoine J., Le Bourhis X., Louis H., Boilly B., Nurcombe V., Révillion F., Peyrat JP., Hondermarck H. Proteomic analysis reveals that 14-3-3 sigma is down-regulated in human breast cancer cells. *Cancer Res* 2001 ; 61 : 76-80.
43. Tan S., Tan HT., Chung MCM. Membrane proteins and membrane proteomics. *Proteomics* 2008 ; 8 : 3924-3932.
44. Yang W., Steen H., Freeman MR. Proteomic approaches to the analysis of multiprotein signaling complexes. *Proteomics* 2008 ; 8 : 832-851.
43. Giddings JC. Concepts and comparisons in multidimensional separation. *J. High Res Chromatogr* 1987 ; 10 : 319-323.
44. Washburn MP., Wolters D., Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001 ; 19 : 242-247.

45. Moore AW., Larmann JP., Lemmo AV., Jorgenson JW. Two-dimensional liquid chromatography-capillary electrophoresis techniques for analysis of proteins and peptides. *Methods Enzymol* 1996 ; 270 : 401-419.
46. Tang W., Harrata AK., Lee CS. Two-dimensional analysis of recombinant E. Coli proteins using capillary isoelectric focusing electrospray ionization mass spectrometry. *Anal Chem* 1997 ; 69 : 3177-3182.
47. Zhu Y., Wu R., Yoo C., Cho KR., Shedden KA., Katabuchi H., Lubman DM. Classifications of ovarian cancer tissues by proteomic patterns. *Proteomics* 2006 ; 6 : 5846-5856.
48. Wisztorski M., Lemaire R., Stauber J., Menguellet SA., Jardin-Mathé O., Day R., Salzet M., Fournier I. Imagerie MALDI : une nouvelle technologie pour découvrir et valider de nouveaux biomarqueurs. *Médecine Sciences* 2007 ; 23 : 31-36.
49. Gygi SP., Rist B., Gerber SA., Turecek F., Gelb MH., Aebersold R. Quantitative analysis of protein mixtures using isotope coded affinity tags. *Nat Biotechnol* 1999 ; 17 : 994-999.
50. Han DK., Eng J., Zhou H., Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol* 2001 ; 19 : 946-951.
51. Chen R., Pan S., Yi EC., Donohoe S., Bronner MP., Potter JD., Goodlett DR., Aebersold R., Brentnall TA. Quantitative proteomic profiling of pancreatic cancer juice. *Proteomics* 2006 ; 6 : 3871-3879.
52. Ross PL., Huang YN., Marchese JN., Williamson B., Parker K., Hattan S., Khainovski N., Pillai S., Dey S., Daniels S., Purkayastha S., Juhasz P., Martin S., Bartlet-Jones M., He F., Jacobson A., Pappin DJ. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004 ; 3 : 1154-1169.
53. Glen A., Gan CS., Hamdy FC., Eaton CL., Cross SS., Catto JW., Wright PC., Rehman I. iTRAQ-facilitated proteomic analysis of human prostate cancer cells identifies proteins associated with progression. *J. Proteome Res* 2008 ; 7 : 897-907.
54. Gevaert K., Impens F., Ghesquière B., Van Damme P., Lambrechts A., Vanderkerckhove J.. Stable isotopic labeling in proteomics. *Proteomics* 2008 ; 8 : 4873-4885.
55. Chapman K. The ProteinChip biomarker system from Ciphergen Biosystems : a novel proteomics platform for rapid biomarker discovery and validation. *Biochem Soc Trans* 2002 ; 30 : 82-87.
56. Petricoin EF., Ardekani AM., Hitt BA., Levine PJ., Furaso VA., Steinberg SM., Mills GB., Simone C., Fishman DA., Kohn EC., Liotta LA. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002 ; 359 : 572-577.
57. Kozak KR., Su F., Whitelegge JP., Faull K., Reddy S., Farias-Eisner R. Characterization of serum biomarkers for detection of early stage ovarian cancer. *Proteomics* 2005 ; 5 : 4589-4596.
58. Moshkovskii SA., Vlasova MA., Pyatnitskiy MA., Tikhonova OV., Safarova MR., Makarov OV., Archarov AI. Acute phase serum amyloid A in ovarian cancer as an important component of the proteome diagnosis profiling. *Proteomics Clin Appl* 2007 ; 7 : 107-117.
59. Li Y., Dang TA., Shen J., Perlaky L., Hicks J., Murray J., Meyer W., Chintagumpala M., Lau CC., Man TK. Identification of a plasma proteomic signature to distinguish pediatric osteosarcoma from benign osteochondroma. *Proteomics* 2006 ; 6 : 3426-3435.
60. Ricolleau G., Charbonnel C., Lodé L., Loussouarn D., Joalland MP, Bogumil R., Jourdain S., Minvielle S., Campone M., Déporte-Fety R., Champion L., Jézéquel P. Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry protein profiling identifies ubiquitin and ferritin light chain as prognostic biomarkers in node-negative breast cancer tumors. *Proteomics* 2006 ; 6 : 1963-1975.
61. Wang JX., Yu JK., Wang L., Liu QL., Zhang J., Zheng S. Application of serum protein fingerprint in diagnosis of papillary thyroid carcinoma. *Proteomics* 2006 ; 6 : 5344-5349.
62. Petricoin EF., Ornstein DK., Paweletz CP., Ardekani AM., Hackett PS., Hitt BA., Velasco A., Trucco C., Wiegand L., Wood K., Simone CB., Levine PJ., Linehan WM., Emmert-Buck MR., Steinberg SM., Kohn EC., Liotta LA. Serum proteomic patterns for detection of prostate cancer. *J. Natl Cancer Inst* 2002 ; 94 : 1576-1578.
63. Poon T., Yip TT., Chan A., Yip C, Yip V, Mok T, Lee C, Leung T, Ho S, Johnson P. Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes. *Clin Chem* 2003; 49 : 752-760.

64. Paradis V., Degos F., Dargère D., Pham N., Belghiti J., Degott C., Janeau JL., Bezeaud A., Delforge D., Cubizolles M., Laurendeau I., Bedossa P. Identification of a new marker of hepatocellular carcinoma by serum protein profiling of patients with chronic liver diseases. *Hepatology* 2005 ; 41 : 40-47.
65. Lee IN., Chen CH., Sheu JC., Lee HS., Huang GT., Chen DS., Yu CY., Wen CL., Lu FJ., Chow LP. Identification of complement C3a as a candidate biomarker in human chronic hepatitis C and HCV-related hepatocellular carcinoma using a proteomics approach. *Proteomics* 2006 ; 6 : 2865-2873.
66. Vlahou A., Schellhammer PF., Mendrinou S., Patel K., Kondylis F., Gong L., Nasim S., Wright GL. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. *Am J. Pathol* 2001; 158 : 1491-1502.
67. Wright GL., Cazares LH., Leung SM., Nasim S., Adam BL., Yip TT., Schellhammer PF., Gong L., Vlahou A. ProteinChip® surface enhanced laser desorption/ionization (SELDI) mass spectrometry : a novel protein biochip technology for detection of prostate cancer biomarkers in complex protein mixtures. *Prostate Cancer Prostate Dis* 1999 ; 2 : 264-276.
68. Adam BL., Qu Y., Davis JW., Ward MD., Clements MA., Cazares LH., Semmes OJ., Schellhammer PF., Yasui Y., Feng Z., Wright GL. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res* 2002 ; 62 : 3609-3614.
69. Cazares LH., Adam BL., Ward MD., Nasim S., Schellhammer PF., Semmes OJ., Wright GL. Normal, benign, preneoplastic, and malignant prostate cells have distinct protein expression profiles resolved by surface enhanced laser desorption/ionization mass spectrometry. *Clin Cancer Res* 2002 ; 8 : 2541-2552.
70. Diamantis EP. Proteomic patterns in biological fluids : do they represent the future of cancer diagnostics ? *Clin Chem* 2003 ; 49 : 1272-1278.



# **Spectrométrie de masse et analyse protéomique**

**Anne-Marie Hesse, Delphine Pflieger,**

**Joëlle Vinh**

**CHAPITRE II**

# Introduction

---

Appréhender le fonctionnement d'une cellule suppose de déchiffrer les mécanismes moléculaires complexes qui sous-tendent les différentes activités cellulaires. L'ensemble des gènes, ou génome, constitue une banque de données statique et caractéristique de cet organisme. A partir d'un génome unique, chaque type cellulaire d'un organisme va exprimer un ensemble donné d'ARN messagers (ARNm), ou transcriptome. Ce dernier est « l'intermédiaire de synthèse » des protéines exprimées, ou protéome, qui dépendent de l'environnement auquel la cellule est confrontée (température, potentiel redox, stimuli émis par d'autres cellules, etc...). Ces deux ensembles dynamiques de bio-molécules renseignent sur l'état d'activation des gènes, et donc sur l'activité des cellules. Comme les protéines constituent les acteurs finaux des processus biologiques, leur étude est susceptible d'offrir la vision la plus pertinente du fonctionnement d'une cellule. Dans une première approche, l'étude protéomique se veut exhaustive. Les analyses visent alors à détecter le plus grand nombre possible des protéines présentes dans des échantillons très complexes, tels que des lysats cellulaires entiers, ou des extraits d'organelles. Ces études peuvent, par exemple, permettre de valider l'existence de protéines dont l'expression était jusqu'alors hypothétique ou de localiser les protéines dans des compartiments cellulaires. Une seconde approche se veut différentielle ; l'étude d'extraits protéiques provenant de deux types cellulaires différents permet de mieux connaître les pathologies, en essayant de discriminer un tissu malade d'un tissu sain. Une dernière approche se veut fonctionnelle et se concentre sur une fonction cellulaire donnée ou une voie métabolique particulière. Les protéines spécifiquement ciblées sont alors celles impliquées dans la fonction étudiée. Souvent la cascade de signalisation est finement modulée au niveau protéique par des modifications post-traductionnelles (MPTs), dont la caractérisation est le dernier degré de description de la protéine que peuvent atteindre les analyses protéomiques, avant les études d'interactions intermoléculaires.

La technique d'électrophorèse bidimensionnelle sur gel (2-DE) a été le premier outil permettant de séparer et de visualiser des milliers de protéines. La chromatographie liquide à polarité de phases inversée (RPLC) s'est imposée pour l'analyse de mélanges complexes de peptides protéolytiques. Les techniques de séparation sont ensuite associées à la spectrométrie de masse (MS) pour l'identification des protéines étudiées. L'identification d'une protéine requiert souvent d'obtenir des informations de masse et de séquence par une analyse en mode spectrométrie de masse tandem (MS/MS). De nombreux logiciels ont été développés en support.

L'analyse protéomique est un défi analytique en raison de la diversité chimique des protéines et de la gamme de concentrations auxquelles elles sont exprimées [1-3].

## L'analyse protéomique « bottom up »

### Principes

La stratégie la plus courante est une approche de type « bottom-up », c'est-à-dire une approche utilisant une série d'identifications de fragments de la protéine (les peptides) qui apportent chacun une information spécifique pour l'identification et la caractérisation globale de la protéine étudiée. Les principales étapes de cette stratégie sont les suivantes :

1. Simplification du mélange protéique (optionnel, électrophorèse ou chromatographie par exemple).
2. Digestion du mélange protéique avec une endoprotéase : les protéines extraites sont réduites et alkylées pour réduire les ponts disulfures, avant d'être protéolysées par des endoprotéases spécifiques. L'endoprotéase spécifique utilisée le plus couramment est la trypsine qui coupe les protéines sur les sites de clivage R-X et K-X (excepté RP-KP). D'autres endoprotéases peuvent être employées.
3. Séparation des peptides par chromatographie en phase liquide (mono ou multidimensionnelle).
4. Couplage avec une analyse en MS : la MS simple permet d'établir avec précision la masse des peptides présents. La MS en tandem permet d'établir des éléments de séquences pour ces peptides. En MS, les peptides se fragmentent majoritairement en rompant ces liaisons peptidiques. L'étude des spectres de fragmentation permet de remonter à la séquence du peptide.
5. Traitement bio-informatique pour l'identification des protéines candidates.

La spécificité des protéases permet de prévoir les cartes de masses des peptides ou carte peptidique (PMF pour « Peptide Mass Fingerprint »), caractéristiques de la séquence protéique. Le PMF peut à lui seul permettre l'identification de la protéine. Des informations structurales peuvent aussi être obtenues. Les masses ou les éléments de séquence des peptides obtenus expérimentalement par MS vont être comparés aux masses et séquences obtenues par calcul à partir de toutes les protéines répertoriées dans les bases de données.

Quand les séquences en AA sont longues (une dizaine d'AA ou plus), elles permettent de faire des recherches par homologies de séquences (BLAST), voir même d'initier des expériences de clonage (lorsque le génome entier est connu, 5-6 AA permettent d'identifier la protéine avec BLAST). Les logiciels de recherche d'homologies peuvent être aussi utilisés pour identifier la protéine lorsque des ambiguïtés subsistent dans la séquence peptidique, ce qui est inévitable pour le couple d'AA de même masse Leucine/Isoleucine, mais aussi pour le couple Glutamine/Lysine séparés uniquement avec des spectromètres de masse travaillant à haute résolution.

Quand ces séquences sont courtes (2-3 AA), il est possible de réaliser une recherche d'identification par « Peptide Sequence Tag » (PST) [4]. Typiquement un PST contient un morceau de séquence d'au moins trois AA dont la localisation dans la séquence peptidique totale a été déterminée par MS/MS. Ces données sont très spécifiques ; l'extraction d'un Tag peut être déduite de la [figure 6](#) par exemple.

## Séparation et identification des protéines

### L'électrophorèse bidimensionnelle sur gel

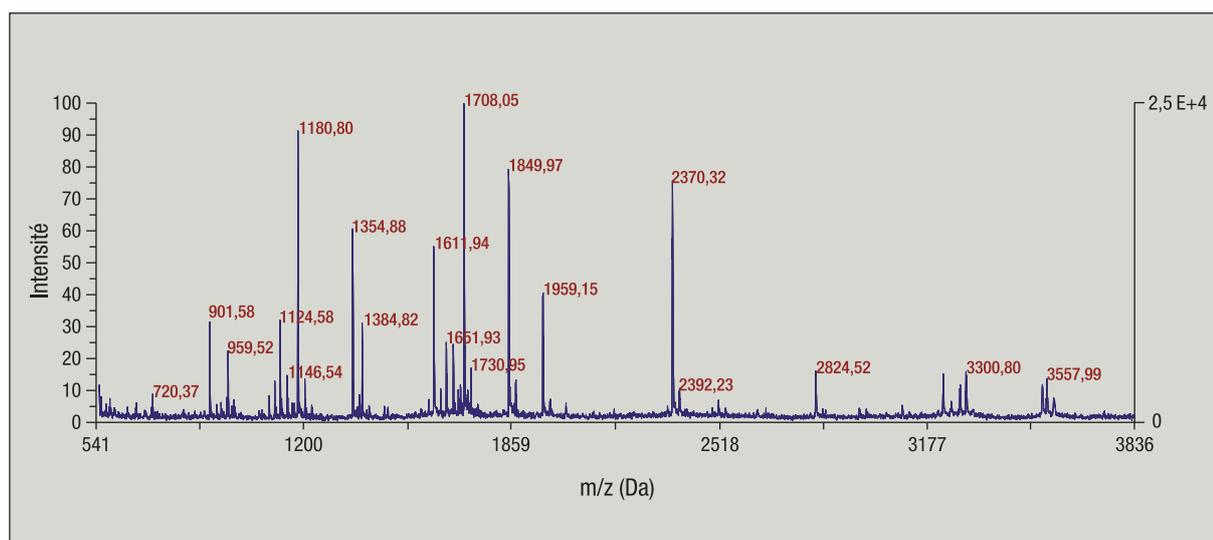
- Séparation et visualisation des protéines :

L'électrophorèse bidimensionnelle sur gel (2-DE) associe deux dimensions de séparation orthogonales, permettant de séparer les protéines suivant leur point isoélectrique (pI) et leur masse moléculaire (MM) [5-7]. Plusieurs milliers de protéines peuvent être résolues sur un même gel [8]. L'électrophorèse bidimensionnelle sur gel est moins adaptée à l'analyse de certaines classes de protéines, telles que les protéines membranaires, ou très basiques, ou encore de faible abondance dans l'échantillon analysé [9-13]. Des protocoles ont été optimisés afin de pouvoir séparer et visualiser correctement ces diverses protéines (voir Chapitre 1).

- Identification d'une protéine par cartographie peptidique :

Dans l'approche protéomique classique, une protéine séparée sur le gel 2D est digérée par une endoprotéase spécifique. L'ensemble des masses des peptides générés (PMF) est confronté à l'ensemble des masses théoriques qui proviendraient de la digestion virtuelle des protéines répertoriées dans les banques, pour obtenir une liste de protéines candidates. Un exemple de spectre de PMF est représenté figure 1.

Figure 1 : Analyse par MALDI-TOF-MS du produit de digestion trypsique de la collicine d'*Escherichia coli*.



Cette approche d'identification échoue lorsque la protéine génère peu de peptides de digestion (protéine de faible MM), lorsque la protéine est fortement modifiée, de sorte qu'un nombre trop important de masses expérimentales de peptides diffèrent des masses théoriques des séquences associées, et lorsque plusieurs protéines de même abondance co-migrent sur le gel (zones hétérogènes).

- Identification d'une protéine par le séquençage de peptides :

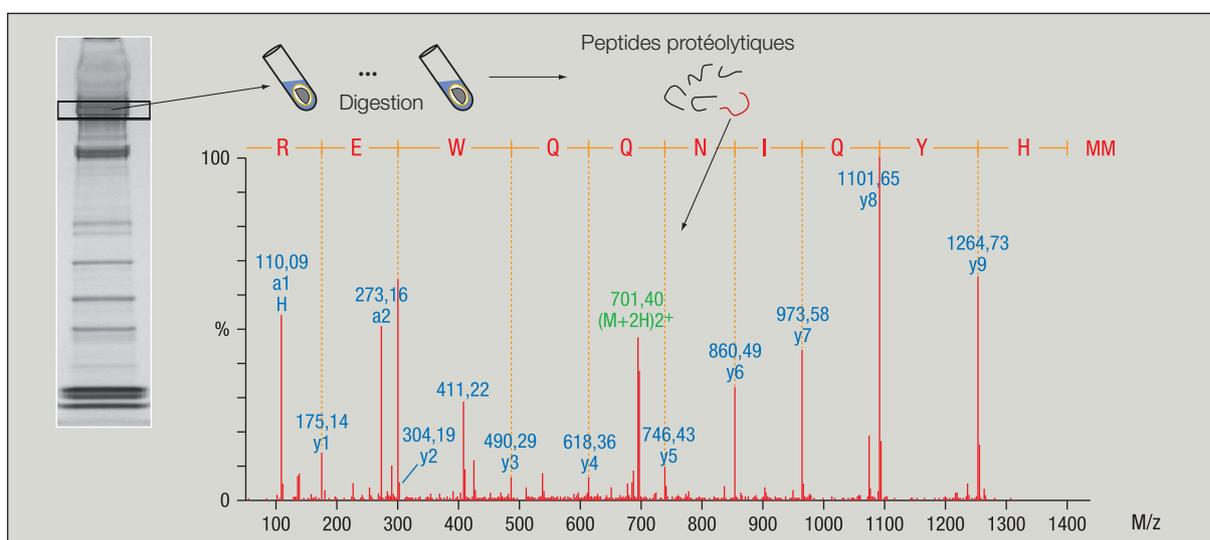
Lorsque le PMF ne permet pas d'identifier la protéine, il est possible de fragmenter quelques peptides protéolytiques en spectrométrie de masse en tandem (MS/MS), afin d'obtenir des informations de séquence. Dans ces conditions, les peptides se fragmentent essentiellement au niveau des liaisons peptidiques. La fragmentation des peptides est revue en détail plus loin.

La fragmentation des peptides protéolytiques peut être réalisée par une analyse LC-MS/MS en mode automatique : la séparation chromatographique permet d'ioniser séquentiellement les peptides avant qu'ils soient détectés puis fragmentés en mode MS/MS. Dès 1998, l'équipe de Yates rapportait une limite d'identification avec cette approche de quelques ng (soit quelques fmol) de protéine détectée sur le gel 2D [14]. Ce type d'analyse est désormais effectué en routine dans de nombreux laboratoires.

### Approche PAGE-SDS et spectrométrie de masse

Afin de s'affranchir des pertes de matériel à ces étapes, une alternative consiste à simplement effectuer une séparation du mélange protéique d'intérêt par électrophorèse monodimensionnelle sur gel de polyacrylamide (PAGE-SDS) [15]. L'identification des composés présents dans une bande nécessite alors d'analyser les peptides protéolytiques associés par séquençage en MS/MS puisque la cartographie peptidique ne permet pas d'identifier les protéines en mélange complexe (protéines membranaires du Golgi [16] et du pore nucléaire [17] par exemple). Généralement l'analyse implique un couplage entre la chromatographie à polarité de phases inversée et la MS/MS après séparation sur gel PAGE-SDS, une stratégie dénommée GeLC-MS/MS [18, 19]. La séparation par PAGE-SDS permet de limiter les pertes de matériel protéique [20]. Néanmoins, on peut supposer que dans le cas de protéines présentes en faible quantité, une fraction plus importante de peptides de digestion n'est pas efficacement extraite du gel de polyacrylamide [21].

**Figure 2 : Schéma de l'association entre une séparation de l'échantillon protéique par PAGE-SDS, la digestion enzymatique de bandes protéiques et l'analyse par MS/MS de peptides protéolytiques. Le gel a été obtenu à partir d'un extrait de matrice extracellulaire.**



### Séparation de protéines entières vers les approches « top-down »

La stratégie de séparation consiste à séparer les protéines intactes par couplage de la MS avec des méthodes de chromatographie ou d'électrophorèse capillaire. Les protéines séparées peuvent ainsi être détectées en ligne par un spectromètre de masse afin de déterminer leur MM, voire directement être fragmentées, mais avec une couverture de séquence plus faible. C'est l'approche « top-down »,

qui se caractérise par une stratégie ciblant la protéine entière directement pour la caractériser. Alternativement, les protéines peuvent être collectées pour être notamment digérées par une enzyme, pour une analyse bottom-up.

### **Séparation par RPLC de protéines digérées**

Afin d'identifier des protéines en mélange complexe, une stratégie analytique largement développée ces dernières années consiste à digérer l'échantillon protéique dans son ensemble, puis à analyser les peptides résultants par un couplage entre une méthode de séparation capillaire en phase liquide et la MS/MS. La chromatographie à polarité de phases inversée (RPLC) est la plus répandue pour la séparation de peptides couplée à la MS [22]. La séparation des peptides est essentiellement dictée par l'hydrophobie [23], [24], et peut faire intervenir un mécanisme de formation de paires d'ions [25]. A pH acide (environ 2-3), les fonctions acides carboxyliques des chaînes latérales sont neutres, et les fonctions amines protonées peuvent s'associer avec le contre-ion du tampon ; ceci a pour effet d'accroître sa rétention.

Les quantités de matériel disponibles pour les études protéomiques sont généralement faibles. Diminuer le diamètre interne de la colonne permet de gagner en sensibilité d'analyse, tout en préservant la capacité de chargement en matériel peptidique pour l'analyse de mélanges complexes. La concentration maximale d'un analyte dans l'éluat chromatographique est, en théorie, inversement proportionnelle au carré du diamètre interne de la colonne [26]. Ceci est particulièrement important dans le cas d'un couplage avec un appareil de type ESI-MS, car le signal détecté en masse est approximativement proportionnel à la concentration de l'analyte dans la source [27-29].

## **L'analyse en spectrométrie de masse**

### **La spectrométrie de masse biologique**

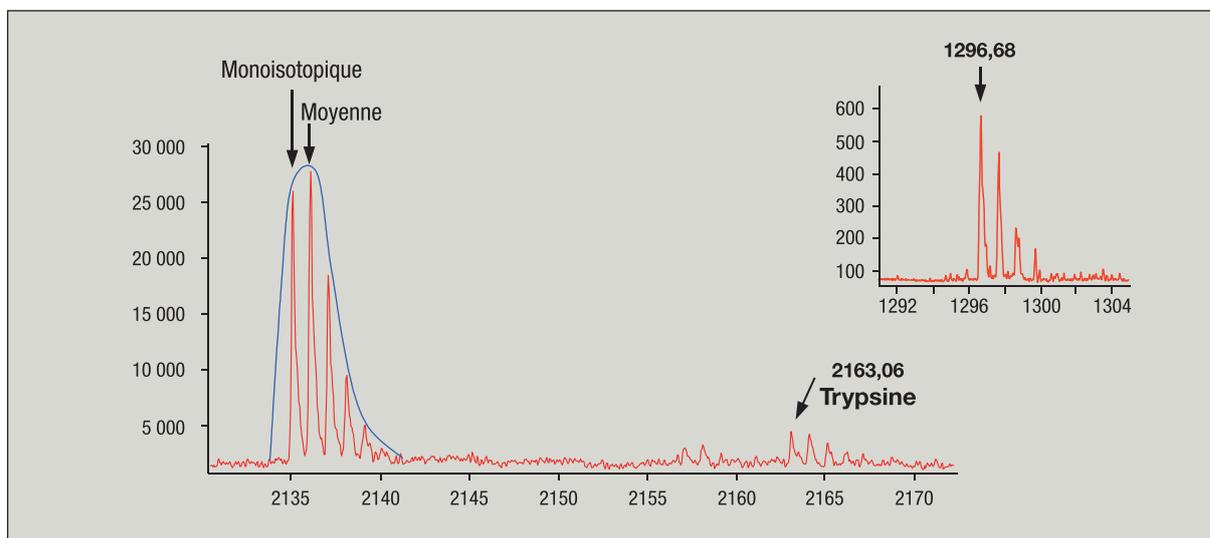
#### **Principe de détermination de la masse moléculaire d'une espèce**

La MS permet de déterminer avec une précision et une sensibilité élevée la masse des molécules. Un spectromètre de masse est un outil analytique qui permet de transformer des molécules dans leur état naturel en ions à l'état gazeux, puis de les séparer en fonction de leur rapport de masse/charge, noté  $m/z$  pour déterminer leur masse moléculaire  $m$ , où  $m$  est la masse du composé et  $z$  sa charge.

La distribution statistique des isotopes naturels dans les macromolécules biologiques a pour conséquence de générer, pour chaque espèce, des massifs de pics équidistants et distants d'environ 1 uma (unité de masse atomique) correspondant à des compositions isotopiques variables (figure 3). Ces compositions résultent principalement de l'abondance naturelle de 1 % de  $^{13}\text{C}$  mais aussi de O, N, etc... selon la composition de l'espèce étudiée. Cette dispersion isotopique varie en fonction du nombre total d'atomes et donc de sa masse. Trois masses peuvent être définies pour une molécule. La masse exacte est calculée à partir des masses isotopiques des isotopes les plus légers. La masse moyenne tient compte de la contribution de l'abondance de chaque isotope. La masse nominale est obtenue à partir des masses isotopiques arrondies à l'unité. Pour pouvoir les

distinguer il faut donc utiliser un spectromètre de résolution suffisante pour séparer les isotopes d'une espèce moléculaire donnée. De même trois unités de mesure peuvent être définies en spectrométrie de masse : l'unité de masse atomique (uma), le dalton (Da) et le thomson (Th).

**Figure 3 : Exemple de profil peptidique obtenu avec un appareil à temps de vol de résolution moyenne permettant de distinguer masse monoisotopique et masse moyenne.**



### Rapide historique

L'histoire de la MS commence avec la découverte de l'électron, en 1897, par le physicien Joseph John Thomson (1856-1940). Il montra l'existence de l'électron et mesura son rapport m/z. Thomson construisit le premier spectromètre de masse au cours des premières années du XX<sup>ème</sup> siècle, et reçut le prix Nobel de Physique en 1906 « en reconnaissance du grand mérite de ces recherches théoriques et expérimentales sur la conduction de l'électricité dans les gaz ».

Les véritables applications biochimiques pour l'étude des peptides et de protéines prirent leur essor vers 1988 avec la découverte de l'ionisation par « électrospray » (ESI) et « désorption/ionisation laser assistée par matrice » (MALDI). Si jusque vers le début des années 90 les appareils à secteurs ont dominé la MS à haute résolution, les spectromètres de masse à temps de vol, quadripôles ou trappe ionique, moins chers et plus simples de mise en œuvre, ont toujours aujourd'hui une place prépondérante pour les applications en biologie et/ou en biochimie, bien que, le coût des appareils à très haute résolution diminuant, ils redeviennent d'actualité (exemple : les appareils à transformée de Fourier).

L'analyse directe de protéines/peptides intacts permet de mettre en évidence des différences de masse moléculaire correspondant à des mutations ponctuelles, à des ponts disulfures, à la présence de MPTs, ou encore à des sites de clivages post-traductionnels dus à une activité protéolytique particulière. La fragmentation en MS permet d'avoir des informations de séquence. Les analyses par MS de mélanges de peptides peuvent être réalisées directement à partir d'organe ou de cellules uniques, par exemple l'analyse de neurones géants, d'hématies ou de cellules d'organe sécréteur d'écrevisse [30].

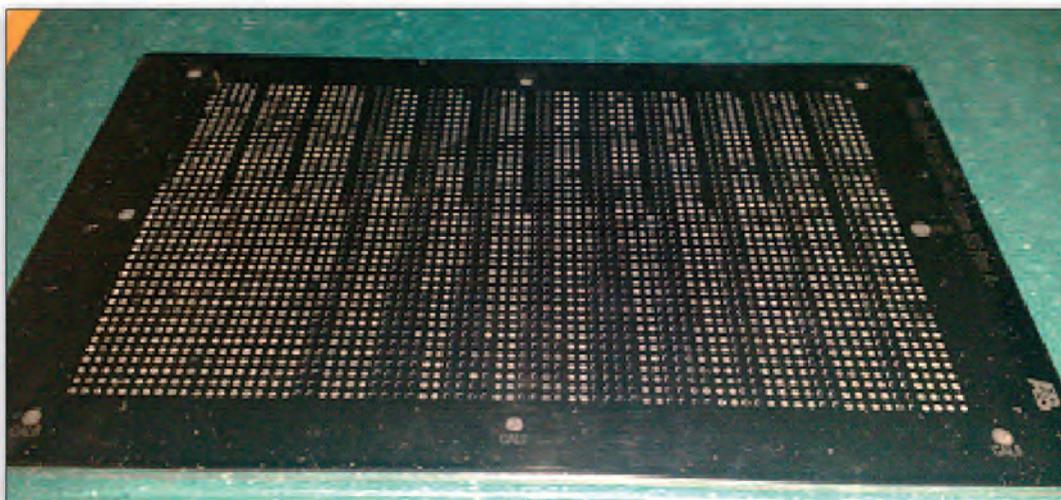
## Aspects techniques de la spectrométrie de masse

### Les différents modes d'ionisation

Lors de l'American Society of Mass Spectrometry Conference de Juin 1988, John Fenn présenta l'application de l'électrospray (ESI) à l'analyse de protéines. Ceci constituait une véritable innovation pour la spectrométrie de masse des biomolécules de haute masse moléculaire. Quelques mois plus tard, Koichi Tanaka et Franz Hillenkamp présentaient le mode d'ionisation MALDI. Hillenkamp montra, lors de l'International Mass Spectrometry Conference, la mesure de masse de protéines allant jusqu'à 117 kDa. Le MALDI semblait aussi prometteur que l'ESI. Fenn et Tanaka ont ensemble reçu le prix Nobel de Chimie en 2002 pour le développement de méthodes pour « le développement de méthodes douces de désorption et d'ionisation pour l'analyse par spectrométrie de masse de macromolécules biologiques ». Nous ne reviendrons pas en détail sur les principes théoriques de ces modes d'ionisation qui ont été traités par ailleurs ([Chapitre 5](#)). En 1988, une question se posait : quelle technique deviendrait la future méthode de choix dans un laboratoire de chimie des protéines ? Il semble aujourd'hui qu'il n'y a pas lieu de trancher. La MS est devenue une approche quasi incontournable dans l'étude et la caractérisation des biomolécules, notamment pour l'étude structurale des peptides et des protéines, indépendamment du mode d'ionisation.

La source ESI est la plus fréquemment associée aux méthodes de séparation en phase liquide, car elle permet une analyse en ligne par MS de l'effluent de la colonne ou du capillaire [31, 32]. Cependant, lors de l'analyse de mélanges complexes, l'échelle de temps imposée par ce couplage en ligne peut limiter la quantité de données rassemblées. En revanche, si l'éluat LC est déposé sur une plaque MALDI ([voir figure 4](#)), l'ensemble des spots déposés peut être analysé plusieurs fois en changeant les paramètres d'acquisition.

**Figure 4 : Photographie d'une cible MALDI préparée en couplage LC-MS. Douze séparations LC ont été collectées en 240 fractions sur cette plaque qui comporte donc 2880 dépôts.**



La source MALDI permet de découpler les analyses MS et MS/MS. Des ions donnant des signaux plus faibles peuvent être sélectionnés, et la durée de fragmentation peut être augmentée, au détriment seulement de la durée d'analyse. A contrario, le couplage LC-ESI-MS/MS consiste à

fragmenter préférentiellement les ions donnant les signaux les plus intenses. Avec ce couplage MALDI-MS/MS hors ligne, il est enfin possible « d'arrêter le temps » : un échantillon donnant des résultats non satisfaisants peut être analysé à nouveau, après un traitement supplémentaire (chimique, enzymatique) du dépôt, ou simplement en insistant sur la fragmentation de certains peptides. Ainsi l'équipe de Moseley a analysé le produit de digestion trypsique de ribosomes de mitochondries de bœuf, en utilisant une source ESI et une source MALDI. Seulement 63 % des peptides ont été caractérisés en commun dans les deux cas, de sorte que chacune des stratégies analytiques apporte des informations spécifiques [33].

### **Les différents types d'analyseurs courants**

Les analyseurs classiques les plus courants en protéomique sont des combinaisons des analyseurs simples suivants : le filtre quadripolaire (Q pour Quadripôle) ; la trappe ionique, tridimensionnelle ou linéaire ; le temps de vol (TOF pour « Time Of Flight ») ; et les trappes ioniques à Transformée de Fourier (à résonance cyclotron ou à potentiel électrostatique « quadralogarithmique »).

Un analyseur est caractérisé par au moins trois paramètres clé :

- la résolution, c'est-à-dire le plus petit écart de masse  $\Delta M$  qui peut être mesuré (ou résolu) à une masse  $M$  donnée, la résolution est le plus fréquemment exprimée comme le rapport  $M/\Delta M$
- la gamme de masse accessible, c'est-à-dire le rapport  $m/z$  maximum transmis par l'analyseur de la source au détecteur
- la sensibilité, ou son pouvoir de transmission, c'est-à-dire la proportion des ions issus de la source effectivement transmis par l'analyseur au détecteur.

La demande en capacités analytiques puissantes dans le domaine de l'analyse biochimique des protéines étant de plus en plus importante, l'instrumentation en MS a évolué très rapidement dans les dernières années. De nouveaux types d'analyseurs et des instruments hybrides ont fait leur apparition, fournissant ainsi de nouvelles opportunités dans l'étude des protéomes complexes [34, 35]. En particulier, les instruments permettant des mesures de masse exacte sont de plus en plus utilisés dans les études protéomiques, non seulement pour identifier des protéines mais aussi pour déterminer des MPTs, ou encore des interactions entre des protéines et d'autres molécules de manière moins ambiguë qu'auparavant.

### **Pourquoi a-t-on besoin d'une mesure de masse exacte ?**

Le niveau de confiance des identifications dépend directement de l'exactitude des mesures de masse [36], c'est-à-dire le degré de conformité de la quantité mesurée avec la vraie valeur. Le nombre de séquences d'AA candidates pour une même masse décroît si la précision de mesure de masse augmente [37-41]. Une précision de mesure de +/- 1 ppm permet d'exclure 99 % des peptides qui ont la même masse nominale mais une composition en éléments chimiques et/ou en AA différente, ce qui résulte en une plus grande fiabilité des identifications des peptides [41]. Ceci s'applique aussi à la mesure des fragments en MS/MS à haute précision [38, 42, 43].

Un haut pouvoir de résolution est nécessaire pour atteindre une précision suffisante et attribuer une masse exacte. Le besoin d'une instrumentation haute résolution est indispensable pour la détermination précise et exacte des masses de massifs isotopiques en abondance significativement

différente et de gamme de rapports m/z.

Les spectromètres de masse capables d'atteindre des précisions de l'ordre de quelques ppm avec une haute résolution lors d'une expérience classique en protéomique se limitent à l'heure actuelle aux appareils à Transformée de Fourier (FTMS) [44, 45] et aux temps de vol (TOF) [46]. La trappe ionique offre quelques caractéristiques intéressantes : une sensibilité très élevée et la possibilité de réaliser avec un seul analyseur des études structurales assez complètes. Les trappes linéaires font partie des spectromètres de masse en tandem les plus populaires utilisés dans les études protéomiques notamment en raison de leur grande vitesse de balayage et de leur faible coût : elles sont capables d'acquérir des milliers de spectres MS/MS durant une séparation chromatographique d'une heure ; cependant, leur précision de mesure de masse est généralement faible [47]. La sensibilité est en général estimée globalement pour l'association d'un type de source d'ions et d'analyseur. Actuellement il est facile de détecter des quantités de l'ordre de la femtomole ( $10^{-15}$  mole) avec tous ces instruments.

## L'analyse en spectrométrie de masse tandem

La spectrométrie en mode MS/MS introduit une méthode d'activation qui provoque une fragmentation des ions. Les fragments se forment par rupture préférentielle de certaines liaisons ; l'analyse de ces fragments permet de déterminer certains éléments de structure. Dans le cas des peptides, les ruptures qui se rompent préférentiellement sont les liaisons peptidiques.

### Spectrométrie de masse tandem (MS/MS)

#### Principes généraux

La spectrométrie de masse en tandem (ou MS/MS) combine en général deux analyseurs. En MS/MS les ions formés dans la source entrent dans le premier analyseur  $MS^1$ . Parmi ces ions, un ion précurseur est sélectionné ( $MS^1$  est utilisé comme filtre d'ions). Seuls les ions ayant le rapport m/z sélectionné pourront sortir de  $MS^1$ . Ils entrent alors dans une cellule de collision où l'énergie interne des ions est augmentée pour les activer, de manière à provoquer leur décomposition. Les ions précurseurs activés se dissocient en ions fragments, en rompant certaines liaisons chimiques dans la molécule. Les rapports m/z de ces fragments sont alors étudiés dans le second analyseur  $MS^2$ . La formation de ces ions fragments suit des règles précises. Les groupes chimiques ont leurs caractéristiques propres de fragmentation, et le profil d'un spectre MS/MS donne en général des informations complètes et fiables sur la structure du composé précurseur.

La fragmentation par dissociation induite par collision (CID) est la méthode de fragmentation des peptides la plus couramment utilisée [48]. La fragmentation a lieu lorsque les peptides entrent en collision avec des molécules de gaz neutre (Argon, Hélium, Néon). Puisque la dissociation intervient lorsque l'énergie de collision est supérieure à l'énergie d'activation, on parle aussi de dissociation activée par collision (CAD). La fragmentation à basse énergie (10-100 eV) intervient dans les quadripôles et les trappes ioniques, ce qui résulte en pertes de neutres et en ruptures des liaisons

peptidiques [49]. La fragmentation à haute énergie (quelques keV) avec les configurations de type MALDI-TOF/TOF engendre la rupture d'autres liaisons du squelette peptidique et des chaînes latérales des AA [50].

En protéomique top-down certains modes de fragmentations spécifiques sont aussi utilisés dans la cellule FT. Il s'agit de l'ECD (« Electron Capture Dissociation ») qui est une fragmentation induite par l'énergie de recombinaison après absorption d'un électron par les ions multichargés, et de l'IR-MPD (« InfraRed Multiphoton Dissociation ») qui est une fragmentation par accumulation de l'énergie par rayonnement IR. Ces deux méthodes de fragmentation permettent aussi l'étude de MPTs (phosphorylation, glycosylation par exemple) en protéomique bottom-up.

### **Analyseurs TOF/TOF**

Les appareils de type MALDI-TOF/TOF permettent conjointement sur une même préparation de réaliser une cartographie peptidique, en mode MALDI-TOF simple, et des analyses structurales en MS/MS pour obtenir des informations de séquence en fragmentant ces précurseurs, en mode MALDI-TOF/TOF. Ce type d'instrument présente une très haute sensibilité, une bonne précision de masse et permet de travailler à haut débit car il est automatisable. Son avantage majeur est l'absence d'une limitation en gamme de masse théorique. Ils sont particulièrement adaptés au mode d'ionisation MALDI qui forme des ions faiblement chargés et de haut rapport m/z.

### **Trappe ionique**

La trappe ionique a l'avantage de pouvoir fonctionner en mode  $MS^n$ , en utilisant un seul analyseur. Il est possible en sélectionnant correctement les valeurs de tension appliquée de ne conserver dans la trappe que les ions de m/z donné que l'on veut fragmenter. L'énergie cinétique des ions est augmentée en superposant à la radiofréquence appliquée à la trappe, une tension de fréquence égale à la fréquence propre des ions, et d'amplitude faible, de manière à ne pas les expulser de la trappe. Les fragments obtenus sont ensuite détectés. Une fois le spectre MS/MS obtenu, il est possible de sélectionner un des fragments mis en évidence et de le traiter comme nouveau précurseur afin d'avoir des informations structurales sur cette espèce : on obtient alors un spectre MS/MS/MS ou  $MS^3$ . En théorie il est possible d'itérer cette expérience pour réaliser des spectres  $MS^n$ , bien que les rendements de chaque étape limitent en pratique cette possibilité à  $n=4$  en protéomique.

### **Analyseur à résonance cyclotronique et à transformée de Fourier (FT-ICR)**

La spectrométrie de masse à résonance cyclotronique et Transformée de Fourier (FT-ICR, « Fourier Transform Ion Cyclotron Resonance ») fut développée en 1974 par Comisarow et Marshall. Cette technique trouve des applications dans des domaines variés, de la chimie analytique [44-51] à la chimie physique des réactions ion/molécule. De récents développements technologiques de cet analyseur ont permis de l'adapter aux sources d'ionisation douces comme l'ESI et le MALDI, et en ont fait un outil extrêmement puissant pour l'analyse de macrobiomolécules [52]. Il faut par exemple citer l'analyse de l'hémoglobine contenue dans un unique globule rouge [53].

Dans un champ magnétique, la trajectoire des ions est incurvée. L'ion peut ainsi être « piégé » sur une trajectoire circulaire : c'est le principe du cyclotron. La fréquence de rotation de l'ion ou fréquence cyclotronique peut alors s'exprimer :

$$v_c = \frac{eB_0}{2\pi m/z} = \frac{\lambda B_0}{m/z}$$

Avec  $v_c$  la fréquence en Hz,  $B_0$  le champ en tesla,  $m$  la masse en kg,  $z$  le nombre de charges de l'ion et  $e$  la charge de l'électron ( $1,6 \cdot 10^{-19}$  C). Si  $m$  est exprimé en Da alors  $\lambda$  est une constante de valeur  $1,535611 \times 10^7$  C.Da.s<sup>-1</sup>. A champ constant uniforme, et à rapport  $m/z$  constant, cette fréquence est donc constante, et indépendante de la vitesse initiale.

La transformée de Fourier est une opération mathématique qui permet de décomposer un signal en une superposition de signaux harmoniques de fréquences distinctes, c'est-à-dire de passer du domaine temporel aux domaines des fréquences. En FTICR MS un courant induit est enregistré dans le temps ; il est considéré provenir de la résonance d'ions de divers  $m/z$  et donc de diverses fréquences de résonance associées. Le signal est donc la superposition de divers signaux sinusoïdaux de ces fréquences. Il est en revanche modifié par l'imperfection du système (atténuation par manque de vide parfait par exemple). C'est pour cela que la cellule FT doit être sous vide (environ  $10^{-10}$  mbar). En première approximation, il est donc possible de remonter directement aux fréquences ICR des ions et donc à leur  $m/z$ .

### **Analyseur à trappe ionique électrostatique à transformée de Fourier (Orbitrap®)**

Depuis son introduction comme deuxième analyseur d'un instrument hybride commercial au 53<sup>ème</sup> ASMS (Annual Meeting on Mass Spectrometry and Allied Topics) à San Antonio en Juin 2005, l'Orbitrap® est devenu un instrument de choix dans de nombreuses applications protéomiques. Le piégeage orbital fut mis en place pour la première fois en 1923 par Kingdon pour capturer les ions en les faisant tourner dans l'orbite d'une électrode centrale [54]. En 2000, Makarov a décrit l'Orbitrap® comme un nouvel analyseur de MS dérivé du piégeage orbital [45]. L'instrument fonctionne sur le principe du piégeage des ions dans un champ électrostatique (pas d'aimant, pas de champ radio fréquence). C'est le premier analyseur de principe réellement nouveau depuis la commercialisation du quadripôle dans les années 80.

Le piège consiste en une électrode externe semblable à un baril et en une électrode centrale de forme fuselée le long de l'axe. Les électrodes axiales symétriques de l'Orbitrap® génèrent un potentiel électrostatique « quadralogarithmique ». Selon l'axe  $z$ , le mouvement est une simple oscillation de pulsation.  $\dot{u}_z = \sqrt{k \frac{e}{m/z}}$  Cette oscillation suit donc

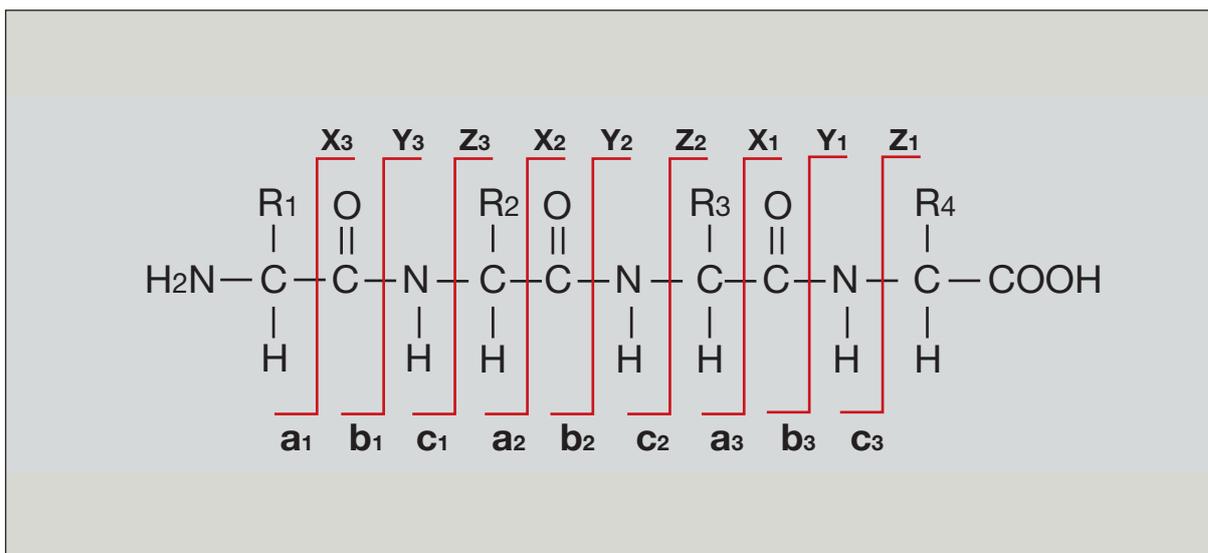
une fréquence qui est inversement proportionnelle à la racine carrée du rapport  $m/z$  de l'ion analysé. Ce rapport  $m/z$  peut ainsi être calculé par transformée de Fourier inverse à partir de l'enregistrement du signal. Les performances de l'Orbitrap® rivalisent avec le FT-ICR. En revanche le temps d'analyse nécessaire à résolution égale est en défaveur de l'Orbitrap®.

## Séquençage des peptides en MS/MS

### Nomenclature

Les travaux pionniers décrivant le séquençage de peptides par MS datent des années 1980 [55]. Un peptide est un copolymère de résidus AA, reliés entre eux par des liaisons amides -CO-NH- ; il possède une extrémité N-terminale (groupe -NH<sub>2</sub>) et une extrémité C-terminale (groupe -COOH).

**Figure 5 : Schéma d'un térapeptide. Rappel des fragments potentiellement formés à partir d'un peptide protoné, par ruptures le long du squelette peptidique.**



Les AA sont les unités structurales de base des protéines. Un acide  $\alpha$ -aminé est constitué d'une fonction amine, d'un groupe carboxyle et d'un groupe caractéristique variable R<sub>n</sub> formant ainsi un répertoire de 20 AA. La masse moléculaire moyenne d'un AA est de 110 Da environ.

**Tableau 1 : Liste et nomenclature des résidus dérivés des 20 acides aminés naturels.**

ACIDE AMINÉ	CODE À 3 LETTRES	CODE À 1 LETTRE	MASSE MOYENNE	MASSE MONOISOTOPIQUE
Alanine	Ala	A	71,0788	71,03711
Arginine	Arg	R	156,1876	156,10111
Asparagine	Asn	N	114,1039	114,04293
Acide aspartique	Asp	D	115,0886	115,02694
Cystéine	Cys	C	103,1448	103,00919
Glutamine	Gln	Q	128,1308	128,05858
Glycine	Gly	G	57,0520	57,02146
Acide glutamique	Glu	E	129,1155	129,04259
Histidine	His	H	137,1412	137,05891

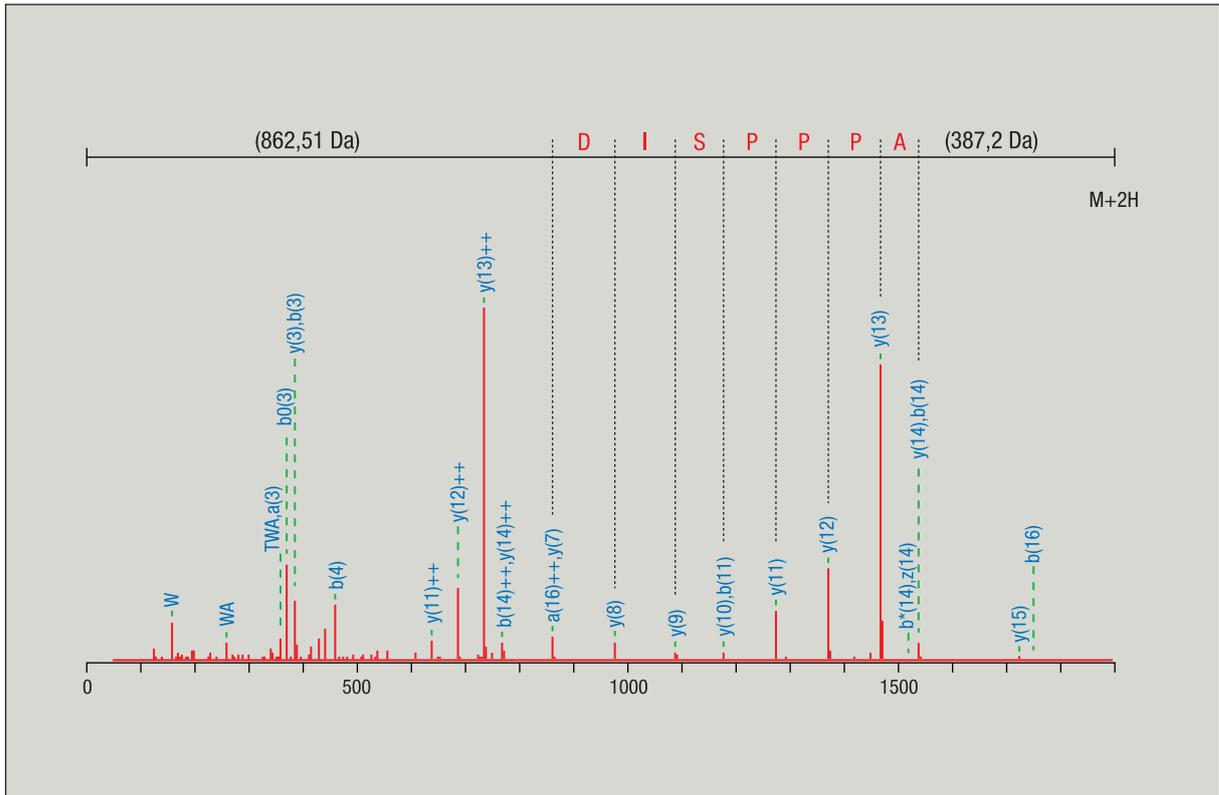
Isoleucine	Ile	I	113,1595	113,08406
Leucine	Leu	L	113,1595	113,08406
Lysine	Lys	K	128,1742	128,09496
Méthionine	Met	M	131,1986	131,04049
Phénylalanine	Phe	F	147,1766	147,06841
Proline	Pro	P	97,1167	97,05276
Sérine	Ser	S	87,082	87,03203
Thréonine	Thr	T	101,1051	101,0468
Tryptophane	Trp	W	186,2133	186,07931
Tyrosine	Tyr	Y	163,1760	163,06333
Valine	Val	V	99,1326	99,06841

Des informations structurales sur les peptides peuvent être obtenues par dissociation des espèces ionisées après activation. En spectrométrie de masse, les peptides se fragmentent majoritairement en rompant les liaisons peptidiques. Roepstorff et Fohlman ont introduit une nomenclature (B, Y, etc), pour représenter conceptuellement les ions formés [56]. Johnson et al ont complété et modifié cette nomenclature (b, y, etc) [50].

La [figure 5](#) rappelle les différents fragments résultant de ruptures le long du squelette peptidique. L'indice accolé en bas à droite de chaque lettre indique le nombre de résidus contenus dans le fragment. Plusieurs mécanismes ont été proposés pour rendre compte des ions fragments observés à partir de précurseurs protonés. Parmi eux, le « modèle du proton mobile » a été vérifié expérimentalement et peut être expliqué par des calculs de mécanique quantique [57-59]. Ce modèle propose la protonation aléatoire des groupes amides du squelette peptidique. Pendant l'ionisation, des espèces chargées sont formées, dans lesquelles les protons sont initialement localisés sur les fonctions basiques (amines), comme c'est le cas en phase liquide. Dans la phase gazeuse, ces protons se délocalisent le long des liaisons peptidiques car l'affinité protonique des groupes amides devient comparable à celle des groupes amines [60]. Lorsque l'azote impliqué dans la liaison peptidique est protoné, la liaison peptidique est affaiblie, et il y a rupture de cette liaison, produisant les ions fragments de types b et y.

## Exemple d'interprétation de spectre MS/MS

**Figure 6 :** Spectre MS/MS issu du précurseur à  $m/z = 963,51$  ( $2+$ ) et correspondant à la séquence VTWAPPP-SIDLTNFLVR de la fibronectine. La série de trois prolines donne lieu aux fragments majoritaires  $y_{11}$  à  $y_{13}$ , qui apparaissent à la fois sous formes mono- et bi-chargées. Le fragment  $y_{13}$  est intense, alors qu' $y_{11}$  et  $y_{12}$  le sont beaucoup moins. On peut aussi noter l'immonium intense à 159 Da, révélant la présence d'un résidu W dans le peptide. Les ions immonium sont caractéristiques des résidus indépendamment de leur position dans la séquence.



## Différents logiciels développés

Différents programmes de recherche dans les banques MS/MS ont été développés (tableau 2) pour les approches de type bottom-up. En revanche très peu d'outils sont disponibles pour les approches de type top-down et les interprétations sont encore réalisées manuellement.

En pratique, les logiciels prennent chaque spectre expérimental en entrée et lui attribuent un score en le comparant aux profils théoriques de fragmentation des peptides de la banque de données. Le nombre de peptides candidats est réduit en fonction des critères spécifiés par l'utilisateur comme la tolérance de masse, l'enzyme de digestion, les modifications chimiques ou post-traductionnelles autorisées. En sortie les logiciels génèrent une liste de séquences peptidiques classées selon leur score de corrélation avec le spectre expérimental. Ce score mesure le degré de similitude entre le spectre expérimental et le spectre théorique et sert de premier paramètre pour discriminer les identifications correctes de celles qui sont fausses.

Tableau 2 : Outils disponibles pour l'interprétation des spectres MS/MS (d'après [61]).

PROGRAMME	RÉF.	SITE WEB
<b>Outils de recherche dans les banques :</b>		
SEQUEST	[62]	<a href="http://www.thermo.com">http://www.thermo.com</a>
Mascot	[63]	<a href="http://matrixscience.com">http://matrixscience.com</a>
ProteinProspector	[37]	<a href="http://prospector.ucsf.edu">http://prospector.ucsf.edu</a>
ProbiD	[64]	<a href="http://tools.proteomecenter.org/wiki/index.php?title=Software:ProbiD">http://tools.proteomecenter.org/wiki/index.php?title=Software:ProbiD</a>
Tandem	[65]	<a href="http://www.thegpm.org">http://www.thegpm.org</a>
SpectrumMill		<a href="http://www.chem.agilent.com">http://www.chem.agilent.com</a>
Phenyx	[66]	<a href="http://www.phenyx-ms.com">http://www.phenyx-ms.com</a>
OMSSA	[67]	<a href="http://pubchem.ncbi.nlm.nih.gov/omssa">http://pubchem.ncbi.nlm.nih.gov/omssa</a>
VEMS	[68]	<a href="http://personal.cicbiogune.es/rmatthiesen">http://personal.cicbiogune.es/rmatthiesen</a>
MyriMatch	[69]	<a href="http://www.mc.vanderbilt.edu/msrc/bioinformatics">http://www.mc.vanderbilt.edu/msrc/bioinformatics</a>
<b>Outils de recherche dans les banques de spectres :</b>		
SpectraST	[70]	<a href="http://www.peptideatlas.org/spectrast">http://www.peptideatlas.org/spectrast</a>
X ! P3	[71]	<a href="http://p3.thegpm.org/tandem/ppp.html">http://p3.thegpm.org/tandem/ppp.html</a>
BibliSpec	[72]	<a href="http://proteome.gs.washington.edu/bibliospec">http://proteome.gs.washington.edu/bibliospec</a>
<b>Outils de séquençage <i>de novo</i> :</b>		
Lutefisk	[73]	<a href="http://www.hairyfatguy.com/lutefisk">http://www.hairyfatguy.com/lutefisk</a>
PepNovo	[74]	<a href="http://peptide.ucsd.edu/pepnovo.py">http://peptide.ucsd.edu/pepnovo.py</a>
PEAKS	[75]	<a href="http://www.bioinformaticssolutions.com">http://www.bioinformaticssolutions.com</a>
Sequit		<a href="http://www.proteomefactory.com">http://www.proteomefactory.com</a>
<b>Approches hybrides ou par séquence tag :</b>		
GutenTag	[76]	<a href="http://fields.scripps.edu/GutenTag">http://fields.scripps.edu/GutenTag</a>
Inspect	[77]	<a href="http://peptide.ucsd.edu/inspect.html">http://peptide.ucsd.edu/inspect.html</a>
Popitam	[78]	<a href="http://www.expasy.org/tools/popitam">http://www.expasy.org/tools/popitam</a>
Paragon	[79]	<a href="https://licensing.appliedbiosystems.com/download/ProteinPilot/2.0/">https://licensing.appliedbiosystems.com/download/ProteinPilot/2.0/</a>

Certaines approches sont basées à la fois sur l'analyse *de novo* et la recherche dans les banques de données. La première étape consiste en la déduction de courts *tags* peptidiques issus des spectres MS/MS, suivie par une recherche dans les banques en mode *error tolerant*. Cela signifie notamment que la recherche autorise une (ou plusieurs) erreur(s) entre la séquence peptidique réelle

ayant produit le spectre expérimental et la séquence dans la banque. Ce type d'approche a d'abord été introduit par le groupe de Mann [4]. La limitation des recherches aux seuls peptides contenant le *tag* (ou les *tags*) permet de réduire considérablement la durée de recherche dans les banques. Ces approches hybrides présentent également un fort potentiel dans l'analyse systématique des peptides modifiés de manière artificielle ou post-traductionnelle. En effet si tous les types de modifications sont autorisés sur n'importe quel site, cela amène à une explosion du nombre des combinaisons possibles, ce qui est peu compatible avec les recherches dans les banques de données. En revanche, l'utilisation de tags de séquences permet de réduire la taille de la banque et donc rend les recherches réalistes. Ainsi la **figure 6** présente un tag de la forme 852.61(DISPPPA)387.2 en y pour un précurseur  $m/z = 963,51$  (doublement chargé).

## Conclusion

---

Ce rapide survol des techniques de MS appliquées à l'analyse protéomique montre qu'il s'agit d'une discipline relativement récente, qui implique une approche intégrative avec des compétences dans divers domaines. Ainsi la préparation des échantillons repose sur une expertise en biologie pour la compréhension des systèmes étudiés, en biochimie pour la préparation des extraits protéiques, en techniques séparatives et en spectrométrie de masse. L'interprétation des données nécessite des développements en bio-informatique et en traitement du signal.

Si les approches de type bottom-up sont bien développées et sont utilisées en routine dans de nombreux laboratoires, il n'en est pas de même pour les approches top-down. Pourtant ces deux approches se révèlent être très complémentaires. L'étude des peptides permet de retrouver les séquences primaires des protéines d'intérêt dans les banques de données, mais l'étude sur la protéine entière permet d'obtenir des informations essentielles sur l'état de maturation de la protéine fonctionnelle et de différencier des isoformes proches. Ainsi l'étude des cascades de signalisation ou de complexes moléculaires peut être traitée comme une étude de mélange complexe en stratégie bottom-up, puis le suivi moléculaire et cinétique des processus de signalisation peut être réalisé sur les protéines entières en stratégie top-down.

## Bibliographie

---

1. Corthals G.L., Wasinger V.C., Hochstrasser D.F., Sanchez J.C. The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis*, 2000, 21(6), 1104-1115.
2. Rabilloud T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics*, 2002, 2(1), 3-10.

3. Hamdan M., Righetti P.G. Assessment of protein expression by means of 2-D gel electrophoresis with and without mass spectrometry. *Mass Spectrom Rev.* 2003, 22(4), 272-84.
4. Mann M. and Wilm M., Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem*, 1994, 66(24), p. 4390-9.
5. Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik.* 1975, 26(3), 231-243.
6. Scheele G.A. Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins. *J Biol Chem.* 1975, 250(14), 5375-5385.
7. O'Farrell P.H. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.* 1975, 250(10), 4007-4021.
8. Pedersen S.K., Harry J.L., Sebastian L., Baker J., Traini M.D., McCarthy J.T., Manoharan A., Wilkins M.R., Gooley A.A., Righetti P.G., Packer N.H., Williams K.L., Herbert B.R. Unseen proteome: mining below the tip of the iceberg to find low abundance and membrane proteins. *J Proteome Res.* 2003, 2(3), 303-311.
9. Santoni V., Molloy M., Rabilloud T. Membrane proteins and proteomics : un amour impossible? *Electrophoresis.* 2000, 21(6), 1054-1070.
10. Luche S., Santoni V., Rabilloud T. Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics.* 2003, 3(3), 249-253.
11. Wacker M., Herrmann M., Klose J. Identification of membrane proteins in 2DE patterns. in 51st Conference on Mass Spectrometry and Allied Topics Proceedings, Poster TPW430 ; Montreal, Quebec, Canada, June 8-12, 2003.
12. Wildgruber R., Reil G., Drews O., Parlar H., Gorg A. Web-based two-dimensional database of *Saccharomyces cerevisiae* proteins using immobilized pH gradients from pH 6 to pH 12 and matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Proteomics.* 2002, 2(6), 727-732.
13. Gygi S.P., Corthals G.L., Zhang Y., Rochon Y., Aebersold R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A.* 2000, 97(17), 9390-9395.
14. Gatlin C.L., Kleemann G.R., Hays L.G., Link A.J., Yates J.R. 3rd. Protein identification at the low femtomole level from silver-stained gels using a new fritless electrospray interface for liquid chromatography-microspray and nanospray mass spectrometry. *Anal Biochem.* 1998, 263(1), 93-101.
15. Gottschalk A., Neubauer G., Banroques J., Mann M., Luhrmann R., Fabrizio P. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snrNP. *EMBO J.* 1999, 18(16), 4535-4548.
16. Bell A.W., Ward M.A., Blackstock W.P., Freeman H.N., Choudhary J.S., Lewis A.P., Chotai D., Fazel A., Gushue J.N., Paiement J., Palcy S., Chevet E., Lafreniere-Roula M., Solari R., Thomas D.Y., Rowley A., Bergeron J.J. Proteomics characterization of abundant Golgi membrane proteins. *J Biol Chem.* 2001, 276(7), 5152-5165.
17. Rout M.P., Aitchison J.D., Suprapto A., Hjertaas K., Zhao Y., Chait B.T. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol.* 2000, 148(4), 635-651.
18. Pflieger D., Chabane S., Gaillard O., Bernard B.A., Ducoroy P., Rossier J., Vinh J. Comparative proteomic analysis of extracellular matrix proteins secreted by two types of skin fibroblasts. *Proteomics.* 2006, 6(21), 5868-5879.
19. Simpson R.J., Connolly L.M., Eddes J.S., Pereira J.J., Moritz R.L., Reid G.E. Proteomic analysis of the human colon carcinoma cell line (LIM 1215): development of a membrane protein database. *Electrophoresis.* 2000, 21(9), 1707-1732.
20. Haynes P.A., Fripp N., Aebersold R. Identification of gel-separated proteins by liquid chromatography-electrospray tandem mass spectrometry: comparison of methods and their limitations. *Electrophoresis.* 1998, 19(6), 939-945.
21. Speicher K.D., Kolbas O., Harper S., Speicher D.W. Systematic Analysis of Peptide Recoveries from In-Gel Digestions for Protein Identifications in Proteome Studies. *J Biomol Tech.* 2000, 11, 74-86.
22. Shen Y., Smith R.D. Proteomics based on high-efficiency capillary separations. *Electrophoresis.* 2002, 23(18), 3106-3124.
23. Mönch W., Dehnen W. High-performance liquid chromatography of peptides. *J Chromatogr.* 1977, 140, 260-262.
24. O'Hare M.J., Nice E.C. Hydrophobic high-performance liquid chromatography of hormonal polypeptides and proteins on alkylsilane-bonded silica. *J Chromatogr.* 1979, 171, 209-226.

25. Hancock W.S., Bishop C.A., Prestidge R.L., Harding D.R., Hearn M.T. Reversed-phase, high-pressure liquid chromatography of peptides and proteins with ion-pairing reagents. *Science*. 1978, 200(4346), 1168-1170.
26. Purnell J.H. *Gas Chromatography*. Wiley, New York. 1962.
27. Whitehouse C.M., Dreyer R.N., Masamichi Y., Fenn J.B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal Chem*. 1985, 57(3), 675-679.
28. Wahl J.H., Goodlett D.R., Udseth H.R., Smith R.D. Use of small-diameter capillaries for increasing peptide and protein detection sensitivity in capillary electrophoresis-mass spectrometry. *Electrophoresis*. 1993, 14(5-6), 448-457.
29. Constantopoulos T.L., Jackson G.S., Enke C.G. Effects of salt concentration on analyte response using electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom*. 1999, 10(7), 625-34.
30. Redeker, V., Toullec J.Y., Vinh J., Rossier J., Soyez D. Combination of peptide profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and immunodetection on single glands or cells. *Anal Chem*, 1998, 70(9), 1805-1811.
31. Cole R.B. Some tenets pertaining to electrospray ionization mass spectrometry. *J Mass Spectrom*. 2000, 35(7), 763-772.
32. Cech N.B., Enke C.G. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom Rev*. 2001, 20(6), 362-387.
33. Bodnar W.M., Blackburn R.K., Krise J.M., Moseley M.A. Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J Am Soc Mass Spectrom*. 2003, 14(9), 971-979.
34. Baldwin M.A., *Mass spectrometers for the analysis of biomolecules*. *Methods Enzymol*, 2005, 402, 3-48.
35. Domon, B. and Aebersold R. *Mass spectrometry and protein analysis*. *Science*. 2006, 312(5771), 212-217.
36. Zubarev R. and Mann M. On the proper use of mass accuracy in proteomics. *Mol Cell Proteomics*. 2007, 6(3), 377-381.
37. Clauser K.R., Baker P., Burlingame A.L. Role of accurate mass measurement ( $\pm 10$  ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem*, 1999, 71(14), 2871-2882.
38. Conrads T.P., Anderson G.A., Veenstra T.D., Pasa-Tolic L., Smith R.D. Utility of accurate mass tags for proteome-wide protein identification. *Anal Chem*. 2000, 72(14), 3349-3354.
39. Sleno L., Volmer D.A., Marshall A.G., Assigning product ions from complex MS/MS spectra: the importance of mass uncertainty and resolving power. *J Am Soc Mass Spectrom*. 2005, 16(2), 183-98.
40. Takach E.J., Hines W.M., Patterson D.H., Juhasz P., Falick A.M., Vestal M.L., Martin S.A. Accurate mass measurements using MALDI-TOF with delayed extraction. *J Protein Chem*. 1997, 16(5), 363-369.
41. Zubarev R.A., Hakansson P., Sundqvist B. Accuracy requirements for peptide characterization by monoisotopic molecular mass measurements. *Anal Chem*. 1996, 68(22), 4060-4063.
42. Haas W., Faherty B.K., Gerber S.A., Elias J.E., Beausoleil S.A., Bakalarski C.E., Li X., Villen J., Gygi S.P., Optimization and use of peptide mass measurement accuracy in shotgun proteomics. *Mol Cell Proteomics*. 2006, 5(7), 1326-1337.
43. He F., Emmett M.R., Hakansson K., Hendrickson C.L., Marshall A.G. Theoretical and experimental prospects for protein identification based solely on accurate mass measurement. *J Proteome Res*. 2004, 3(1), 61-67.
44. Comisarow M.B. and Marshall A.G., Theory of Fourier-Transform Ion-Cyclotron Resonance Mass-Spectroscopy .1. Fundamental Equations and Low-Pressure Line-Shape. *J. Chem Phys*. 1976, 64(1), 110-119.
45. Makarov, A., Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem*, 2000, 72(6), 1156-1162.
46. Dawson J. and Guilhaus M. Orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun Mass Spectrom*. 1989, 3(5), 155-159.
47. Gorshkov M.V. and Zubarev R.A. On the accuracy of polypeptide masses measured in a linear ion trap. *Rapid Commun Mass Spectrom*. 2005, 19(24), 3755-3758.

48. Hunt D.F., Buko A.M., Ballard J.M., Shabanowitz J., Giordani A.B., Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. *Biomed Mass Spectrom.* 1981, 8(9), 397-408.
49. Papayannopoulos I.A., The Interpretation of Collision-Induced Dissociation Tandem Mass-Spectra of Peptides. *Mass Spectrom Rev.* 1995, 14(1), 49-73.
50. Johnson R.S., Martin S.A., Biemann K., Stults J.T., Watson J.T., Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal Chem.* 1987, 59(21), 2621-2625.
51. Comisarow M., Marshall A.G. Fourier-Transform Ion-Cyclotron Resonance Spectroscopy. *Chem Phys Lett.* 1974, 25(2), 282-283.
52. Senko M.W., Canterbury J.D., Guan S., Marshall A.G. A high-performance modular data system for Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun Mass Spectrom.* 1996, 10(14), 1839-1844.
53. Hofstadler S.A., Severs J.C., Smith R.D., Swanek F.D., Ewing A.G., Analysis of single cells with capillary electrophoresis electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun Mass Spectrom.* 1996, 10(8), 919-922.
54. Kingdon K.H. A method for the neutralization of electron space charge by positive ionization at very low gas pressures. *Physical Review.* 1923, 21(4), 408-418.
55. Hunt D.F., Yates J.R. 3rd, Shabanowitz J., Winston S., Hauer C.R. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A.* 1986, 83(17), 6233-6237.
56. Roepstorff P., Fohlman Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *J. Biomed Mass Spectrom.* 1984, 11(11), 601.
57. Dongre A.R., Somogyi A., Wysocki V.H. Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *J Mass Spectrom.* 1996, 31(4), 339-350.
58. Johnson R.S., Krylov D., Walsh K.A. Proton mobility within an electrospray peptide ion. *J. Mass Spectrom.* 1995, (30), 386-387.
59. Csonka I.P., Paizs B., Lendvai G., Suhai S. Proton mobility in protonated peptides: a joint molecular orbital and RRKM study. *Rapid Commun Mass Spectrom.* 2000, 14(6), 417-431.
60. Tang X.J., Thibault P., Boyd R.K. Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation. *Anal Chem.* 1993, 65(20), 2825-2834.
61. Nesvizhskii A.I., Vitek O., Aebersold R., Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat Methods.* 2007, 4(10), 787-797.
62. Eng J.K., McCormack A.L., Yates J.R. An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. *J Am Soc Mass Spectrom.* 1994, 5(11), 976-989.
63. Perkins D.N., Pappin D.J., Creasy D.M., Cottrell J.S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis.* 1999, 20(18), 3551-3567.
64. Zhang N., Aebersold R., Schwikowski B. ProbiD : a probabilistic algorithm to identify peptides through sequence database searching using tandem mass spectral data. *Proteomics.* 2002, 2(10), 1406-1412.
65. Craig R. and Beavis R.C., TANDEM : matching proteins with tandem mass spectra. *Bioinformatics.* 2004, 20(9), p. 1466-1467.
66. Colinge J., Masselot A., Giron M., Dessingy T., Magnin J., OLAV : towards high-throughput tandem mass spectrometry data identification. *Proteomics.* 2003, 3(8), 1454-1463.
67. Geer L.Y., Markey S.P., Kowalak J.A., Wagner L., Xu M., Maynard D.M., Yang X., Shi W., Bryant S.H. Open mass spectrometry search algorithm. *J Proteome Res.* 2004, 3(5), 958-964.
68. Matthiesen R., Trelle M.B., Hojrup P., Bunkenborg J., Jensen O.N. VEMS 3.0 : algorithms and computational tools for tandem mass spectrometry based identification of post-translational modifications in proteins. *J Proteome Res.* 2005, 4(6), 2338-2347.

69. Tabb D.L., Fernando C.G., Chambers M.C. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res.* 2007, 6(2), 654-661.
70. Lam H., Deutsch E.W., Eddes J.S., Eng J.K., King N., Stein S.E., Aebersold R. Development and validation of a spectral library searching method for peptide identification from MS/MS. *Proteomics.* 2007, 7(5), 655-667.
71. Craig R., Cortens J.P., Beavis R.C., The use of proteotypic peptide libraries for protein identification. *Rapid Commun Mass Spectrom.* 2005, 19(13), 1844-1850.
72. Frewen B.E., Merrihew G.E., Wu C.C., Noble W.S., MacCoss M.J. Analysis of peptide MS/MS spectra from large-scale proteomics experiments using spectrum libraries. *Anal Chem.* 2006, 78(16), 5678-84.
73. Johnson R.S. and Taylor J.A. Searching sequence databases via de novo peptide sequencing by tandem mass spectrometry. *Mol Biotechnol.* 2002, 22(3), 301-315.
74. Frank A. and Pevzner P. PepNovo: de novo peptide sequencing via probabilistic network modeling. *Anal Chem.* 2005, 77(4), 964-973.
75. Ma B., Zhang K., Hendrie C., Liang C., Li M., Doherty-Kirby A., Lajoie G. PEAKS : powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2003, 17(20), 2337-2342.
76. Tabb D.L., Saraf A., Yates J.R. 3rd, GutenTag : high-throughput sequence tagging via an empirically derived fragmentation model. *Anal Chem.* 2003, 75(23), 6415-6421.
77. Tanner S., Shu H., Frank A., Wang L.C., Zandi E., Mumby M., Pevzner P.A., Bafna V. InsPecT : identification of posttranslationally modified peptides from tandem mass spectra. *Anal Chem.* 2005, 77(14), 4626-4639.
78. Hernandez P., Gras R., Frey J., Appel R.D., Popitam: towards new heuristic strategies to improve protein identification from tandem mass spectrometry data. *Proteomics.* 2003, 3(6), 870-8.
79. Shilov I.V., Seymour S.L., Patel A.A., Loboda A., Tang W.H., Keating S.P., Hunter C.L., Nuwaysir L.M., Schaeffer D.A., The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics.* 2007, 6(9), 1638-1655.



**Exemple d'une chaîne  
d'analyse protéomique :  
de l'électrophorèse  
bidimensionnelle  
à la LC-MS/MS**

**Florence Poirier, Céline Boursier**

# Introduction

---

L'analyse protéomique, qui découle du terme protéome défini dans le Chapitre 1, est une retombée directe du séquençage des génomes [1]. Bien que la génomique puisse donner la liste complète des protéines pouvant être exprimées par un génome, elle ne présume pas, pour une situation donnée, des gènes exprimés dans la cellule et à quels niveaux. En cela, l'analyse transcriptomique, fondée sur l'analyse des ARNm exprimés, étudie d'un point de vue qualitatif et quantitatif l'expression de l'ensemble des ARNm contenus dans la cellule [2]. Cette approche conduit à des informations fonctionnelles à partir d'analyses comparatives entre plusieurs états mais ne permet pas cependant de savoir si la protéine correspondante est exprimée. En effet, le niveau d'expression protéique n'est pas un simple reflet des niveaux d'expression des ARNm et les durées de vie de ces deux entités cellulaires ne sont pas corrélées [3]. De plus, comme vu dans un précédent chapitre, les protéines peuvent subir une série de modifications post-traductionnelles (MPTs) modulant leur activité, et qui ne sont pas identifiables à partir de leur séquence nucléotidique. Par conséquent, seule une analyse protéique directe peut donner une image globale des systèmes bio-moléculaires dans leur complexité. Notons cependant que l'analyse protéomique ne permet pas à elle seule, de déterminer si la régulation de l'expression protéique s'effectue au niveau traductionnel ou transcriptionnel. Afin d'avoir accès à cette information, il est donc souvent nécessaire de combiner les données issues de la protéomique à celles de la transcriptomique.

L'analyse protéomique met donc l'accent sur la description dynamique de la régulation des gènes et ce faisant, sur la régulation moléculaire étudiée au travers de l'expression protéique. Il s'agit, au travers de la comparaison de cartographies protéiques, d'établir une analyse différentielle et sans a priori des modifications survenant lors d'un processus biologique donné. La possibilité de comparer quantitativement différents protéomes (cellules saines/cellules pathologiques, cellules traitées par une substance/cellules non traitées, cellules différenciées/cellules non différenciées) fait donc de cette approche analytique globale un outil puissant pour identifier des protéines impliquées dans les processus physiologiques ou pathologiques. En effet, cette approche conduit non seulement à améliorer notre connaissance sur ces mécanismes, mais également à l'identification de bio-marqueurs d'états pathologiques ainsi qu'à celle de nouvelles cibles thérapeutiques. L'analyse d'un protéome ou, plus simplement, l'identification d'une ou plusieurs protéines associées à un état d'une cellule ou d'un organisme, nécessite donc un ensemble de compétences et d'instruments regroupés dans une « plate-forme » telle que la plate-forme TransProt à la Faculté de Pharmacie de Châtenay-Malabry.

## Présentation générale de la plate-forme Transcriptome/Protéome (TransProt)

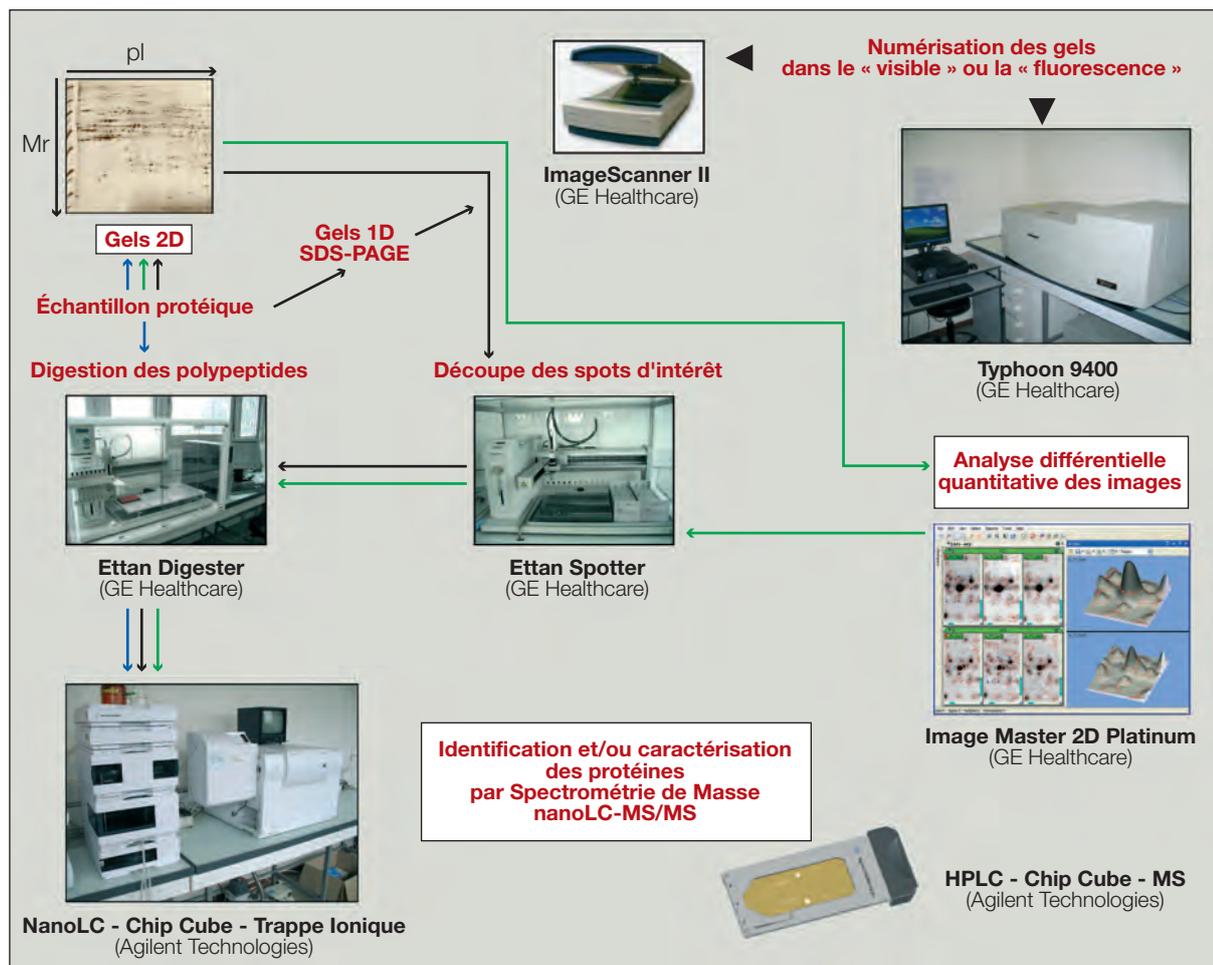
L'IFR-141 (<http://ifr141.ccp.u-psud.fr/>) de la Faculté de Pharmacie de Châtenay-Malabry (Université Paris-Sud 11) regroupe différentes équipes de recherche aux savoir-faire complémentaires,

notamment dans les domaines de la chimie, de l'immunologie, de la biochimie, de la biologie moléculaire, cellulaire et structurale, sans oublier la génétique et la pharmacologie. L'IFR-141 dispose également de différentes plates-formes technologiques proposant une expertise solide dans des disciplines variées (protéomique, transcriptomique, imagerie cellulaire, spectrométrie de masse pour l'analyse de petites molécules, criblage moléculaire, animalerie). La plate-forme Transcriptome/Protéome (TransProt) de l'IFR-141 est située au 1<sup>er</sup> étage de la Tour E1 de la Faculté de Pharmacie. Elle est née en 2008 du regroupement de deux plates-formes : les plates-formes « transcriptome » et « protéome » mises en place respectivement en 2002 et 2005. La mission principale de TransProt est d'offrir à la communauté scientifique des secteurs public et privé des outils technologiques en protéomique et en transcriptomique permettant de mener à bien des projets de recherche sur des modèles biologiques variés (plantes, animaux, micro-organismes, parasites, tissus, cellules, fluides biologiques...). Le rôle du personnel de la plate-forme est multiple : il aide les équipes intéressées dans la mise en œuvre de stratégies pertinentes en transcriptomique et protéomique ; il les forme aux différentes méthodes analytiques disponibles sur la plate-forme et les aide à l'interprétation des résultats. Depuis 2009, la plate-forme TransProt est dotée d'un site web qui regroupe diverses informations concernant le laboratoire, les conditions d'accès et de travail sur la plate-forme, la charte de la plate-forme, ainsi que les formulaires à consulter pour toute nouvelle collaboration (<http://ifr141.cep.u-psud.fr/transcriptome-proteome/index.html>). L'activité de la plate-forme TransProt est actuellement assurée par Claudine Deloménie (Ingénieur de Recherche, INSERM) pour la partie transcriptomique, par Céline Boursier (Ingénieur d'Etudes, Université Paris-Sud 11) et Florence Poirier (Chercheur, INSERM) pour la partie protéomique. Depuis 2008, Céline Boursier est Présidente du Club Jeune de la Société Française d'Electrophorèse et d'Analyse Protéomique (<http://sfeap.free.fr/>) et, depuis 2009, Florence Poirier siège au conseil d'administration de la Fondation Sciences Citoyennes (<http://www.sciencescitoyennes.org/>). Le personnel « protéomique » de TransProt est par ailleurs membre, avec d'autres partenaires, du Réseau UniverSud Paris Protéome (RUSP2) qui rassemble la communauté scientifique de la région Ile-de-France Sud agissant dans le domaine de la protéomique.

## **Les outils de l'analyse protéomique sur TransProt**

Le laboratoire de protéomique de la plate-forme TransProt dispose de toutes les approches analytiques nécessaires à l'analyse qualitative et quantitative de protéomes variés (figure 1). Ces approches font appel à deux principales techniques utilisées en analyse protéomique : l'électrophorèse bidimensionnelle (2-DE) et la spectrométrie de masse (MS). Une chaîne robotisée de préparation des échantillons (découpe des gels, digestion des polypeptides et dépôt sur cible MALDI) permet des études menées à grande échelle.

Figure 1 : Equipement et stratégies analytiques en protéomique sur TransProt.



## L'électrophorèse bidimensionnelle

La 2-DE sur gel de polyacrylamide (2D-PAGE), qui tient compte à la fois de la charge des protéines et de leur masse moléculaire, est encore actuellement la méthode de séparation la mieux adaptée à l'étude du protéome. Ses principes ont été vus dans le Chapitre 1. Elle permet de mettre en évidence dans un même gel plusieurs centaines de protéines exprimées par la cellule y compris leurs différentes isoformes résultant de MPTs (spots). Cependant, l'analyse par 2-DE présente quelques limitations. En effet, dans l'impossibilité d'utiliser des détergents à fort pouvoir solubilisant comme le SDS pour ne pas modifier le pI des protéines de l'échantillon, la séparation de 1<sup>ère</sup> dimension se limite à l'emploi conjoint d'une forte concentration d'agents chaotropes et de détergents non ioniques ou zwitterioniques. En conséquence, les protéines présentant un caractère plus hydrophobe comme les protéines à ancrage membranaire, les protéines nucléaires ou les protéines de très haute masse moléculaire, ont beaucoup plus de mal à se solubiliser convenablement et restent difficilement analysables sur les gels 2D [4]. Une autre limitation de la 2-DE concerne les protéines situées aux pH extrêmes qui ont du mal à focaliser en 1<sup>ère</sup> dimension, et les protéines de très faible abondance qui ne sont généralement pas visibles sur un gel 2D sans méthode d'enrichissement préalable. Enfin, l'automatisation de cette étape n'est à l'heure actuelle toujours pas résolue du fait de la complexité de la méthode. Ces limitations peuvent être en partie contournées grâce au développement d'approches complémentaires faisant appel à la MS [5, 6]. L'adaptation de la MS

à l'analyse qualitative et/ou quantitative des protéines a permis non seulement d'augmenter le débit de l'analyse mais également la sensibilité des méthodes analytiques traditionnelles. Cependant, ces approches d'analyse globale du protéome se heurtent également aux limites imposées par la MS. La première limite provient de la nature biochimique des peptides de l'échantillon et/ou de l'effet de certains composés chimiques [7]. Il en résulte que certains peptides donneront un signal faible et d'autres ne seront jamais analysés. La deuxième limite se situe au niveau des peptides de très faible abondance présents dans l'échantillon [8]. Bien que les spectromètres de masse soient de plus en plus performants en termes de sensibilité et de résolution, ils peuvent néanmoins exclure de l'analyse certains peptides de l'échantillon situés en deçà de ce seuil. Enfin, contrairement aux gels 2D, la quantification différentielle des différentes isoformes protéiques ne peut se réaliser qu'au niveau peptidique, nécessitant une étape d'enrichissement préalable en peptides d'intérêt pouvant induire des biais quantitatifs et/ou des pertes d'information liés à la faible spécificité de la méthode d'enrichissement [9-11]. Les analyses par 2-DE et MS restent donc à l'heure actuelle deux approches analytiques complémentaires pour l'étude qualitative et quantitative des protéomes.

## **La spectrométrie de masse**

L'étape finale de l'analyse protéomique est l'identification des polypeptides d'intérêt séparés sur le gel de polyacrylamide. Lorsqu'il s'agit d'identifier sur un gel 2D un nombre de protéines sur lesquelles on ne possède aucune information a priori, l'établissement de cartes peptidiques massiques et/ou l'établissement par micro-séquençage de courtes séquences peptidiques terminales ou internes (tags) permet la plupart du temps l'identification de la protéine. Dans ce domaine, la MS suivie d'une analyse bio-informatique (interrogation des banques de données protéiques, analyse des MPTs) est devenue une approche incontournable [12].

Nous pouvons distinguer deux modes d'ionisation par MS : le mode MALDI (désorption/ionisation laser assistée par matrice) et le mode ESI (ionisation électrospray). En mode MALDI, l'échantillon protéique est co-cristallisé avec une matrice chimique puis bombardé par un laser. Les peptides désorbés et ionisés sont analysés le plus souvent par un analyseur en temps de vol (TOF). La MS MALDI-TOF est la plus largement utilisée pour l'identification des protéines par « empreinte peptidique massique » (« peptide mass fingerprinting » ou PMF). Dans cette approche, la protéine d'intérêt séparée sur gel, est enzymatiquement ou chimiquement clivée dans le gel, puis la masse de chaque peptide généré est mesurée par MS. Ces masses peptidiques expérimentales sont ensuite comparées aux masses peptidiques théoriques obtenues par clivage virtuel de toutes les séquences protéiques répertoriées dans les banques de données. À l'issue de cette recherche, l'analyse pertinente des résultats permet d'aboutir à l'identification du polypeptide. Cette stratégie basée sur la mesure des masses, permet également de rechercher des MPTs portées par les peptides, par comparaison de leur masse expérimentale avec leur masse théorique non modifiée. En mode ESI, l'échantillon protéique en solution est pulvérisé en fines gouttelettes par action d'un champ électrique intense. Le plus souvent, ces peptides sont séparés en amont de la source ESI par chromatographie liquide (LC) à nano-débits (0,2 µl/min) dotée de colonnes analytiques adaptées (75 µm de diamètre interne). Les peptides ionisés sont analysés le plus souvent en mode MS/MS. Un peptide sélectionné est fragmenté par collision avec un gaz résiduel puis les masses des

fragments obtenus sont ensuite mesurées. La fragmentation ayant lieu au niveau des liaisons peptidiques entre chaque acide aminé, la différence de masse entre chaque fragment permet de remonter à une séquence peptidique partielle. Cette séquence en acides aminés qui est comparée aux séquences protéiques répertoriées dans les banques de données, permet d'aboutir à l'identification du polypeptide. Cette stratégie permet non seulement de rechercher des MPTs portées par ces peptides, mais également de localiser ces modifications au sein de leur séquence.

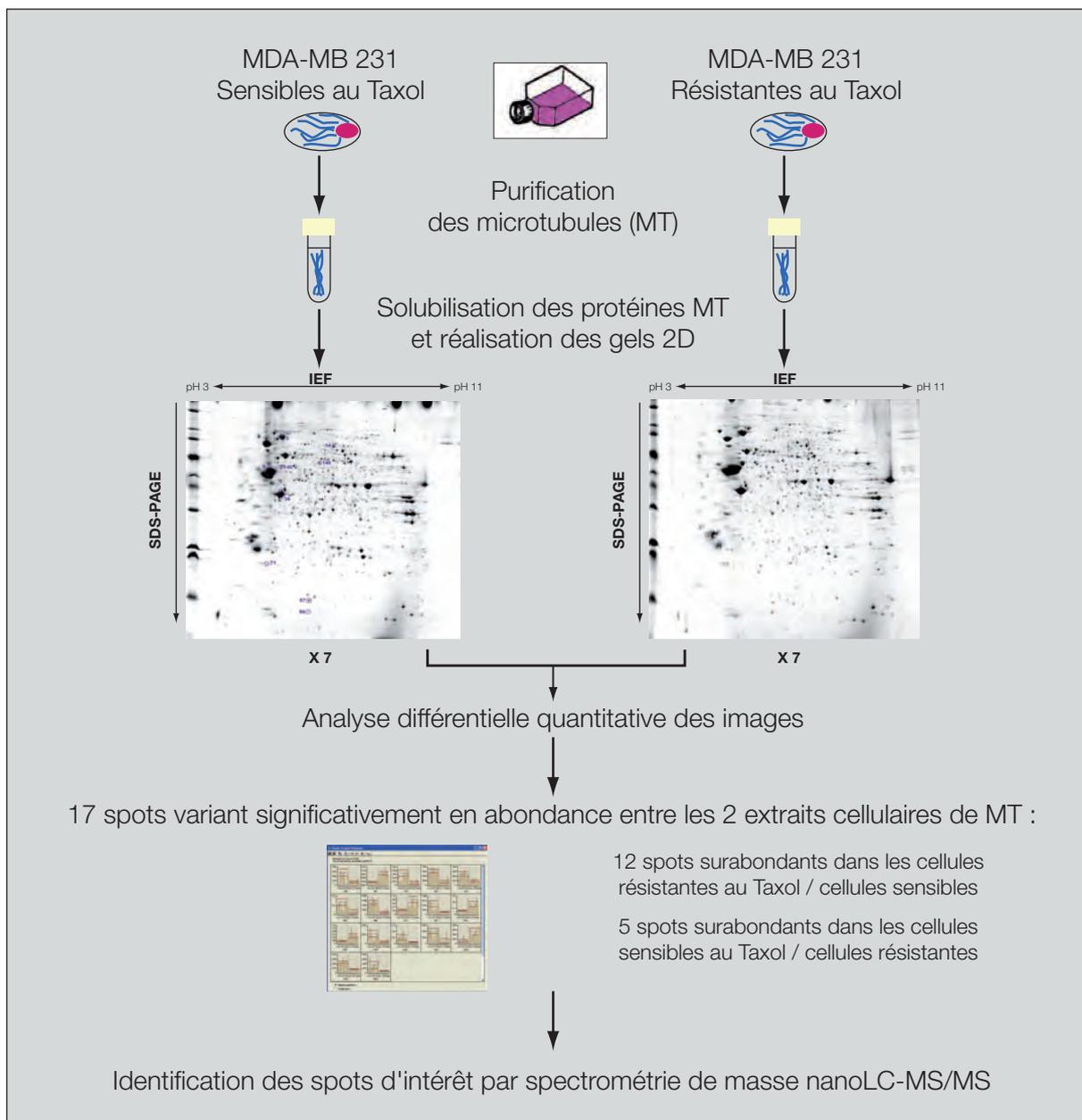
## Exemple d'une analyse protéomique sur TransProt

Trois principales stratégies en protéomique sont appliquées sur la plate-forme ([figure 1](#)) : l'analyse différentielle quantitative des gels 2D suivie de l'identification des spots d'intérêt par MS, l'identification systématique de protéines séparées par gels 1D ou 2D, et la caractérisation de protéines cibles par MS. Nous ne détaillerons ici qu'un seul exemple, celui de l'analyse différentielle par 2-DE qui est actuellement l'approche la plus fréquemment utilisée sur TransProt.

### Analyses différentielles par 2-DE

L'analyse protéomique différentielle permet d'étudier les mécanismes moléculaires mis en jeu dans la cellule lorsque celle-ci est soumise à des conditions physiologiques particulières. Classiquement, les profils d'expression protéiques obtenus par 2-DE avant et après exposition des cellules à cette condition sont comparés, et les protéines variant en expression sont identifiées. En recherche pharmacologique par exemple, la connaissance des effets néfastes des médicaments sur l'organisme peut être obtenue en comparant des profils d'expression protéiques de cellules normales exposées ou non à ces substances [13]. Dans la recherche biomédicale, isoler des marqueurs protéiques de maladies présente l'avantage de pouvoir développer des méthodes efficaces en matière de pronostic, diagnostic et de concevoir des thérapies pertinentes adaptées à ces marqueurs [14]. L'analyse protéomique différentielle permet également d'identifier des marqueurs protéiques associés à l'évolution des tumeurs humaines en vue d'établir leur classification [15]. La [figure 2](#) permet d'illustrer les différentes étapes de l'analyse différentielle par 2-DE : la réalisation des gels 2D, l'analyse d'images et l'identification des protéines d'intérêt par MS. Cette figure montre l'exemple d'une analyse comparative effectuée entre des extraits microtubulaires de cellules tumorales mammaires résistantes ou non au Taxol (paclitaxel), un médicament utilisé dans le traitement du cancer du sein. Le but de l'étude était d'identifier des protéines associées aux microtubules qui pourraient empêcher l'action du Taxol et expliquer que certaines cellules tumorales mammaires puissent devenir résistantes à cette chimiothérapie [16].

**Figure 2 : Analyse différentielle quantitative réalisée entre des gels 2D d'extraits microtubulaires de cellules mammaires sensibles ou résistantes au Taxol [16].**



### Réalisation des gels 2D

La bonne réalisation des gels 2D est une étape cruciale pour la réussite de l'analyse différentielle. Il faut non seulement maîtriser la reproductibilité quantitative (quantité de protéines chargée) et qualitative (migration électrophorétique des protéines) de la 2-DE, mais également réaliser un nombre suffisamment important de gels par condition expérimentale afin de rendre l'analyse quantitative plus robuste d'un point de vue statistique. La reproductibilité quantitative des gels 2D dépend essentiellement de l'habileté du manipulateur lorsque celui-ci prélève les échantillons protéiques et les charge dans les gels de première dimension. La reproductibilité qualitative de la 2-DE est généralement assurée par l'usage d'appareils dédiés qui permettent la migration simultanée

d'un grand nombre de gels en 1<sup>ère</sup> et 2<sup>nde</sup> dimensions. En revanche, le dernier point peut être plus délicat à réaliser. Idéalement, pour chaque condition expérimentale, nous préconisons sur la plateforme de réaliser au moins 6 gels 2D comportant des répliquas de l'échantillon protéique obtenu à partir d'expériences indépendantes. Ainsi les gels 2D, pour une condition donnée, sont un reflet plus exact de la variabilité intra-expérience (répétabilité) et inter-expériences (reproductibilité au sens strict). Ceci est bien entendu réalisable lorsqu'il s'agit d'échantillons biologiques dont la concentration protéique n'est pas une limitation (cellules en culture, sérums etc.). En revanche, tout devient plus compliqué lorsqu'il s'agit d'échantillons précieux et/ou de faible concentration protéique ne permettant pas de réaliser un grand nombre de gels. Dans ce dernier cas, une des solutions consiste à se tourner vers des méthodes de détection des protéines les plus sensibles possibles, de manière à diminuer la quantité protéique par gel tout en ne dégradant pas la qualité du profil 2D (nombre de spots).

La détection des polypeptides sur le gel de deuxième dimension est en effet une étape importante conditionnant les résultats de l'analyse protéomique [17]. C'est sur la coloration que repose l'analyse d'images et par conséquent l'analyse différentielle. Il est souhaitable de disposer d'une méthode de coloration la plus sensible et la plus linéaire possible de manière à accéder aux protéines minoritairement exprimées, consommer le moins d'échantillon biologique et augmenter la précision de l'analyse quantitative. La méthode fréquemment employée est celle utilisant le nitrate d'argent mais d'autres méthodes de détection plus linéaires fondées sur la radioactivité, la fluorescence ou l'emploi de colorants peuvent également être appliquées [18-20]. La plate-forme TransProt disposant d'un scanner capable de lire la fluorescence (figure 1), c'est ce mode de coloration qui est le plus souvent utilisé du fait de sa grande sensibilité et de sa large gamme linéaire [20]. Le marquage des protéines par des fluorochromes peut être réalisé de façon globale sur tout l'échantillon protéique (cyanines, SyproRuby®...), ou de façon plus ciblée si l'on souhaite étudier certaines catégories de protéines comme les phosphoprotéines et les glycoprotéines, ou encore des protéines pouvant être la cible de molécules chimiques [21-23].

### **Analyse d'images des gels 2D**

Étant donné que le simple examen d'un gel 2D après coloration peut faire apparaître plusieurs centaines de taches polypeptidiques (spots), l'analyse différentielle qui consiste en la détermination des différences qualitatives et quantitatives des protéines de deux échantillons biologiques différents n'est rendue possible que par analyse d'images à l'aide de logiciels spécialisés comme ImageMaster 2-D Platinum® (GE Healthcare/Amersham Biosciences) disponible sur TransProt. Les logiciels de dernière génération comme le logiciel Delta2D® (Decodon) ou le logiciel Progenesis SameSpots® (Nonlinear Dynamics) ont grandement optimisé la durée de réalisation de l'analyse d'images. En effet, ces logiciels permettent non seulement un alignement numérique des images des gels 2D mais également la réalisation d'une image de fusion entre les différents profils obtenus dont le profil commun pourra ensuite être repositionné sur chacune des images 2D. Ces avantages arrivent à réduire de façon considérable les erreurs de correspondance des spots d'une image à l'autre qui contraignaient l'utilisateur à rectifier manuellement ces erreurs pour ne pas s'exposer à des biais d'interprétation lors de l'analyse différentielle quantitative [17].

La robustesse de l'analyse d'images dépendra du nombre et de la qualité des gels réalisés, mais également des performances du logiciel de traitement des images et de la stratégie d'analyse choisie. L'analyse d'images comprend généralement trois principales étapes : la préparation des images 2D, l'appariement automatique des spots entre les images (« matching ») et l'analyse statistique des données. Pour la première étape, le logiciel d'analyse d'images détecte et délimite chaque spot suivant des paramètres densitométriques choisis. Cette détection automatique s'effectuant sur le niveau de densité optique des spots, le logiciel peut être confronté à certaines limites. En effet, il détecte en plus des taches polypeptidiques, certains artéfacts de coloration entrant dans la gamme densitométrique choisie et n'est pas capable, dans certains cas, de discriminer deux taches adjacentes ayant des niveaux de gris trop proches. Afin de palier à cet inconvénient, il est souvent nécessaire, par intervention manuelle sur l'image, d'éliminer les artéfacts de coloration et de redessiner les contours de certains spots. La préparation des images 2D peut également inclure une étape de normalisation permettant d'atténuer, au sein d'un même groupe de gels (mêmes conditions expérimentales), des différences d'intensité des spots essentiellement liées à la coloration. Pour la seconde étape, le logiciel va apparier de façon automatique chaque spot de même nature entre les différents gels. Cette étape cruciale de l'analyse, appelée matching des spots, peut être de qualité variable selon le type de logiciel utilisé. Les erreurs rencontrées s'expliquent généralement par l'existence de légères différences de migration des spots entre les différents gels qui ne sont pas rectifiées de façon numérique par certains logiciels lors de l'alignement des images. L'existence de spots mal focalisés, de zones riches en spots ou de spots absents sur certains gels peuvent également accentuer ces erreurs de matching, qui peuvent être évitées de façon importante par l'utilisation de logiciels de dernière génération, ou bien rectifiées manuellement sur les images 2D [17]. Les erreurs d'appariement des spots dues aux légères différences de migration inter-gels peuvent également être évitées en utilisant la technologie DIGE® (« Differential Gel Electrophoresis », GE Healthcare/Amersham Biosciences), qui consiste à faire migrer, au sein d'un même gel, différents échantillons protéiques après les avoir couplés covalentement à différents fluorochromes [23]. Les échantillons ayant été soumis aux mêmes contraintes électrophorétiques dans les gels de 1<sup>ère</sup> et 2<sup>nde</sup> dimensions, les images 2D acquises aux longueurs d'ondes de chaque fluorochrome pourront se superposer parfaitement sans qu'il y ait de doute sur la position des spots de même nature entre les différentes images. Bien que cette approche soit très efficace lorsqu'il s'agit de comparer quantitativement deux échantillons au sein d'un même gel, le problème de matching des images inter-gels se posera inévitablement si l'on souhaite comparer les images obtenues à partir de plusieurs gels DIGE. Dans ce cas, la bonne correspondance des spots inter-gels DIGE reposera uniquement sur les performances du logiciel d'analyse d'images. L'analyse statistique des données est la dernière étape de l'analyse d'images. Elle permet de mettre en évidence sur les gels 2D, les spots polypeptidiques dont l'abondance varie de façon significative dans l'un des deux échantillons. La stratégie de l'analyse statistique n'est pas universelle et peut varier en fonction des conditions expérimentales (nombre de gels, nombre de spots à analyser...) mais également des outils statistiques disponibles (via le logiciel ou après exportation des données brutes).

### **Identification des spots d'intérêt par spectrométrie de masse (MS)**

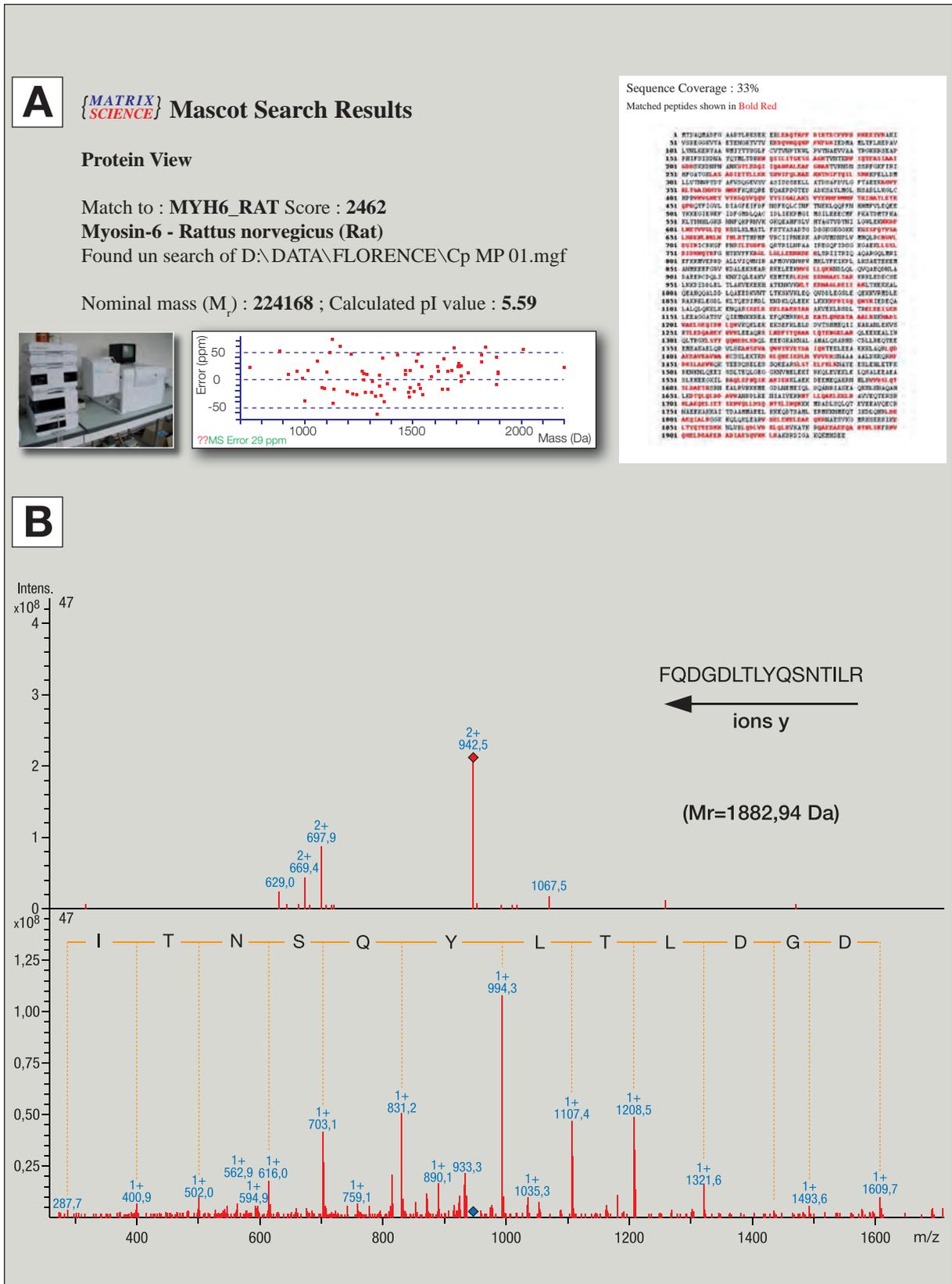
Pour leur identification, les polypeptides d'intérêt mis en évidence par analyse d'images, sont excisés des gels 2D, puis soumis à une protéolyse en utilisant généralement de la trypsine (figure 1). Cette étape peut être réalisée de façon manuelle ou semi-automatique en utilisant les robots dédiés de la plate-forme. Les peptides sont ensuite analysés par MS et les polypeptides correspondants sont identifiés après consultation des banques de données protéiques telle que la banque Swiss-Prot (<http://www.expasy.ch/>).

La plate-forme TransProt est équipée d'un spectromètre de masse ESI de type quadripôle à trappe ionique couplé au système micro-fluidique nanoLC Chip Cube. La technologie HPLC-Chip/MS (Agilent Technologies) est un système de chromatographie liquide miniaturisé où les colonnes et les vannes sont intégrées dans une puce composée d'un film semi-rigide de faible surface (environ 6 x 2,5 cm), avec, en sortie de colonne analytique (75 µm x 45 mm ou 150 mm), une pointe intégrée faisant office d'aiguille nanométrique. L'utilisation de cette puce requiert un adaptateur (HPLC-Chip Cube®) permettant sa connexion avec la pompe HPLC et le spectromètre de masse quadripôle à trappe ionique. La configuration puce apporte trois principaux avantages : une réduction de la taille des connexions entre les colonnes et en sortie de colonne analytique, une rapidité accrue des analyses, et une facilité de réglage du spray à la nano-source de la MS. En revanche, lorsque l'une des composantes est obstruée, toute la « chip » doit être remplacée avec le coût que cela engendre. Avec une telle performance chromatographique, la plate-forme TransProt s'est équipée d'un spectromètre de masse quadripôle à trappe ionique de haute capacité capable d'atteindre des vitesses de balayage (jusqu'à 26000 m/z par seconde) permettant la réalisation de plusieurs cycles d'acquisition MS et MS/MS pendant la courte durée d'élution des pics chromatographiques. Deux exemples d'analyses MS/MS réalisées sur la plate-forme sont présentés dans la figure 3. La plate-forme TransProt a également accès à un spectromètre de masse de type MALDI-TOF grâce à un accord établi avec Pierre Le Maréchal et Christophe Marchand de l'Institut de Biochimie et Biophysique Moléculaire et Cellulaire (UMR-CNRS-8619, Université Paris-Sud 11, Orsay).

Figure 3 : Exemples d'analyses par nanoLC-MS/MS (HPLC - Chip Cube - Trappe ionique).

A) Identification de la Myosine-6 via le logiciel Mascot (<http://www.matrixscience.com>)

B) Spectres MS (en haut) et MS/MS (en bas) du peptide tryptique di-charge (942.5 m/z) de la Glutathion S-transférase P.



## Exploitation des données 2D

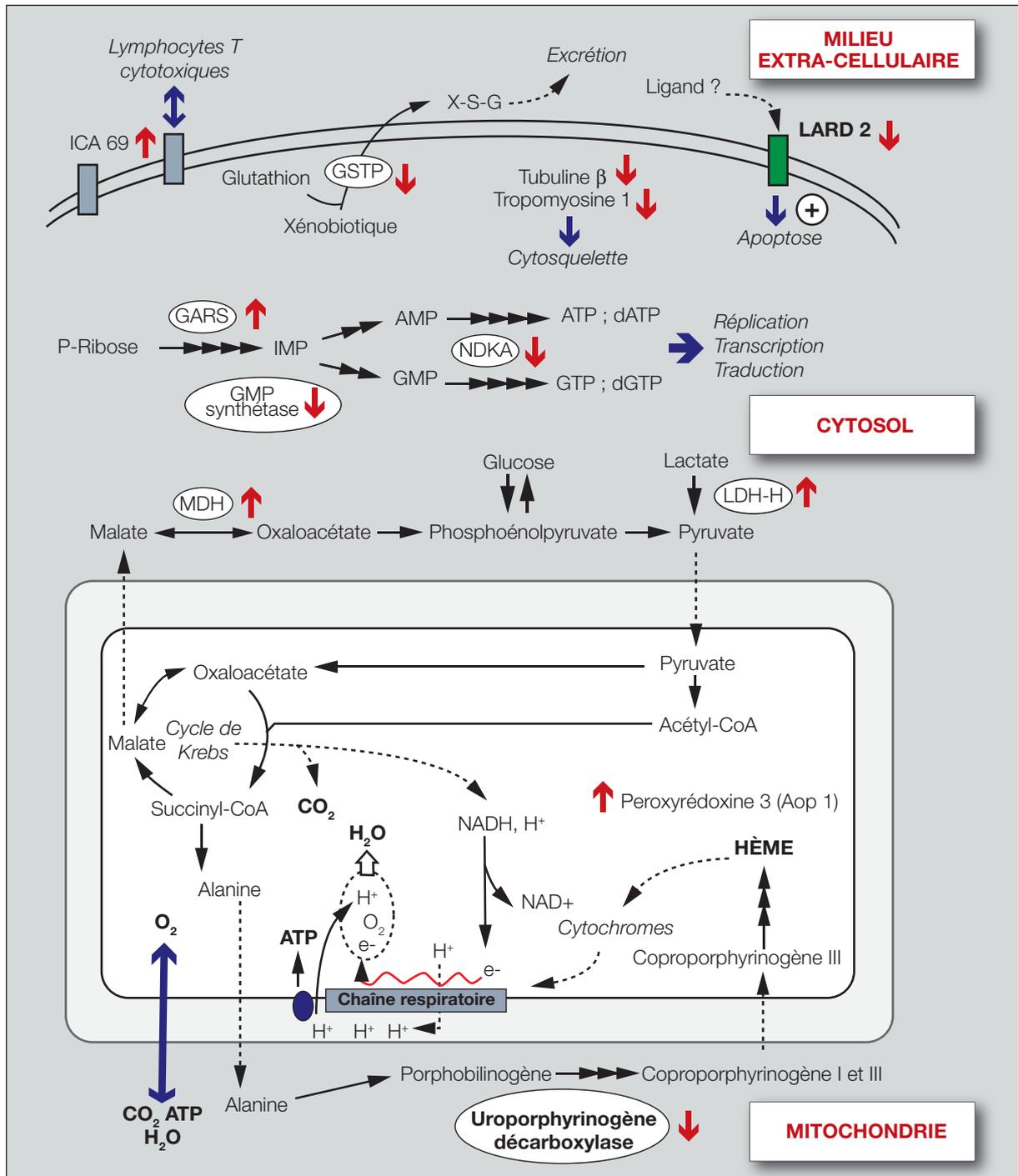
### *La protéomique fonctionnelle*

La suite logique donnée à l'analyse différentielle rentre dans le cadre de la protéomique fonctionnelle. Après identification, les protéines d'intérêt sont replacées dans les voies métaboliques et fonctionnelles dans lesquelles elles interviennent afin d'interpréter les conséquences de ces variations d'expression tant sur le plan moléculaire et cellulaire qu'au niveau de l'organisme concerné (figure 4). Les données de la littérature ont déjà montré l'utilité de ce type d'approche. Chez la levure par exemple, son adaptation au stress oxydant est notamment caractérisée par une forte sur-expression de protéines anti-oxydantes et de protéines du choc thermique, ainsi que par l'induction et la répression de certaines voies métaboliques [25]. Sur la lignée humaine de lymphome de Burkitt DG75, la 5'-azacytidine utilisée dans le traitement des hémopathies malignes affecte l'expression de protéines impliquées dans l'apoptose, dans le maintien du cytosquelette, ou encore dans certaines voies métaboliques comme la synthèse des nucléotides puriques [26, 27]. Plus récemment, des variations d'expression de protéines impliquées notamment dans le stress oxydant, le métabolisme des sucres, la compartimentation vacuolaire et la dégradation des protéines, ont été mises en évidence chez *Saccharomyces cerevisiae* (levure) cultivée en présence d'un herbicide de type sélectif [28].

### *Les cartographies protéiques*

L'ensemble des identifications polypeptidiques obtenues par analyse protéomique conduit généralement à la construction de cartographies protéiques 2D de référence correspondant à des modèles biologiques particuliers. Ces cartographies peuvent ainsi servir à identifier rapidement sur des gels 2D des polypeptides d'intérêt localisés au cours d'études protéomiques par analyse différentielle [29]. A cette fin, différentes cartographies 2D sont élaborées sur la plate-forme TransProt comme celle réalisée pour la lignée tumorale mammaire MDA-MB-231 où 142 spots ont déjà pu être identifiés [16]. Ces identifications étant réalisées sur la totalité des spots 2D détectés, le recours aux robots dédiés de la plate-forme est devenu indispensable pour la réalisation de ces cartographies. Depuis quelques années, des banques de données de cartographies protéiques 2D comme SWISS 2D-PAGE (<http://www.expasy.ch/ch2d/>) sont accessibles sur Internet et donnent accès à différents protéomes correspondant à des types cellulaires et des conditions physiologiques particuliers, comme le protéome des cellules endothéliales de la veine ombilicale de nouveau-nés (<http://www.huvec.com>) [30, 31]. Ces banques permettent ainsi de mettre à disposition de la communauté scientifique de nombreux résultats expérimentaux exploitables [32]. Dans cette contribution, les cartographies 2D réalisées sur TransProt peuvent être consultées via le site web de la plate-forme (<http://ifr141.cep.u-psud.fr/transcriptome-proteome/index.html>).

**Figure 4 : Voies métaboliques et fonctionnelles altérées au cours du traitement par la 5'-azacytidine dans la lignée de lymphome de Burkitt DG75 [26, 27]. Protéines quantitativement augmentées ou diminuées par le traitement.**



# Bibliographie

---

1. Wilkins MR., Sanchez JC., Gooley AA., Appel RD., Humphery-Smith I., Hochstrasser DF. and Williams KL. Progress with proteome projects : why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 1996 ; 13 : 19-50.
2. Schena M. Genome analysis with gene expression microarrays. *Bioassays* 1996 ; 18 : 427-431.
3. Anderson L. and Seilhamer J. A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 1997 ; 18 : 533-537.
4. Rabilloud T. Two-dimensional gel electrophoresis in proteomics : Old, old fashioned, but it still climbs up the mountains. *Proteomics* 2002 ; 2 : 3-10.
5. Link AJ., Eng J., Schieltz DM., Carmack E., Mize GJ., Morris DR., Garvik BM. and Yates JR. 3rd. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* 1999 ; 17 : 676-682.
6. Chen CH. Review of a current role of mass spectrometry for proteome research. *Anal. Chim. Acta* 2008 ; 624 : 16-36.
7. Annesley TM. Ion suppression in mass spectrometry. *Clin. Chem.* 2003 ; 49 :1041-1044.
8. Delahunty C. and Yates JR. 3rd. Protein identification using 2D-LC-MS/MS. *Methods* 2005 ; 35 : 248-255.
9. Niya-Lazar A., Saito-Benz H. and White FM. Quantitative phosphoproteomics by mass spectrometry : Past, present, and future. *Proteomics* 2008, 8 : 4433-4443.
10. Pan S., Aebersold R., Chen R., Rush J., Goodlett DR., McIntosh MW., Zhang J. and Brentnall TA. Mass spectrometry based targeted protein quantification: methods and applications. *J. Proteome Res.* 2009 ; 8 : 787-797.
11. Tsai CF., Wang YT., Chen YR., Lai CY., Lin PY., Pan KT., Chen JY., Khoo KH. and Chen YJ. Immobilized metal affinity chromatography revisited : pH/acid control toward high selectivity in phosphoproteomics. *J. Proteome Res.* 2008 ; 7 : 4058-4069.
12. Lahm HW. and Langen H. Mass spectrometry : a tool for the identification of proteins separated by gels. *Electrophoresis* 2000 ; 21 : 2105-2114.
13. Steiner S., Aicher L., Raymackers J., Meheus L., Esquer-Blasco R., Anderson NL., Cordier A. Cyclosporine A. decreases the protein level of the calcium-binding protein calbindin-D 28kDa in rat kidney. *Biochem. Pharmacol.* 1996 ; 51 : 253-258.
14. Vercoutter-Edouart AS., Lemoine J., Le Bourhis X., Louis H., Boilly B., Nurcombe V., Révillion F., Peyrat JP. and Hondermarck H. Proteomic analysis reveals that 14-3-3 sigma is down-regulated in human breast cancer cells. *Cancer Res.* 2001 ; 61 : 76-80.
15. Celis JE., Celis P., Ostergaard M., Basse B., Lauridsen JB., Ratz G., Rasmussen HH., Orntoft TF, Hein B., Wolf H. and Celis A. Proteomics and immunohistochemistry define some of the steps involved in the squamous differentiation of the bladder transitional epithelium: a novel strategy for identifying metaplastic lesions. *Cancer Res.* 1999 ; 59 : 3003-3009.
16. Froidevaux L., Poirier F., Crépin R., Boursier C., Poüs C., Baillet A. and Baudin B. Comparative proteomic analysis of the microtubule-associated proteins involved in Taxol resistance in the mammalian breast cancer cell line MDA-MB 231. Présentation par affiche. Congrès de la Société Française d'Electrophorèse et d'Analyse Protéomique, Tours, 6-8 octobre 2008.
17. Voss T. and Haberl P. Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels : consequences for comprehensive data analysis. *Electrophoresis* 2000 ; 21 : 3345-3350.
18. Rabilloud T. Mechanisms of protein silver staining in polyacrylamide gels : A 10-year synthesis. *Electrophoresis* 1990 ; 11 : 785-794.

19. Chevallet M., Luche S. and Rabilloud T. Silver staining of proteins in polyacrylamide gels. *Nat. Protocols* 2006 ; 1 : 1852-1858.
20. Patton WF. A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis* 2000 ; 21 : 1123-1144.
21. Kulathingal J., Ko LW., Cusack B. and Yen SH. Proteomic profiling of phosphoproteins and glycoproteins responsive to wild-type alpha-synuclein accumulation and aggregation. *Biochim. Biophys. Acta* 2009 ; 1794 : 211-224.
22. Cazorla O., Lucas A., Poirier F., Lacampagne A. and Lezoualc'h F. The cAMP binding protein Epac regulates cardiac myofilament function. *Proc. Natl. Acad. Sci. U.S.A.* 2009 ; 106 : 14144-14149.
23. Derbré S., Gil S., Taverna M., Boursier C., Nicolas V., Demey-Thomas E., Vinh J., Susin SA., Hocquemiller R. and Poupon E. Highly cytotoxic and neurotoxic acetogenins of the Annonaceae: New putative biological targets of squamocin detected by activity-based protein profiling. *Bioorg. Med. Chem. Lett.* 2008 ; 18 : 5741-5744.
24. Timms JF. and Cramer R. Difference gel electrophoresis. *Proteomics* 2008 ; 8 : 4886-4897.
25. Godon C., Lagniel G., Lee J., Buhler JM., Kieffer S., Perrot M., Boucherie H., Toledano MB. and Labarre J. The H<sub>2</sub>O<sub>2</sub> stimulon in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 1998 ; 273 : 22480-22489.
26. Poirier F., Pontet M., Labas V., Le Caër JP., Sghiouar-Imam N., Raphaël M., Caron M. and Joubert-Caron R. Two-dimensional database of a Burkitt lymphoma cell line (DG 75) proteins : protein pattern changes following treatment with 5'-azacytidine. *Electrophoresis* 2001 ; 22 : 1867-1877.
27. Poirier F., Joubert-Caron R., Labas V. and Caron M. Proteomic analysis of a lymphoma-derived cell line (DG75) following treatment with a demethylating drug : modification of membrane-associated proteins. *Proteomics* 2003 ; 3 : 1028-1036.
28. Teixeira MC., Santos PM., Fernandes AR. and Sá-Correia I. A proteome analysis of the yeast response to the herbicide 2,4-dichlorophenoxyacetic acid. *Proteomics* 2005 ; 5 : 1889-1901.
29. Toda T., Sugimoto M., Omori A., Matsuzaki T., Furuichi Y. and Kimura N. Proteomic analysis of Epstein-Barr virus-transformed human B-lymphoblastoid cell lines before and after immortalization. *Electrophoresis* 2000 ; 21 : 1814-1822.
30. Bruneel A., Labas V., Mailloux A., Sharma S., Vinh J., Vaubourdolle M. and Baudin B. Proteomic study of human umbilical vein endothelial cells in culture. *Proteomics* 2003 ; 3 : 714-723.
31. Bruneel A., Labas V., Mailloux A., Sharma S., Royer N., Vinh J., Pernet P., Vaubourdolle M. and Baudin B. Proteomics of human umbilical vein endothelial cells applied to etoposide-induced apoptosis. *Proteomics* 2005 ; 5 : 3876-3884.
32. Poirier F., Imam N., Pontet M., Joubert-Caron R. and Caron M. The PBP (protein biochemistry and proteomics) two-dimensional electrophoresis database. *J. Chromatogr. B* 2001 ; 753 : 23-28.



# **Pré- et post-analytique en protéomique clinique**

**Pr Sylvain Lehmann**

**CHAPITRE IV**

# Introduction

---

En biologie médicale, les outils analytiques sont en plein développement grâce aux apports, entre autres, des nanotechnologies, de la microfluidique et aux évolutions récentes de la spectrométrie de masse, en particulier associée à la protéomique clinique [1]. Cette dernière vise à rechercher, valider, puis utiliser de nouveaux bio-marqueurs pour le diagnostic ou le suivi de patients atteints de pathologies de toutes les spécialités médicales (voir chapitres précédents) [2]. Ceci nous conduit, en tant que biologiste, à revisiter les notions classiques des traitements pré- et post-analytiques qui conditionnent la qualité et la validité des résultats obtenus. Nous sommes en effet confrontés à une nouvelle façon d'aborder la recherche et la biologie clinique dans laquelle des dizaines, voire des milliers de paramètres, connus ou inconnus, peuvent être mesurés en parallèle. Ces données fournies en grandes quantités nécessitent des traitements informatiques spécifiques, d'autant plus qu'elles peuvent être associées à celles d'autres disciplines, comme la génomique, afin d'avoir une vision d'une situation pathologique à la fois de plus en plus globale, mais aussi individuelle et pertinente [3].

Dans ce chapitre, nous passerons tout d'abord en revue les éléments importants pour la mise en place de programmes de recherche en protéomique clinique ce qui nous conduira aux meilleures conditions et recommandations pré-analytiques pour les laboratoires d'analyse biologique.

## Choix de la question clinique et du protocole

Les projets de protéomique clinique sont souvent de grande envergure et leur mise en place peut durer plusieurs années en impliquant un grand nombre d'intervenants et de ressources. Le choix et la pertinence de la question clinique sont donc particulièrement cruciaux pour ne pas mener de recherches en vain. Dans ces projets qui impliquent des prélèvements biologiques humains, le « protocole » (de soins courants ou de recherche) qui leur est associé va obligatoirement passer devant un Comité de Protection des Personnes (CPP) en vertu de la loi du 6 août 2004 relative à la bioéthique. Ce filtre est particulièrement important car il va évaluer l'intérêt du protocole selon les principes de la bioéthique ; ces principes trouvent leur origine dans le procès de Nuremberg à la fin de la deuxième guerre mondiale [4]. Ils comprennent principalement les 5 éléments suivants concernant l'expérimentation humaine :

- 1) le consentement de la personne est une obligation ;
- 2) le sujet doit subir un minimum de contraintes ;
- 3) la qualification des expérimentateurs doit être établie ;
- 4) le protocole doit être scientifique ;
- 5) les risques encourus par le sujet doivent être inférieurs aux bénéfices escomptés.

Le CPP peut ainsi évaluer la partie scientifique du projet et sa pertinence par rapport à l'état de l'art dans le domaine ; il s'agit d'une question cruciale que l'on doit se poser à chaque mise en place d'un protocole. Deux éléments sont déterminants pour, par exemple, un projet de recherche sur un nouveau bio-marqueur dans une pathologie humaine. Premièrement, il faut que ce bio-marqueur puisse apporter un réel « plus » par rapport aux moyens cliniques et para-cliniques déjà existants. Il peut s'agir d'un marqueur diagnostique plus précoce ou plus précis, du remplacement par ce marqueur d'une approche coûteuse ou invasive, de la capacité à évaluer par ce marqueur un pronostic ou une réponse à un traitement. S'il existe déjà une possibilité de répondre à la question de façon satisfaisante (avec des seuils de sensibilité et de spécificité de 80 % ou plus), établir un nouveau projet pour encore augmenter ces seuils semble un pari osé. L'autre élément à prendre en compte est la possibilité, grâce aux moyens cliniques et para-cliniques déjà existants, de définir des groupes phénotypiques caractérisés permettant la recherche de nouveaux bio-marqueurs. Il s'agit là d'un paradoxe important, car d'un côté on veut être dans une situation où l'on a pas ou peu de moyens de différencier des phénotypes différents (voir élément précédent), mais de l'autre, il faut que l'on puisse en amont définir les groupes aussi précisément que possible, afin d'avoir le plus de probabilité de trouver des marqueurs différentiels. On peut résoudre ce dilemme en formant des groupes un peu différents pour la partie de découverte et la partie validation de l'étude ; ce qui risque cependant d'introduire des biais statistiques. On remarquera que les marqueurs pronostiques ou de choix thérapeutique sont le sujet de nombreux projets actuels. En effet, on est ici dans une situation où des prélèvements, généralement pris au moment du diagnostic, vont servir à la recherche d'une information que l'on ne disposera théoriquement que plusieurs années après (pronostic) et/ou après différents moyens thérapeutiques (chimiothérapie, chirurgie...). Le bénéfice lié à la découverte d'un nouveau bio-marqueur est donc ici très important car il permet d'assurer un traitement et un suivi adapté à chaque patient.

En conclusion de cette partie, il nous faut insister sur l'importance d'une réflexion pluridisciplinaire entre médecins, biologistes, spécialistes concernés (imagerie..) afin de choisir au mieux la question clinique posée au départ du programme.

## **Calcul du nombre de prélèvements en fonction de l'approche protéomique**

Dans les programmes de recherche clinique, une grande attention est portée au choix de la taille des effectifs des groupes de patients à inclure afin de garantir une puissance suffisante de l'étude. Ce choix est classiquement basé sur quatre éléments :

- la variabilité du critère de jugement (exactitude et précision de l'analyse biologique, reproductibilité, sensibilité et spécificité par rapport à la pathologie concernée)
- la différence cliniquement intéressante que l'on veut mettre en évidence (entre deux traitements par exemple)
- le risque  $\alpha$ , dit de première espèce, généralement fixé à 5 %, qui représente le risque de conclure à une différence significative alors qu'en réalité cette différence n'existe pas

- le risque  $\beta$ , dit de deuxième espèce, qui est de ne pas conclure à une différence significative alors qu'en réalité cette différence existe. La valeur  $(1 - \beta)$  est la puissance souhaitée de l'essai et correspond à la probabilité de mettre en évidence une différence significative sachant qu'elle est vraie.

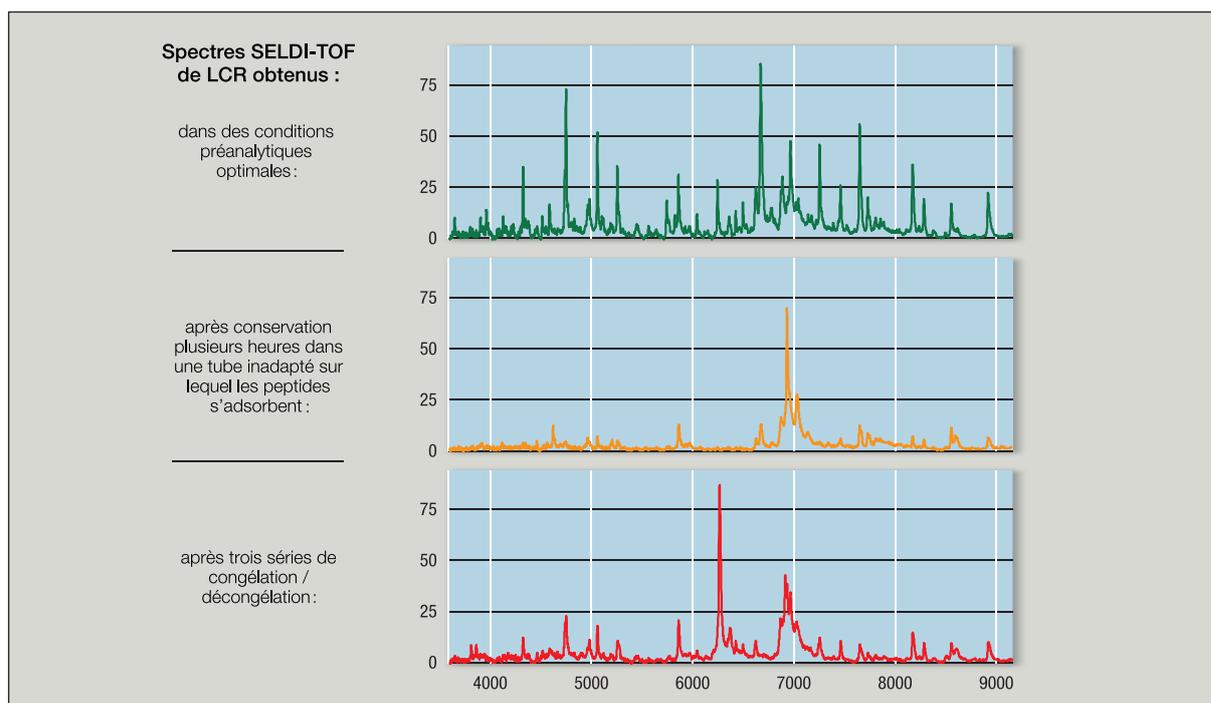
Dans le domaine de la recherche de nouveaux bio-marqueurs on est en fait confronté à une première incertitude sur la variabilité de l'analyse. En effet si on peut connaître la variabilité de l'approche protéomique utilisée, ce qui donne déjà une estimation de la capacité de découverte, on n'a pas encore de valeur sur le bio-marqueur et ses variations dans les populations. La situation est compliquée par le fait que les recherches protéomiques sont par essence multivariées à la fois dans leur phase de découverte (« training ») où plusieurs milliers de variables peuvent être prises en compte, que dans leur phase de validation (« testing ») où il n'y a généralement pas plus d'une dizaine de variables. Pour la phase de découverte, tout comme dans les analyses génomiques sur puces, un ajustement des « p-values » est généralement réalisé en utilisant des méthodes de correction mathématique (Bonferroni, Holm, Hochberg...) ou empiriquement (Bayes) [5]. Cela permet de prendre en compte le taux de faux positif (« False Discovery Rate » ou FDR). Compte tenu de ces incertitudes, la réponse classique dans le domaine de la protéomique clinique à la question cruciale du nombre de patients à inclure est généralement « autant que possible »... à condition comme nous le verrons ci-après que les conditions pré-analytiques de tous les prélèvements inclus soient bien contrôlées. Dans les phases de validation qui permettent d'éliminer beaucoup de faux positifs, on utilise généralement des effectifs par groupe au moins dix fois supérieur au nombre de paramètres pris en compte. Enfin, il est fortement recommandé une validation des bio-marqueurs avec une méthode différente de celle utilisée dans la phase de découverte (passage de la spectrométrie de masse à l'ELISA par exemple). Ces considérations sur le nombre de prélèvements à utiliser sont en fait fortement dépendantes du type d'approche protéomique utilisée. Le débit d'analyse est en effet très variable selon la technologie et souvent inversement proportionnelle à la quantité de données générées. Autrement dit, des approches de type multiplexe (en milieu liquide ou sur puce) ne présentent souvent pas de limitation de débit au niveau analytique [6]. Un grand nombre d'analyses pouvant en effet être réalisées en peu de temps avec des systèmes semi-automatisés et au pré-analytique relativement simple. Un facteur limitant peut être alors le coût de l'analyse, qui peut aller de quelques dizaines à plus de mille euros (puce à protéines) par prélèvement [7]. Par contre, dans une méthode d'analyse comme l'électrophorèse bidimensionnelle, le temps expérimental (1 à 2 jours par petites séries de gels) et la nature de l'information (images des gels en colorimétrie ou en fluorescence) font que cette approche est peu compatible avec de grandes séries de prélèvements (> 100). Enfin, dans les analyses par spectrométrie de masse, il existe des approches haut débit comme avec le SELDI-TOF®, qui réalise un « profil » protéique mais sans identification des candidats [8] (voir Chapitre 1). Les méthodes permettant une exhaustivité et une résolution importante (LC-MS/MS, Orbitrap®, FT-ICR..) sont, quant à elles, difficilement utilisables sur de grands nombres de prélèvements car le temps analytique/post-analytique est très élevé et la quantification des candidats entre les prélèvements reste très délicate à réaliser (utilisation de tags, de peptides rapporteurs..).

## Recommandations pré-analytiques

Les approches de protéomique clinique portent principalement sur les fluides biologique tels que le sang (sérum, plasma), le liquide céphalorachidien (LCR) et les urines. Pour chacun de ces fluides, par leur origine, les quantités disponibles et sa composition protéique nécessitent une attention pré-analytique particulière. Par ailleurs, en raison de la complexité de la composition de ces échantillons, des méthodes de pré-fractionnement (déplétion des protéines majoritaires, fractionnement chromatographique...) sont souvent utilisées pour les approches visant à identifier de nouveaux bio-marqueurs [9]. L'utilisation de méthodes analytiques de plus en plus performantes, comme la spectrométrie de masse, met en première ligne le contrôle du pré-analytique qui assure l'intégrité des échantillons par la standardisation des conditions de prélèvement, de conservation et de stockage. Les conditions pré-analytiques risquent donc d'entraîner des modifications plus ou moins importantes de la composition protéique des prélèvements ce qui peut conduire à des erreurs importantes (figure 1). Le contrôle des conditions pré-analytiques est d'autant plus crucial que l'augmentation du nombre de patients par groupe pour les phases de recherche et de validation implique souvent la nécessité d'effectuer des recrutements multicentriques, voire de travailler en réseau au niveau des différentes plates-formes analytiques. Dans les analyses multiplexes il existe également une multiplication des risques d'erreurs liées aux conditions pré-analytiques avec la possibilité que, sur un panel, certaines analyses soient affectées et d'autres pas. Ceci rend d'autant plus complexe l'interprétation des résultats ou le calcul de score au moyen d'algorithmes multiparamétriques.

**Figure 1 : Exemple de l'impact des conditions pré-analytiques sur une analyse protéomique.**

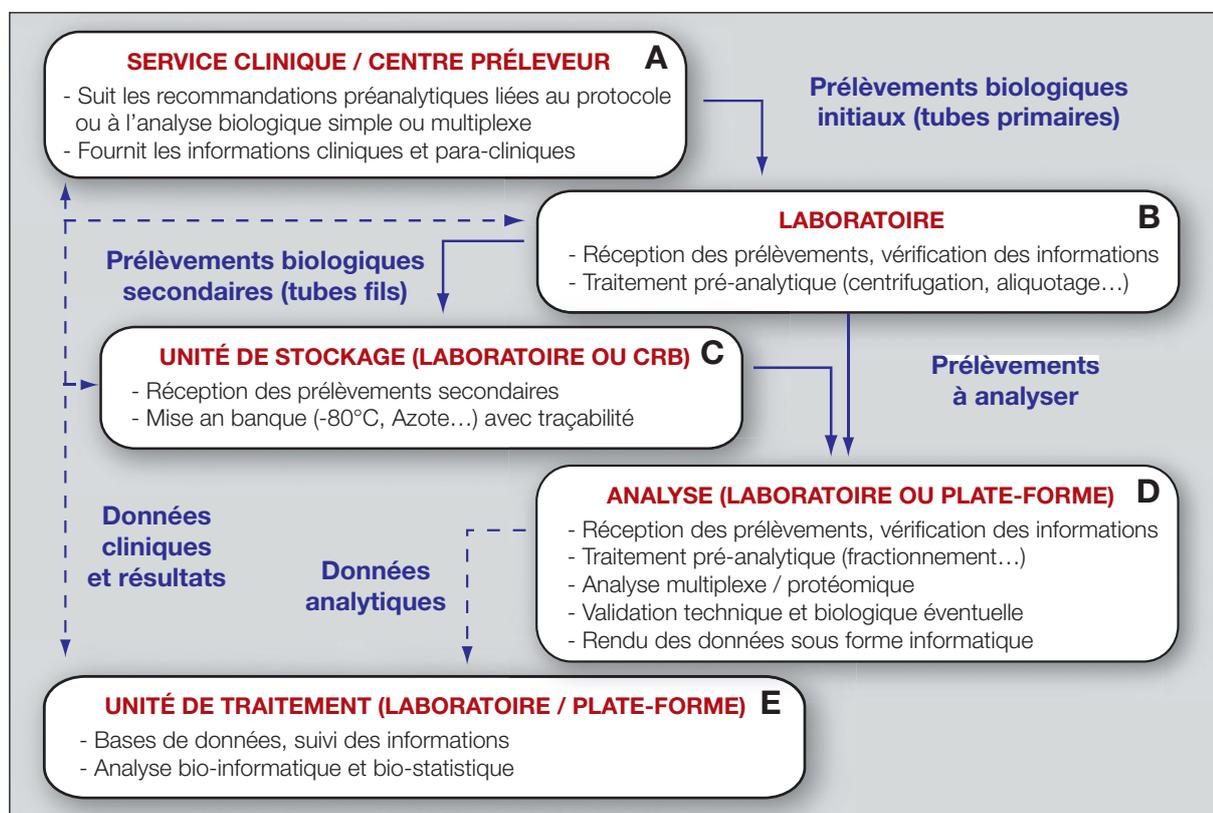
La technologie SELDI-TOF permet d'obtenir un profil protéique à partir d'un fluide biologique (ici du LCR) après capture sur une surface chromatographique et spectrométrie de masse. La complexité d'un profil réalisé dans des conditions pré-analytiques optimales est fortement diminuée si le prélèvement n'est pas conservé dans des tubes appropriés, ou après plusieurs séries de congélation/décongélation.



Afin d'éviter tous biais d'analyse, il est important d'envisager les modifications pouvant être provoquées aux différentes étapes de gestion, de traitement et d'analyse des échantillons (figure 2). La connaissance des étapes sensibles permet de mettre en place des actions correctives afin de diminuer les écarts entre les centres et/ou de les contrôler.

**Figure 2 : Relation entre les structures impliquées dans les projets de protéomique clinique.**

Les centres de prélèvements (A), généralement des structures cliniques, sont impliqués dans la collection des prélèvements selon les recommandations d'inclusion (dans le cas d'une étude clinique) et de types d'échantillons (plasma...) prévus pour la réalisation des analyses. Le laboratoire qui réceptionne ces prélèvements (B) est tout d'abord responsable du traitement des échantillons (centrifugation...) et du suivi pré-analytique (vérification du délai d'acheminement...), de la présence d'un consentement associé au prélèvement. Ensuite il va faire parvenir les prélèvements à l'unité de stockage pour mise en bio-banque (C) et/ou faire suivre les prélèvements au laboratoire (D) responsable des analyses qui pourront alors être traitées en lots. Ces actions sont organisées/supervisées par une unité de traitement (E) qui comprend des compétences en bio-statistique pour traiter et analyser les données.



On portera une attention particulière :

- au recueil des informations médicales et biologiques concernant le patient et le prélèvement (type de tube, identification du tube et lien avec le dossier du patient)
- à la gestion de l'échantillon entre le prélèvement et sa congélation (délais, conditions de centrifugation et de congélation...)
- aux traitements pré-analytiques additionnels (pré-fractionnement) et aux méthodes de décongélation avant utilisation
- aux analyses bio-informatiques et bio-statistiques.

Pour chacune de ces étapes, il est nécessaire et même indispensable de détenir des protocoles validés permettant une homogénéité des échantillons et des analyses afin de permettre, à long terme, la comparaison d'un grand nombre de résultats sans nécessiter obligatoirement l'analyse de l'ensemble des échantillons. A titre d'illustration sont listés dans la **figure 3** quelques exemples du laboratoire jugeant de la qualité du prélèvement biologique et de la conservation du sérum, du plasma, du LCR et des urines.

**Figure 3 : Conditions pré-analytiques en protéomique clinique.**

Ces conditions sont celles utilisées dans le laboratoire pour les programmes de recherche de nouveaux bio-marqueurs en protéomique clinique. Elles sont compatibles avec la plupart des analyses cliniques. Une importance particulière doit être portée sur la traçabilité et sur l'homogénéité pré-analytique de tous les prélèvements (procédures, consommables..).

Sang plasma	Sang sérum	Urine	LCR
<b>A - PRÉLÈVEMENT</b>			
TUBE : Utilisation d'un tube EDTA : "Spray coated K <sub>2</sub> EDTA"	Utilisation d'un tube sec sans activateur de la coagulation	Contenant stérile large, d'au moins 100 mL, avec bouchon vissant, en polypropylène	Tube en polypropylène (type conique 15ml)
PRÉLÈVEMENT : Intraveineux par ponction veineuse directe. Inverser doucement le tube 8 x	Intraveineux par ponction veineuse directe. Garder si possible en position verticale à 4°C	Inclure qu'un seul type d'échantillon (miction du matin, aléatoire dans la journée, ou urines de 24h)	Prélever au moins 3 ml directement dans le tube. Ne pas prendre les 3 premières gouttes
<b>B - TRANSPORT</b>			
Transport à température ambiante, délais < 4h	Transport à température ambiante, délais < 4h	Transport à 4°C ou à température ambiante (conserver une seule condition), délais < 4h	Mettre le prélèvement à 4°C et prévoir le transport à 4°C au laboratoire, délais < 4h
<b>C - TRAITEMENT AU LABORATOIRE</b>			
Centrifuger le tube dans une centrifugeuse climatisée à 20°C à 2000g, 10 min, sans frein. Une deuxième centrifugation pourra être prévue	Centrifuger le tube dans une centrifugeuse climatisée à 20°C à 2000g, 10 min, sans frein	Transfert sur tubes secs de 50 mL, à fond conique en polypropylène, centrifugation 10 min à 2000 g, à température ambiante ou 4°C, sans frein	Centrifuger le LCR dans le tube original de prélèvement centrifugation 10 minutes à 1000 g, à 4°C, sans frein
<b>D - STOCKAGE</b>			
Aliquoter le plasma par 500µl dans des tube en polypropylène ou des paillettes ou cryotubes appropriées	Aliquoter le sérum par 500µl dans tube en polypropylène ou des paillettes ou cryotubes appropriées	Aliquoter en tubes de 5 mL en polypropylène	Aliquoter par 500µl dans des tubes en polypropylène. Veiller à prélever le LCR sans « perturber » le culot
Congeler à - 80°C directement, ou conserver à 4°C avant la mise en congélation (maximum 1h) en vapeur d'azote	Congeler à -80°C directement, ou dans l'azote	Congélation et conservation immédiate à -80°C. Eviter toute conservation temporaire à -20°C	Conservation immédiate à - 80°C ou dans l'azote. Eviter toute conservation temporaire à -20°C

## Gestion et analyse des données

Les laboratoires d'analyse médicale utilisent différentes solutions logicielles qui assurent de façon sécurisée le lien entre les prélèvements, l'identité des patients, les analyses prescrites, les résultats et leurs validations technique et biologique. Dans le domaine de la protéomique clinique, les analyses multiplexes de paramètres cliniques suivent bien évidemment les mêmes règles marquées notamment par l'évolution vers la norme ISO 15189 : 2007.

Dans les laboratoires de recherche et les plates-formes protéomiques, il existe souvent des systèmes spécifiques de gestion informatisée de laboratoire (SGL) souvent appelé LIMS pour « Laboratory Information Management System ». La normalisation suivie est d'ailleurs plutôt de type ISO 17025 : 2005 qui est destinée aux laboratoires d'essai et/ou d'étalonnage. Les LIMS assurent idéalement la traçabilité des échantillons, la gestion des utilisateurs, des appareils et instruments, des stocks et surtout ils font le lien avec les données issues des analyses protéomiques. Ces dernières sont de différents types selon les technologies mises en œuvre (images pour les gels 2D, spectres pour la spectrométrie de masse...) et de format informatique également variable selon les fournisseurs des appareils. Les données sont d'ailleurs souvent prises en charge par des logiciels spécialisés couplés aux appareils utilisés qui assurent en premier lieu la détection du signal et le traitement des données brutes (élimination du bruit de fond, harmonisation, recherche des variables). Ce premier niveau de traitement assure une analyse qualitative des données permettant d'écarter les résultats douteux, c'est-à-dire là où l'influence de facteurs autres que celui étudié a introduit un biais notable (conditions pré-analytiques altérées par exemple). L'analyse des données fait ensuite appel à des outils bio-statistiques comme des algorithmes de classification [10].

## Conclusion

---

La protéomique clinique est une nouvelle discipline, originale par la notion d'exhaustivité et les analyses bio-informatiques qu'elle contient. Les laboratoires d'analyse médicale sont au premier plan de cette discipline par leur triple implication dans le domaine de la découverte de nouveaux biomarqueurs, dans leur validation, puis finalement dans leur utilisation. Cela nécessite de porter une attention particulière aux conditions pré-analytiques afin de s'adapter à la mesure de dizaines, voire des milliers de paramètres connus ou inconnus en parallèle.

**Note :** ce texte s'inspire du travail réalisé sur le pré-analytique en protéomique dans des groupes de travail de l'INCa (Institut National du Cancer) et de la Société Française de Biochimie Clinique (SFBC).

# Bibliographie

---

1. Domon B. and Aebersold R. Mass spectrometry and protein analysis. *Science* 2006 ; 312 : 212-217.
2. Lehmann S., Dupuy A., Peoc'h K., Roche S., Baudin B., Quillard M., Berger F., Briand G., Chwetzoff S., Dine G., Gonzalo P., Dastugue B., Sève M., Siest G. et Beaudoux J. Present possibilities and future development of Clinical Proteomics. *Ann. Biol. Clin. (Paris)* 2007 ; 65 : 463-471.
3. Fung ET., Weinberger SR., Gavin E. and Zhang F. Bioinformatics approaches in clinical proteomics. *Expert. Rev. Proteomics* 2005 ; 2 : 847-862.
4. Katz J. The Nuremberg Code and the Nuremberg Trial. A reappraisal. *JAMA* 1996 ; 276 : 1662-1666.
5. Rice TK., Schork NJ. and Rao DC. Methods for handling multiple testing. *Adv. Genet.* 2008 ; 60 : 293-308.
6. Dupuy AM., Lehmann S. and Cristol JP. Protein biochip systems for the clinical laboratory. *Clin. Chem. Lab. Med.* 2005 ; 43 : 1291-1302.
7. Roche S., Dauvilliers Y., Tiers L., Couderc C., Piva MT., Provansal M., Gabelle A. and Lehmann S. Autoantibody profiling on high-density protein microarrays for biomarker discovery in the cerebrospinal fluid. *J. Immunol. Methods* 2008 ; 338 : 75-78.
8. Rubin RB. and Merchant M. A rapid protein profiling system that speeds study of cancer and other diseases. *Am. Clin. Lab.* 2000 ; 19 : 28-29.
9. Roche S., Tiers L., Provansal M., Piva MT. and Lehmann S. Interest of major serum protein removal for Surface-Enhanced Laser Desorption/Ionization - Time Of Flight (SELDI-TOF) proteomic blood profiling. *Proteome Sci.* 2006 ; 4 : 20.
10. Reynès C., Sabatier R., Molinari N. and Lehmann S. A new genetic algorithm in proteomics: feature selection for SELDI-TOF data. *Computational Statistics and Data Analysis* 2008 (in press).



***La spectrométrie  
de masse et applications  
aux petites molécules***

***Marie-Claude Menet***

**CHAPITRE V**

# Introduction

---

La spectrométrie de masse est une méthode d'analyse qui permet la détermination des masses moléculaires des composés analysés ainsi que leur identification et leur quantification. Elle est fondée sur la séparation et la détection d'ions formés dans une source d'ionisation, ces ions provenant de la molécule à analyser. Dans le cas de méthodes d'ionisation dites douces, l'ion formé peut être consécutif à l'addition d'un ion ( $H^+$ ,  $Na^+$ ,  $K^+$ ,  $NH_4^+$ ) ou d'un électron, ou, à la soustraction d'un électron ou d'un hydrogène ionisé (hydrure ou proton) à la molécule. Cet ion est appelé ion pseudo-moléculaire ou ion moléculaire. Dans le cas de méthodes d'ionisation moins douces, les ions pseudo-moléculaires ou moléculaires se fragmentent pour donner des ions fils en suivant des règles de fragmentation connues et caractéristiques des structures des molécules à analyser. Il est donc possible avec cette méthode d'obtenir des informations structurales. Actuellement, la spectrométrie de masse est facilement couplée aux méthodes séparatives telles les chromatographies en phase gazeuse et en phase liquide. Par le comportement chromatographique des molécules, une autre dimension est alors apportée à l'identification.

Dans ce chapitre, seront tout d'abord données les définitions de plusieurs termes classiquement utilisés en spectrométrie de masse. La description du spectromètre de masse sera ensuite abordée en détail, tant au niveau des sources que des analyseurs. Puis seront développées quelques applications afin de montrer les possibilités qu'apporte la spectrométrie de masse dans l'analyse qualitative et quantitative des molécules d'intérêts pharmacologique et toxicologique.

## Définitions

### Le spectromètre de masse

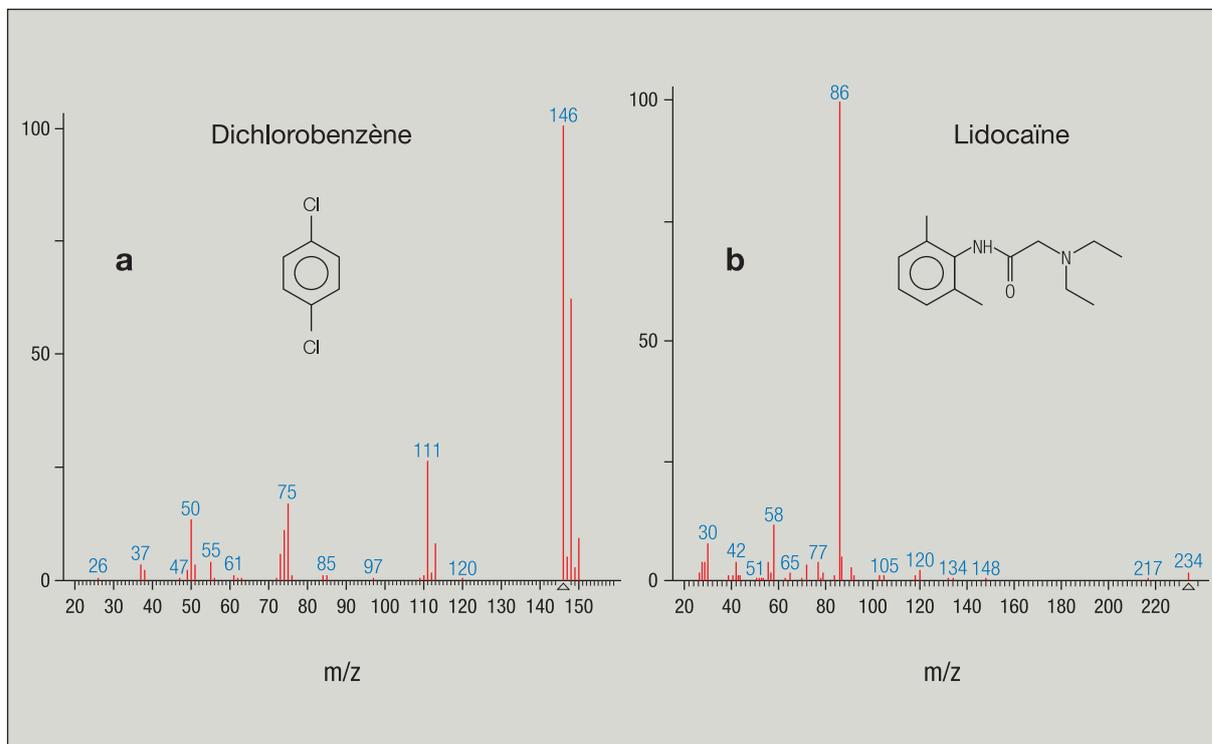
Le spectromètre de masse est un instrument qui comprend différents constituants placés en série et qui permettent successivement, après introduction de l'échantillon, l'évaporation et l'ionisation des molécules de l'échantillon (source), l'accélération des ions formés, la séparation de ces ions en fonction de leur rapport masse sur charge ( $m/z$ ) (analyseurs) et enfin leur détection.

### Le spectre de masse

Le spectre de masse est un graphique ou un tableau à double entrée produit par un spectromètre de masse. Il présente le courant d'ions mesuré en fonction du rapport masse (dalton) sur la charge ( $m/z$  en Thomson ou Th, unité généralement omise). Le graphique peut se présenter sous la forme de deux tracés différents. Le premier, qualifié de vrai spectre de masse ou spectre continu (mode « profile ») est le reflet fidèle du courant ionique que reçoit le détecteur en fonction du rapport  $m/z$  et se présente sous la forme d'un ensemble de pics plus ou moins larges. Le deuxième, spectre

« barres » (mode « centroïde ») est obtenu après traitement informatique du vrai spectre de masse, chaque pic étant présenté sous la forme d'un trait vertical (**figure 1**). Le pic le plus intense du spectre, correspondant au type d'ion le plus abondant, est appelé le pic de base, son intensité est chiffrée à 100 %. Les intensités des autres pics s'expriment en % de celle du pic de base.

**Figure 1 : Spectres obtenus par impact électronique : (a) du dichlorobenzène ; (b) de la lidocaïne.**



## La masse

Les unités de masse usuelles sont :

- *l'unité de masse atomique (uma)* : qui est 1/12 de la masse de l'isotope 12 du carbone, donc  $1 \text{ uma} = 10^{-3} \text{ kg.mol}^{-1}/N = 1,661.10^{-27} \text{ kg}$  (N étant le nombre d'Avogadro)
- *le dalton (Da)* : qui est la masse de l'isotope 1 de l'hydrogène, donc  $1 \text{ Da} = 1,673.10^{-27} \text{ kg} = 1,007 \text{ uma}$ .

Pour une molécule, trois types de masse sont définis : les masses nominale, isotopique et moyenne.

- *La masse nominale* : est la masse calculée à partir de la composition élémentaire utilisant les masses atomiques nominales (en nombre entier) des isotopes les plus abondants.
- *La masse isotopique* : est la masse calculée à partir de la composition élémentaire utilisant pour chaque atome contenu dans la molécule, la masse atomique exacte des isotopes les plus abondants.
- *La masse moyenne ou masse chimique* : est la masse calculée à partir de la composition élémentaire utilisant les masses atomiques moyennes (qui tient compte de l'abondance relative des isotopes).

## L'exactitude (justesse) en masse (« accuracy »)

La justesse en masse est exprimée en erreur relative (%) ou en ppm (partie par million). Par exemple, la masse isotopique de la bradykinine étant 1060,5692, si la masse mesurée est 1060,6201, la justesse en masse est déterminée par :

$$(1060,5692-1060,6201)/1060,5692 = 4,8 \cdot 10^{-5} = 48 \text{ ppm} = 0,0048 \%$$

## La résolution

La résolution (R) est la capacité que présente un spectromètre de masse à distinguer des ions présentant des rapports m/z voisins mais différents. Elle est définie par :  $R = m/\Delta m$  où  $\Delta m$  est la différence de masse entre deux pics adjacents résolus (c'est-à-dire que l'intensité de la vallée entre les deux pics n'excède pas 10 % de la hauteur du pic le plus faible) et m est la masse du premier pic ou la moyenne des deux. Une autre formule permet de calculer la résolution sur un pic isolé. Dans ce cas m est la masse du pic et  $\Delta m$  sa largeur à mi-hauteur ; on parle alors de « Full Width at Half Maximum » (FWHM). Seuls les spectres obtenus en mode « profile » permettent d'effectuer ce calcul.

## Les ions

- *L'ion moléculaire* : c'est l'ion produit quand une molécule introduite dans la source d'un spectromètre de masse perd ( $M^+$ ) ou acquiert un électron ( $M^-$ ). L'ion moléculaire est toujours un ion à nombre impair d'électrons, de masse égale à celle de la substance neutre introduite dans la source. Il permet la détermination de la masse moléculaire.

Cette masse est impaire si la molécule présente dans sa formule un nombre impair d'azote (comme par exemple  $\text{NH}_3$  de masse moléculaire 17 Da) ; elle est paire si la molécule ne présente pas d'azote ou contient dans sa formule un nombre pair d'azote (comme par exemple  $\text{C}_6\text{H}_6$  de masse moléculaire 78 Da et  $\text{N}_2\text{H}_4$  de masse moléculaire 32 Da). La « règle de l'azote » énoncée ci-dessus s'applique parce que les isotopes les plus abondants des éléments usuels ont une masse et une valence de même parité exceptée l'azote (masse 14 Da, valence 3).

- *L'ion pseudo-moléculaire* : c'est l'ion résultant de la perte ou de l'addition d'atomes chargés à la molécule pour former un ion (comme par exemple  $M+H^+$  pour donner  $MH^+$ , ou  $M+Na^+$  pour donner  $MNa^+$ ). Il permet la détermination de la masse moléculaire.

- *L'ion parent* : c'est un ion qui engendre un ion fragment. Ce terme est applicable à tout ion susceptible de se fragmenter.

- *L'ion fragment* (ou ion fils,  $F^+$  ou  $F^{+*}$ ) : c'est l'ion produit par la perte d'un ou plusieurs fragments d'un ion parent. Si cet ion est issu d'un réarrangement (rétro-Diels-Alder, Mc Lafferty) qui conduit à la perte d'un fragment neutre ( $\text{H}_2\text{O}$  de 18 Da ou  $\text{NH}_3$  de 17 Da, par exemple), sa parité est identique à celle de l'ion parent à condition que le fragment perdu ne contienne pas d'atome d'azote.

## Principes et appareillages

### Les sources

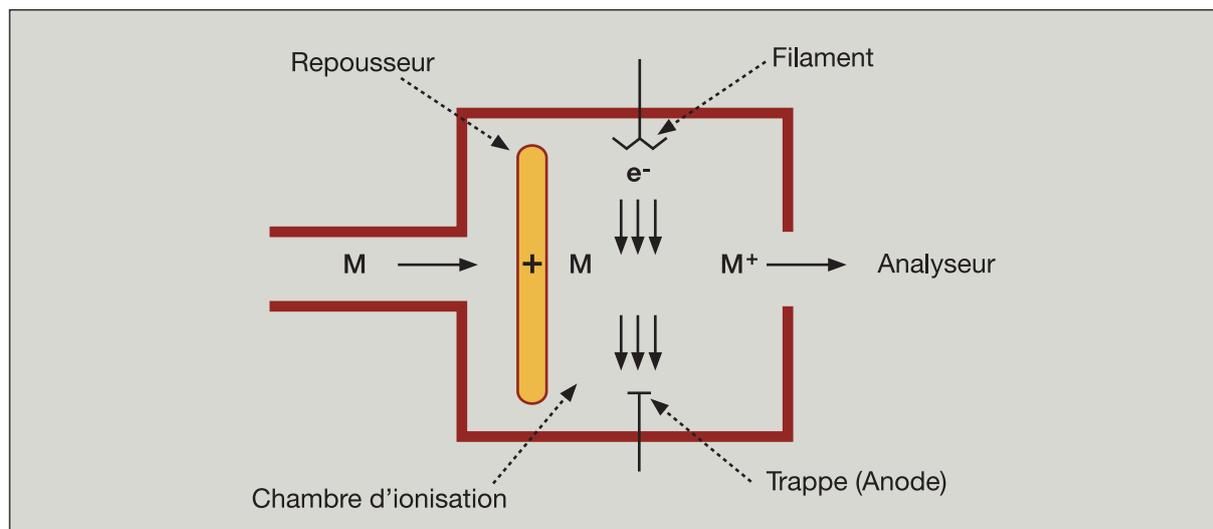
La source est l'élément du spectromètre de masse qui permet l'évaporation et l'ionisation des analytes contenus dans l'échantillon. Plusieurs méthodes d'ionisation sont disponibles et leur choix d'utilisation dépend des propriétés physicochimiques des molécules à étudier et des renseignements désirés. L'ionisation par impact électronique et l'ionisation chimique à pression réduite sont les plus anciennes et sont réservées aux composés facilement volatilisables et de faible masse moléculaire. Elles sont souvent couplées à la chromatographie en phase gazeuse.

D'autres sources d'ionisation dites «douces» ont été développées depuis une vingtaine d'année. Elles permettent l'ionisation de molécules sans provoquer (ou peu) leur fragmentation. La désorption et ionisation par impulsion laser et assistées par une matrice (source MALDI) est fréquemment couplée à un spectromètre de masse à temps de vol (TOF) pour l'analyse de molécules présentant une masse moléculaire élevée (au moins supérieure à 500 Da). L'électro-nébulisation (ESI) et l'ionisation chimique à pression atmosphérique (APCI) font parties des méthodes d'ionisation à pression atmosphérique (API) et sont idéalement utilisées pour l'ionisation de molécules de faible masse, en couplage avec la chromatographie en phase liquide en particulier. L'ESI permet aussi l'analyse de macromolécules (protéines, acides nucléiques) puisque le processus d'électronébulisation est capable de générer des ions multichargés, présentant des rapports  $m/z$  inférieurs à 2000 Th et donc facilement séparés et détectés par un spectromètre de masse classique.

### *Ionisation sous faible pression*

- *Ionisation et fragmentation par impact électronique* : l'ionisation par impact électronique est la plus ancienne et la mieux connue des méthodes d'ionisation (figure 2). La source comprend un filament chauffé émettant des électrons qui sont accélérés vers une anode, appelée piège ou trappe. Le faisceau d'électrons entre alors en collision avec les molécules sous forme gazeuse présentes dans la source et leur arrache un électron pour former l'ion moléculaire  $M^+$ , lui-même expulsé vers l'analyseur grâce à une plaque portée à un potentiel positif et appelée repousseur. Généralement, le potentiel d'ionisation (PI) (énergie nécessaire à l'ionisation d'une molécule par expulsion d'un électron de valence) se situe autour de 10 eV. Les composés à analyser étant bombardés par un faisceau de 70 eV, la différence augmente l'énergie interne de la molécule et intervient dans les processus de fragmentation si l'énergie apportée à la molécule dépasse le potentiel d'apparition (PA) (énergie nécessaire à l'apparition de fragments). Les principaux processus de fragmentation sont décrits dans les applications en fin de chapitre.

**Figure 2 : Source d'ions par impact électronique (d'après « Appareils et méthodes en biochimie et biologie moléculaire », B. Hainque, B. Baudin et P. Lefebvre. Flammarion Médecine-Sciences, 2008).**



Cette méthode d'ionisation-fragmentation est particulièrement adaptée aux composés de masse moléculaire faible (inférieure à 1000 Da), facilement volatilisables et stables à température élevée ; ce n'est pas une méthode douce. Les spectres de masse présentent de nombreux pics, celui de plus haut rapport  $m/z$  correspondant à l'ion moléculaire, les autres de  $m/z$  plus faibles, à des ions fragments. Ce mode d'ionisation-fragmentation est très reproductible et fréquemment utilisé pour obtenir des renseignements structuraux ou pour identifier formellement une molécule en comparant le spectre obtenu à ceux contenus dans une bibliothèque comme celle du NIST (« National Institute of Standards and Technology »).

- *Ionisation chimique positive* : la source d'ionisation chimique est souvent identique à celle utilisée en impact électronique (source combinée) mais le mécanisme d'ionisation est indirect (figure 2). Un gaz réactant R (ammoniac, méthane, isobutane) introduit dans la source à la pression de 1 mbar entre en collision avec un faisceau d'électrons générés par un filament et accélérés grâce à une différence de potentiel. Une énergie cinétique de 10 eV permet l'ionisation du gaz et la formation dans la chambre d'ionisation d'un plasma (mélange d'ions positifs  $R^+$  et d'électrons). Les molécules  $M$  à analyser, introduites sous forme gazeuse dans la source, sont alors très peu voire non ionisées par les électrons provenant du filament. Elles réagissent avec les ions positifs du plasma. Trois types de réactions peuvent intervenir :

- le transfert de proton ( $M + RH^+ \rightarrow MH^+ + R$ )
- la formation d'adduits ( $M + R^+ \rightarrow MR^+$ )
- l'abstraction d'hydrure ( $MH + R^+ \rightarrow M^+ + RH$ ) (cas des hydrocarbures).

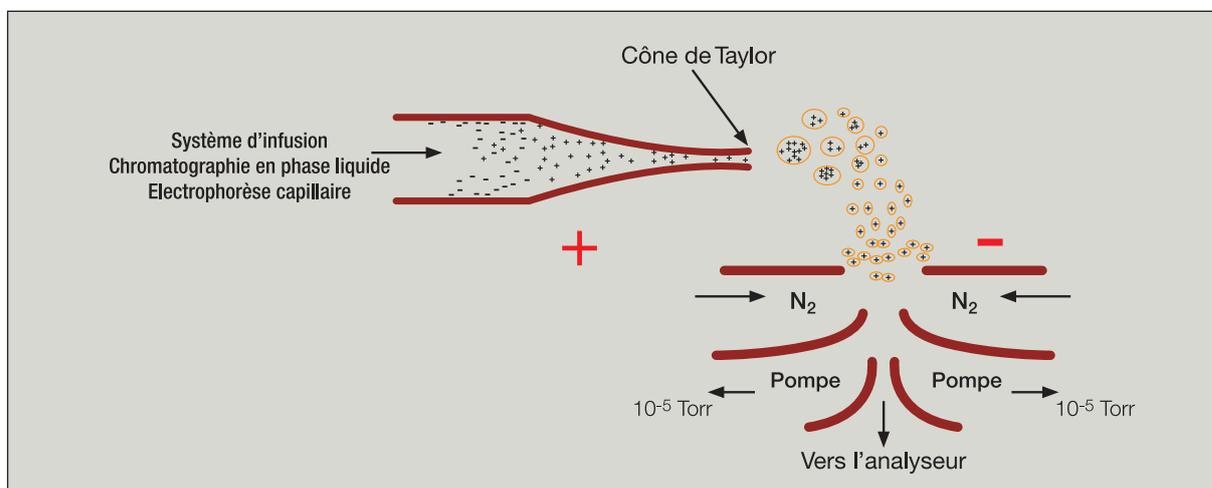
Ce mode d'ionisation est beaucoup plus doux que l'impact électronique. Le processus de fragmentation est pratiquement inexistant (sauf quand le méthane est utilisé comme réactant). L'ionisation de la molécule  $M$  conduit donc surtout à la formation d'ions pseudo-moléculaires qui permettent de déterminer la masse moléculaire du composé analysé. Comme pour l'impact électronique, cette méthode d'ionisation est particulièrement adaptée aux composés de masse moléculaire faible (inférieure à 1000 Da), facilement volatilisables et stables à température élevée.

- Ionisation - désorption par impulsion laser et assistées par une matrice (MALDI, « Matrix Assisted Laser Desorption Ionization ») : la désorption et l'ionisation assistées par laser consistent à vaporiser et ioniser des espèces co-cristallisées au sein d'un solide appelé matrice. Pour obtenir ce cristal, le composé à analyser doit être préalablement dissout dans une solution contenant la molécule matrice à une concentration de 100 à 10 000 fois supérieure. Une gouttelette (0,5 à 1  $\mu\text{L}$ ) de cette solution est déposée sur une plaque métallique dite plaque MALDI et évaporée à l'air. Le co-cristal est introduit dans la source du spectromètre de masse et irradié par un rayon laser (par exemple 337 nm en UV-MALDI). La matrice absorbe une grande partie de l'énergie lumineuse qu'elle transmet aux analytes, permettant ainsi leur désorption et leur ionisation et prévenant leur décomposition. Cette méthode d'ionisation est dite « douce » avec généralement formation d'un ion pseudo-moléculaire de la forme  $\text{MH}^+$ , sans fragmentations. Plusieurs composés « matrices » sont disponibles ; l'ionisation ayant lieu par transfert de protons, les matrices sont fréquemment des acides carboxyliques, solides à température ambiante. Elles présentent en solution un coefficient d'absorption moléculaire élevé à la longueur d'onde d'utilisation (entre 3000  $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$  et 16000  $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ). Toutes les matrices montrent des pics à des rapports  $m/z$  inférieurs à 500. Corrélativement, les molécules de faible masse moléculaire (< 500) sont parfois difficiles à analyser par cette méthode. Cette source est donc peu utilisée pour l'analyse de petites molécules organiques.

### **Ionisation à pression atmosphérique**

- Ionisation par électro-nébulisation (ESI, « Electrospray Ionization ») : elle consiste à pulvériser une solution, à ioniser les espèces qui y sont présentes et à en assurer l'évaporation, ceci en appliquant à pression atmosphérique un champ électrique sur un liquide traversant un capillaire (nébuliseur). On applique une différence de potentiel de quelques milliers de volts entre le nébuliseur et une contre-électrode (figure 3a).

**Figure 3a : Sources d'ionisation à pression atmosphérique (API) : (a) source d'ions ESI (électronébuliseur) - Processus d'électronébulisation en mode positif. La configuration orthogonale augmente le rapport signal/bruit (d'après « Appareils et méthodes en biochimie et biologie moléculaire », B. Hainque, B. Baudin et P. Lefebvre. Flammarion Médecine-Sciences, 2008)**

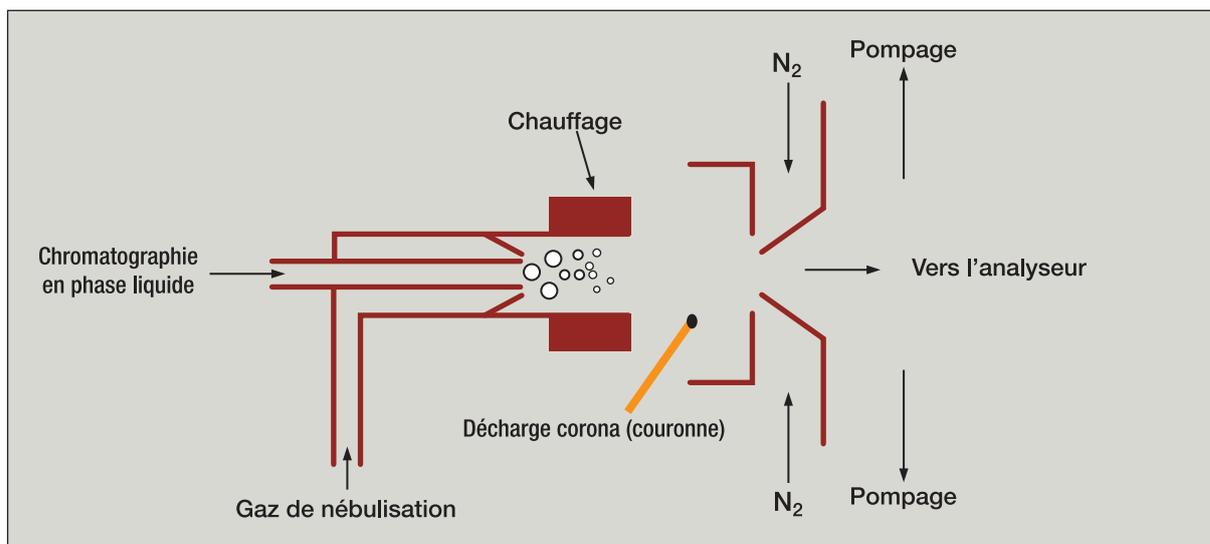


En mode positif par exemple (cas où l'embout du nébuliseur joue le rôle de l'anode), les ions chargés positivement ou négativement présents dans l'électrolyte vont migrer et se séparer à l'extrémité du capillaire, les ions positifs se plaçant à la surface du liquide (au niveau du ménisque) et les ions négatifs s'accumulant au niveau du capillaire. Les répulsions coulombiennes entre les charges positives tendent à étirer le ménisque et à former le cône de Taylor qui se transforme alors en filament liquide puis en fines gouttelettes au fur et à mesure que le champ augmente. L'échantillon est ainsi pulvérisé sous forme de fines gouttelettes à la sortie du nébuliseur. Le solvant contenu dans les gouttelettes est ensuite progressivement évaporé à l'aide d'un gaz neutre à contre-courant (azote). Les ions sont formés par la simple évaporation d'une gouttelette ne contenant qu'une molécule en présence de protons ou par éjection de molécules chargées de la gouttelette du fait des répulsions électrostatiques. Puis ces ions entrent dans le spectromètre de masse. Pour des petites molécules organiques, en mode positif, l'ion formé est ionisé une seule fois ; pour un peptide ou une protéine, le nombre de charges positives par ion dépend du nombre d'acides aminés basiques contenus dans la molécule, en général une charge élémentaire est présente pour 1000 Da. Plusieurs propriétés physicochimiques du liquide influent sur la qualité du nébulisât, en particulier la température d'ébullition ( $T_{eb}$ ), la tension superficielle, la conductivité des solvants et électrolytes. Les liquides introduits dans la source d'électro-nébulisation doivent être facilement volatilisables, les tampons minéraux (phosphate) sont en principe à proscrire. Ce sont pour les solvants présentant la plus faible tension superficielle que le cône de Taylor apparaîtra le plus facilement et que la nébulisation sera la meilleure (méthanol, acétonitrile, isopropanol). Ce n'est pas le cas des liquides riches en eau. Si la conductivité est trop élevée, un arc électrique peut apparaître avec pour conséquence une baisse de la sensibilité ; si la conductivité est trop faible le cône de Taylor sera trop court (cas de l'hexane et du cyclohexane, non utilisables lors de l'ionisation par ESI).

D'autres paramètres, comme le pH du solvant et la présence ou non de contre-ions en solution conjointement aux molécules, interviennent dans le degré d'ionisation des molécules. Peuvent être utilisés : l'acide acétique, l'acide formique, l'acide trifluoroacétique, l'acétate d'ammonium, la triéthylamine, l'acide heptafluorobutyrique (en mode positif) ou l'hydroxyde d'ammonium, l'ammoniaque (en mode négatif)... De plus, la présence de contre-ions influe sur l'intensité du signal, sur l'état de charge moyen ainsi que sur la répartition des charges. En effet, des paires d'ions neutres se forment en solution avec neutralisation, par le contre-ion négatif, des sites chargés positivement de la molécule à analyser, les trifluoroacétates formant des paires d'ions beaucoup plus stables que les formiates et les acétates. Ainsi la concentration conseillée d'utilisation de l'acide trifluoroacétique en solution est inférieure à 0,1 % (v/v) alors que celle en acides formique et acétique peut dépasser 1 % (v/v).

- *Ionisation chimique à pression atmosphérique (APCI, « Atmospheric Pressure Chemical Ionization »)* : elle peut être utilisée en aval d'un chromatographe en phase liquide. Elle consiste, dans un premier temps, à nébuliser à l'aide d'un gaz (azote), une solution contenant les analytes d'intérêt (figure 3b).

Figure 3b : Sources d'ionisation à pression atmosphérique (API) : (b) Source d'ions APCI (ionisation chimique à pression atmosphérique).



Cet aérosol est évaporé après son introduction dans un manchon chauffé. Le mélange des gaz de nébulisation et ceux issus de l'évaporation est entraîné vers une électrode à décharge couronne (ou « corona »). Celle-ci, source constante d'électrons crée un plasma en ionisant le gaz ambiant pour former des ions réactifs (dits « réactants »). Il est communément admis que l'ionisation des molécules d'intérêt (M) provient de réactions acido-basiques, en phase gazeuse, entre les ions réactifs du plasma et M. La nature des ions réactants présents dépend du solvant solubilisant initialement les analytes. En APCI en mode positif, ce sont des ions réactifs dits acides tels  $\text{H}_3\text{O}^+$ ,  $\text{CH}_3\text{OH}_2^+$ ,  $\text{CH}_3\text{CNH}^+$  et  $\text{NH}_4^+$  qui permettent la formation d'ions de type  $\text{MH}^+$  ; en APCI en mode négatif, ce sont des ions réactifs dits basiques tels  $\text{OH}^-$ ,  $\text{CH}_3\text{O}^-$ ,  $\text{CH}_2\text{CN}^-$ ,  $\text{CH}_3\text{COO}^-$  et  $\text{Cl}^-$  qui permettent la formation de  $(\text{MH})^-$ . L'ionisation des composés est très favorisée dans la source par la fréquence de collisions élevée à pression atmosphérique.

Les sources ESI et APCI sont complémentaires. La source d'ionisation chimique à pression atmosphérique est peu compatible avec les nano-débits utilisés en nano-chromatographie en phase liquide contrairement à la source ESI. La source APCI permet cependant l'ionisation mono-chargée de composés de faible masse moléculaire et présentant une polarité moyenne.

Les constructeurs développent actuellement des sources mixtes (ESI/APCI) ou facilement démontables pour passer d'une source d'ionisation à l'autre. Lors de la mise au point d'une méthode d'analyse par chromatographie en phase liquide couplée à la spectrométrie de masse, on ne peut que recommander d'évaluer la réponse relative de ces deux modes d'ionisation.

## Les analyseurs de masse

L'analyseur de masse correspond à la partie du spectromètre de masse qui sépare les ions formés dans la source en fonction de leur rapport  $m/z$ . Plusieurs types d'analyseurs sont disponibles.

Leur mode de fonctionnement repose sur l'utilisation de champs électriques et/ou magnétiques. L'analyseur à temps de vol (TOF) permet d'obtenir précisément des mesures de masses moléculaires élevées. Ce spectromètre de masse est fréquemment couplé à une source MALDI (MALDI-TOF). Le filtre quadripolaire et le piège à ions, fréquemment couplés aux sources API, allient coût et encombrement faibles mais présentent des performances beaucoup plus modestes que les spectromètres électromagnétiques, comme par exemple l'analyseur à résonance cyclotronique. Analyseur à temps de vol et quadripôle peuvent être utilisés en série et ainsi donner naissance à des appareils dits « hybrides ». Les analyseurs permettent la séparation des ions mais pas leur détection qui est assurée par un dispositif de multiplication d'électrons.

### **Analyseurs utilisant un champ électrique constant - Analyseurs à temps de vol (TOF, « Time Of Flight »)**

Le principe de fonctionnement du TOF est simple : les ions sont formés dans une région relativement limitée de la source en présence d'un champ électrique E (généré par une différence de potentiel U) qui accélère les ions sur la longueur s de la source avant leur parcours dans une région sans champ et de longueur D beaucoup plus élevée, jusqu'au détecteur. Un ion de masse m et de charge z aura une vitesse v et une énergie cinétique ( $E_c$ ) à l'arrivée au détecteur telle que :  $E_c = 1/2 [m \cdot v^2] = z U$ . Il parcourra la distance D en un temps t tel que :  $t = D/v$  soit  $t^2 = [m/z] [D^2/2U]$ . Le temps de vol t est ainsi directement lié au rapport m/z. La mesure du temps de vol permet alors une détection différentielle des espèces ayant un rapport m/z différent. Actuellement, avec les perfectionnements apportés et les développements des systèmes d'acquisition, les valeurs de résolution rencontrées en analyse par un TOF couplé à une source MALDI peuvent être très supérieures à 5000.

### **Analyseurs utilisant un champ électrique oscillant**

Le fonctionnement de ce type d'analyseurs (filtre quadripolaire, piège à ions) est fondé sur le mouvement d'ions dans un champ électrique oscillant. En effet, si l'on considère le trajet d'un ion dans une dimension, placé entre deux électrodes, l'ion se déplace et est accéléré vers l'électrode de signe opposé. Si la polarité de l'électrode est inversée, l'ion se trouvera accéléré dans l'autre direction. Lorsque les modifications de l'amplitude et de la fréquence du champ électrique sont suffisantes, l'ion reste confiné entre les deux électrodes. Les champs électriques, dans un filtre quadripolaire et un piège à ions, sont générés par des électrodes dont la géométrie et la disposition dans l'espace diffèrent d'un analyseur à l'autre. Dans le filtre quadripolaire, ces champs permettent aux ions de maintenir une trajectoire dite stable jusqu'au détecteur. Dans le piège à ions, ils assurent le confinement des ions dans l'analyseur, leur détection s'effectuant après leur éjection sélective.

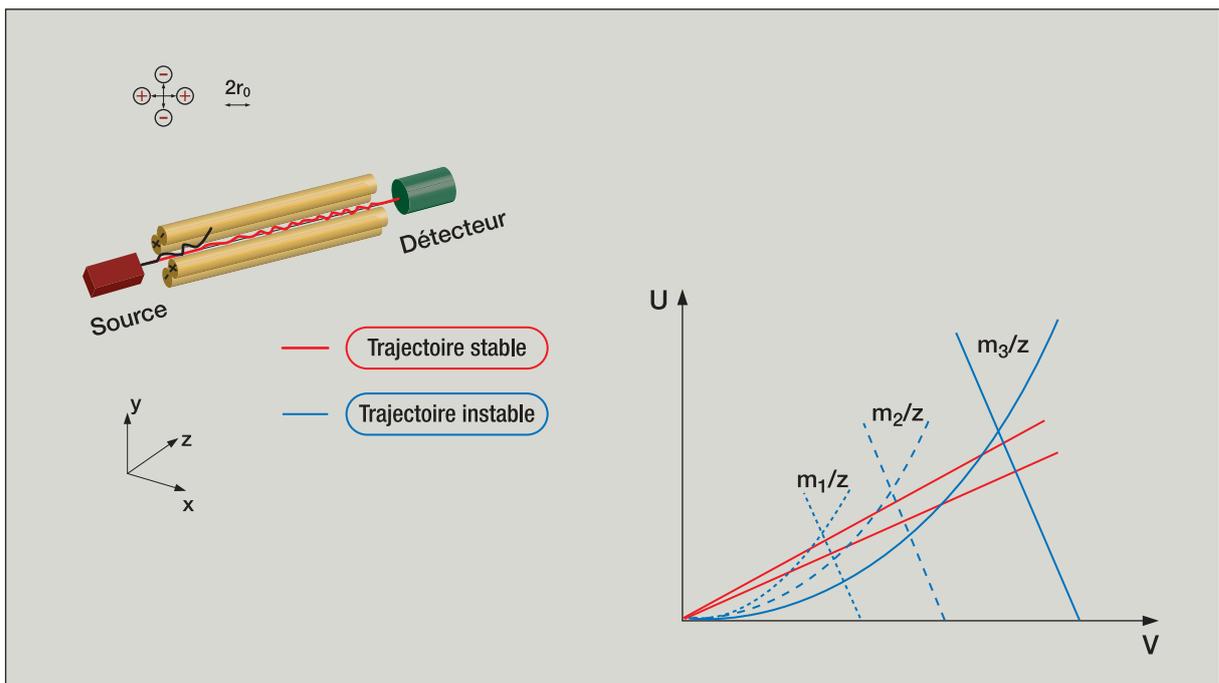
- *Analyseur filtre quadripolaire* : il est constitué successivement d'une série de lentilles électroniques permettant la convergence et l'accélération des ions sortant de la source, puis d'un quadripôle formé le plus communément de quatre barres entre lesquelles sont établies des différences de potentiel, l'ensemble débouchant sur un détecteur (figure 4). L'ion entrant dans le quadripôle est soumis à l'influence d'un champ électrique total, constitué d'un champ alternatif quadripolaire

superposé à un champ constant, résultant de l'application sur les barres, deux à deux, des potentiels :

$$+ (U - V\cos\omega t) \text{ et } - (U - V\cos\omega t)$$

où,  $U$  est la tension continue (en volt),  $V$  et  $\omega$  respectivement l'amplitude de la tension alternative (en volt) et la pulsation de radiofréquence ( $\omega = 2\pi f$  en  $\text{rad}\cdot\text{s}^{-1}$ ,  $f$  étant la fréquence en Hz). L'ion de masse  $m$ , de charge  $z_e$  (cette notation est employée pour éviter la confusion avec l'axe  $z$ ) accéléré suivant l'axe  $z$ , pénètre entre les barres du quadripôle et conserve sa vitesse selon cet axe. Si les tensions et la fréquence imposées sont telles que cet ion entre en contact avec une barre, il sera déchargé et ne pourra pas sortir du quadripôle ni par conséquent être détecté (il aura ainsi été « filtré »). On dit qu'il a présenté une trajectoire instable. Si le potentiel appliqué à une barre change de signe avant que l'ion ne s'y décharge, l'ion changera de direction. Ainsi il oscillera en cheminant selon l'axe  $z$  jusqu'au détecteur. Seuls seront détectés les ions ayant présenté une trajectoire dite stable, c'est-à-dire ceux qui ont traversé le quadripôle sans jamais se heurter aux barres. Tant que simultanément  $x$  et  $y$ , qui mesurent la distance par rapport au centre des barres, restent inférieurs à  $r_0$ , l'ion pourra traverser le quadripôle sans toucher les barres. Pour différents ions, les zones de stabilité, dites « zones A » à l'intérieur des courbes bleues présentées (figure 4), ou zones de valeurs de  $U$  et  $V$  telles que  $x$  et  $y$  n'atteignent pas, au cours du temps, des valeurs supérieures ou égales à  $r_0$ , peuvent être superposées sur un diagramme (figure 4). Ce dernier permet d'expliquer le principe de la séparation des ions, en superposant, les zones A obtenues à différentes masses et des droites dites de « fonctionnement » (droites en rouge figure 4).

**Figure 4 : Schéma de principe d'un triple quadripôle - Zones A (à l'intérieur des zones bordées de bleu) obtenues à différentes valeurs du rapport  $m/z$  - Droites de fonctionnement (en rouge) (d'après « Appareils et méthodes en biochimie et biologie moléculaire », B. Hainque, B. Baudin et P. Lefebvre. Flammarion Médecine-Sciences, 2008).**



Ainsi, un balayage des potentiels U et V suivant un rapport U/V constant permet de laisser passer à travers le quadripôle jusqu'au détecteur, successivement, les ions présentant des rapports m/ze différents. Plus la pente de la droite de fonctionnement est élevée plus la résolution est importante mais moins d'ions sont détectés. Plus la pente est faible moins la résolution est bonne mais plus d'ions sont détectés. Si la valeur de U est nulle, la résolution est nulle et tous les ions ont une trajectoire stable tant que la valeur de V est dans les limites de leur zone de stabilité.

L'analyseur quadripolaire peut être utilisé selon deux modes. En mode balayage ou « scan », qui permet la détection de tous les ions, les valeurs de U et V varient selon une droite de fonctionnement. En mode SIM (« Single Ion Monitoring ») les valeurs de U et V sont fixées pour ne détecter spécifiquement qu'un ou plusieurs ions.

Plusieurs quadripôles peuvent être couplés entre eux en série et constituer un spectromètre de masse « tandem ». Ce dispositif permet, après la sélection d'un ion par le premier quadripôle, de le fragmenter dans le deuxième et d'analyser les ions fils apparus dans le troisième. Le premier et le troisième quadripôles (Q<sub>1</sub> et Q<sub>3</sub>) sont des spectromètres de masse susceptibles de séparer les ions. Le deuxième (Q<sub>2</sub>) est un quadripôle à radiofréquence uniquement (valeur de U nulle) et n'a pour rôle que de focaliser les ions au centre des barres puisque c'est le lieu de formation des ions fragments par collisions (CID, « Collision Induced Dissociation »). Un gaz de collision y est introduit à une pression telle qu'un ion parent entrant dans ce quadripôle Q<sub>2</sub> subira une ou plusieurs collisions. Si le gaz est inerte, une partie de l'énergie de collision est convertie en énergie interne. L'ion parent se fragmente et les ions fragments fils sont analysés par le troisième quadripôle. Le quadripôle Q<sub>2</sub> est appelé « chambre de collision ». Plusieurs configurations peuvent être mises en œuvre dans l'utilisation d'un tel couplage en fonction du réglage des deux spectromètres de masse Q<sub>1</sub> et Q<sub>3</sub> : elles sont résumées dans le [Tableau 1](#). Lorsqu'un quadripôle est utilisé en mode statique (SIM), il ne détecte qu'un ion à un rapport m/z<sub>e</sub> donné. Lorsqu'un quadripôle est en mode balayage, il détecte tous les ions dans une gamme de rapport m/z<sub>e</sub> prédéfinie.

**Tableau 1 : Modes d'utilisation d'un triple quadripôle.**

	<b>Q<sub>3</sub></b>		
<b>Q<sub>3</sub></b>		<b>Mode statique</b>	<b>Mode scan</b>
	<b>Mode statique</b>	(1) «Selected reaction monitoring»	(2) Spectre de masse des ions fils
	<b>Mode scan</b>	(3) Spectre de masse des ions parents	(4) Perte de fragment neutre

La combinaison de ces deux modes permet de réaliser 4 types de procédures :

1) « *Selected Reaction Monitoring* » (SRM) ou « *Multiple Reaction Monitoring* » (MRM) : il y a détection, par  $Q_3$ , d'un ion ou de plusieurs ions fils fragment(s) spécifique(s) formé(s) à partir de la fragmentation d'un ion parent spécifique.  $Q_1$  est réglé pour la sélection exclusive de l'ion parent,  $Q_3$  celle du ou des ion(s) fils. Ce mode est utilisé lors de problèmes de quantification. La double sélection permet de diminuer le bruit de fond ; le rapport signal/bruit obtenu est alors favorable à une diminution de la limite de quantification.

2) *Spectre de masse des ions fils* : il y a détection, par  $Q_3$ , de tous les ions fils obtenus après fragmentation dans  $Q_2$  d'un ion parent préalablement sélectionné par  $Q_1$ . Ce mode est utilisé lors de la détermination de structure de petites molécules organiques ou biologiques.

3) *Spectre de masse des ions parents* : il y a détection par  $Q_1$  des ions parents présentant un ion fils fragment commun (repéré par  $Q_3$ ). Ce mode est utilisé pour le criblage d'une classe d'ions parents présentant un fragment en commun (par exemple, les phospholipides).

4) *Perte de fragment neutre* :  $Q_1$  et  $Q_3$  sont en mode balayage en maintenant entre eux une différence constante en rapport  $m/z_e$  (masse du fragment neutre perdu).  $Q_1$  sélectionne tous les ions parents qui présentent un ion fils de rapport  $m/z_e$  correspondant au rapport  $m/z_e$  de l'ion parent moins la masse du fragment neutre. Ce mode est utilisé lors de la mise en évidence de modifications post-traductionnelles de protéines (par exemple, la phosphorylation est caractérisée par la perte de  $HPO_3$ ).

- *Analyseur à piège à ions ou « trappe d'ions »* : il est constitué de trois électrodes : une électrode annulaire et deux électrodes « chapeau », ces dernières étant quasiment identiques et de géométrie hyperbolique. L'électrode chapeau supérieure possède en son centre une ouverture qui permet l'introduction des ions issus de la source. L'électrode chapeau inférieure possède une série de trous permettant l'éjection des ions vers le détecteur. Les ions entrant dans la trappe sont piégés sur une trajectoire oscillante, de fréquence spécifique pour chaque ion, formant une sorte de huit en trois dimensions. Ce mouvement provient de l'application d'un potentiel alternatif  $\Phi_0$  appliqué sur les électrodes, de la même forme générale que celui employé dans un quadripôle ( $\pm \Phi_0 = \pm V \cos \omega t$ ), la tension continue  $U$  étant dans ce cas de valeur nulle. Alors que dans les quadripôles on règle les potentiels de manière que seuls les ions d'un rapport  $m/z$  déterminé traversent les barres, ici les ions de différents rapports  $m/z_e$  sont présents simultanément dans la trappe. Le spectre est obtenu en les expulsant en fonction de leur rapport  $m/z_e$  en déstabilisant leur trajectoire suivant l'axe  $z$ , soit par augmentation progressive de la tension  $V$  (modulation de l'amplitude) soit en appliquant un potentiel oscillant supplémentaire (éjection par résonance ou modulation axiale). De plus, il est nécessaire d'introduire systématiquement dans l'analyseur un gaz neutre (hélium), pour garantir un confinement efficace et focaliser le nuage d'ions au centre de la trappe, puisque ces ions, s'ils présentent la même charge, peuvent se repousser mutuellement et déformer leurs trajectoires, voire les déstabiliser.

Un analyseur de type trappe d'ions peut être utilisé en mode « tandem », mais contrairement au triple quadripôle il ne nécessite pas une mise en série de plusieurs pièges. La trappe, successivement, isole l'ion parent, le fragmente, sépare et fait sortir vers le détecteur les ions fils apparus. L'ion parent

est sélectionné par expulsion de tous les autres ions en jouant sur les tensions ou la radiofréquence. Cet ion isolé est ensuite fragmenté par collision (CID). Il est d'abord excité en appliquant entre les électrodes une radiofréquence de faible amplitude, dite de résonance, caractéristique de l'ion. L'ion n'est pas expulsé, mais voit son énergie cinétique augmenter ; il entre alors en collision avec des molécules d'hélium présentes dans le piège. L'énergie cinétique est convertie en énergie interne qui, si elle est suffisante, permet la fragmentation de l'ion. Après fragmentation, les ions fils formés, qui ne présentent pas la même fréquence de résonance que l'ion parent, sont immédiatement piégés au centre de la trappe. Ils sont ensuite expulsés du piège en fonction de leur rapport  $m/z_e$  puis détectés.

Une variante de la trappe d'ions, la trappe linéaire, a été développée récemment. Elle présente le même type de fonctionnement que les pièges à ions mais les ions, dans ce cas, sont piégés entre les barreaux d'un quadripôle à radiofréquence (valeur de U nulle) et expulsés ensuite en fonction de leur rapport  $m/z_e$ . L'espace de confinement des ions est plus grand que dans un piège classique. Les ions y étant stockés en plus grand nombre, ce type d'analyseur conduit à une limite de détection plus basse que celle rencontrée pour l'analyseur à trappe d'ions classique.

### **Analyseurs hybrides**

Ils comprennent au moins deux types d'analyseurs séparant les ions selon deux processus différents.

- *Couplage source ESI - analyseurs quadripolaire et à temps de vol (ESI-Q-TOF)* : cet analyseur est composé d'une source ESI reliée à une série de quadripôles ( $Q_0$  qui guide les ions,  $Q_1$  qui les sépare,  $Q_2$  qui est utilisé comme chambre de collision) puis à un TOF placé perpendiculairement pour obtenir une bonne résolution et une bonne précision dans la mesure des temps de vol.

- *Couplage source ESI - quadripôle linéaire (LTQ®) et Orbitrap®* : depuis peu a été développé le LTQ-Orbitrap® associant une trappe d'ions linéaire (« Linear Ion Trap Quadripole », LTQ®) qui a pour rôle de stocker les ions produits par la source, puis de les éjecter latéralement pour qu'ils soient détectés, ou co-axialement pour qu'ils entrent dans l'analyseur couplé en ligne, l'Orbitrap®. Les ions y sont alors soumis à un champ électrique particulier et suivent une trajectoire circulaire autour et le long d'une électrode cylindrique. La fréquence du mouvement est inversement proportionnelle à la racine carrée de leur rapport masse sur charge. Plusieurs ions de rapport différent étant présents en même temps dans l'Orbitrap®, c'est une somme de fréquences qui est détectée. Celle-ci permet le tracé du spectre de masse après calcul par transformée de Fourier. Les performances obtenues avec ce système, en terme de limite de détection, de résolution (jusqu'à 100 000) et de justesse en masse (< 5 ppm en calibrage externe), sont supérieures à celles observées lors de l'utilisation du système hybride, Q-TOF dont le fonctionnement est décrit ci-dessus.

### **Analyseurs utilisant un champ magnétique**

La propriété des champs magnétiques à dévier la trajectoire des ions a été mise à profit dès le début du vingtième siècle pour déterminer la masse des ions. Par exemple en 1933, le spectromètre de

Bainbridge combinait champs électriques et champ magnétique. Les champs électriques accélèrent les ions et homogénéisent leur énergie cinétique, le champ magnétique les dévie sur des trajectoires circulaires dont le rayon dépend de leur rapport m/z. Actuellement l'action de ces champs est exploitée dans le spectromètre de masse à résonance cyclotronique ionique (FTICR). Cet analyseur est peu usité mais les valeurs de résolution très élevées qu'il présente (de l'ordre de 80 000) permettent la visualisation des massifs isotopiques et une précision en masse extrême (de l'ordre de 1 ppm) utiles dans la détermination de la composition élémentaire des composés ( $C_xH_yN_zO_n$ ).

## Applications : analyse de molécules d'intérêt pharmacologique ou toxicologique

La spectrométrie de masse possède actuellement un large choix de sources et d'analyseurs. Chaque configuration présente ses propres caractéristiques, mises à profit lors de problèmes analytiques qualitatifs ou quantitatifs. Les premiers, consistant essentiellement en la détermination de structures, peuvent être en partie résolus par l'analyse isotopique et la fragmentation des ions parents. Les deuxièmes, consistant à déterminer la concentration d'un composé connu dans un mélange, exploitent la capacité de certaines sources à être couplées à la chromatographie.

### Isotopes et détermination de la composition élémentaire des molécules

La spectrométrie de masse permet facilement de déterminer le nombre de carbones d'une petite molécule organique et de déceler la présence ou non d'halogènes dans sa formule brute. Pour cela, on utilise la notion d'abondance isotopique (AI) qui peut être définie par la probabilité de rencontrer dans la molécule un atome A+1 ( $^{13}C$ ,  $^{15}N$  ou  $^2H$ , [tableau 2](#)) ou un atome A+2 ( $^{37}Cl$ ,  $^{81}Br$ , [voir tableau 2](#)), soit :  $AI = (Abondance\ du\ pic\ M+1\ ou\ M+2) / (Abondance\ du\ pic\ M)$ , si c'est un ion moléculaire  $M^+$  de rapport m/z égal à M (masse moléculaire) qui est formé.

Dans le cas du pic à M+1, les probabilités d'abondance des différents éléments étant indépendantes on peut écrire :  $AI = (1,10\ nombre\ de\ C)\% + (0,015\ nombre\ d'H)\% + (0,37\ nombre\ d'N)\%$ . Généralement l'abondance isotopique du pic M+1 donne directement le nombre de carbones présents dans la molécule. L'abondance de l'ion isotopique M+1 peut être supérieure à l'abondance de l'ion moléculaire M si le nombre de carbones est supérieur à 100. Plusieurs pics peuvent être apparents ; on parle alors d'amas isotopique.

Si la molécule contient dans sa formule un atome de chlore, le spectre de masse présente deux ions moléculaires séparés de 2 Th et d'intensités relatives 3 ( $^{35}Cl$ ) et 1 ( $^{37}Cl$ ). Si l'ion moléculaire contient dans sa formule n (>1) atomes de chlore, le spectre présente (n+1) ions  $M^+$  séparés de 2 Th et dont l'intensité relative est obtenue en développant le polynôme  $(a + b)^n$ , où a et b sont les intensités relatives des pics (a = 3 et b = 1 pour le chlore). Si par exemple la formule brute de l'ion  $M^+$  contient 2 chlores ([Fig. 1a](#)), le nombre d'ions formant le massif moléculaire est de 3 et les intensités relatives de ces ions sont :  $3^2(a^2)$ ,  $2 \times 3 \times 1$  (2 ab),  $1^2(b^2)$  soit, 9, 6 et 1. La probabilité de trouver 2  $^{35}Cl$  dans

la molécule est de 32, la probabilité de trouver 1  $^{35}\text{Cl}$  et 1  $^{37}\text{Cl}$  est de  $2 \times 3 \times 1$ , la probabilité de trouver 2  $^{37}\text{Cl}$  est de  $1^2$ .

Si la molécule contient dans sa formule un atome de brome, le spectre de masse présente deux ions moléculaires séparés de 2 Th et d'intensité relative identique. Si la molécule contient plusieurs atomes de brome le raisonnement décrit pour le chlore peut être appliqué. Dans ce cas les intensités relatives des deux isotopes sont égales ( $a = b = 1$ ).

**Tableau 2 : Masses nominale, exacte, moyenne (uma).**

Atome	Masse nominale	Masse exacte	Abondance relative (%)	Masse moyenne
H	1	1,0078	99,985	1,008
	2	2,0141	0,015	
C	12	12,0000	98,900	12,011
	13	13,0034	1,100	
N	14	14,0031	99,630	14,007
	15	15,0001	0,370	
O	16	15,9949	99,760	15,999
	17	16,9991	0,040	
	18	17,9992	0,200	
P	31	30,9738	100	30,974
S	32	31,9721	94,750	32,066
	33	32,9715	0,789	
	34	33,9679	4,440	
	36	35,9671	0,021	
Cl	35	34,9689	75,770	35,453
	37	36,9659	24,230	
Br	79	78,9183	50,690	79,904
	81	80,9163	49,310	

## Apport de la fragmentation dans la détermination de la structure moléculaire

Le spectre de masse obtenu grâce à la source d'impact électronique ou dans une chambre de collision comporte des pics d'intensité différente à des rapports  $m/z$  différents. Les ions fils sont formés, soit après perte logique de fragments neutres ou radicalaires (tableau 3), soit suivant des mécanismes de fragmentation bien connus comme les ruptures dites  $\alpha$  ou  $i$  (figure 5) :

- *la rupture  $\alpha$*  : est une rupture homolytique de la liaison initiée par le radical, avec migration du site radicalaire et sans migration de la charge (figure 5a). La réactivité des sites suit l'ordre des basicités de Lewis ( $N > S, O, C_6H_5 > F, Cl, Br, I$ ). En illustration, la figure 1b montre le spectre de la lidocaïne obtenu par impact électronique. Le pic de base, au rapport  $m/z$  86, apparaît après une rupture  $\alpha$  initiée par la charge positive positionnée sur l'azote (figure 5a).

- *la rupture  $i$*  : est une rupture hétérolytique de la liaison initiée par la charge, avec déplacement de la charge. L'effet inductif est d'autant plus prononcé que l'atome Y est électronégatif ( $F > Cl > Br > I > O, S \gg N, C$ ) (figure 5b). Cette rupture est moins favorable que la rupture  $\alpha$  car elle nécessite la migration de la charge. Dans ce cas, les ruptures qui interviennent préférentiellement sont celles qui conduisent à la formation des ions fils stabilisés par effet mésomère ou inducteur ou par perte de l'électron célibataire.

De plus, l'apparition de fragment neutre (potentiel d'ionisation élevé) par réarrangement de la molécule est largement favorisée aussi bien en impact électronique qu'en mode CID :

- *le réarrangement d'hydrogène* : peut intervenir en présence d'une fonction alcool. Il aboutit à la perte d'une molécule d'eau (figure 5c).

- *le réarrangement de Mc Lafferty* : consiste dans le transfert d'un atome d'hydrogène sur un groupement polaire. Il nécessite la présence d'une double liaison et d'un hydrogène en  $\gamma$  de cette insaturation (figure 5d).

- *le réarrangement rétro-Diels-Alder (RDA)* : intervient quand une double liaison est présente dans un cycle à 6 atomes de carbone, comme le montre la figure 5e. Si le potentiel d'ionisation (PI) de R est élevé ( $R = H, PI = 10,5$ ), on dit qu'il y a rétention de la charge ; dans ce cas, la fragmentation par RDA intervient après deux ruptures  $\alpha$  successives. Dans le cas contraire ( $R = C_6H_5, PI = 9,1$ ), on dit qu'il y a migration de la charge ; la fragmentation par RDA intervient après une rupture  $\alpha$  suivie d'une rupture  $i$ . La perte de fragment neutre s'accompagne, en principe, d'une conservation de la parité entre l'ion parent et l'ion fils, à condition que le fragment neutre ne comporte pas d'atome d'azote dans sa structure (règle de l'azote).

Tableau 3 : Pertes communes de radicaux et molécules neutres.

Perte (uma)	Nature	Commentaires
1	H	Hydrogène radicalaire
15	CH <sub>3</sub>	Radical méthyle
18	H <sub>2</sub> O	Alcool, acide
19	H <sub>2</sub> O + H	Alcool, acide
27	HCN	Amine aromatique simple
28	CO	Quinone, phénol simple, hétérocycle oxygéné
29	C <sub>2</sub> H <sub>5</sub>	Dérivé éthylé
29	CO + H	Quinone
31	OCH <sub>3</sub>	Ester méthylique
33	HS	Dérivés soufrés
34	H <sub>2</sub> S	Thiols
35	<sup>35</sup> Cl	Chlorure d'alkyle, d'acide, chlorhydrate
36	H <sup>35</sup> Cl	Chlorure d'alkyle, d'acide, chlorhydrate

Figure 5a : Principaux ruptures et réarrangements : (a) rupture  $\alpha$  survenant dans la source d'ionisation par impact électronique pour la lidocaïne, apparition de l'ion fragment à  $m/z$  86.

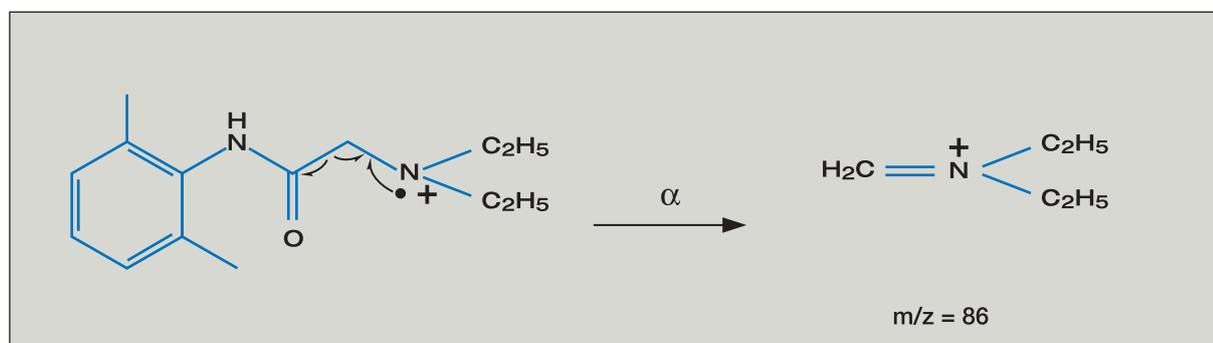


Figure 5b : Principaux ruptures et réarrangements : (b) rupture i.

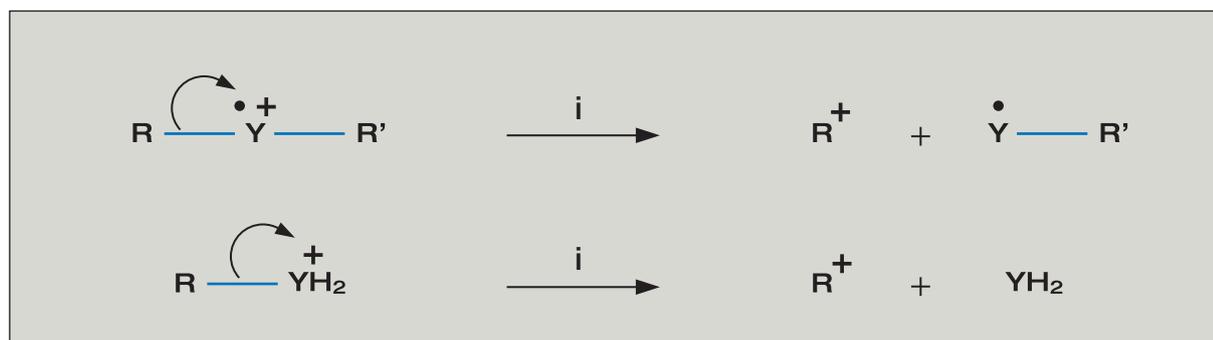


Figure 5c : Principaux ruptures et réarrangements : (c) réarrangement d'hydrogène, élimination d'une molécule d'eau.

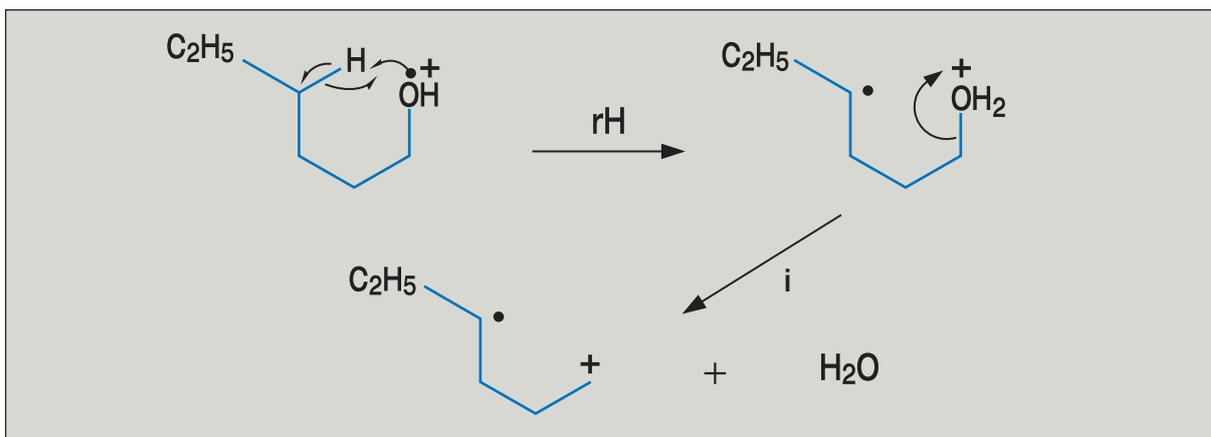


Figure 5d : Principaux ruptures et réarrangements : (d) réarrangement de McLafferty, élimination de molécule neutre.

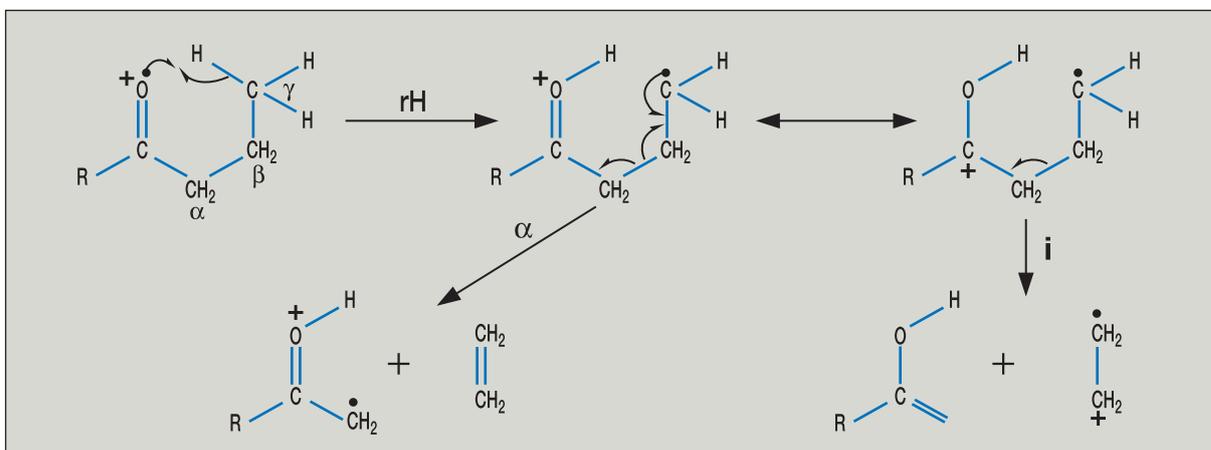
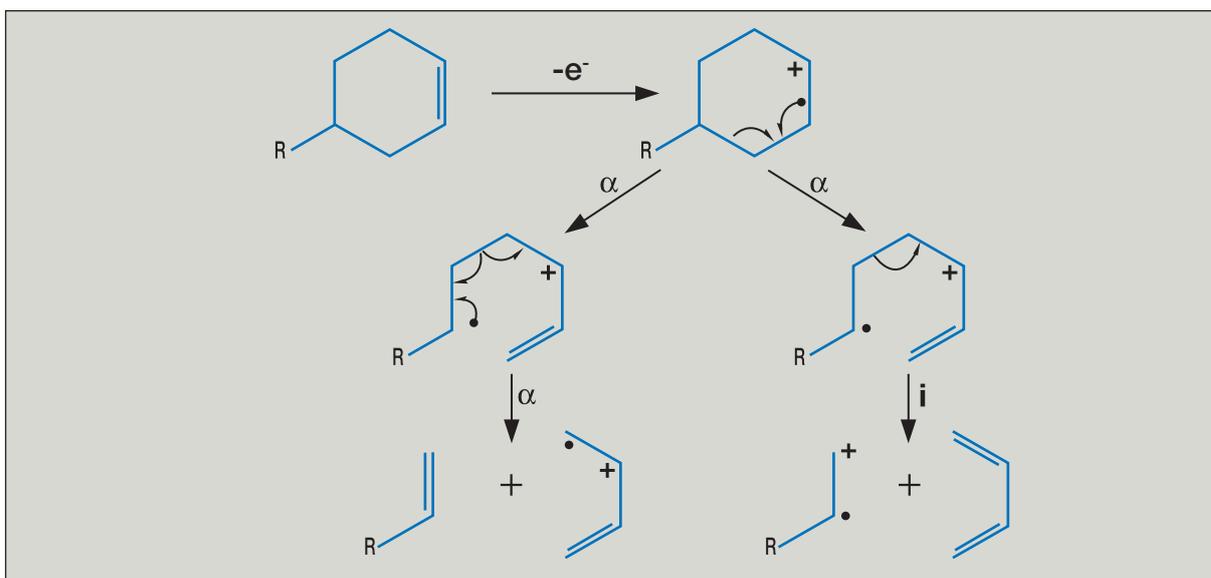


Figure 5e : Principaux ruptures et réarrangements : (e) réarrangement «rétro-Diels-Alder», élimination de molécule neutre.



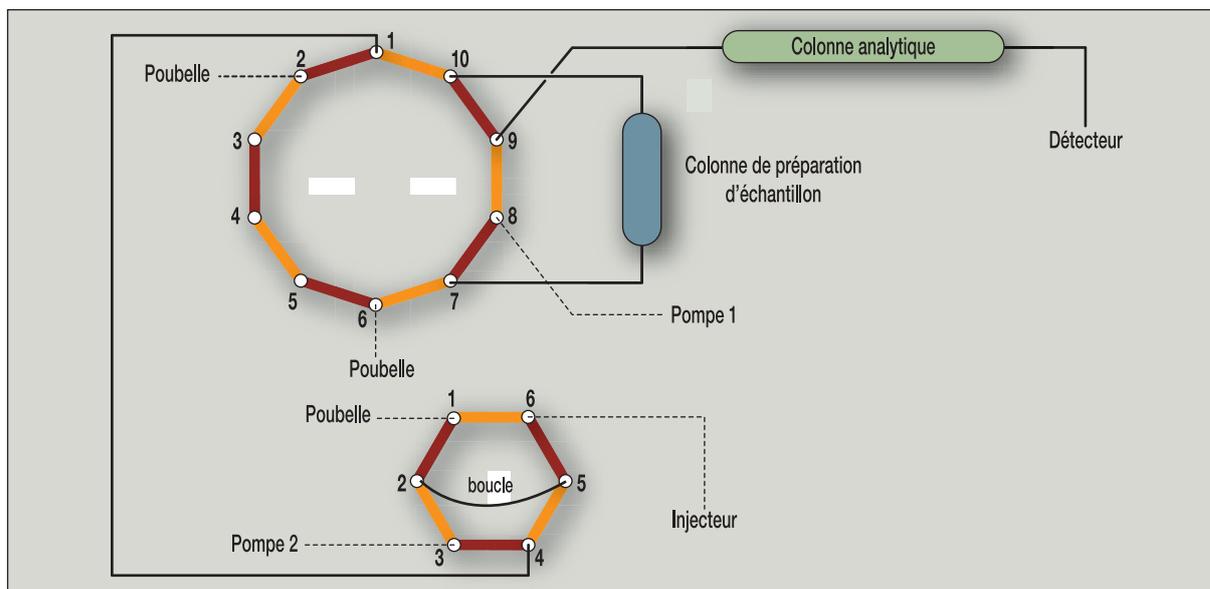
## Couplage des méthodes séparatives avec la spectrométrie de masse

Les méthodes séparatives (chromatographies en phases liquide et gazeuse) sont fréquemment couplées à la spectrométrie de masse. Elles permettent d'introduire dans le spectromètre de masse un échantillon relativement pur et exempt de composés qui pourraient perturber l'ionisation des solutés d'intérêt dans la source. Le couplage avec la chromatographie en phase gazeuse est le plus facile. La colonne capillaire est introduite directement dans la source (impact électronique ou ionisation chimique) du spectromètre de masse puisque l'échantillon est élué de la colonne déjà sous forme gazeuse. Les spectres obtenus par ionisation par impact électronique peuvent être comparés à ceux contenus dans une bibliothèque (NIST) pour l'identification formelle du composé. La chromatographie en phase liquide (CLHP), quant à elle, est fréquemment couplée à la spectrométrie de masse par l'intermédiaire d'une source API qui permet de transformer l'échantillon sous forme gazeuse et d'ioniser les composés à analyser (figure 3). Les liquides (phases mobiles) arrivant dans la source doivent donc être facilement volatilisables et présenter des propriétés physicochimiques telles qu'ils permettent une bonne ionisation des produits. Les tampons ou réactifs dits « volatils » les plus utilisés dans les phases mobiles lors de l'utilisation d'une source ESI sont des solutions d'acétate et de formiate d'ammonium, de triéthylamine, d'acide heptafluorobutyrique, d'acide formique, d'acide acétique, d'acide trifluoroacétique (en mode positif) ou d'hydroxyde de tétraéthyl(ou tétrabutyl)-ammonium (en mode négatif), à la concentration de 10 à 100 mmol.l<sup>-1</sup>. L'utilisation d'agents de cationisation (acétate de sodium ou de potassium 20 à 50 µmol.l<sup>-1</sup>) permet la formation d'adduits sodique ou potassique dans la source. La stabilité des phases stationnaires contenues dans les colonnes de CLHP doit être compatible avec la composition des phases mobiles.

Les développements technologiques actuels en CLHP permettent d'envisager aisément une préparation d'échantillon, en ligne, par extraction liquide-solide. La figure 6 montre le type de montage habituellement utilisé. Une colonne contenant de la phase solide d'extraction liquide-solide (« colonne de préparation de l'échantillon ») est positionnée au niveau d'une vanne 10 voies. Dans la position 1-10 de cette vanne, la « pompe 2 » pousse l'échantillon dans la colonne d'extraction retenant les composés à analyser. Dans cette position, les deux circuits, « pompe 1-colonne analytique » et « pompe 2-colonne de préparation de l'échantillon » sont indépendants. La phase mobile pompée par la pompe 2 peut être judicieusement choisie afin de permettre une bonne rétention des composés à analyser et une élution (vers la « poubelle ») des impuretés contenues dans l'échantillon. Lorsque la vanne 10 voies passe en position 1-2, les composés retenus dans la colonne d'extraction sont poussés dans la colonne analytique qui les séparera. Ce système peut être couplé en aval à un spectromètre de masse de type « triple quadripôle » qui isolera les ions de rapport m/z spécifiques (en mode SIM ou MRM) des composés séparés à leur temps de rétention respectif. Ce couplage « colonne de préparation d'échantillon-colonne analytique-spectromètre de masse triple quadripôle » est classiquement utilisé en analyse quantitative en présence d'étalons internes deutérés.

Des applications de ces couplages en biologie clinique sont données dans le [Chapitre 6](#).

**Figure 6 : Couplage de l'extraction liquide-solide en ligne avec la chromatographie en phase liquide et la spectrométrie de masse.**



## Pour en savoir plus

### Ouvrages de référence

- Analyse chimique - Méthodes et techniques instrumentales modernes. F. Rouessac, A. Rouessac, D. Cruché. 6<sup>ème</sup> éd. Science Sup, Dunod, 2004.
- Appareils et méthodes en biochimie et biologie moléculaire, B. Hainque, B. Baudin et P. Lefebvre. Flammarion Médecine-Sciences, 2008.
- Current Practice of Gas Chromatography - Mass Spectrometry. W. M. A. Niessen. Chromatographic Science Series (86). CRC, 2001.
- Interpretation of Mass Spectra. F. W. McLafferty and J. Choi. 4<sup>ème</sup> éd. University Science Books, 1993.
- Liquid Chromatography : Mass Spectrometry. W. M. A. Niessen. Chromatographic Science Series (79). 2<sup>ème</sup> éd. CRC, 1999.
- Liquid Chromatography : Mass Spectrometry. W. M. A. Niessen. Chromatographic Science Series (97). 3<sup>ème</sup> éd. CRC, 2006.
- Spectrométrie de masse : Cours et exercices corrigés. E. de Hoffmann, J. Charrette et V. Stroobant. 2<sup>ème</sup> éd. Dunod, 1999.

## Ressources Internet

Agilent : <http://www.home.agilent.com/agilent/home.jsp?cc=FR&lc=fr>

Applied Biosystems : <http://www.appliedbiosystems.com/?abhomepage=eur>

Brücker : <http://www.brucker.com/>

Ciphergen : <http://www.ciphergen.com/>

Dionex : <http://www1.dionex.com/en-us/index.html> ou <http://www.lcpackings.nl/>

Millipore : <http://www.millipore.com/>

NIST : <http://www.nist.gov/index.html>

Shimadzu : <http://www.shimadzu.fr/home/>

Thermo Electron SA : <http://www.thermo.com/>

Varian : <http://www.varianinc.com/cgi-bin/nav?worldwide/websites/france/index&cid=JKPJMPIQFIH>

Waters : <http://www.waters.com/WatersDivision/International/france/index.asp>

ou <http://www.waters.com/watersdivision/Contentd.asp?ref=CEAN-5KUSS8>

**Apport de la spectrométrie  
de masse dans l'exploration  
des maladies héréditaires  
du métabolisme**

**Jean-François Benoit**

**CHAPITRE VI**

# Introduction

---

La spectrométrie de masse (SM ou MS, « Mass Spectrometry ») fait aujourd'hui partie des méthodologies reconnues et essentielles pour le diagnostic et le suivi des maladies héréditaires du métabolisme (MHM). Dans la majorité des cas, le diagnostic biochimique de ces maladies repose sur une analyse semi-quantitative la plus exhaustive possible de l'ensemble des composés appartenant à une même voie métabolique. Cette approche méthodologique permet d'établir un profil de métabolites (ou « métabolite profiling ») qui peut-être considéré comme la « signature biochimique » du déficit d'une enzyme donnée dans la voie métabolique étudiée. Par la suite, le déficit ainsi suspecté peut être confirmé par la mesure de l'activité enzymatique spécifique et/ou par une étude génétique.

Depuis le milieu des années 1960, aussitôt après la commercialisation des premiers spectromètres de masse alors couplés à la chromatographie en phase gazeuse (GC-MS pour « Gaz Chromatography coupled to Mass Spectrometry »), de nombreuses MHM ont été découvertes. La SM sépare, sous l'action de champs électriques ou magnétiques, les molécules ionisées et à l'état de gaz en fonction de leur rapport masse sur nombre de charges ( $m/z$ ). Cette technique est particulièrement adaptée à l'analyse des grandes voies du métabolisme intermédiaire dont les métabolites, présents dans les fluides biologiques, sont le plus souvent des molécules de faible masse moléculaire (< 1500 Da), facilement volatilisables et qui peuvent être aisément ionisées.

L'analyse de molécules polaires non volatiles a été rendue possible avec l'apparition de nouveaux modes d'ionisation au début des années 1980, comme le bombardement d'atomes rapides (FAB pour « Fast Atom Bombardment ») et l'ionisation à pression atmosphérique (ESI, APCI) élargissant ainsi le champ d'investigation biologique des maladies métaboliques. Enfin, dans les années 1990, la commercialisation des premiers spectromètres de masse en tandem (MS/MS) associée à l'accès de plus en plus facile aux isotopes stables a ouvert la voie à l'analyse de familles de molécules avec une préparation d'échantillon minimale et parfois même sans séparation analytique préalable. Actuellement, cette méthodologie est à la base du dépistage néonatal systématique de nombreuses MHM à partir d'une simple goutte de sang séchée déposée sur un papier buvard de type Guthrie (papier utilisé pour la détection de la phénylcétonurie).

Longtemps réservée aux structures de recherche et aux industries du pétrole et pharmaceutiques, la SM s'est miniaturisée ; son coût a considérablement diminué ; son usage s'est simplifié par le développement d'outils informatiques adaptés au milieu hospitalier, et c'est ainsi que de nos jours elle a conquis une place prépondérante et irremplaçable dans les laboratoires pour l'exploration des MHM.

## Le couplage à la chromatographie en phase gazeuse

### Historique

Le principe de la SM est de séparer sous l'influence de champs électromagnétiques des composés chargés à l'état de gaz en fonction de leur rapport  $m/z$ , c'est donc naturellement la chromatographie en phase gazeuse (GC) qui, séparant des molécules vaporisées par effet thermique et entraînées par un gaz neutre, fût la première méthode de séparation analytique à être couplée à la SM. De nos jours, la GC-MS occupe encore une place prépondérante dans nos laboratoires. Cette technique hautement résolutive autorise l'analyse directe de tous les métabolites facilement volatilisables. La première et principale application de la GC-MS dans un laboratoire de biochimie métabolique est l'analyse des acides organiques. Il existe cependant de nombreuses autres applications de cette technique parmi lesquelles on peut citer l'analyse des stéroïdes et des polyols ; une liste non exhaustive est donnée dans le [Tableau 1](#). Nous avons choisi de ne développer ici que l'étude des acides organiques.

**Tableau 1 : Liste des principales molécules analysées par spectrométrie de masse pour l'exploration biochimique des maladies héréditaires du métabolisme (P : Plasma, U : Urines, LCR : Liquide céphalo-rachidien, LA : Liquide Amniotique).**

Molécules analysées	Milieux biologiques	Pathologies explorées	GC-MS	MS/MS
Acides organiques	P, U, LCR, LA	Une quarantaine d'aciduries organiques	X	
Acides aminés	P, U, LCR, LA	Une dizaine d'aminocidopathies		X
Acylcarnitines et carnitine	P, U, LA, sang total, tissus	Une vingtaine de déficits de la $\beta$ -oxydation des acides gras et aciduries organiques		X
Acylglycines	U, LCR, LA	Certaines aciduries organiques	X	

Acides gras à très longues chaînes et plasmalogènes	P	Une dizaine de pathologies du peroxyosome	X	X X
Polyols et oses-phosphate	P, U, LCR, LA	2 pathologies de la voie des pentoses	X	X
Bases puriques et pyrimidiques	P, U, LCR, LA	Une dizaine de pathologies du métabolisme des bases puriques et pyrimidiques		X
Stérols	P, LA	6 pathologies sur la voie de synthèse du cholestérol	X	
Porphyrinogène	cellules	3 types de porphyries		X
Oligosaccharides	U, LA	Une dizaine de pathologies du lysosome		X
Neurotransmetteurs	(P), U, LCR	Pathologies du métabolisme des catécholamines et de la sérotonine		X
Vitamères des folates	P, tissus	Anomalies du métabolisme des folates		X
Cardiolipides	tissus	Syndrome de Barth		X
Coenzyme Q10	P, tissus	Déficit I ou II en coenzyme Q10		X
Homocystéine, bétaine, SAM-SAH	P, LCR, LA	Pathologies de la reméthylation et de la transulfuration		X
Créatine et guanidinoacétate	U, LCR, LA	3 pathologies du métabolisme de la créatine		X
Globotriasyl-céramide	U	Maladie de Fabry		X
Méthylmalonate	P, U, LCR, LA	Acidurie méthylmalonique	X	X

L'analyse des acides organiques (AO) par GC-MS a historiquement démontré l'intérêt des techniques de mesure de masses pour l'établissement de profils métaboliques. Dès la commercialisation des premiers spectromètres de masse couplés à la GC, en 1966, Tanaka découvrit une maladie jusqu'alors inconnue : l'acidurie isovalérique (déficit en isovalérylCoA-déshydrogénase). Il l'identifia et en comprit le mécanisme physiopathologique en caractérisant, dans les urines de patients, l'isovalérylglycine, forme conjuguée de l'isovalérylCoA, lui-même produit du catabolisme de la leucine. L'évolution exponentielle du nombre de nouvelles aciduries organiques découvertes par la suite illustre les apports et l'intérêt de cette technique. Les aciduries organiques constituent un groupe hétérogène d'erreurs innées du métabolisme qui ont comme caractéristique commune l'accumulation dans les urines, et à un moindre degré dans tous les liquides biologiques, d'AO, métabolites produits du catabolisme des acides aminés (AA). Le [Tableau 2](#) reprend la liste des principales aciduries organiques. Un AO est un composé qui possède un groupement carboxylique de formule générale R-COOH où R correspond à un radical organique possédant éventuellement d'autres groupements fonctionnels tels que alcool, cétone ou encore amine. Les AO sont des métabolites intermédiaires de nombreuses voies métaboliques. Leur analyse explore principalement le catabolisme des AA, mais aussi, par leur structure chimique commune, tout ou partie du métabolisme des oses, des acides gras (AG), des acides nucléiques ou encore des stéroïdes.

**Tableau 2 : Les principales aciduries métaboliques.**

Maladies	Voie métabolique	Enzyme déficitaire
Acidurie propionique Acidurie méthylmalonique	Catabolisme des valine, isoleucine, méthionine, thréonine, et des acides gras à nombre impair d'atomes de carbone	PropionylCoA-carboxylase MéthylmalonylCoA-mutase et coenzyme
Acidurie isovalérique 3-méthylcrotonylglycinurie Acidurie 3-méthylglutaconique Acidurie 3-hydroxy 3-méthylglutarique	Catabolisme de la leucine	IsovalérylCoA-déshydrogénase 3-méthylcrotonylCoA-carboxylase 3-méthylglutaconylCoA-hydratase HMGC0A-lyase

Acidurie méthylbutyrique	Catabolisme de l'isoleucine	2-méthylbutyryl-CoA déshydrogénase
Déficit en $\beta$ -cétotiolase	Catabolisme de l'isoleucine et des corps cétoniques	2-méthylacétoacétylCoA-thiolase
Isobutyrylgycolurie	Catabolisme de la valine	IsobutyrylCoA-déshydrogénase
Acidurie glutarique de type I	Catabolisme des lysine, hydroxyllysine et tryptophane	GlutarylCoA-déshydrogénase
Acidurie homogentisique (alcaptonurie)	Catabolisme de la tyrosine	Homogentisique-dioxygénase
Maladie de Canavan	Métabolisme du N-acétylaspartate	N-asparto-acylase
Acidurie fumarique Acidurie $\alpha$ -cétoglutarique	Cycle de Krebs	Fumarase $\alpha$ -cétoglutarate-déshydrogénase
Acidurie pyroglutamique	Cycle du glutathion	Glutathion-synthase
Hyperoxaluries de types I et II	Métabolisme du glyoxylate	Glyoxylate-aminotransférase et D-glycérate-déshydrogénase
Acidurie malonique	Déficit mitochondrial mal connu	MalonylCoA-décarboxylase
Acidurie mévalonique	Synthèse des isoprényles	Mévalonate-kinase
Acidurie D-2-hydroxyglutarique Acidurie L-2-hydroxyglutarique	Métabolisme de l' $\alpha$ -cétoglutarate	D-2-hydroxyglutarate-déshydrogénase L-2-hydroxyglutarate-déshydrogénase
Déficit multiple en carboxylases Déficit en biotinidase	Métabolisme de la biotine	Holocarboxylate-synthase Biotinidase
Acidurie glutarique de type II	Déficit multiple en déshydrogénases à FAD	ElectronTransfert Flavoprotéine (ETF) et ETF-déshydrogénase

## La Chromatographie des acides organiques (CAO)

### Préparation de l'échantillon

Les urines constituent un prélèvement de choix pour l'analyse des AO, de préférence celles émises à la suite d'un épisode aigu. Si les urines ne sont pas disponibles, tout autre liquide biologique peut être analysé, le plasma ne constituant généralement pas une matrice d'intérêt et ne devrait être analysé, sauf situations exceptionnelles, qu'en ultime recours. Des diagnostics prénataux d'aciduries organiques sont réalisés par analyse du liquide amniotique. Les urines (une miction ou un recueil partiel) doivent être recueillies sans conservateur et conservées à 4°C si l'analyse est rapidement réalisée ; au-delà de quelques heures il convient de les congeler. La mauvaise conservation et notamment la prolifération microbienne peuvent considérablement perturber le profil chromatographique. L'analyse est réalisée après extraction et dérivation des AO contenus dans les urines.

Classiquement, 1 à 2 ml d'urines (QSP 1 mmole de créatinine) sont extraits par un volume équivalent de solvant organique de polarité intermédiaire, comme l'acétate d'éthyle, par extraction liquide/liquide. Après acidification des urines (pH < 2) et saturation en sels, les AO présents sous forme neutre (carboxyle protoné) sont préférentiellement extraits par la phase organique. L'extraction est ainsi répétée plusieurs fois ; les phases organiques collectées et réunies sont évaporées à sec avant d'être dérivées puis analysées. L'évaporation du solvant organique est obtenue par un chauffage modéré (< 45°C pour ne pas perdre les AO volatils), combiné à l'action du vide ou plus simplement à celle d'un courant d'azote. La réaction de dérivation qui suit la mise à sec confère un caractère volatil et thermostable à tous les composés à analyser, caractère indispensable pour l'analyse en GC. La méthode de dérivation du groupement carboxyle la plus utilisée est la formation d'éther de triméthylsilyle (TMS) obtenu par le chauffage des AO avec le N,O-bis(triméthylsilyl)trifluoroacétamide à 1 % de triméthylchlorosilane en milieu anhydre (pyridine). La formation d'éther de TMS n'est pas la méthode de dérivation optimale pour tous les composés d'une CAO, ainsi les acylglycines (forme de conjugaison des AO par la glycine) ont un faible rendement de dérivation et donnent deux pics chromatographiques (mono et di-TMS). D'autres méthodes de dérivation comme la formation d'ester de méthyle sont possibles et couramment utilisées dans le cas de l'étude des AG par exemple.

### Appareillage

Les colonnes analytiques utilisées pour la CAO sont des colonnes capillaires d'une longueur de 30 m x 0,25 mm de diamètre interne, de type DB-5MS, dont la phase stationnaire est constituée de 5 % phényle et 95 % diméthylpolysiloxane. La GC sépare les AO qui, en sortie de colonne, sont ionisés et fragmentés par impact électronique. L'ensemble des ions ainsi produits à partir d'une molécule pure donne un spectre de masse de référence caractéristique de cette molécule. La production de spectres de masse est utilisée pour l'identification absolue des composés par comparaison avec un spectre de référence. L'impact électronique est la source d'ionisation utilisée pour cette application ; son principe est le suivant : des électrons émis par un filament et accélérés par une différence de potentiel (énergie cinétique optimale d'environ 70 eV) percutent le flux des

molécules séparées en fin de colonne et déjà volatilisées. L'impact provoque l'arrachement d'un électron périphérique de la molécule pour former un ion  $M^+$  instable qui se fragmente alors en ions secondaires stables et en fragments neutres. Cette méthode d'ionisation directe offre l'avantage d'un bon rendement d'ionisation (environ 1 molécule sur 1000) et d'une fragmentation importante, ce qui fournit une source d'information structurale importante. En standardisant l'énergie cinétique des électrons, l'impact électronique procure une détection universelle des composés, permettant d'un laboratoire à l'autre l'utilisation des mêmes bibliothèques de spectres, spectres de masse alors assimilables à une empreinte structurale de la molécule analysée. Si l'interprétation des spectres de masse requiert une formation pointue, pour l'usage courant il n'est pas nécessaire de connaître les mécanismes fins de fragmentation des molécules, celles-ci pouvant être identifiées par comparaison avec des spectres de référence (ensemble des  $m/z$  et leur proportion relative). Des logiciels exploitant des bibliothèques commerciales ou privées peuvent assister l'opérateur pour extraire à chaque temps de rétention les spectres de masse et proposer une identification des composés en fonction d'un score de similitude, par exemple avec la bibliothèque du « National Institute of Standards and Technology » et son programme de déconvolution et d'identification automatique (AMDIS).

De nombreux fabricants de GC proposent pour cette application des spectromètres de masse de paillasse qui sont soit de simples quadripôles (Agilent, Shimadzu, Thermo, Waters...), soit des trappes d'ions (Thermo, Varian...), généralement d'encombrement inférieur à celui de la GC et pilotés par le même logiciel.

### **Analyse qualitative et quantitative**

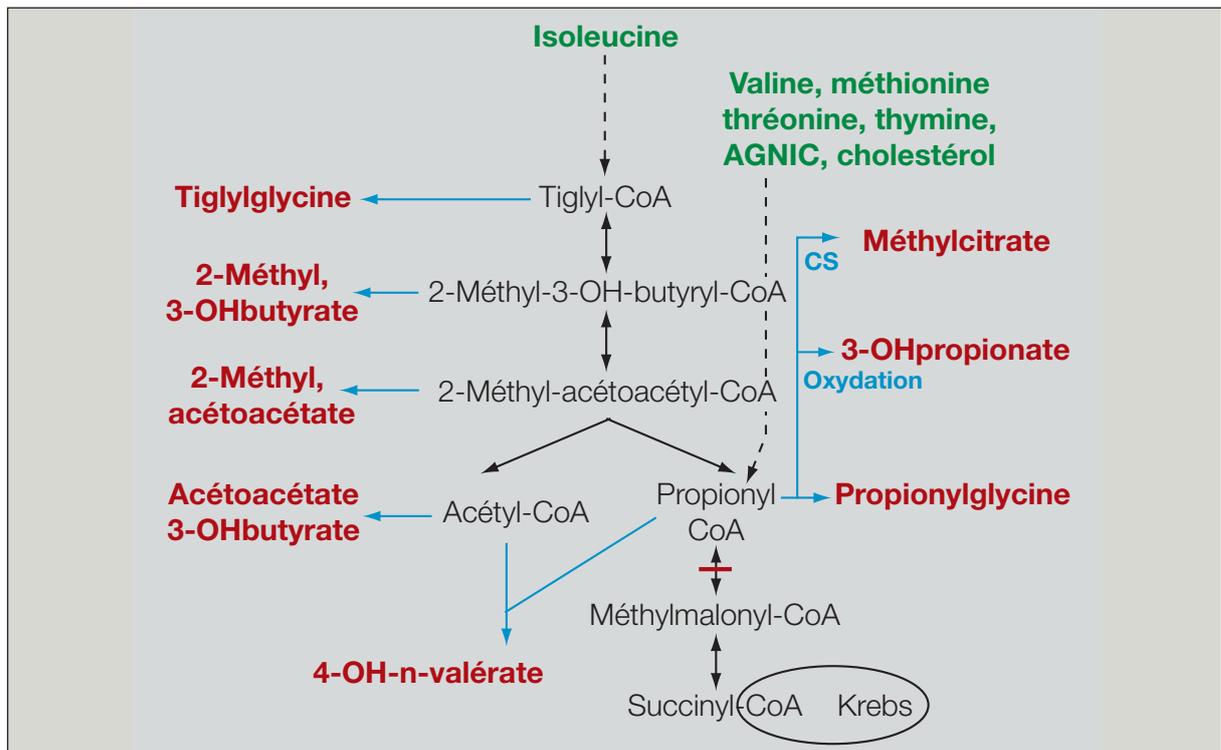
L'analyse de recherche non ciblée de métabolites anormaux (« metabolite profiling ») requiert le maximum d'information structurale sur les composés analysés afin de pouvoir les identifier. Dans ce cas où l'approche qualitative est privilégiée, le spectromètre de masse est réglé sur un mode de balayage de la gamme de masses définie, permettant l'analyse successive et très rapidement répétée de chacune des masses ( $m/z$ ) de la gamme à étudier afin d'obtenir les spectres complets de chaque composé élué. Les analyseurs quadripolaires qui possèdent une vitesse de balayage très rapide sont particulièrement adaptés pour ce mode d'analyse. L'identification des composés repose alors sur le temps de rétention et le spectre de masse. L'ajout d'un étalon interne couplé à des gammes externes des composés d'intérêt permet une approche semi-quantitative suffisante pour l'orientation du diagnostic. Si l'analyse qualitative est toujours réalisée en première intention dans un but diagnostique, le dosage quantitatif restreint à quelques métabolites ciblés est le plus souvent indispensable à la confirmation du diagnostic et au suivi biologique de certaines pathologies (ex : dosage de l'acide méthylmalonique dans les acidémies méthylmaloniques). Dans ce cas, c'est l'approche quantitative qui est privilégiée. Idéalement, l'étalon interne est alors un isotope stable de la molécule à doser (méthode de type dilution isotopique) et seuls quelques ions sont analysés. Généralement ils sont au moins au nombre de deux par molécule, un ion très spécifique de la molécule (ion qualifiant) pour identifier la molécule et un second généralement le plus intense du spectre de masse sert à la quantification (ion quantifiant). Il en va de même pour l'étalon interne. Ce mode d'analyse ne ciblant qu'un ou deux ions par molécule (mode SIM, « Single Ion Monitoring ») entraîne inéluctablement une perte considérable d'information structurale et ne peut

s'appliquer que pour des métabolites connus et une méthode chromatographique robuste. L'identification ne repose plus alors que sur le temps de rétention (co-élution du composé à doser et de son isotope stable) et sur le rapport ion qualifiant/ion quantifiant qui doit être constant. La quantification est fonction du rapport des surfaces du composé et de son étalon interne (isotope stable). Le choix de la méthode d'extraction et de l'étalon interne est prépondérant pour l'approche quantitative.

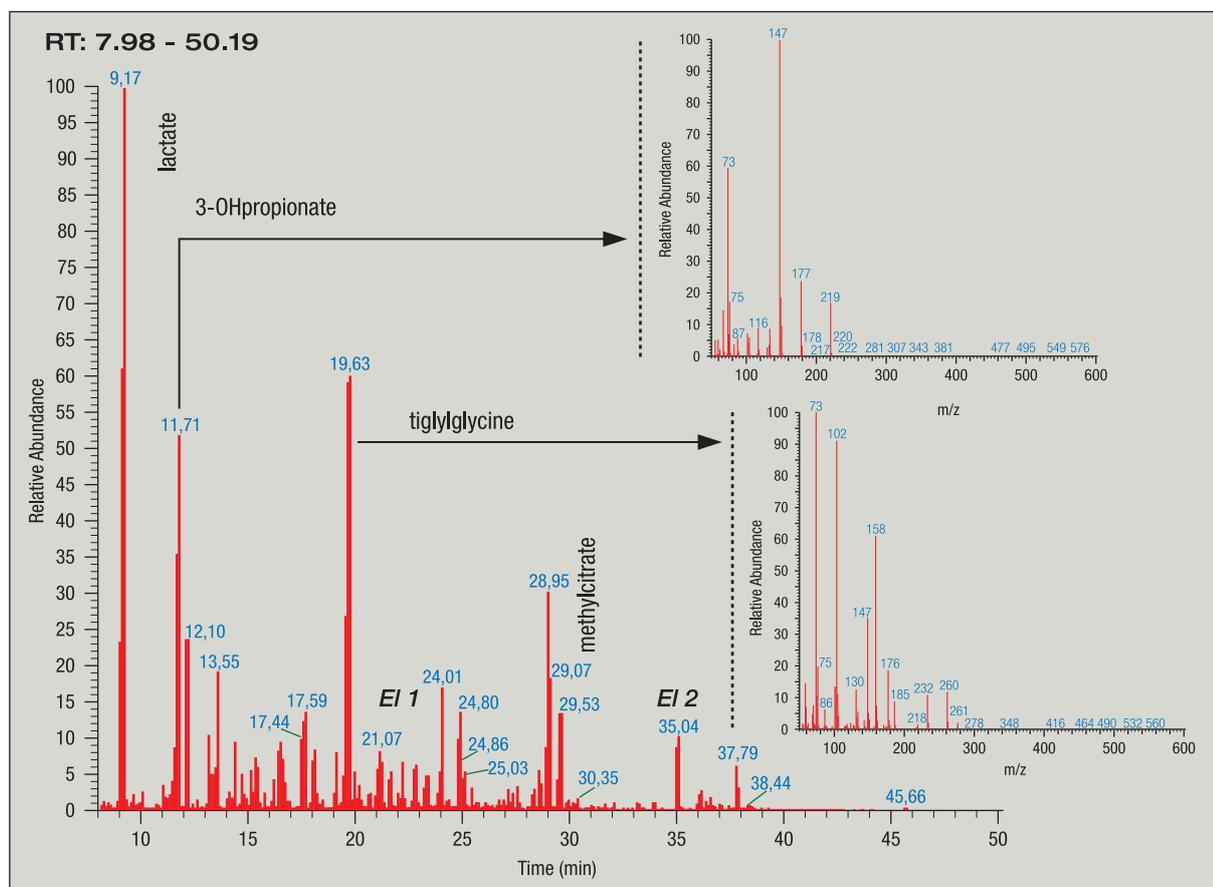
### Interprétation d'une CAO

Certains AO, pratiquement indétectables chez le sujet sain, n'apparaissent qu'à la suite d'un déficit enzymatique, soit parce qu'ils constituent des formes d'élimination majoritaire de métabolites accumulés comme les dérivés conjugués par la glycine (ex : la propionylglycine ou la tiglylglycine dans l'acidurie propionique), soit parce qu'ils constituent des culs de sac métaboliques produits par des voies secondaires (ex : méthylcitrate produit par la citrate-synthase dans l'acidurie propionique, voir la figure 1). De nombreux traitements médicamenteux sont susceptibles d'interférer avec le résultat, parmi lesquels on peut citer l'exemple de l'acide valproïque qui génère une dizaine de pics interférant et perturbe le métabolisme du coenzyme A et de la carnitine. De même, les apports alimentaires sont susceptibles de modifier considérablement le profil de la CAO (exemple des laits enrichis en triglycérides à chaînes moyennes). Le Tableau 2 regroupe les principales aciduries organiques à ce jour identifiées et dont le diagnostic repose, au moins en partie, sur l'analyse des AO par GC-MS. Ce tableau n'inclut pas toutes les MHM pour lesquelles les résultats du profil des AO constituent un élément d'orientation du diagnostic tels que les déficits de la bêta-oxydation des AG.

Figure 1A : La voie du propionate et le mécanisme de formation des acides organiques observé au cours d'une acidurie propionique par déficit en propionylCoA-carboxylase.



**Figure 1B : Exemple d'un profil d'acides organiques obtenu par GC-MS au cours d'une acidurie propionique.**



Les laboratoires effectuant ces analyses sont théoriquement affiliés à l'ERNDIM (« European Research Network for Evaluation and Improvement of Screening, Diagnosis and Treatment of Inherited Disorders of Metabolism ») qui propose des contrôles de qualité des AO, qualitatifs et quantitatifs, ainsi que des tests d'évaluation globale de la compétence analytique et de l'interprétation biologique à l'origine d'une mise à niveau régulière des outils analytiques et des compétences des biologistes.

## Le couplage à la chromatographie en phase liquide

De nombreux métabolites ne sont cependant pas analysables par GC-MS comme par exemple les composés polaires, peu volatils ou thermiquement fragiles ; de plus les méthodes d'analyse par GC-MS présentent comme désavantage la lourdeur de la préparation des échantillons (extraction, dérivation), très consommatrice de temps technique. C'est pourquoi les couplages à la chromatographie en phase liquide (LC-MS) ou à l'électrophorèse capillaire (EC-MS) ont été développés.

## La source d'ionisation en LC-MS

Elle doit non seulement ioniser les molécules mais également les faire passer à l'état de vapeur, le tout à partir d'un liquide dans un environnement à pression atmosphérique. Pour cela, les sources d'ionisation font appel au phénomène de désorption, c'est-à-dire d'ionisation et de vaporisation simultanées par un transfert extrêmement rapide d'une grande quantité d'énergie à la molécule ; le phénomène doit être très rapide pour éviter la dégradation thermique. Les premières sources éprouvées pour l'étude du métabolisme par ces méthodes d'ionisation dites indirectes utilisaient le bombardement des molécules dissoutes dans une matrice susceptible de céder des protons par des atomes de gaz rare à haute énergie cinétique (FAB). Cependant, la préparation de l'échantillon ne permet pas de couplage direct et reste fastidieuse, limitant de beaucoup l'usage de cette source pour des analyses dites de routines à haut débit. L'électronébulisation (ESI, « ElectroSpray Ionization ») a révolutionné l'usage de la SM en permettant un couplage direct avec une chromatographie en phase liquide (LC, « Liquid Chromatography »). Le principe, en apparence simple, consiste à faire passer l'échantillon en sortie de colonne analytique au travers d'un capillaire métallique, très fin, porté à un très haut potentiel. Le mécanisme exact permettant la production d'ions est encore largement discuté. Le champ électrique intense (quelques kV) appliqué entre le capillaire par lequel arrive le liquide et l'entrée du spectromètre de masse proprement dit va entraîner la formation d'un spray (théorie du cône de Taylor). La tension imposée au capillaire métallique crée un mouvement électrophorétique des espèces chargées dans la solution concentrant les charges à la surface du liquide. Les forces électrostatiques ainsi générées sont à l'origine de la formation des gouttelettes du spray. La désolvatation progressive par l'action combinée d'un gaz neutre à contre-courant et d'un chauffage ainsi que l'explosion en de multiples micro-gouttelettes par répulsion coulombienne font partie des hypothèses évoquées pour aboutir finalement, par désorption, à la production d'un ion en phase gazeuse. Les ions ainsi générés à pression atmosphérique entrent à travers un orifice de taille très réduite (< 1 mm) dans une première enceinte où règne un vide partiel puis via un système de focalisation dans le premier quadripôle qui se trouve lui dans un vide poussé, vide assuré par de puissantes pompes turbo-moléculaires. Contrairement à l'impact électronique précédemment cité, avec ce mode d'ionisation dit « doux », il n'y a pratiquement pas de fragmentation en source et il est possible d'obtenir des espèces multichargées. Une variante, l'ionisation chimique à pression atmosphérique (APCI, « Atmospheric Pressure Chemical Ionization ») est aussi utilisée. Son principe est le suivant : l'éluat de la colonne analytique est nébulisé par un gaz et chauffé à haute température ; les vapeurs sont ionisées par une décharge de corona générée par une électrode. Un début de fragmentation en source peut toutefois survenir avec l'usage d'une source APCI. Très vite l'énorme potentiel des sources ESI et APCI placées en ligne derrière une LC et couplées à un spectromètre de masse a ouvert un très vaste champ d'analyse de métabolites. En effet, les avantages de ces deux sources sont d'autoriser un couplage direct avec un système de séparation en phase liquide (LC ou EC), de générer des ions à l'état de vapeur physiquement intacts (sans fragmentation en source) et potentiellement multichargés. En revanche, leurs inconvénients sont liés à l'efficacité d'ionisation qui est molécule-dépendante avec un risque de compétition pour prendre la charge entre différents composés d'où une influence parfois très importante de la matrice biologique.

## La spectrométrie de masse en tandem (MS/MS)

Au cours de ces dix dernières années, la SM en tandem couplée à la chromatographie liquide (LC-MS/MS) est devenue la méthodologie la plus usitée pour l'analyse en séries de petites molécules d'intérêt biologique (ou pharmacologique) dans les liquides biologiques (une liste non exhaustive des applications est donnée dans le [Tableau 1](#)). La MS/MS offre trois avantages majeurs : 1) sa capacité à étudier de nombreuses molécules appartenant ou non à la même famille structurale en une seule analyse avec une préparation d'échantillon minimale et parfois sans séparation analytique préalable ; 2) sa capacité à détecter de nouvelles maladies par la mise en évidence de métabolites jusqu'alors non détectables par d'autres méthodologies ; 3) le caractère automatisé de cette technologie offrant la possibilité de grandes séries d'analyses. Cependant, il subsiste des difficultés pour le dosage des métabolites dans les liquides biologiques, difficultés liées surtout au mode d'ionisation et à l'effet propre de la matrice biologique. Les spectromètres de masse en tandem qui équipent nos laboratoires sont essentiellement des triples quadripôles, résultant de l'association de deux analyseurs quadripolaires en série, séparés par une cellule de collision souvent constituée d'un quadripôle plus court. Cet outil offre quatre modes d'analyse différents selon l'information recherchée : 1) génération d'ion(s) fils (mode SIM) : l'ion caractéristique (généralement l'ion pseudo-moléculaire) de la molécule à analyser est sélectionné dans le 1<sup>er</sup> quadripôle (Q1) en focalisant sur son m/z ; il est fragmenté dans la chambre de collision et l'ensemble des ions fils générés sont analysés par un balayage d'une gamme de masses sur le Q3. Ce mode est plus spécifiquement dédié aux développements de méthodes afin de sélectionner les ions fils caractéristiques qui seront retenus pour définir les transitions spécifiques utilisées en mode MRM (cf. point 4) ; 2) le profil d'ion parents (mode PIS, « Precursor Ion Scan ») : le Q1 balaie une gamme de masses tandis que le Q3 est focalisé sur un seul rapport m/z. Tous les ions générés en source et capables de donner, après collision en Q2, un fragment commun à la famille des molécules étudiées, seront donc pris en compte pour constituer un spectre de masse. C'est le mode d'analyse utilisé pour étudier des molécules présentant une parenté de structure et générant un ion fils commun comme dans le cas des acylcarnitines ; 3) la perte de neutre (mode « Neutral Loss ») : le Q1 et le Q3 balayent une gamme de masses simultanément et avec un décalage de masse constant. Le spectre établi présentera alors tous les ions parents capables de se fragmenter en générant un neutre de masse égale au décalage imposé. Une application de ce mode est l'analyse des AA neutres et acides dont la fragmentation génère une perte de neutre commune ; 4) le mode MRM (« Multiple Reaction Monitoring ») : des couples ion parent/ion fils caractéristiques des molécules à analyser sont préalablement sélectionnés. Le Q1 et le Q3 sont simultanément focalisés pendant une fraction de seconde sur le couple d'ions (ion parent en Q1 et ion fils en Q3 spécifique de chaque molécule à analyser. Contrairement au mode SIM utilisé avec un simple quadripôle, le mode MRM donne une sensibilité supérieure et davantage de garantie d'identification par la sélection d'une transition (ion précurseur/ion fils) spécifique de la molécule étudiée ; ceci autorise des chromatogrammes plus courts ainsi qu'une préparation d'échantillon moins élaborée. C'est le mode le plus couramment utilisé. Il est donc possible d'obtenir, en utilisant les modes PIS ou Neutral Loss, des profils métaboliques de familles complètes de molécules. C'est le principe utilisé pour le dépistage néonatal systématique des MHM par l'analyse des acylcarnitines en PIS, et des AA en Neutral Loss. Ainsi, environ trente MHM

sont potentiellement dépistées à partir d'une simple goutte de sang séché déposée sur un papier buvard de type Guthrie en moins de 5 minutes. De nombreux pays proposent des programmes nationaux de dépistage néonatal des MHM comprenant de deux (phénylcétonurie et déficit en acylCoA-déshydrogénase des AG à chaîne moyenne) à une vingtaine de pathologies différentes par l'analyse quantitative des acylcarnitines et des AA sans séparation analytique préalable.

### **Analyse de la carnitine et des acylcarnitines par LC-MS/MS**

L'oxydation des AG dans la mitochondrie ne se déroule qu'après activation sous forme d'acylCoA ; de même, les acylCoA sont des intermédiaires de l'oxydation de la copule carbonée d'AA. Les acylCoA ne peuvent s'échanger librement entre le compartiment intramitochondrial où ils sont produits et le compartiment cytosolique par lequel ils doivent transiter pour être éliminés de la cellule en cas d'accumulation anormale. L'échange ne peut se faire qu'après conjugaison avec la carnitine (acylcarnitines) ou la glycine (acylglycines). La L-carnitine (3-hydroxy-4-triméthylammonio-butanoate) est une molécule en partie synthétisée dans le foie et le rein à partir d'AA, mais elle est majoritairement d'origine exogène, apportée par l'alimentation. Les déficits de la  $\beta$ -oxydation des AG et certaines aciduries organiques entraînent l'accumulation dans la mitochondrie d'acylCoA. La carnitine joue alors un rôle essentiel pour le transfert de ces intermédiaires dans le cytosol et leur élimination de la cellule donc pour l'homéostasie, ici fortement perturbée, du pool des acylCoA et du CoA libre. Les acylcarnitines sont ainsi retrouvées dans tous les liquides biologiques. Les différents déficits enzymatiques de l'oxydation des AG ou des AA donnent des profils d'acylcarnitines plus ou moins caractéristiques. La réalisation du profil d'acylcarnitines est l'analyse la plus spécifique pour l'orientation diagnostique des déficits de la  $\beta$ -oxydation.

Les prélèvements biologiques les plus pertinents pour l'exploration des anomalies de l'oxydation des AG sont le plasma (à conserver congelé) et le sang total absorbé sur buvard. S'il n'est pas possible d'obtenir un prélèvement sanguin (prélèvement post-mortem par exemple), des dépôts de liquide biliaire sur buvard sont réalisables. Les buvards de sang ou de bile doivent être séchés à l'air libre (au moins deux heures) avant d'être acheminés au laboratoire, sous enveloppe et à température ambiante. Il est toutefois conseillé de les congeler en cas de stockage prolongé afin d'éviter la perte des acylcarnitines à chaîne courte. Ces prélèvements doivent être réalisés prioritairement au décours d'un épisode aigu ; en effet, certains déficits de l'oxydation des AG donnent des profils d'acylcarnitines peu informatifs en période d'équilibre métabolique. A défaut, l'analyse sera réalisée après un jeûne physiologique. Les acylcarnitines du plasma (ou plus rarement du sérum) sont extraites par une solution de méthanol ou d'acétonitrile acidifié contenant des étalons internes, isotopes stables d'acylcarnitines de différentes longueurs de chaîne et à des concentrations définies. La phase organique dénature les protéines et solubilise les acylcarnitines à longue chaîne. Plusieurs méthodes d'analyses peuvent être utilisées, soit une analyse directe sans dérivation, soit une analyse après dérivation. La butylation est le mode de dérivation le plus largement utilisé ; elle est réalisée par une solution d'HCl 3 N dans du butanol à 60-65°C pendant 15-20 minutes. La méthylation en méthanol/HCl peut également être réalisée dans les mêmes conditions. Dans tous les cas l'échantillon ainsi préparé est évaporé à sec sous un courant d'azote avant d'être repris par

un mélange acétonitrile/acide formique à 0,1 % dans de l'eau ultrapure (80/20 v/v).

Qu'elles soient dérivées ou non, les acylcarnitines sont analysées en mode positif (protonation des acylcarnitines) le plus souvent sans séparation analytique (pas de colonne analytique), la chaîne CLHP ne servant qu'à délivrer un débit constant de la phase mobile et à injecter les échantillons. L'échantillon (10 à 20 µl environ) est acheminé sous un faible débit (50 µl/min environ) jusqu'à la source. En l'absence de séparation analytique, l'analyse dure moins d'une minute ; pendant la vingtaine de secondes durant laquelle l'échantillon passe dans le spectromètre de masse, la rapidité de balayage du Q1 permet de balayer à de multiples reprises la gamme de masses à analyser ( $m/z$  de 200-500 amu). Comme pour les AO, lors d'une approche diagnostique, une première analyse qualitative est obtenue en réalisant un profil d'ions parents. Toutes les acylcarnitines produisent un même ion fils, caractéristique variable en fonction du mode de dérivation utilisé (exemple  $m/z = 85$  amu pour les acylcarnitines non dérivés ou butylées) sur lequel est focalisé le Q3. Le système informatique, remontant aux ions parents filtrés dans le Q1 qui ont donné naissance à cet ion fils (figures 2 et 3), construit un profil qui s'apparente à un profil chromatographique (intensité du signal détecté en fonction des  $m/z$ ). Une approche semi-quantitative est obtenue par le calcul du rapport des surfaces de l'acylcarnitine à quantifier sur celle de son isotope stable de concentration connue utilisé comme étalon interne. Dans le cadre d'une analyse quantitative plus précise, le même échantillon est analysé en mode MRM par une méthode de dilution isotopique utilisant là aussi un isotope stable de la molécule à doser comme étalon interne. L'interprétation, comme pour les AO, ne peut se faire sans renseignements sur l'état clinique et nutritionnel du patient ainsi que sur les traitements reçus. L'ERNDIM a mis en place un contrôle de qualité inter-laboratoire qui permet, comme pour les AO, de confronter ses performances analytiques et biologiques avec les autres laboratoires spécialisés. Il faut noter toutefois que, dans le cas des acylcarnitines à chaîne courte (4 à 6 atomes de carbone), le recours à la LC permet de séparer les isobares indissociables sans séparation analytique, par exemple la succinylcarnitine et la méthylmalonylcarnitine, ou encore la butyrylcarnitine issue du catabolisme des AG et l'isobutyrylcarnitine issue du catabolisme de la valine.

La quantification précise d'acylcarnitine ou de tout autre composé en MS/MS se heurte à plusieurs problèmes comme l'existence d'isobares envisagée précédemment ou davantage encore le problème de l'effet matrice que l'on peut résumer comme un effet de compétition pour prendre la charge (ici protonation) entre le composé à analyser et l'ensemble des composés présents au même moment dans la source. Cet effet matrice peut complètement éteindre le signal d'un ion pour peu que celui-ci soit difficilement ionisable (effet de suppression spectrale). Le choix de l'étalon interne est donc crucial ; idéalement ce doit être un isotope stable de la molécule à analyser avec si possible 4 à 5 Da de plus. Il est parfois nécessaire de recourir à une séparation chromatographique afin de corriger les difficultés rencontrées en cas de molécules isobares ou d'effet matrice. Cependant l'objectif de la chromatographie est différent de celui de la CLHP traditionnelle ; la résolution parfaite du pic d'intérêt n'est pas toujours nécessaire pour la MS/MS. La séparation chromatographique limite considérablement l'effet matrice en limitant le nombre de molécules se présentant simultanément dans la source.

Figure 2 : Mode de fonctionnement d'un tripe quadripôle lors de la réalisation d'un profil d'acylcarnitines après butylation (ion fils commun  $m/z = 85$  amu).

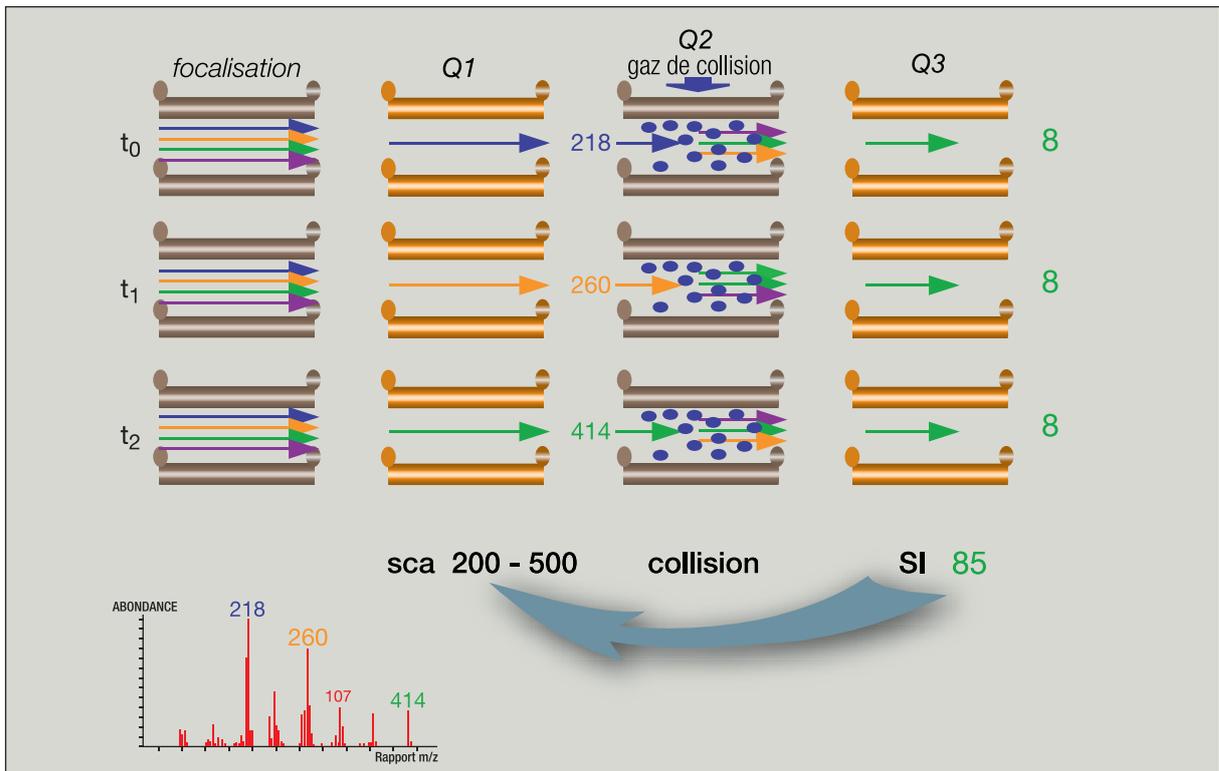
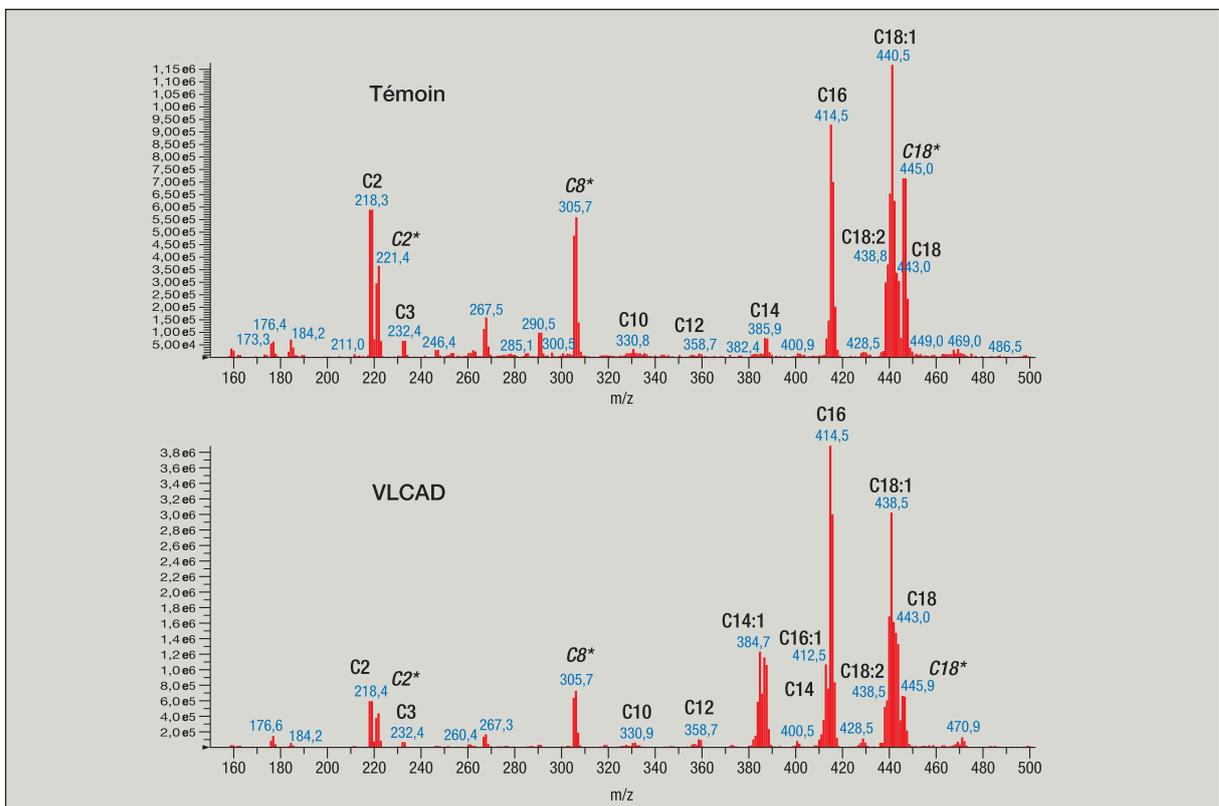


Figure 3 : Exemple de profils d'acylcarnitines obtenus chez un sujet sain (témoin) et un malade porteur d'un déficit en « Very Long Chain Acyl-CoA Dehydrogenase ».



Pour conclure, il est clair que la SM est un outil indispensable pour l'exploration biochimique des MHM. Ses principaux avantages sont en premier lieu sa spécificité et sa sensibilité, ce qui permet d'analyser de multiples métabolites à partir de faibles quantités de prélèvement et, dans le cas de la LC-MS/MS, avec une préparation de l'échantillon très simple et une grande rapidité d'analyse adaptée aux grandes séries. Les industriels ont fait évoluer les appareils afin de les rendre plus facilement utilisables, de réduire leur encombrement et leur coût tout en augmentant leurs performances (sensibilité et gamme de masses analysée). Il ne faut cependant pas croire que l'on peut tout doser dans n'importe quelle matrice biologique par les techniques de masse. Chacune a ses limites, par exemple les molécules thermolabiles pour la GC-MS et les molécules non détectables car non ionisables pour la MS/MS. Enfin, il est important de souligner que l'entretien des appareils et le développement de nouvelles méthodes analytiques requièrent un personnel (techniciens, ingénieurs et biologistes) performant et formé à la spectrométrie de masse.

## Pour en savoir plus

---

Chalmers RA. and Lawson AM. Organic acids in man. Chapman and Hall, London, 1982.

Matern D. Acylcarnitines, including in vitro loading tests, pp.171-207 in Laboratory Guide to the Methods in Biochemical Genetics. Blau N., Duran M., Gibson KM. (Eds), 2008.

Bakhtiar R. and Majumdar TK. Tracking problems and possible solutions in the quantitative determination of small molecule drugs and metabolites in biological fluids using liquid chromatography-mass spectrometry. J. Pharmacol Toxicol Methods 2007 ; 55 : 227-243.

Chace DH. Mass spectrometry in newborn and metabolic screening: historical perspective and future directions. J. Mass Spectrom. 2009 ; 44 : 163-70.

***Analyses multiples  
des protéines : multiplexage  
en phase solide (puces  
à protéines) ou liquide ;  
applications en biologie  
médicale et perspectives***

***Pr Sylvain Lehmann, Stéphane Roche,  
Anne-Marie Dupuy***

# Introduction

---

Les méthodes d'analyse et d'investigation des phénomènes biologiques bénéficient depuis plusieurs années d'approches permettant de suivre un grand nombre de paramètres biologiques en parallèle (dit en multiplexe), grâce à des systèmes en milieu liquide ou sur des supports solides communément appelés biopuces (« biochips » en terminologie anglo-saxonne)[1-3]. Ces technologies, et en particulier les biopuces, ont d'abord été développées dans le domaine de la génomique permettant à ce jour de détecter plusieurs centaines de milliers de séquences en une seule analyse [4]. Dans ce domaine comme dans celui des protéines on peut constater des améliorations spectaculaires d'échelle et de complexité grâce à l'apport des nanotechnologies, de la micro-fluidique ou encore de la bio-informatique essentielle pour traiter la grande quantité de données ainsi générées.

L'application des analyses multiplexes aux paramètres protéiques fait partie du champ de la protéomique au sens large. Ce terme d'analyse protéomique, déjà défini dans les chapitres précédents, recouvre l'ensemble des approches, en particulier destinées à la recherche de bio-marqueurs, par l'étude du protéome, c'est-à-dire tous les peptides et protéines présents dans une cellule, un tissu ou un prélèvement biologique, à un instant donné et dans une situation physiopathologique donnée. L'exhaustivité dans ce domaine est limitée par les technologies disponibles (spectrométrie de masse...) et par la complexité des prélèvements, en particulier celle des fluides biologiques, comme le sang et le LCR, qui comportent plusieurs milliers de constituants avec des différences de concentrations couvrant dix ordres de magnitude [5,6]. Dans le domaine de la protéomique on peut classiquement séparer deux types d'approches ; la première est une recherche reposant sur des méthodes comme l'électrophorèse bidimensionnelle, les profils des gels 2D et le séquençage des peptides présents ou issus de la digestion des protéines d'un échantillon biologique ([voir Chapitre 1](#)). Elle permet la recherche de bio-marqueurs sans a priori car on ne connaît pas au départ l'identité des candidats qui seront sélectionnés. Inversement, dans un deuxième groupe d'approches, on part d'un ensemble de marqueurs dont on connaît l'identité, mais sans savoir lequel ou lesquels sont d'intérêt pour répondre à une question physiopathologique, diagnostique.... Les analyses multiplexes présentées dans ce chapitre font partie de ce dernier groupe d'analyses dont les résultats devraient permettre à terme l'établissement de profils et d'algorithmes présentant un intérêt clinique. Leurs applications concernent des maladies largement répandues : cancers, maladies cardiovasculaires, maladies métaboliques et maladies neurodégénératives. Actuellement, on distingue deux types d'analyses protéiques multiplexes, celles réalisées en milieu liquide et celles sur support solide. Le format des analyseurs va de systèmes de haut débit avec de nombreux paramètres jusqu'à des applications sur un faible nombre de paramètres, au lit du patient (« point of care ») et sur des systèmes intégrés (« lab-on-chips ») [1]. Nous nous concentrerons sur les applications tournées vers la clinique des analyses en milieu liquide ou avec des biopuces dédiées à la détection de peptides et de protéines.

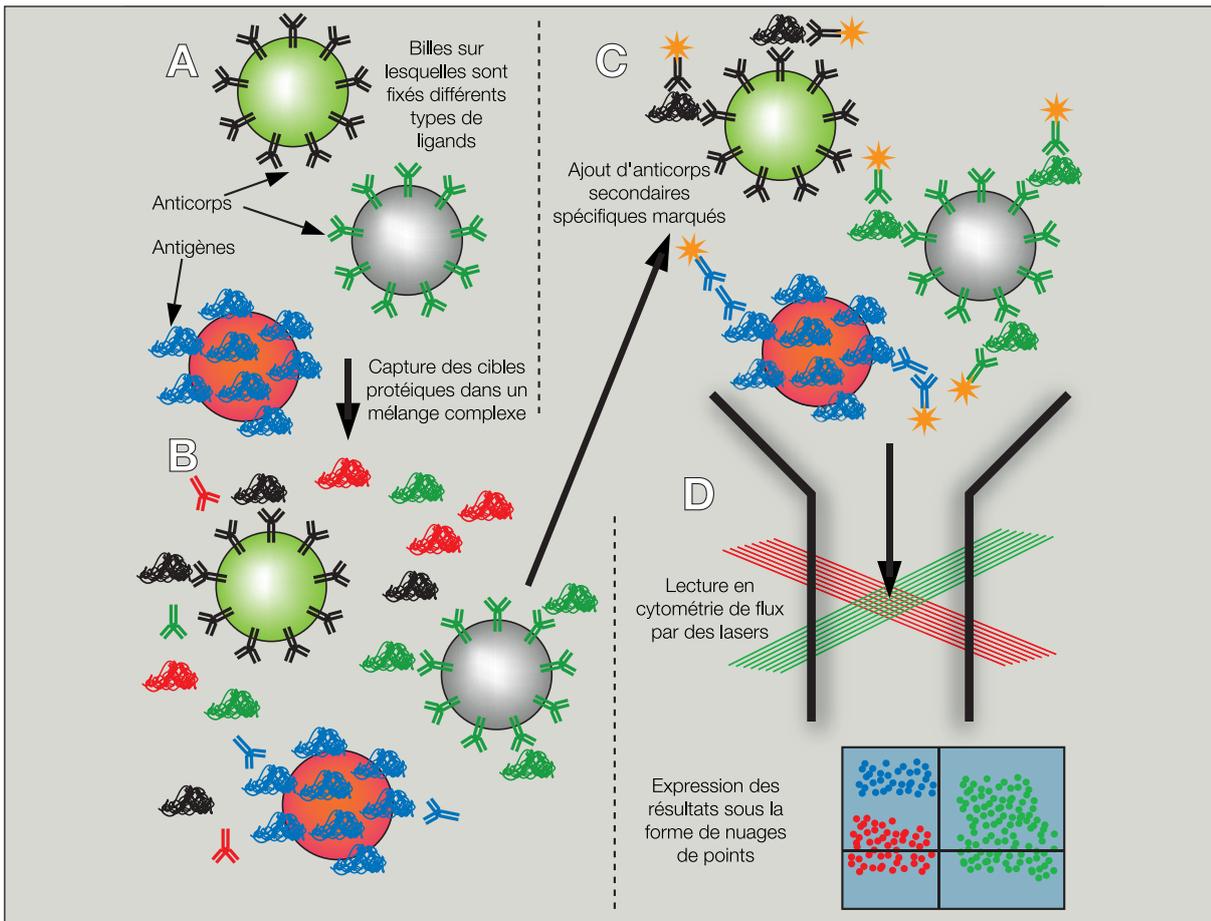
## Analyses multiplexes en milieu liquide

Cette approche repose principalement sur la technologie de cytométrie de flux dans laquelle des cellules, mais également des billes, circulent dans une veine liquide permettant de récupérer des informations individuelles de taille, de nombre ou de fluorescence grâce à un ou plusieurs lasers intégrés au système [2, 7]. L'exemple le plus achevé de cette approche est représenté par le système Luminex® associant deux faisceaux laser appliqués à la détection et au dosage de molécules de différents types : des peptides (cytokines...), des complexes antigènes-anticorps, d'autres protéines, des acides nucléiques ou encore des xénobiotiques. Cette technologie utilise des microsphères en polystyrène sur lesquelles sont fixées des ligands (anticorps, antigènes...) qui les rendent spécifiques des molécules à doser (figure 1). Ceci est rendu possible par la présence sur les billes activées de groupements carboxylate, amine ou thiol appropriés pour fixer de façon covalente les ligands. Le mode opératoire classique de cette approche est simple, rapide (3 à 4 h incluant la préparation des échantillons, l'analyse et l'obtention des résultats) et automatisable. Il comprend une étape réactionnelle initiale dans des plaques à 96 puits de type ELISA (« Enzyme-Linked ImmunoSorbent Assay ») avec des volumes réactionnels faibles (10 à 50 µl) ce qui est intéressant pour réduire le volume de prélèvement nécessaire à la réalisation de nombreuses analyses. Au cours de l'analyse, qui est réalisée en phase liquide pour l'ensemble des étapes réactionnelles, les billes sont détectées et comptées en temps réel. Les résultats sont visualisables sous la forme de nuages de points répartis sur un graphe en fonction généralement du type et de la fluorescence des billes (intensité et longueurs d'onde utilisées). Le positionnement de ce nuage de points se fait au sein d'un masque pré-établi ce qui permet de s'assurer rapidement de la qualité de l'analyse (quantité et qualité des prélèvements et des réactifs...).

Il est possible de combiner les analyses multiplexes, par exemple jusqu'à 100 types de microsphères avec l'approche Luminex®. En effet, les billes ont incorporé au cours de leur fabrication deux fluorochromes différents permettant de les identifier avec l'un des lasers du module d'analyse micro-fluidique. L'identification des marqueurs fixés sur les billes dépend de la méthode utilisée ; le plus souvent dans le cas de marqueurs protéiques, il s'agit de la détection de fluorochromes couplés à un anticorps de révélation ou au produit d'amplification de détection qui aura été préalablement ajouté au milieu réactionnel (figure 1) [2]. Cette technologie a été installée dans de nombreux analyseurs comme les systèmes xMAP® (Luminex), Bio-Plex® (Bio-Rad) et QUANTA-Plex® (Inova/Menarini) pour lesquels des différences résident principalement dans les suites logicielles et les analyses disponibles. Hors la technologie Luminex®, des systèmes basés sur la cytométrie de flux capables d'analyser à la fois des cellules et des microbilles de différentes taille et/ou fluorescentes sont également disponibles. Il s'agit par exemple du système FACSArray® développé par BD Biosciences ([www.bdbiosciences.com](http://www.bdbiosciences.com)) dont le principe général d'analyse reste inchangé. En conclusion, les analyses multiplexes en milieu liquide sont des approches très versatiles jouissant d'une grande capacité d'analyse dans de nombreuses disciplines biologiques.

### Figure 1 : Analyses multiplexes en milieu liquide.

(A) : Des microbilles sont préparées de telle sorte que des ligands spécifiques (anticorps, antigènes...) soient fixés à leur surface. Ces billes peuvent différer les unes des autres par leur fluorescence ou leur diamètre ; (B) : Les billes sont mises en contact avec un échantillon biologique afin de réaliser une capture des cibles dans des conditions optimales ; (C) : Après rinçage, les billes sont généralement incubées avec un réactif secondaire (anticorps marqués...) qui reconnaît la cible, puis elles sont à nouveau rincées ; (D) : Les billes sont dirigées à travers une veine liquide permettant de récupérer des informations individuelles de taille, de nombre et de fluorescence grâce à un ou plusieurs lasers intégrés au système. Le comptage en temps réel des billes permet une représentation en nuages de points en fonction généralement du type et de la fluorescence des billes (intensité et longueurs d'onde utilisées).



## Analyseurs multiparamétriques miniaturisés

Il s'agit d'analyseurs qui ont la particularité d'être transportables, utilisables au lit du patient en urgence et/ou de façon délocalisée pour mesurer en une fois un petit nombre de bio-marqueurs. Ils font ainsi partie du groupe d'analyseurs dit « Point of Care Testing » (POCT) utilisés principalement en biologie délocalisée. Ce type d'approche est très répandu à l'étranger et en particulier aux Etats-Unis. En France les paramètres principalement concernés par la biologie délocalisée sont : glycémie, coagulation, gaz du sang et pH, électrolytes, hémoglobine et acide lactique. Le développement de la délocalisation en France est prudent en raison des précautions prises en terme de contrôle de qualité et de validation biologique. Dans le domaine des analyses multiples de paramètres protéiques, le développement de la miniaturisation et des micro- ou nanotechnologies (micro-fluidique...) permet théoriquement d'adapter de nombreuses analyses

à cette approche. En pratique il n'y a que peu d'analyseurs disponibles pour la biologie clinique. Le plus démonstratif et versatile semble être le système Triage® (Biosite) qui permet de doser en même temps plusieurs paramètres sous la forme de panels adaptés à une situation pathologique donnée ([tableau 1](#)) [8].

**Tableau 1 : Exemples de panels pour la biologie clinique.**

Domaine	Paramètres	Fabriquant	Technologie
Cancérologie	MIF, ostéopontine, leptine, CA-125, prolactine, IGF-II	Millipore	Cytométrie
Cardiovasculaire	Créatine-kinase MB (CK-MB), Fatty Acid Binding Protein (FABP), myoglobine, troponine I	Randox	Puce
Cardiovasculaire	CK-MB, myoglobine, troponine I, BNP, D-dimère	Triage	Microfluidique
Coagulation	$\alpha$ 1-antitrypsine, antithrombine, apolipoprotéine H, F.V, F.VII, F.IX, F.XI, F.XII, F.XIII, protéine C, protéine S, prothrombine, vWF	Bender System	Cytométrie
Cytokines et chimiokines	IL-8, CXCL-5, FGF basique, G-CSF, GM-CSF, IFN $\gamma$ , IL-1F1, IL-1F3, IL-1 $\beta$ , IL-2, IL-4, IL-5, IL-6, IL-10, IL-17, TNF $\alpha$ , TPO, VEGF	R&D System	Cytométrie
Cytokines et chimiokines	IFN $\gamma$ , TNF $\alpha$ , IL-1 $\beta$ , IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12 p70, IL-13, IL-17, GM-CSF, G-CSF, MCP-1, MIP $\beta$	Bio-Rad	Cytométrie
Cytokines et chimiokines	C5a, CD-40 ligand, G-CSF, GM-CSF, GRO1-309, sICAM-1, IFN $\gamma$ , IL-1 $\beta$ , IL-1ra, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12 p70, IL-13, IL-16, IL-17, IL-17E, IL-23, IL-27, IL-32 $\alpha$ , IP-10, I-TAC, MCP-1, MIF, MIP-1 $\alpha$ & $\beta$ , serpine E1, RANTES, SDF-1, TNF $\alpha$ , sTREM-1	R&D System	Membranes
Cytokines et chémokines	EGF, IL-1 $\alpha$ , IL-1 $\beta$ , IL-2, IL-4, IL-6, IL-8, IL-10, IFN $\gamma$ , MCP-1, TNF $\alpha$ , VEGF	Randox	Puce
Fertilité	LH, FSH, progestérone	Mesocale	Plaques ELISA
Fertilité	Œstradiol, Follicle Stimulating Hormone (FSH), Luteinizing Hormone (LH), progestérone, prolactine (PRL), testostérone	Randox	Puces
Immunologie	HLA IgG : anticorps contre les antigènes de classe I et/ou de classe II	Gen-Probe	Cytométrie
Infectiologie	Recherche d'anticorps contre 14 sérotypes de pneumocoques	Luminex	Cytométrie
Inflammation	$\alpha$ 2-macroglobuline, CRP, ferritine, fibrinogène, procalcitonine, haptoglobine, SAA, SAP, TPA	Bio-Rad	Cytométrie
Métabolisme osseux	PAL, ostéocalcine, ostéonectine, ostéoprotégérine, sclérostine	Mesocale	Plaques ELISA
Métabolisme osseux	ACTH, insuline, leptine, ostéocalcine, ostéopontine, ostéoprotégérine, PTH	Millipore	Cytométrie
Neurologie	Peptides Ab37, Ab38, Ab39, Ab40, Ab42...	Bio-Rad	SELDI-TOF
Neurologie	Peptides Ab42, tau, P-tau (181)	Innogenetics	Cytométrie
Neurologie	Peptides Ab38, Ab40, Ab42	Mesocale	Plaques ELISA

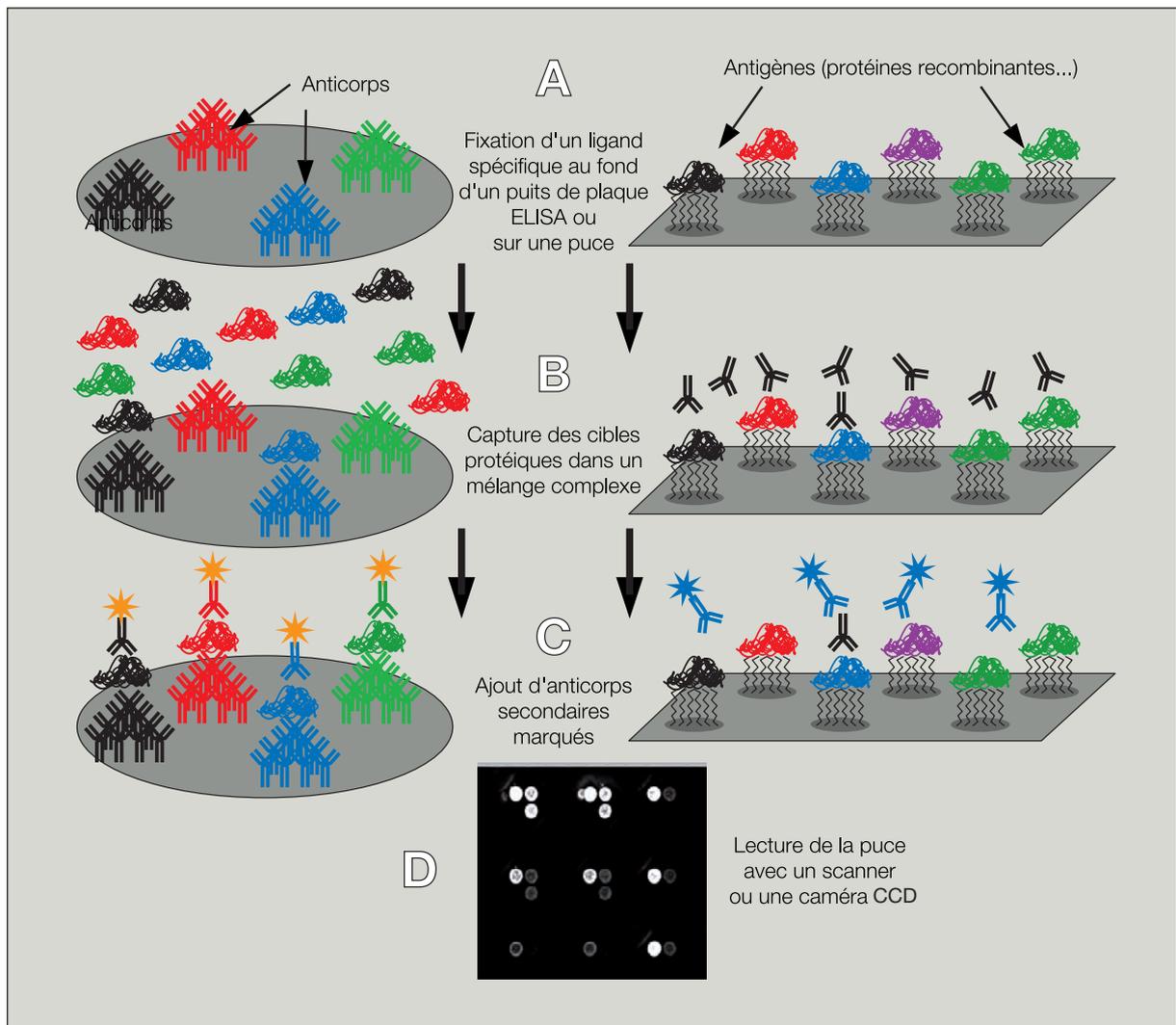
## Analyses multiplexes en phase solide

Le principe général de ces analyses est la détection et la quantification d'une interaction entre un ligand fixé à une surface et un bio-marqueur à analyser. Utilisée en multiplexe et pouvant donc concerner de nombreux paramètres protéiques, il s'agit de fait d'une approche protéomique. Les applications vont de la caractérisation d'une fonction protéique spécifique (phosphorylation...), ou de l'étude d'affinité moléculaire ( $K_D$ ...) avec des ligands divers (protides, glycanes, acide nucléiques...), à la détection de molécules déjà connues (notamment pour la biologie clinique). Les biopuces présentent des surfaces de différentes natures (nitrocellulose, silice, polymères, particule d'or, support chromatographique...) qui, comme pour les billes décrites précédemment, vont être traitées ou « fonctionnalisées » pour retenir un premier composant qui sert de ligand. La miniaturisation du dépôt de ligand permet de produire des puces avec plusieurs milliers de dépôts en goutte (ou « spots ») par centimètre carré ; on parle alors de « microarray ». Les formats les plus classiques sont en fait des « puces à anticorps » (figure 2). Dans ces biopuces, des anticorps spécifiques de cibles peptidiques ou protéiques sont déposés avec des densités variables, jusqu'à plusieurs milliers dans le domaine de la recherche. En biologie clinique ces systèmes ne comportent souvent que quelques dizaines d'anticorps différents souvent regroupées sous la forme de panels thématiques (tableau 1). Il est en effet délicat de combiner avec les contraintes de la clinique (importance de la reproductibilité intra et inter-expérience, des réactions croisées, de la linéarité) des marqueurs multiples en utilisant des anticorps aux spécificités différentes. En biologie clinique, le prélèvement (sérum, LCR...) est classiquement utilisé sans marquage préalable des protéines et les cibles capturées sont détectées grâce à un deuxième anticorps marqué (enzymatique, fluorescent, biotinylé...). La lecture se fait ensuite par une méthode adaptée au marquage : chromogénique, par électro-chimioluminescence, transfert d'énergie de fluorescence ou en simple fluorescence, avec une lecture par des systèmes de type scanner ou caméra CCD (« Charge Coupled Device ») (figure 2). Ce type de puces utilisant deux anticorps (méthode sandwich) réalise ainsi des dosages comparables à des ELISA en multiplexe (voir la préparation des anticorps, Chapitre 8).

On retrouve cela notamment sur le système Investigator® de la société Randox qui propose différents panels déposés sur des puces de silice dans un format propriétaire et avec des seconds anticorps fluorescents (tableau 1). La société MSD® propose des ELISA multiplexes en plaques 96 puits avec une révélation par électro-chimioluminescence. Enfin un dernier exemple qui se différencie par sa méthode de détection des marqueurs capturés est représenté par l'analyseur SELDI-TOF® (« Surface Enhanced Laser Desorption/ Ionization Time-Of-Flight Mass Spectrometry ») de la société Bio-Rad (voir Chapitre 1). Cet appareil permet la détection de marqueurs par spectrométrie de masse ce qui amène à quantifier des variants du paramètre recherché comme dans le cas des peptides A, dans la maladie d'Alzheimer [9].

**Figure 2 : Analyses multiplexes en phase solide.**

(A) : Les surfaces (biopuces, membranes, lame de verre..) sont traitées pour retenir un premier type de composé qui sert de ligand. Il peut s'agir d'anticorps ou d'autres types de ligands (antigènes, récepteur..) ; (B) : Les surfaces sont mises en contact avec un échantillon biologique afin de réaliser une capture des cibles dans des conditions optimales ; (C) : Après rinçage, les surfaces sont généralement incubées avec un réactif secondaire (anticorps marqué..) qui reconnaît la cible, puis elles sont à nouveau rincées ; (D) : Les puces ou membranes sont ensuite analysées avec un système de lecture de type scanner ou caméra CCD en fonction de la technologie. Une image de la puce permettant la quantification des spots est alors obtenue.



Enfin, notons parmi les autres types de puces utilisées en biologie clinique, celles destinées à la détection d'anticorps présents dans les liquides biologiques grâce à des molécules antigéniques fixées sur la puce (figure 2). Ces approches sont particulièrement intéressantes pour la détection d'anticorps dans les pathologies auto-immunes et dans le domaine de l'infectiologie [10-12].

Dans toutes ces approches, l'analyse des résultats est effectuée à l'aide d'outils informatiques dédiés qui prennent en compte les valeurs de standards internes mesurés sur la puce. A terme, l'intérêt de tels dosages multiplexes sera dans l'utilisation d'algorithmes qui permettront le calcul de scores ou d'indices donnant des probabilités diagnostiques ou pronostiques (voir Chapitre 4).

## Panels multiplexes

Il existe une grande diversité d'approches multiplexes qui regroupent les cibles protéiques d'intérêt en fonction des applications. En recherche, ce sont principalement les panels de détection des phosphoprotéines et des acteurs de voies spécifiques (apoptose, protéasome...) qui apportent un bénéfice considérable aux études des modulations physiologiques, pathologiques ou par des agents pharmacologiques des multiples voies métaboliques ou de signalisation cellulaire, qui sont souvent très complexes. Dans le domaine des panels cliniques qui nous intéressent ici, les marqueurs proposés visent une même pathologie, ces marqueurs pouvant être très divers, ou inversement une même famille de marqueurs peuvent jouer un rôle physiopathologique coordonné ([tableau 1](#)). Le but est de tendre vers une exploration cohérente dans le cadre du suivi d'un état pathologique et de permettre l'utilisation d'algorithmes pour le calcul de scores comme dit précédemment.

Actuellement, les grands domaines concernés par les méthodes d'analyses multiplexes sont l'immunologie, la microbiologie, la virologie et la biochimie. En microbiologie et virologie, sont par exemple disponibles des multiplexes permettant de déterminer la présence et le titre des anticorps dirigés contre différentes protéines de nombreux agents infectieux (HIV, HPV, VRS...). Une des premières applications concerne l'analyse multiplexe d'interleukines et sous de multiples formats présentés par plusieurs sociétés ([tableau 1](#)). De six à plusieurs dizaines d'interleukines et de cytokines peuvent être ainsi simultanément identifiées et quantifiées grâce à des standards appropriés. Ceci est particulièrement intéressant en immunologie définissant ainsi une nouvelle approche des maladies auto-immunes et inflammatoires. Par ailleurs, dans le domaine de l'histocompatibilité, de plus en plus de laboratoires ont recours à la technologie Luminex® pour le typage HLA (HLA-A, HLA-B...), le dépistage et l'identification en routine des anticorps anti-HLA de classe I et de classe II d'isotype IgG car la méthode est simple, sensible, reproductible et standardisée. Comme listés dans le [Tableau 1](#), on retrouve en fait sur le marché de nombreux panels couvrant divers états physiopathologiques : inflammation, diabète, infertilité, dysfonctionnement hormonal, cancer, maladies cardiovasculaires, neurodégénérescence...

## Conclusion

---

Les analyses multiples de paramètres protéiques présentent a priori un grand potentiel dans le domaine clinique car elles permettent la mesure simultanée de nombreux paramètres et ceci avec un faible volume d'échantillon. Elles ouvrent la porte également à l'utilisation d'algorithmes permettant sur la base des résultats obtenus de calculer des indices ou des scores indiquant, avec

une probabilité plus ou moins grande, la présence d'une pathologie, d'un type d'évolution ou de réponse à un traitement.

Dans le domaine de la protéomique fondamentale ou de la recherche clinique ce type d'approche est très largement utilisé, notamment pour faire l'étude moléculaire d'une situation donnée, ou pour la recherche de nouveaux bio-marqueurs. Par contre, en biologie clinique, l'utilisation reste limitée, sauf dans des domaines particuliers comme l'immunologie. Cette dissociation est probablement en relation avec, dans le cas de la recherche, la possibilité de faire une étude de grande ampleur à un moment donné et souvent dans un modèle intégré (cellules, animaux...). La solidité des résultats obtenus est alors dépendante de la précision analytique au sein d'une même expérience, avec les mêmes lots de réactifs et les mêmes conditions pré-analytiques. Dans le domaine clinique, il est nécessaire de réaliser des analyses et de comparer les valeurs au cours du temps et dans une population composée de nombreux patients/témoins différents. Il existe ainsi une très grande exigence sur la solidité des mesures en fonction des lots de réactifs et de puces/billes/membranes, pour l'utilisation d'un contrôle de qualité, et pour la comparaison à des méthodes de référence. Or, dans les analyses multiplexes telles que décrites précédemment la modification d'un seul paramètre risque d'altérer l'ensemble des mesures. On a donc une multiplication des contraintes et des risques d'erreurs. De même, si un paramètre nécessite occasionnellement une modification de protocole (dilution par exemple..) cela va finalement obliger à doser à nouveau tous les paramètres. Du point de vue pratique, les panels pré-établis par les sociétés ne répondent pas toujours aux besoins ([tableau 1](#)). On y retrouve en effet de nombreux marqueurs dont l'intérêt clinique n'a pas été à ce jour démontré, et inversement, des marqueurs intéressants pour une pathologie donnée sont souvent absents. Ce manque de cohérence global est probablement en relation avec le manque de disponibilité des réactifs ou leur incompatibilité (réactions croisées). Ces difficultés expliquent certainement le faible taux de pénétration actuel de ces technologies multiplexes au sein des laboratoires, avec cependant quelques exceptions comme l'immunologie, montrant qu'à maturité ces approches seront très intéressantes.

## Bibliographie

---

1. Dupuy AM., Lehmann S. and Cristol JP. Protein biochip systems for the clinical laboratory. Clin. Chem. Lab. Med. 2005 ; 43 : 1291-1302.
2. Lizard G., Monier S., Prunet C., Duviillard L. and Gambert P. Microspheres, nanospheres and flow cytometry : from cellular to molecular analysis. Ann. Biol. Clin. (Paris) 2004 ; 62 : 47-52.
3. Stoevesandt O., Taussig MJ. and He M. Protein microarrays: high-throughput tools for proteomics. Expert. Rev. Proteomics 2009 ; 6 : 145-157.
4. Dufva M. Introduction to microarray technology. Methods Mol. Biol. 2009 ; 529 : 1-22.
5. Roche S., Gabelle A. and Lehmann S. Clinical Proteomics of the Cerebrospinal Fluid : towards the discovery of new biomarkers. Proteomics: Clin. Appl. 2008 (in press).

6. States DJ., Omenn GS., Blackwell TW., Fermin D., Eng J., Speicher DW. and Hanash SM. Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat. Biotech.* 2006 ; 24 : 333-338.
7. Nolan JP. and Mandy F. Multiplexed and microparticle-based analyses : quantitative tools for the large-scale analysis of biological systems. *Cytometry A* 2006 ; 69 : 318-325.
8. Apple FS., Christenson RH., Valdes R. Jr, Andriak AJ., Berg A., Duh SH., Feng YJ., Jortani SA., Johnson NA., Koplen B., Mascotti K. and Wu AH. Simultaneous rapid measurement of whole blood myoglobin, creatine kinase MB, and cardiac troponin I by the Triage cardiac panel for detection of myocardial infarction. *Clin. Chem.* 1999 ; 45 : 199-205.
9. Lewczuk P., Esselmann H., Groemer TW., Bibl M., Maler JM., Steinacker P., Otto M., Kornhuber J. and Wiltfang J. Amyloid beta peptides in cerebrospinal fluid as profiled with surface enhanced laser desorption/ionization time-of-flight mass spectrometry : evidence of novel biomarkers in Alzheimer's disease. *Biol. Psych.* 2004 ; 55 : 524-530.
10. Lalvani A., Meroni PL., Millington KA., Modolo ML., Plebani M., Tincani A., Villalta D., Doria A. and Ghirardello A. Recent advances in diagnostic technology: applications in autoimmune and infectious diseases. *Clin. Exp. Rheumatol.* 2008 ; 26 : S62-66.
11. Roche S., Dauvilliers Y., Tiers L., Couderc C., Piva MT., Provansal M., Gabelle A. and Lehmann S. Autoantibody profiling on high-density protein microarrays for biomarker discovery in the cerebrospinal fluid. *J. Immunol. Methods* 2008 ; 338 : 75-78.
12. Tozzoli R. Review : The diagnostic role of autoantibodies in the prediction of organ-specific autoimmune diseases. *Clin. Chem. Lab. Med.* 2008 ?

**Les anticorps recombinants :  
un répertoire élargi  
source de nouvelles  
applications en biologie  
médicale**

**Pr Philippe Billiald**

**CHAPITRE VIII**

# Résumé

---

La découverte de la technologie des hybridomes dans les années 1980 a conduit à un essor considérable de l'immunoanalyse parce qu'elle a permis de remplacer les immun-sérums polyclonaux hétérogènes par des anticorps monoclonaux qui constituent des traceurs uniques, spécifiques d'un antigène donné, souvent d'affinité élevée et toujours produits à l'identique. Aujourd'hui, les connaissances acquises sur la structure des immunoglobulines, la possibilité de cloner leurs gènes et les progrès accomplis dans le domaine des biotechnologies sont à l'origine d'un renouveau de l'immunoanalyse. Il devient ainsi possible de créer de nouveaux outils, des formats d'anticorps inédits adaptés à des applications particulières et de modifier leur affinité pour une cible, d'ajuster leur spécificité à un antigène donné. La technologie d'exposition à la surface d'un phage (phage-display) autorise même la sélection *in vitro* de fragments d'anticorps spécifiques d'un antigène sans avoir recours à l'immunisation préalable d'animaux.

Dans ce contexte, les anticorps recombinants constituent une source unique de sondes qui devraient contribuer à un nouvel essor de l'immunodiagnostic et connaître des applications biomédicales encore insoupçonnées.

## Introduction

---

En 1975, G. Köhler et C. Milstein (Prix Nobel de Médecine, 1984) décrivaient la technologie des hybridomes qui consiste à immortaliser une cellule sécrétrice d'un anticorps unique, monoclonal, spécifique d'un antigène donné. Assez simple à réaliser avec des plasmocytes de souris ou de rat, cette technologie a conduit à la sélection d'une multitude d'anticorps monoclonaux de rongeurs dirigés contre des antigènes variés avec des applications très diverses. Chaque mois, chaque semaine a vu paraître de nouveaux anticorps monoclonaux aux propriétés bien établies, spécifiques d'un antigène particulier et pouvant être produits de façon stable et continue. Ces anticorps ont permis l'identification de marqueurs cellulaires (CD) ; ils ont contribué à la compréhension des relations structure/fonction de protéines complexes et à élucider les mécanismes d'interaction entre biomolécules. Le secteur de l'immunoanalyse en a beaucoup profité: une simple recherche du terme « immunoassay » dans la banque de données PubMed fait ressortir plus de  $3 \times 10^5$  articles originaux et le marché de l'« immunodiagnostic » représenterait aujourd'hui plus de 6 milliards de dollars.

En revanche, l'utilisation thérapeutique de ces anticorps monoclonaux d'origine animale a connu un échec cinglant. Seuls trois anticorps ont reçu une autorisation de mise sur le marché de la Food and Drug Administration (FDA) : une IgG anti-CD3 (OKT3<sup>®</sup>, Muromomab) utilisée en prévention du rejet aigu d'allogreffes, et deux anticorps anti-CD20 radiomarqués (Bexxar<sup>®</sup>, <sup>131</sup>I-Tositumomab<sup>®</sup> et Zevalin<sup>®</sup>,

$^{111}\text{In}$ - et  $^{90}\text{Y}$ - Ibritumomab) prescrits en deuxième intention dans le cadre du traitement de certains lymphomes. Cette déconvenue a plusieurs origines : la nature xénogénique des anticorps induit une réponse immune de l'hôte et l'apparition d'anticorps humains anti-anticorps de souris (« HAMA » pour « Human Anti-Mouse Antibodies ») ; leur demi-vie plasmatique est courte, et enfin, même s'il n'y a pas de barrière d'espèce au sens strict, les propriétés effectrices des anticorps de souris administrés à l'homme ne sont pas optimales. Ce n'est que très récemment qu'ont été créées des souris transgéniques dans le génome desquelles sont introduits des gènes d'anticorps humains. La technologie des hybridomes peut leur être appliquée, mais le répertoire des anticorps humains produits reste limité.

Dans ce contexte, la recherche d'anticorps optimisés s'est poursuivie de façon soutenue, et l'expansion, dans le même temps, des biotechnologies a permis de concevoir, dès le début des années 1990, des protéines recombinantes capables de lier spécifiquement un antigène avec des caractéristiques fonctionnelles souvent améliorées et mieux adaptées à des usages spécifiques que celles des anticorps conventionnels (Laffly and Sodoyer, 2005). C'est ainsi qu'ont été créés les fragments d'anticorps recombinants dont la structure se limite aux domaines variables (scFv pour fragment variable simple chaîne). Ils constituent l'unité fonctionnelle minimale qui lie un antigène et à partir de laquelle il devient possible d'inventer des formats d'anticorps plus évolués ou encore de moduler les caractéristiques fonctionnelles (affinité, spécificité). Fin 2009, il y avait sur le marché une vingtaine de biomédicaments à activité anticorps issus de ces nouvelles technologies. Plusieurs centaines étaient en phase d'essai clinique et les Entreprises du Médicament (LEEM) ont estimé que les anticorps représenteraient le segment des médicaments d'origine biologique au plus fort taux de croissance, avec une part de marché de 35 % en 2008 et une progression annuelle supérieure à 20 %. Au moins sept de ces anticorps ont acquis le statut de « blockbusters » générant chacun un chiffre d'affaires annuel supérieur à 1 milliard de dollars.

Dans le secteur de l'immunoanalyse, les retombées de ces technologies sont encore très modestes, probablement parce que les anticorps conventionnels poly- ou monoclonaux, affins, spécifiques et issus de cellules d'origines animales, répondent aux besoins actuels de ce secteur. Leur immunogénicité n'est pas un handicap pour un usage *in vitro* et leur coût de production reste modéré. Toutefois, les technologies de l'ADN recombinant offrent de nouvelles ressources et laissent entrevoir le développement de réactifs innovants avec des champs d'applications encore insoupçonnés.

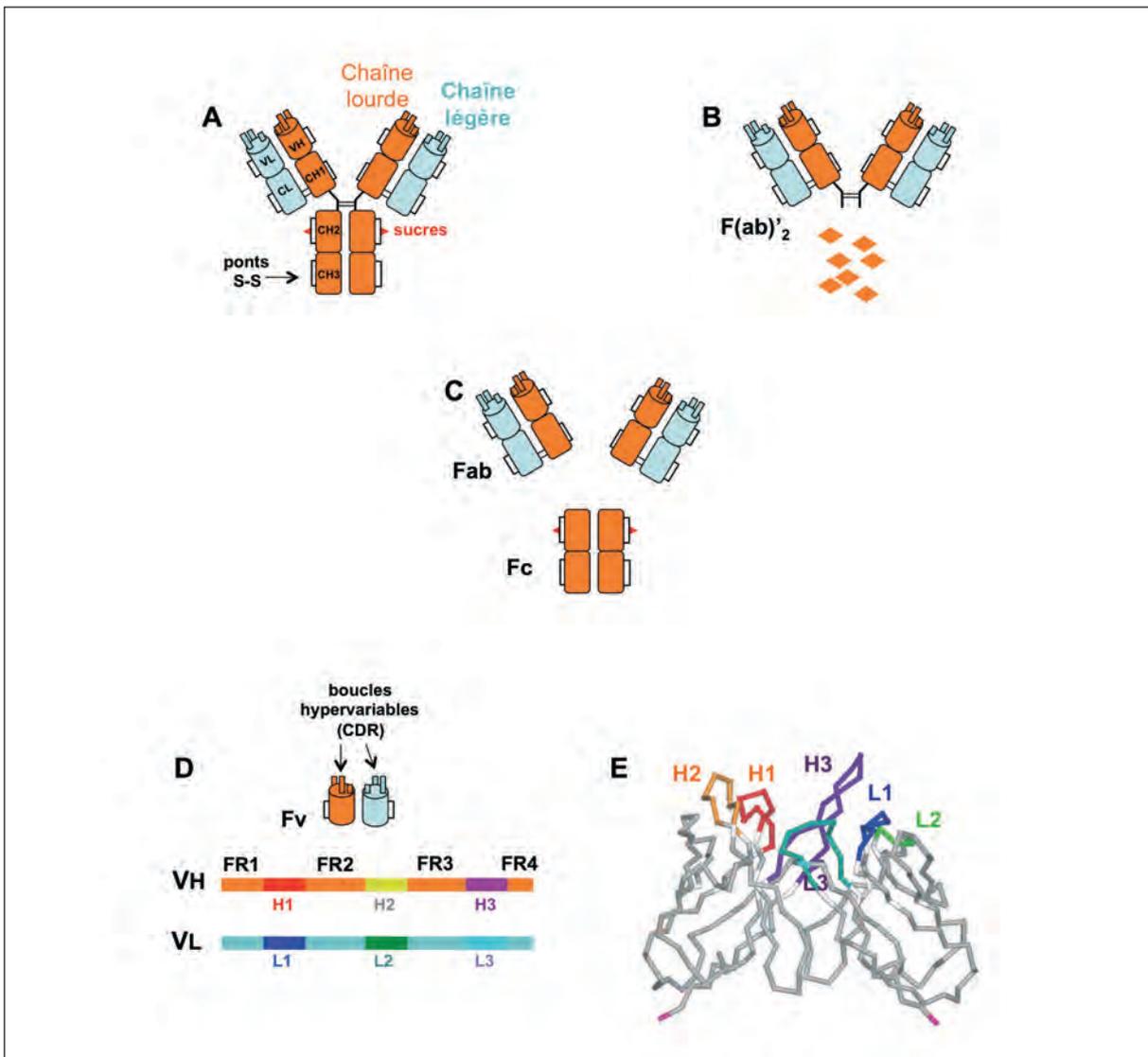
## Anticorps naturels, structure et applications

En réponse à l'injection d'un antigène étranger, les Vertébrés produisent des anticorps, protéines sécrétées par les plasmocytes, qui servent d'effecteurs de l'immunité humorale. Ces anticorps sont habituellement hétérogènes, chaque clone de cellule B produisant un anticorps monoclonal qui reconnaît un motif structural particulier (épitope). L'ensemble de ces anticorps monoclonaux constitue la réponse anticorps polyclonale du sérum. On regroupe les anticorps sous le terme générique d'immunoglobuline (Ig) dont on distingue 5 classes (IgA, IgD, IgE, IgG, IgM) qui partagent

certaines caractéristiques structurales mais diffèrent biochimiquement et fonctionnellement. Les IgG constituent la classe d'immunoglobulines la plus abondante dans le sérum. Ils sont principalement sécrétés par des cellules à mémoire et ont acquis, consécutivement à des processus de commutation de classe et de maturation, une affinité et une spécificité élevées pour l'antigène cible. La structure des IgG est bien comprise depuis les études réalisées dans les années 1950-60 par R. Porter et G. Edelman (Prix Nobel de Médecine, 1972). L'IgG est un monomère d'immunoglobuline constitué de quatre chaînes polypeptidiques identiques deux à deux : deux chaînes polypeptidiques dites lourdes (H) d'environ 50 kDa, chacune associée à une chaîne légère (L) (25kDa) par un pont disulfure (figure 1).

**Figure 1 : Immunoglobuline G et fragments fonctionnels.**

(A) : Monomère d'IgG. (B) : Fragment bivalent  $F(ab)'_2$  obtenu par pepsinolyse. (C) : Fragments Fab et Fc obtenus par papainolyse. (D) : Fragment variable Fv instable constitué des domaine VH et VL. (E) : Structure tridimensionnelle du Fv. Les boucles hypervariables L1, L2, L3 et H1, H2, H3 ou CDR constituent une poche (paratope) qui interagit avec l'antigène cible.



Des différences de structures et de fonctions au niveau des domaines constants des chaînes lourdes permettent de distinguer 4 sous-classes d'IgG chez l'homme. Les chaînes légères sont de deux types ( $\kappa$  ou  $\lambda$ ). Porter a tout d'abord montré qu'une protéolyse ménagée IgG par la papaine conduit à deux fragments identiques (50 kDa), appelés Fab (« fragment antigen-binding ») en raison de leur activité de liaison à l'antigène et un fragment Fc (50 kDa), cristallisable, responsable de l'activation du complément ou des cellules immunocompétentes induites par la formation du complexe antigène-anticorps. Un peu plus tard, A. Nisonoff a montré que la digestion d'une IgG par la pepsine produit des peptides inactifs qui proviennent de la fragmentation du Fc et un fragment F(ab)'<sub>2</sub> (100 kDa) capable de précipiter l'antigène cible parce qu'il possède deux sites de liaisons. Enfin, c'est la réduction des IgG par le 2-mercaptoéthanol qui a conduit G. Edelman et R. Porter, dans les années 1960, à établir leur structure multi-chaînes.

Au sein d'une IgG, chaque chaîne polypeptidique lourde (H) ou légère (L) est organisée en domaines: un domaine variable d'environ 110 résidus, à l'extrémité N-terminale des chaînes légères (VL) et lourdes (VH), associé à un domaine constant (CL) pour la chaîne légère ou 3 domaines constants pour la chaîne lourde (CH1, CH2, CH3). Dans chaque chaîne lourde, un peptide joint les domaines CH1 et CH2. Des liaisons disulfure entre les deux peptides créent une structure, dite région charnière, qui confère à la molécule une grande flexibilité et lui donne la capacité de lier, par exemple, deux épitopes qui seraient dans des orientations contraires l'un par rapport à l'autre. Le site de liaison d'un anticorps à son antigène cible est formé par l'association des domaines variables VH et VL. La variabilité des séquences des domaines V n'est pas distribuée de manière égale. On trouve quatre régions dites régions charpentes ou framework (FR) dans lesquelles des résidus bien conservés déterminent la structure du domaine. Entre les régions charpentes d'un même domaine, existent trois « boucles » dites hypervariables ou « CDR » en raison de la variabilité particulière de leur séquence. Le rapprochement spatial des boucles d'un domaine VL et d'un domaine VH constitue le paratope, d'une surface d'environ 500 Å<sup>2</sup>, qui interagit avec l'épitope de l'antigène cible. Moins d'une vingtaine de résidus d'acides aminés de l'anticorps sont directement impliqués dans ces interactions non covalentes (figure 1).

L'idée d'utiliser les anticorps en dehors de la réponse immune humorale naturelle est née avant même leur identification formelle. D'abord à des fins thérapeutiques, lorsque E. von Behring, premier prix Nobel de médecine (1901) et S. Kitasato ont découvert la notion de sérothérapie, c'est-à-dire le transfert d'immunité à un patient auquel on injecte un immun sérum produit à volonté chez un animal. Les premiers sérums antidiphtériques, antitétaniques et antivenimeux seront commercialisés dans les années 1940. L'expansion des anticorps comme outil d'immunoanalyse, ou de diagnostic en biologie clinique a débuté dans les années 1960 avec l'obtention d'anticorps polyclonaux par immunisation d'animaux (lapins, chèvres, moutons) contre des antigènes d'origine humaine.

La possibilité de produire à l'infini une population homogène d'anticorps (anticorps monoclonaux) par la technologie des hybridomes a ensuite permis, dès les années 1980, de générer des anticorps d'origine murine de spécificité définie et d'affinité constante. L'anticorps est ainsi devenu un produit industriel et de nouveaux concepts, de nouvelles technologies ont émergé en sciences biologiques, en recherche et dans le diagnostic *in vitro*. Ainsi, l'immunocytométrie est devenue un outil de diagnostic

essentiel en hématologie pour l'analyse multiparamétrique des cellules sanguines associant caractéristiques morphologiques (taille, granularité relatives) et présence de marqueurs immunophénotypiques. Ces mêmes anticorps sont exploités en immunohistochimie pour mettre en évidence un antigène particulier dans un tissu. Les méthodes immunoenzymatiques qui font appel à des anticorps couplés à une enzyme « colorimétrique » ont également connu une extraordinaire expansion qui s'explique par leur grande sensibilité, la nécessité d'un appareillage simple et la possibilité d'être robotisées.

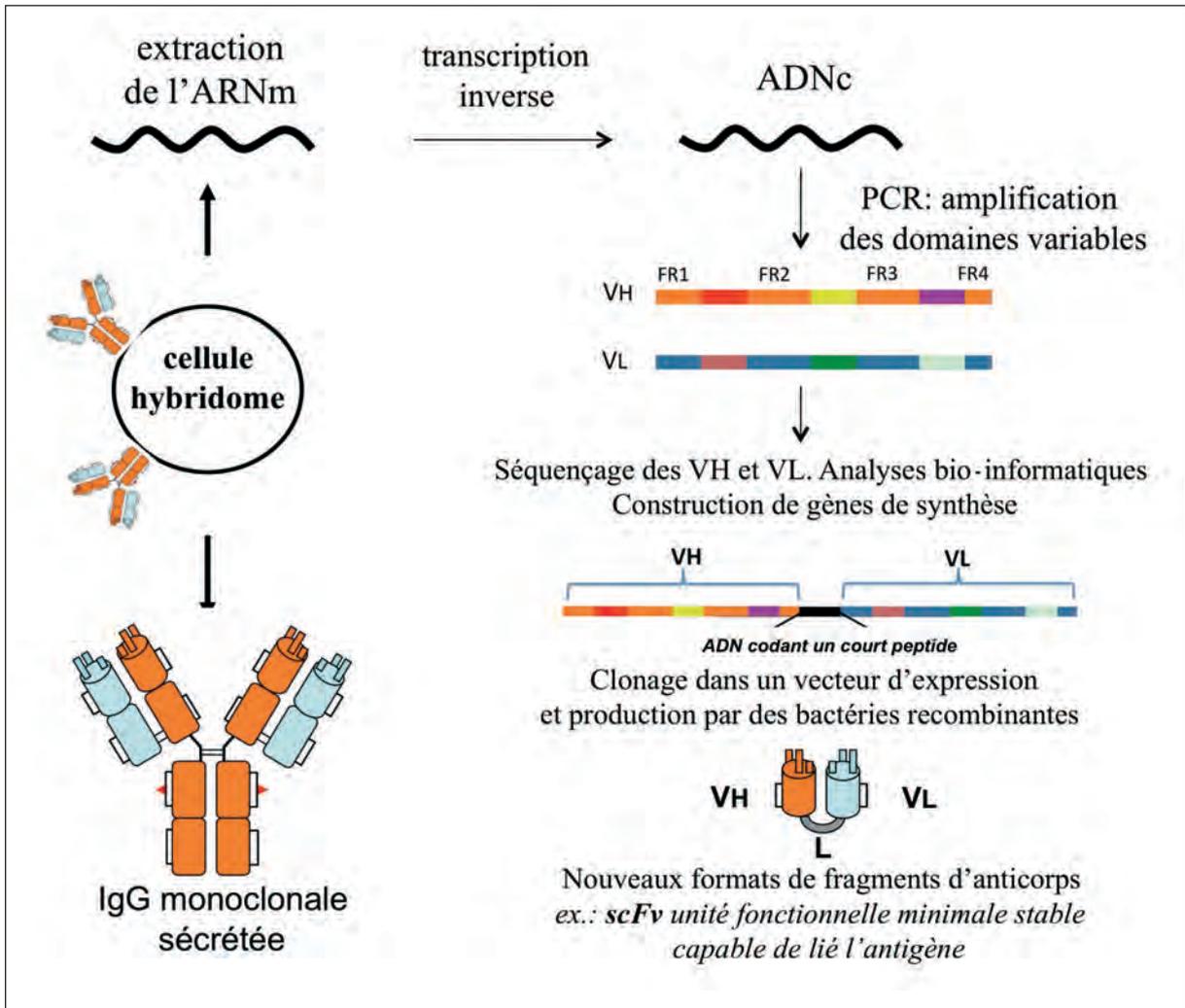
Le champ d'application des anticorps monoclonaux est très vaste. Toutefois, il existe des situations auxquelles ils ne sont pas adaptés. Dans le domaine du diagnostic *in vivo*, comme en thérapie, les anticorps monoclonaux murins seront rapidement reconnus comme substance étrangère par le système immunitaire du patient. La présence du fragment Fc, qui n'a pas de rôle dans la reconnaissance de l'antigène et dont les fonctions effectrices ne sont pas efficaces chez l'homme, constitue une gêne. Les fragments issus de la protéolyse ménagée des anticorps (Fab et F(ab)'<sub>2</sub>), tout comme les IgG (150 kDa) comprennent toujours des domaines constants inutiles à la reconnaissance de l'antigène. Leur masse moléculaire reste élevée, limitant ainsi leur capacité à diffuser dans les tissus. Enfin, les qualités fonctionnelles de ces anticorps restent limitées par la réponse immune de l'animal qui les a produits et ne peuvent être modifiées, altérées ou renforcées.

## Les anticorps recombinants : une multitude de formats

Dès le début des années 1990, le développement des techniques de clonage et d'amplification de gènes par la réaction de polymérisation en chaîne (PCR) de l'ADN a donné une impulsion nouvelle conduisant à l'identification des gènes d'anticorps et à leur expression par des organismes génétiquement modifiés (figure 2). Ces technologies ont rapidement abouti à une grande variété de fragments susceptibles d'être adaptés à de nouvelles applications (figure 3) (Hudson and Souriau, 2003).

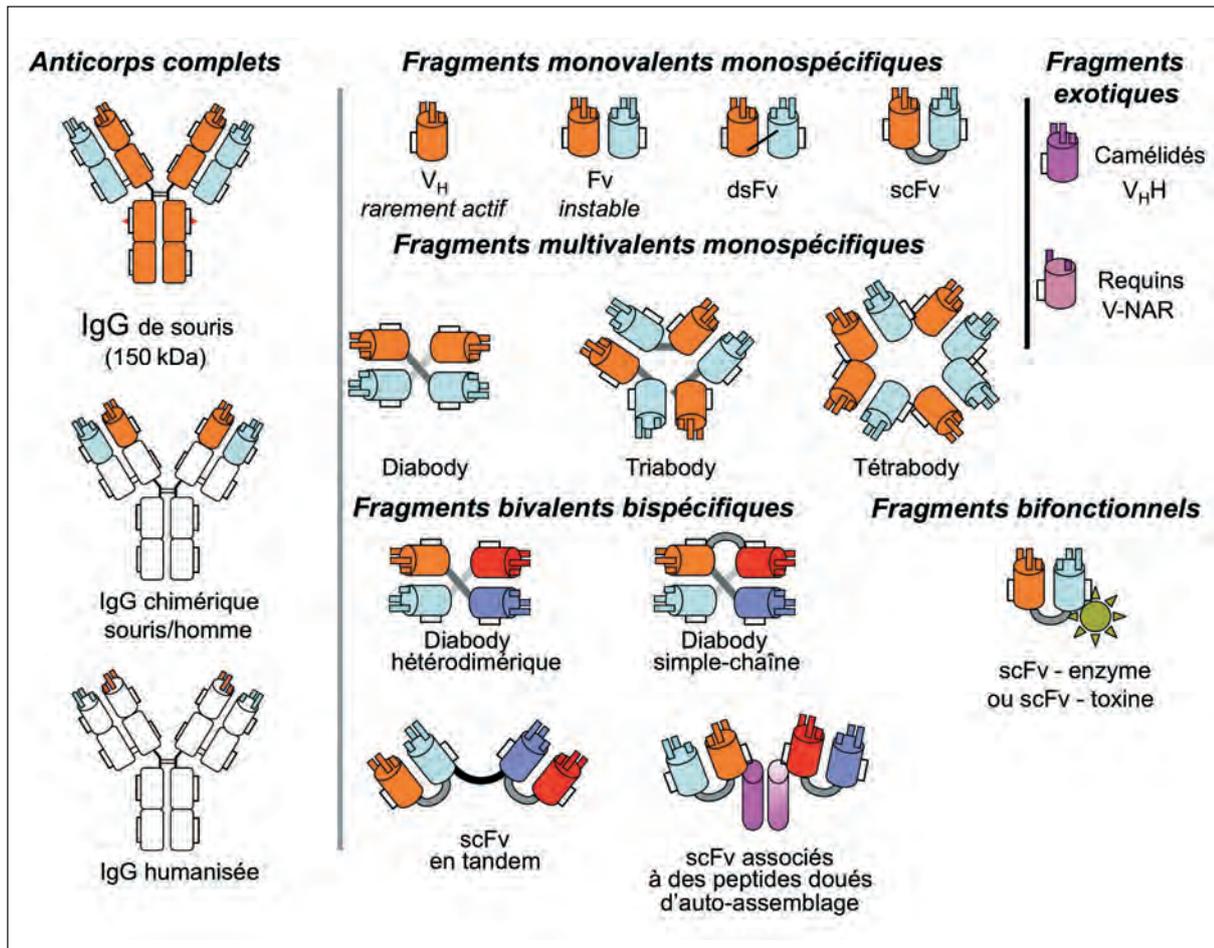
**Figure 2 : Obtention de fragments d'anticorps recombinants à partir d'hybridomes.**

Après avoir isolé l'ARNm des cellules hybridome, l'ADNc codant les domaines variables d'anticorps est amplifié par PCR, cloné et séquencé. Un gène codant pour un scFv constitué des domaines variables de l'anticorps associés par un lien peptidique L est synthétisé et cloné dans un vecteur d'expression. Des bactéries recombinantes sont utilisées pour produire le fragment d'anticorps scFv.



**Figure 3 : Les anticorps recombinants: des structures inédites pour de nouvelles applications.**

L'ingénierie moléculaire permet de créer des anticorps chimériques dans lesquels les domaines constants d'origine murine ont été remplacés par des domaines d'anticorps humains. Dans les anticorps humanisés, seules les boucles hypervariables restent d'origine murine. De multiples formats constitués uniquement des domaines variables de l'anticorps peuvent également être inventés. Enfin, chez certaines espèces animales (camélidés et requins) le site de reconnaissance de l'antigène n'est porté que par un domaine d'anticorps.



## De simples fragments d'anticorps monovalents

Les fragments variables (Fv) monovalents sont constitués exclusivement des domaines  $V_H$  et  $V_L$  d'un anticorps qui, en s'associant l'un à l'autre, forment la plus petite unité fonctionnelle capable de lier un antigène. Toutefois, ces Fv n'ont pratiquement pas connu d'application en raison de leur instabilité due à l'absence de lien covalent entre les deux domaines d'anticorps. Dès lors, de nouvelles structures d'anticorps monovalents ont été inventées. Tout d'abord, les dsFv dans lesquels un résidu de Cys est introduit par mutagenèse dirigée dans chacun des domaines variables afin d'induire un pont disulfure inter-domaine, sans altérer la spécificité de reconnaissance antigénique. Aujourd'hui, beaucoup plus répandus sont les fragments variables simple-chaîne (scFv) dans lesquels les domaines variables d'un anticorps sont produits par des organismes génétiquement modifiés sous la forme d'une protéine de fusion dans laquelle ils sont associés l'un à l'autre par un lien peptidique. Dans cette structure, l'orientation des domaines ( $V_H/V_L$  ou  $V_L/V_H$ ) n'a que peu d'importance. Ce sont la séquence et la longueur du lien

peptidique (L) qui doivent être adaptées pour concilier stabilité et fonctionnalité de la structure. Un peptide de 15 résidus (Gly<sub>4</sub>Ser)<sub>3</sub> satisfait généralement à ces critères pour des scFv qui portent le VH à leur extrémité N-terminale. Ainsi, le scFv est le plus décrit des fragments d'anticorps recombinants monovalents. Il représente le module de base à partir duquel il devient possible de créer des formats d'anticorps mono- ou multivalents originaux, libres ou produits sous la forme de protéines de fusion adaptées à un large choix d'applications thérapeutiques ou diagnostiques. Il convient aussi de signaler l'existence de quelques fragments variables isolés VL ou VH, doués d'une capacité de liaison spécifique à un antigène. On a même décrit chez les camélidés et les requins des formes d'anticorps encore inconnus dont le site de liaison est exclusivement constitué d'un domaine dont la structure est proche de celle de VH d'anticorps humains. Ces nouveaux domaines VHH (camélidés) ou V-NAR (requins) sont de petite taille et très stables. Ils présentent de grandes similitudes structurales avec les VH humains et peuvent être produits par des bactéries ou des levures recombinantes avec un rendement élevé.

### **Des fragments d'anticorps multivalents mono-spécifiques**

Dans un scFv, la longueur et la flexibilité du lien peptidique sont deux facteurs importants pour que l'association des domaines d'anticorps conduise à une structure fonctionnelle. Ainsi, remplacer le pentadécapeptide (Gly<sub>4</sub>Ser)<sub>3</sub> par son pentapeptide homologue (Gly<sub>4</sub>Ser) favorise la formation de dimères dans lesquels le domaine VH de chaque sous-unité s'associe au VL de l'autre sous-unité conduisant ainsi à des structures encore appelées « diabody », mono-spécifiques mais bivalentes, et dont la taille équivaut à celle d'un Fab (50 kDa). Réduire à un résidu, voire éliminer complètement le lien peptidique entre les deux domaines conduit à des trimères, trivalents (« triabody »), voire des tétramères, tétravalents (« tetrabody ») avec comme principale conséquence une avidité accrue pour l'antigène cible. L'analyse fonctionnelle de nombreux diabody dans des études pré-cliniques a montré un gain d'affinité apparente pour l'antigène, une meilleure rétention au niveau des tissus cibles, une plus faible clairance systémique.

### **Des fragments d'anticorps bispécifiques**

L'idée d'élaborer des anticorps bispécifiques capables de lier entre elles différentes cellules ou molécules avec des applications potentielles particulièrement dans le domaine de l'immunothérapie anticancéreuse date des années 1980. Les méthodes décrites à l'époque consistaient principalement à développer des hybrides d'hybridomes (« quadromes ») ou à coupler chimiquement deux Fab issus de la protéolyse ménagée d'anticorps monoclonaux différents, mais toutes ont connu un échec. L'ingénierie moléculaire a permis de contourner les principales difficultés. Elle donne la possibilité de créer des fragments d'anticorps bispécifiques constitués uniquement des domaines variables de deux anticorps différents. Plusieurs agencements sont possibles : des diabody hétérodimériques dans lesquels l'une des sous-unités est constituée du VH d'un premier anticorps associé au VL du deuxième anticorps de spécificité différente. La seconde sous-unité est constituée du VH du deuxième anticorps associé au VL du premier. Une telle structure est très compacte et rigide. Beaucoup plus flexibles sont les scFv en tandem, structures dans lesquelles deux scFv sont associés par un court lien peptidique. D'autres techniques font appel à des

domaines moléculaires qui ont une propension à l'auto-assemblage. Ainsi des scFv peuvent être produits en fusion avec des hélices  $\alpha$  amphipathiques, des structures de type « coil-coil » ou des domaines constants d'anticorps qui, en s'associant à leur partenaire, vont conduire à des molécules bivalentes.

## **Les fragments d'anticorps bifonctionnels ou anticorps fusionnés**

La plupart des tests d'immunoanalyse pratiqués à l'heure actuelle font appel à des anticorps préalablement couplés par voie chimique à une molécule traceuse (une protéine, une enzyme). Les réactions chimiques mises en oeuvre font intervenir des agents réticulant non-spécifiques qui conduisent à des immunoconjugés de structure hétérogène qui ont tendance à s'agréger et dans lesquels l'un ou l'autre des deux partenaires voit parfois ses propriétés fonctionnelles altérées. C'est ainsi que la luciférase n'a jamais pu être couplée de façon efficace à un anticorps alors que ses propriétés de luminescence sont régulièrement utilisées dans les techniques de gène rapporteur. C'est également l'une des raisons de l'échec des immunotoxines à visée thérapeutique. Aujourd'hui, le couplage génétique d'un fragment d'anticorps de type scFv à une protéine traceuse est possible et offre de nombreux avantages : le couplage est total, équimolaire et homogène. De cette homogénéité de structure découlent des procédés de purification simples et rapides. La technologie est reproductible et enfin, les propriétés fonctionnelles de chacune des deux entités peuvent ultérieurement être améliorées par mutagenèse.

## **Les principales sources d'anticorps recombinants : hybridomes, banques combinatoires**

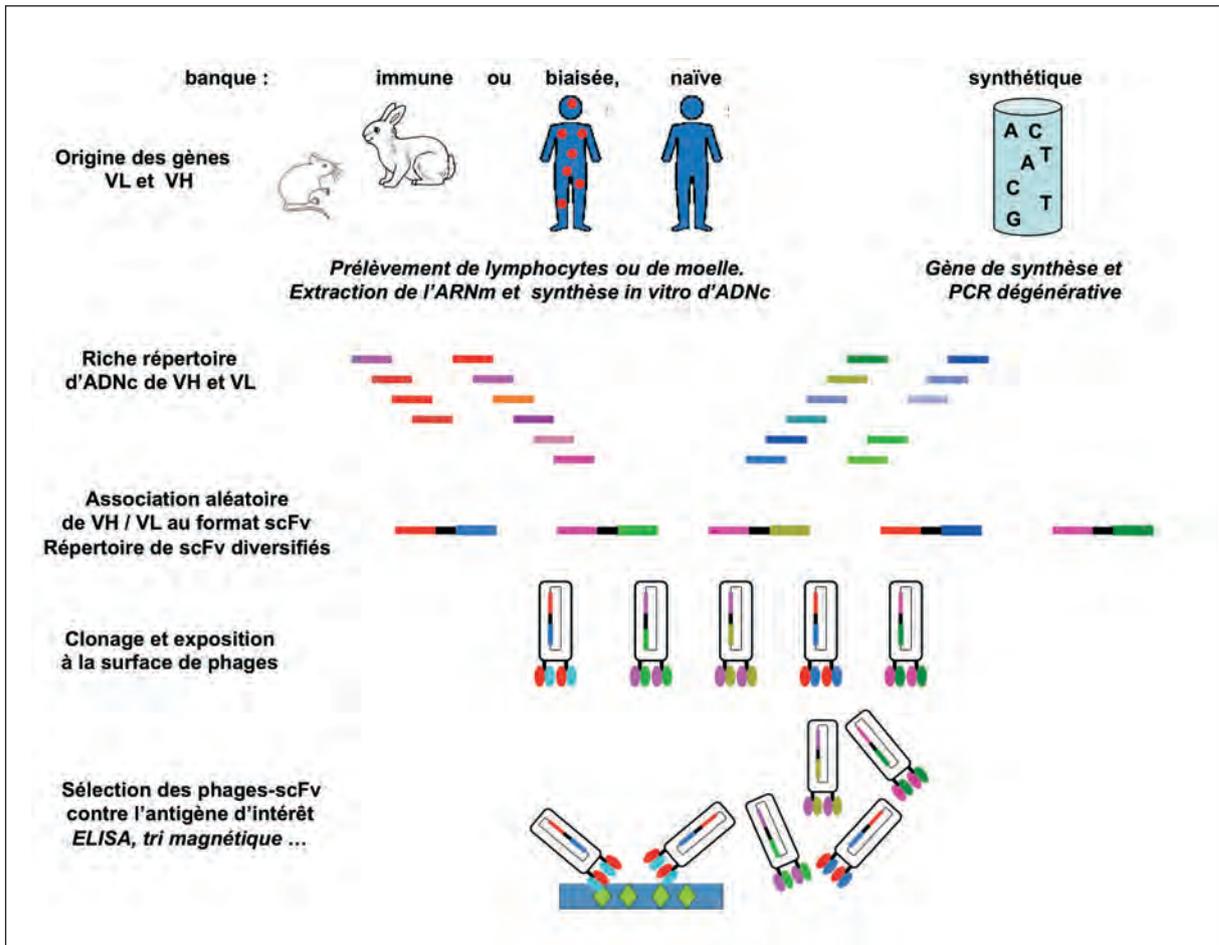
Il existe principalement deux stratégies pour obtenir des anticorps recombinants. La première consiste à utiliser comme matériel de départ l'hybridome qui sécrète un anticorps monoclonal pré-défini et dont on va cloner l'ADNc afin de produire cet anticorps sous un nouveau format mieux adapté à des applications particulières que l'anticorps naturel ou ses fragments fonctionnels issus de la protéolyse ménagée (figure 2). Même si cette approche connaît un grand succès, elle reste limitée à des anticorps d'origine murine pour lesquels il a fallu cloner l'hybridome sécréteur (Weisser and Hall, 2009). Couplée à des technologies d'évolution dirigée, elle permet de moduler l'affinité et la spécificité antigénique, de "chimériser" ou d'"humaniser" l'anticorps. La plupart des anticorps thérapeutiques sur le marché sont issus de ces technologies qui restent toutefois souvent fastidieuses et difficiles à mettre en oeuvre.

Une autre technologie, maintenant très répandue, consiste à créer une banque combinatoire de fragments d'anticorps exposés à la surface de bactériophages puis à sélectionner les phages qui expriment le fragment d'anticorps spécifique d'un antigène d'intérêt (figure 4). Cette technique, à l'origine mise au point pour la présentation et la sélection de peptides a été adaptée aux fragments d'anticorps recombinants (Bradbury and Marks, 2004). Elle consiste dans un premier temps à créer un répertoire de fragments d'anticorps, le plus souvent des scFv dans lesquels les domaines variables sont associés dans des combinaisons aléatoires, un VH pouvant être marié à n'importe

quel VL et *vice et versa*. On recrée ainsi la diversité du répertoire naturel. Selon l'origine des domaines variables qui vont constituer les fragments d'anticorps on distingue trois types de répertoires (banques). Les banques dites "naïves" sont construites à partir d'ADNc codant les domaines variables d'anticorps de donneurs non-immunisés. Outre le fait que ces répertoires peuvent être constitués de fragments d'anticorps aussi bien d'origine humaine qu'animale, leur intérêt est d'offrir une très grande complexité capable d'atteindre  $10^{11}$  combinaisons. Les répertoires « biaisés » sont construits à partir d'ADNc codant les domaines variables d'anticorps de sujets préalablement immunisés chez lesquels les gènes d'immunoglobulines spécifiques sont sur-représentés. Ces répertoires sont souvent d'origine animale; ils peuvent également provenir de sujets atteints de pathologie ou qui produisent des auto-anticorps. Enfin, les répertoires synthétiques ou semi-synthétiques sont construits à partir de gènes de synthèse dans lesquels la diversité des régions charpentes est réduite.

**Figure 4 : La technologie de l'exposition de fragments d'anticorps à la surface de phages.**

Les couples VH/VL sont constitués de façon aléatoire. Ils proviennent d'animaux immunisés, d'hommes sains ou atteints d'une pathologie, voire de la synthèse chimique. Le répertoire de scFv est cloné dans des phages ou des phagémides et sera propagé par des bactéries. Les phages qui présentent à leur surface un scFv affin pour l'antigène d'intérêt seront sélectionnés par criblage (ELISA ...). L'ADNc codant le scFv en sera extrait, puis cloné dans des bactéries non-suppressives pour produire le scFv libre.



Le répertoire de scFv ainsi constitué est cloné dans des phages ou des phagémides en phase de lecture avec le gène de l'une des protéines d'enveloppe du phage. Une banque de phages exposant chacun à leur surface un type de scFv est alors constituée et propagée dans des bactéries suppressives. Les phages produits par les bactéries contiennent le phagémide qui code pour le fragment d'anticorps exposé à la surface de la protéine d'enveloppe. Le répertoire de bactériophages ainsi produits est ensuite criblé contre l'antigène d'intérêt habituellement immobilisé sur un support plastique, comme dans une réaction ELISA. Les bactériophages présentant à leur surface un fragment d'anticorps capable d'interagir avec l'antigène sont retenus. Ils sont élués après lavage de la plaque et les clones ainsi sélectionnés sont récupérés en infectant de nouvelles bactéries suppressives. Cette étape, également appelée « biopanning », répétée quatre à cinq fois, permet d'enrichir la population en bactériophages porteurs de fragments d'anticorps affins pour l'antigène cible ; elle mime la réponse immune naturelle. Une fois le criblage achevé, les phagémides contenus dans les bactériophages sélectionnés sont clonés dans des bactéries non-suppressives pour produire le fragment d'anticorps d'intérêt libre. C'est la présence d'un codon « Ambre » à l'extrémité du gène du fragment d'anticorps qui permet cette manipulation.

Les avantages de la technologie des phages-anticorps sont multiples. Cette technologie offre la possibilité d'obtenir des anticorps d'origine animale mais aussi humaine sans nécessairement requérir une immunisation préalable. La diversité des cibles antigéniques est alors théoriquement infinie.

## **Les anticorps recombinants : de futurs outils pour la recherche et le diagnostic**

C'est au secteur thérapeutique que la découverte des anticorps recombinants a jusqu'à présent le plus profité avec, depuis 1994, la mise sur le marché d'une vingtaine d'anticorps, la plupart étant des immunoglobulines complètes, chimériques ou humanisées. Dans les secteurs de l'immunoanalyse et du diagnostic l'utilisation des anticorps recombinants n'a pas encore connue la même croissance. L'origine hétérologue des anticorps monoclonaux murins ne pose aucun problème pour des applications *in vitro*. Leur spécificité, la possibilité de les produire toujours à l'identique et en grande quantité, de les fixer sur des supports solides ou de les coupler à des traceurs radiomarqués, fluorescents ou enzymatiques ont conduit à un large éventail de tests qui autorisent le dosage *in vitro* d'une grande variété de substances. Ces tests rapides, souvent très sensibles, reproductibles et automatisables sont aujourd'hui utilisés en routine en biologie clinique pour le dosage d'hormones, la recherche de marqueurs tumoraux, d'agents infectieux et de toxiques. Ils sont également exploités pour le diagnostic vétérinaire, le contrôle de qualité de l'environnement, en agroalimentaire et dans bien d'autres secteurs. Toutefois des améliorations restent possibles. Grâce à la technologie des phages-anticorps qui permet d'obtenir des scFv ou des Fab spécifiques de n'importe quel antigène, sans dépendre de la réponse immune d'animaux, les champs d'application de l'immuno-analyse pourraient être élargis. Les difficultés habituellement rencontrées pour obtenir un couplage homogène anticorps/protéine traceuse par les méthodes chimiques conventionnelles pourraient être contournées en adoptant la technologie de fusion de

gènes. Le gène d'un fragment d'anticorps fonctionnel peut être fusionné en phase de lecture avec celui d'une enzyme, cloné dans un vecteur d'expression et introduit dans une cellule hôte qui produira la protéine hybride. On peut recourir ici à des bactéries, des levures, des cellules d'insectes ou de mammifères.

La sur-expression de la protéine de fusion dans des bactéries est rapide et certainement, d'un point de vue pratique et économique, la plus intéressante. La sécrétion dans le périplasme bactérien conduit à des protéines de fusion solubles, correctement repliées et bi-fonctionnelles avec un site de reconnaissance d'antigène spécifique et une enzyme colorimétrique active. Plusieurs exemples de fragments d'anticorps, scFv ou Fab dirigés contre des antigènes d'intérêt clinique (toxines, hormones) ont été décrits, exprimés en fusion avec des enzymes (phosphatase alcaline, acétylcholinestérase), des peptides affins pour la streptavidine ou des protéines fluorescentes (GFP). La sensibilité des tests d'immunodosage pratiqués avec ces protéines de fusion atteint volontiers quelques picomoles. Cette sensibilité peut certainement être encore augmentée et atteindre celle des radioimmunodosages (RIA) les plus sensibles, après avoir amélioré d'une part, la spécificité et l'affinité des anticorps et d'autre part, les propriétés physiques et catalytiques des enzymes. En combinant des étapes de sélection *in vitro* (mutagenèse dirigée ou aléatoire) ou *in vivo*, en utilisant des souches bactériennes provoquant des mutations, l'affinité de certains anticorps a pu être augmentée d'un facteur 1000. Un anticorps monoclonal très affiné pour la progestérone mais également doué de réactivité croisée avec des analogues structuraux stéroïdiens a été reformaté par ces techniques qui ont conduit à le rendre strictement spécifique de la progestérone, sans perte d'affinité, lui conférant ainsi un fort potentiel pour développer un kit de dosage rapide, sensible et spécifique (Ducancel and Dubreuil, 2006).

L'ingénierie des anticorps offre également la perspective de développer des biocapteurs spécifiques capables de reconnaître un antigène dans un milieu hétérogène tel que les liquides biologiques et de transduire instantanément l'événement en un signal directement mesurable. C'est ainsi que Bedouelle *et al* (2007) ont établi un procédé de transformation de fragments d'anticorps recombinants en biocapteurs fluorescents. Ils ont d'abord identifié les résidus d'un scFv au voisinage du site de fixation de l'antigène susceptibles d'être mutés en cystéine sans affecter la fonction de reconnaissance antigénique. La cystéine introduite dans les mutants répondant à ces critères a ensuite été couplée à un fluorophore et les mutants pour lesquels la fixation de l'antigène provoque un changement de l'environnement structural du fluorophore ont été sélectionnés. Ce changement structural induit immédiatement une augmentation de l'intensité de fluorescence proportionnelle à la concentration en antigène. Le procédé a été breveté et pourrait s'appliquer à la détection, au dosage et au suivi de nombreux antigènes, hormones, marqueurs ou agents infectieux dans les fluides biologiques en utilisant le biocapteur libre en solution ou immobilisé sur une puce à protéines.

Les anticorps utilisés dans le secteur du diagnostic *in vivo* restent aujourd'hui encore des anticorps murins, même s'ils sont peu nombreux. A l'origine, ce sont leur grande spécificité associée à la possibilité de les conjuguer à des émetteurs  $\gamma$  qui ont stimulé leur emploi en imagerie médicale pour cibler *in vivo* des antigènes spécifiques associés à des tumeurs, une caméra gamma étant utilisée pour détecter les sites auxquels se concentrent les sondes radio-marquées. Une multitude

d'anticorps ont été développés pour l'immunoscintigraphie, mais moins d'une dizaine ont connu un développement commercial avec des applications limitées au suivi de métastases de cancers colorectaux, des ovaires ou de la prostate et à l'évaluation des nécroses myocardiques consécutives à un infarctus. Ce n'est pas l'immunogénicité potentielle de ces molécules qui en a limité l'exploitation mais plutôt le faible contraste généralement observé entre tissus sains et pathologiques. La taille des sondes, qu'il s'agisse des anticorps entiers ou de leurs fragments Fab, en limite la diffusion et la capacité à atteindre un antigène cible enfoui dans un tissu ou une tumeur. Leur demi-vie plasmatique est relativement élevée, y compris pour les fragments Fab qui peuvent être filtrés par le rein de sorte que le rapport signal spécifique/bruit de fond soit faible (Wu and Senter, 2005). Des études récentes menées sur le modèle animal ont montré les avantages potentiels des fragments d'anticorps recombinants beaucoup plus diversifiés en taille et en valence. Ainsi, des formats d'anticorps bivalents de type diabody sont rapidement éliminés du compartiment plasmatique et diffusent mieux vers les tissus cibles. Le rapport signal/bruit s'en trouve augmenté et les images observées plus contrastées. La sensibilité des tests pourrait être améliorée en utilisant les techniques de tomographie par émission de positons (TEP) ou encore des anticorps couplés à des réactifs fluorescents. Toutefois, aucune molécule n'a encore été mise sur le marché, peut-être en raison des coûts de développement et des habitudes qui font que l'immunoscintigraphie n'est jamais utilisée en première ligne mais seulement dans le cas de récidives et quand les résultats des examens classiques (scanner, échographie) sont douteux.

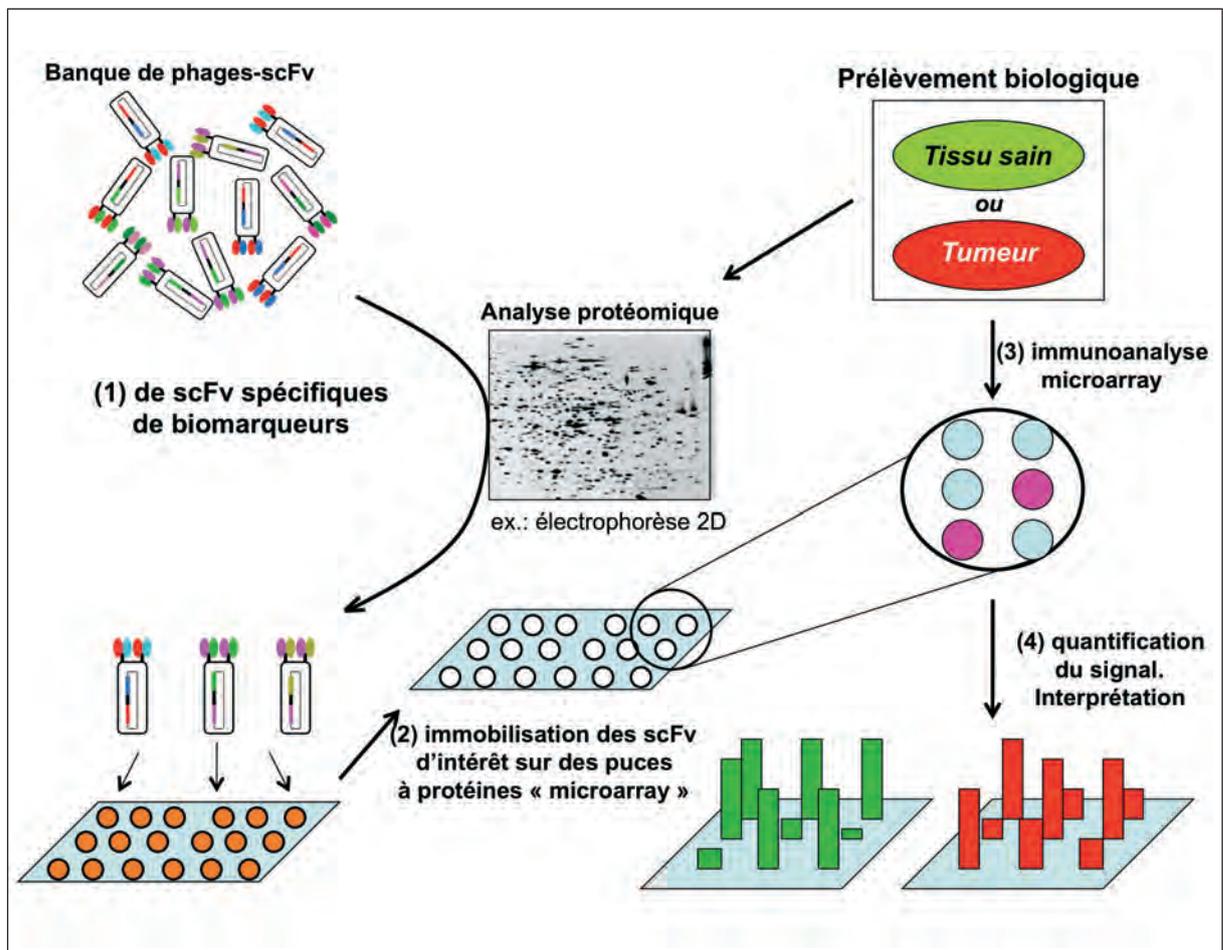
Un autre domaine dans lequel les anticorps recombinants nourrissent de forts espoirs est celui de l'étude protéomique d'échantillons biologiques (voir chapitres précédents). Ces espoirs sont nourris d'une part par l'immense diversité des anticorps que véhiculent les banques combinatoires et les possibilités offertes par les technologies de « microarray » qui permettent de fixer des protéines ou des anticorps sur des nanosurfaces de verre siliciné et de développer des analyses multiplexes rapides, sensibles et robotisables (Wingren and Borrebaeck, 2006). Le criblage de banques combinatoires non-immunes a d'ores et déjà démontré la possibilité de cloner des fragments d'anticorps dirigés contre des épitopes conservés ou des antigènes non-immunogéniques qui n'avaient jamais été obtenus par les méthodes d'hybridomes conventionnelles. Des anticorps dirigés contre la protéine BiP du réticulum endoplasmique des cellules B ou encore des anticorps capables de distinguer les formes activées post-traductionnellement de protéines intracellulaires comme Rab6 ou GTP-tubuline en sont des exemples significatifs (Dimitrov *et al.*, 2008). Le criblage de banques d'anticorps contre des protéines extraites d'un tissu sain ou pathologique par électrophorèse bidimensionnelle et caractérisées par spectrométrie de masse pourrait conduire à la sélection d'anticorps d'intérêt (figure 5). Ces anticorps pourraient être immobilisés à des positions précises sur des nano-surfaces, capturer les molécules d'intérêt contenues dans un prélèvement biologique et, après lavage, divers systèmes de révélation des immuncomplexes pourraient être mis en place, la position du signal sur la puce et son intensité indiquant le résultat pour chacun des paramètres analysés. Plusieurs systèmes de détection sont actuellement évalués parmi lesquels des nano-électrodes, des réactions sandwich faisant intervenir un second anticorps fluorescent, la résonance plasmonique de surface (RPS) ou la spectrométrie de masse MALDI-TOF. L'analyse

différentielle de prélèvements sains et pathologiques apporterait une valeur ajoutée dans le diagnostic précoce, le pronostic ou le suivi thérapeutique de pathologies complexes, ou l'établissement de critères indiquant une probabilité plus ou moins grande de la présence d'une pathologie ou de facteurs de prédisposition.

Ces objectifs pourraient également être atteints en utilisant des variantes technologiques, par exemple des approches micro-fluidiques dans lesquelles les molécules naturelles en solution sont capturées par des anticorps couplés à des billes colorées puis entraînées dans des micro-capillaires où d'autres anticorps spécifiques ont été immobilisés. Ces derniers retiendront les immuns complexes colorés à une position pré-établie autorisant ainsi leur détection (voir Chapitre 7).

**Figure 5 : Application potentielle des anticorps recombinants à l'analyse protéomique de tissus.**

Des fragments d'anticorps spécifiques d'antigènes de tissus sains ou de tumeurs sont identifiés par criblage d'une banque de scFvs (1). Les scFv d'intérêt sont ensuite immobilisés sur une puce à protéines (2) qui va servir à analyser des prélèvements biologiques (3). Le profil de reconnaissance antigénique est interprété comme une signature d'un tissu sain ou pathologique (4). Adapté de Borrebaeck (2000).



# Anticorps et immuno-analyse au fil du temps.

---

**1960** : Premiers dosages radioimmunologiques (RIA) d'hormones circulantes. Le principe consiste à précipiter, à l'aide d'anticorps polyclonaux, l'hormone plasmatique mise en compétition avec une faible concentration d'hormone radiomarquée. Ces tests sont sensibles mais longs (24 heures).

**1971** : Premiers immunodosages sandwich dans lesquels l'antigène cible est capturé par un anticorps immobilisé sur une surface plastique. Les immunocomplexes formés sont détectés par un anticorps secondaire couplé chimiquement à une enzyme (peroxydase, phosphatase alcaline,  $\beta$ -galactosidase).

**1975** : La technologie des hybridomes va révolutionner le secteur de l'immunoanalyse. Il devient possible d'immortaliser une cellule de rongeurs, sécrétrice d'un anticorps monoclonal, unique. Des milliers d'anticorps monoclonaux murins sont ainsi produits contre des antigènes. Simultanément, de nouveaux supports sont mis au point pour immobiliser les anticorps (plaques de microtitrage, billes). Les dosages se simplifient, deviennent plus rapides et autorisent l'analyse simultanée de plusieurs échantillons biologiques.

**1980-1990** : Les méthodes de couplage des anticorps à des traceurs non-isotopiques se généralisent et deviennent plus efficaces. De nouveaux traceurs luminescents ou fluorescents amplifient les signaux et contribuent à accroître la sensibilité des tests. Ils remplacent rapidement les isotopes. Dans les années 1990, l'automatisation de l'ensemble de ces techniques d'immunoanalyse se généralise.

**Début 2000** : De nouveaux formats d'anticorps mono- ou multivalents, mono- ou bispécifiques sont créés par ingénierie moléculaire. Des méthodes d'évolution dirigée permettent de moduler la spécificité de ces anticorps et de l'adapter à des immunodosages hautement spécifiques. La technologie des phages-anticorps rend possible la sélection *in vitro* d'anticorps spécifiques contre n'importe quel antigène sans même avoir recours à l'immunisation préalable d'animaux.

# Conclusion et perspectives

---

Les possibilités offertes par l'ingénierie moléculaire d'adapter le format et la valence des anticorps recombinants à des applications spécifiques, d'accroître leur affinité et de moduler leur spécificité vont considérablement élargir l'éventail de leurs applications. Tout d'abord comme agent thérapeutique : une vingtaine de molécules ont reçu une autorisation de mise sur le marché dans des domaines aussi variés que l'infectiologie, l'oncologie, l'hématologie, le traitement de maladies cardiovasculaires, inflammatoires ou auto-immunes. Près de 500 seraient actuellement en phase d'essais cliniques.

Dans le secteur de l'immunoanalyse, même si les anticorps monoclonaux conventionnels restent des outils incontournables, le panel de leurs cibles est limité à quelques milliers de protéines. Ce secteur devrait donc rapidement profiter des progrès technologiques et voir s'élargir ses champs d'application. Moduler les propriétés intrinsèques de l'anticorps (taille, affinité, spécificité) est aujourd'hui possible même si les méthodes mises en œuvre sont encore difficilement généralisables. La préparation d'immunoconjugués bifonctionnels recombinants, homogènes et parfaitement adaptés aux exigences de l'immunoanalyse est une réalité. Les recherches dans ce secteur sont intenses et des conjugués immunoenzymatiques recombinants vont rapidement apparaître sur le marché du diagnostic. Grâce aux technologies de sélection *in vitro* à partir de banques combinatoires, des sondes spécifiques d'antigènes variés peuvent être isolées, sans même immuniser des animaux. Ainsi, le répertoire des anticorps disponibles s'accroît considérablement, quelles que soient les protéines cibles, y compris des protéines peu immunogènes. La microfluidique et la possibilité de créer des puces à anticorps recombinants qui autorisent la détection et le dosage rapide, sensible et simultané de chacun des antigènes contenus dans un prélèvement biologique ouvrent de nouveaux horizons. De véritables laboratoires sur puces pourraient ainsi être constitués pour établir le profil d'expression protéique de cellules saines ou cancéreuses, découvrir de nouveaux biomarqueurs, comprendre les mécanismes moléculaires de pathologies ou encore identifier de nouvelles cibles thérapeutiques.

## Bibliographie

---

Bedouelle H., Renard M., Belkadi L., England P. Biosensors, process for obtaining them and their applications. Brevet n°EP1742055, 2007.

Borrebaeck C.A.K. Antibodies in diagnostics. From immunoassays to protein chips. *Immunology Today* 2000 21 : 379-381.

Bradbury A.R.M. and Marks. J. D. Antibodies from phage antibody libraries. *J. Immunol. Methods* 2004, 290 : 29-49.

Dimitrov A., Quesnoit M., Moutel S., Cantaloube I., Poüs C., Perez F. Detection of GTP-tubulin conformation *in vivo* reveals a role for GTP remnants in microtubule rescue. *Science* 2008, 322 : 1353-6.

Ducancel F. and Dubreuil O. Method for preparing specific affinity anti-progesterone antibodies, the obtained antibodies and the use thereof for *in vitro* diagnosis. Brevet n° WO2006048576, 2004.

Hudson P.J. and Souriau C. Engineered antibodies. *Nature Medicine* 2003, 9 : 129-134.

Laffly E. and Sodoyer R. Monoclonal and recombinant antibodies, 30 years after ... *Human Antibodies* 2005, 14 : 33-55.

Weisser N.E. and Hall J.C. Applications of single-chain variable fragment antibodies in therapeutics and diagnostics. *Biotechnology Advances* 2009, 27: 502-520.

Wingren C. and Borrebaeck C.A.K. Antibody microarray : Current status and key technological advances. *OMICS* 2006, 10 : 411-27.

Wu A.M. and Senter P.D. Arming antibodies : prospects and challenges for immunoconjugates. *Nature Biotechnology* 2005, 23 : 1137-46.



ISSN : 1293-2892  
ISBN : 2-913633-58-7  
SOUS-TITRE  
19, avenue d'Italie 75013 Paris  
Dépôt légal : Décembre 2010

